JCTC Journal of Chemical Theory and Computation

Introduction: Coarse-Graining in Molecular Modeling and Simulation

This partial special issue of the *Journal of Chemical Theory and Computation* is devoted to the topic of coarse-graining in molecular modeling and simulation. To my knowledge, it is the first time that coarse-graining concepts and techniques from the application areas of biomolecular, materials, and liquid-state systems have been published together in the same place. Not every current researcher working on this particular problem is represented in our partial special issue, but it contains an excellent cross-section of much of the important work being undertaken at the present time.

The impetus for coarse-grained molecular modeling and simulation primarily derives from the need to bridge the atomistic and mesoscopic scales. Typically speaking, there are 2−3 orders-of-magnitude in length and time separating these scales. At the mesoscopic scale, one sees the emergence of critically important phenomena (for example, complex self-assembly in biomolecular or materials systems). Modern molecular simulations, especially as they seek to make increasing contact with experimental results on complex systems, can and will play a crucial role in the exploration of mesoscopic phenomena and, in turn, the behavior of real biomolecular and materials systems. Coarse-graining promises to be a large step forward for the molecular simulation community overall, provided the many challenges faced by the technique can be overcome.

What then are the major challenges of coarse-graining? In many ways the answer to this question depends on the questions one wishes to ask. For example, one common and often legitimate approach is to develop reductionist or "toy" coarse-grained (CG) models that reveal the essential physics of a given class of systems. These models will of course not be quantitative, but they are usually very computationally efficient and informative. Another common approach is to develop CG models using experimental, thermodynamic, and/or average structural properties. A third approach is to bridge atomistic information upward in scale to the CG level in a "multiscale" fashion. All of these approaches have their strengths and weaknesses, and they are certainly complementary to each another.

An absolutely key question for all coarse-graining methodologies is the degree of transferability of the resulting CG models between various systems and from one set of thermodynamic conditions to another. By all rights, a CG model cannot be completely transferable because it is a reduced description of a complex system and some amount of information has been effectively averaged out in one way or another for the given conditions. On the other hand, some aspects of the CG model will certainly be transferable. The key goal is to understand what is and what is not transferable in the CG model and why it is that way. This is not merely a technical issue but instead a very deep problem rooted at the foundations of statistical mechanics. It is a problem that has not yet been solved, despite the many beautiful papers published in this issue and elsewhere. There is also the issue of CG dynamics (i.e., time-dependent behavior), which are not the same as real molecular dynamics. How different are the CG dynamics and why? Can a connection to real dynamics be made? Overall, how much are we to believe from these CG models anyway and how much can they be used to predict and explain unknown phenomena? As I am sure all of the authors in this partial special issue will agree, it is very exciting to be a part of this emerging conceptual challenge and to be able to make a contribution to finding the answer to these critical questions. A satisfactory solution will provide no less than a revolutionary step forward for the field of molecular modeling and simulation.

Gregory A. Voth
Guest Editor
Distinguished Professor of Chemistry
University of Utah

CT600124A

10.1021/ct600124a CCC: $33.50
Published on Web 04/15/2006

JCTC Journal of Chemical Theory and Computation

A New Method for Coarse-Grained Elastic Normal-Mode Analysis

Mingyang Lu,[†] Billy Poon,[‡] and Jianpeng Ma*[,†,‡]

Department of Biochemistry and Molecular Biology, Baylor College of Medicine, One Baylor Plaza, BCM-125, Houston, Texas 77030, and Department of Bioengineering, Rice University, Houston, Texas 77005

Received December 4, 2005

Abstract: In this paper, we report a new method for coarse-grained elastic normal-mode analysis. The purpose is to overcome a long-standing problem in the conventional analysis called the tip effect that makes the motional patterns (eigenvectors) of some low-frequency modes irrational. The new method retains the merits of a conventional method such as not requiring lengthy initial energy minimization, which always distorts structures, and also delivers substantially more accurate low-frequency modes with no tip effect for proteins of any size. This improvement of modes is crucial for certain types of applications such as structural refinement or normal-mode-based sampling.

I. Introduction

In biomolecular simulations, normal-mode analysis (NMA) has been playing a very important role in analyzing structural fluctuations around a well-defined molecular conformation. In recent years, elastic normal-mode analysis (eNMA) has been popularized for its simplicity and wide-range applicability.[1−20] The potential function[1] used in the conventional eNMA[5] is a highly coarse-grained elastic network that has a form of

$$V = (\gamma/2)\sum_i\sum_j \sigma_{ij}(|\mathbf{r}_{ij}| - |\mathbf{r}_{ij}^{\,0}|)^2, \quad \sigma_{ij} = \begin{cases} 1 & |\mathbf{r}_{ij}^{\,0}| \le r_c \\ 0 & |\mathbf{r}_{ij}^{\,0}| > r_c \end{cases} \quad (1)$$

where $|\mathbf{r}_{ij}|$ and $|\mathbf{r}_{ij}^{\,0}|$ are the instantaneous and equilibrium values (or initial values from the coordinates) of pairwise distance between the ith and the jth Cα atoms, respectively. The value of Heavidide step function σ_{ij} specifies the effect of cutoff. The absolute value of force constant γ is irrelevant to the calculation and is often set to 1.0. This elastic potential can be applied to either representative atoms from atomic structures such as Cα traces[4,5] or points placed in continuous density maps from low-resolution experiments such as cryogenic electron microscopy (cryo-EM).[6,8] Numerous studies have shown that eNMA is effective in extracting patterns of low-frequency normal modes (for reviews, see refs 21−23). The success of it hinges on a very important fact that, for compact biomolecular structures such as those of globular proteins, the patterns of low-frequency modes, often referred to as deformational modes, are only sensitive to the shape of the molecules, rather than to the detailed atomic structures, which has been quantitatively demonstrated in our recent study.[24]

Another important advantage of coarse-grained eNMA is that it does not require the initial lengthy energy minimization because the minimum of potential function in eq 1 is the current structure. The initial minimization almost always significantly distorts the structure as in NMA based on molecular mechanics force fields such as CHARMM.[25,26]

Despite its enormous success and widespread popularity in application, the current version of eNMA[5] has an inherent weakness, which we refer to as the "tip effect". In systems with structural components, the "tips", protruding out of the main body, e.g., an isolated surface loop or simply a thinner region in the density map; the tip effect can lead to pathological behavior in motions of points near those regions, presumably due to an imbalance of elastic forces among neighboring harmonic oscillators due to lighter packing

* Corresponding author e-mail: jpma@bcm.tmc.edu. Corresponding author mailing address: One Baylor Plaza, BCM-125, Baylor College of Medicine, Houston, TX 77030.

[†] Baylor College of Medicine.

[‡] Rice University.

10.1021/ct050307u CCC: $33.50
Published on Web 03/29/2006

around tip regions. In the modes tinted with the tip effect, the magnitudes of displacement of the points in eigenvectors at or around the tip regions are much larger than those of the rest of the system but with more or less normal-looking eigenvalues. Since the eigenvectors of normal modes are normalized, the abnormally large magnitudes of displacement in the tip regions make the rest of the system have much less movement and sometimes even static. To make matters worse, so far there has been no systematic method to predict which modes have tip effect, and they can occur in even very low-frequency modes. Although on a case-by-case basis, one might be able to bypass such an effect, to our best knowledge, a systematic way is unavailable for overcoming the tip effect. In practice, if one only wants very few lowest-frequency modes for functional interpretation, the tip effect may not be so severe since the modes picked for functional study may not have tip effects. But in cases when a set of continuous low-frequency modes are needed, such as in structural refinement, this could become a major issue because some of the modes in the chosen set will definitely have some degrees of tip effect.

In this paper, we report a modified eNMA for eradicating the tip effect. Our main strategy is to make the overly soft degrees of freedom around the tip regions stiffer. In our recent study,[24] we demonstrated that a Hessian matrix with randomized off-diagonal elements still gives correct low-frequency eigenvectors. The Hamiltonian used in the conventional eNMA and molecular mechanics such as CHARMM[25,26] can be viewed just as two special cases. Therefore, we hypothesize that somewhere between the two extreme cases, one must be able to find a new Hamiltonian in which the degrees of freedom around the tip regions are stiff enough to eradicate the tip effect and yet still retain the merit of not requiring the initial minimization. Such a Hamiltonian should still deliver correct eigenvectors for low-frequency modes.

The results in this paper show that it is indeed possible to find a Hamiltonian so that the normal modes calculated from it not only retain the merits of original eNMA but also do not contain the tip effect. This was true for single polypeptide proteins in their compact native states or extended denatured states. It was also true for multisubunit supramolecular complexes. The overall patterns of eigenvectors of modes for the new Hamiltonian, in comparison with those of conventional eNMA, were also found to be closer to the ones computed by standard molecular mechanics force fields such as CHARMM. We therefore hope that the new eNMA will be a complementary tool to the existing ones in analyzing protein dynamics, especially in certain applications such as structural refinement in which a continuous set of low-frequency modes would be needed.

II. Theory

Internal Coordinate System. The new eNMA is designed to work with Cα traces or a subset of Cα traces. To keep the length of the Cα−Cα pseudobond fixed, the internal coordinate (IC) system[27] was implemented (Figure 1). Although the formalism seems to be more complex than the original eNMA,[1,5] we found it is essential to work in an IC system because it effectively avoids the stretching of the bond length, which is one of the major sources of the tip effect in elastic network potential.

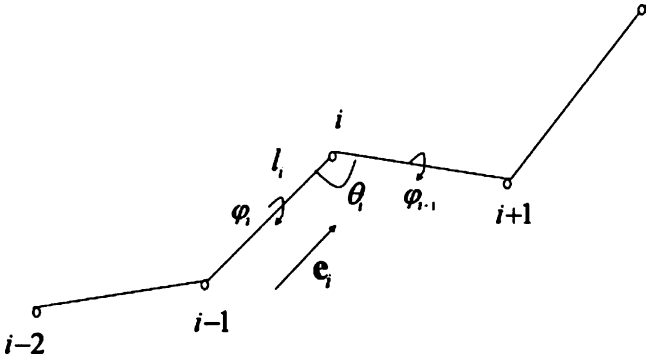


Figure 1. Internal coordinate system used in the new eNMA for eliminating tip effect.

For simplicity, we first describe the case with a single chain. The multichain case can be easily generalized. For a system with N Cα atoms ($i = 1,2,...,N$), the degrees of freedom in IC are $N - 2$ bond angles $\{\theta_i\}$ and $N - 3$ pseudo dihedral angles $\{\varphi_i\}$. Note, the bonds are indexed the same as the second Cα atom, the angles $\{\theta_i\}$ as the middle Cα atom, and the dihedrals as the third Cα atom. The relationship between IC and Cartesian coordinates (CC) is given in the following ways. From IC to CC

$$\mathbf{r}_1 = (0,0,0)^T; \quad \mathbf{e}_2 = (1,0,0)^T; \quad \mathbf{e}_3 = (\cos\theta_2, \sin\theta_2, 0)^T$$

$$\mathbf{e}_{i+1} = \cos\theta_i\,\mathbf{e}_i + \sin\theta_i\cos\varphi_i\frac{(\mathbf{e}_{i-1}\times\mathbf{e}_i)\times\mathbf{e}_i}{||(\mathbf{e}_{i-1}\times\mathbf{e}_i)\times\mathbf{e}_i||} + \sin\theta_i\sin\varphi_i\frac{\mathbf{e}_{i-1}\times\mathbf{e}_i}{||\mathbf{e}_{i-1}\times\mathbf{e}_i||}$$

$$\mathbf{r}_{i+1} = \mathbf{r}_i + l_{i+1}\mathbf{e}_{i+1} \tag{2}$$

where l_i is the ith Cα−Cα pseudo bond length, and $\mathbf{e}_i$ is the directional vector for that bond. The first Cα atom is at the origin, the first bond on the x-axis, and the first bond angle in the plane of XY. From CC to IC, we have

$$\begin{cases} l_i = ||\mathbf{r}_i - \mathbf{r}_{i-1}||; \quad \mathbf{e}_i = \dfrac{\mathbf{r}_i - \mathbf{r}_{i-1}}{l_i}; \quad \theta_i = \cos^{-1}(\mathbf{e}_i\cdot\mathbf{e}_{i+1}) \\ \varphi_i = \begin{cases} \cos^{-1}\left(\left(\dfrac{\mathbf{e}_{i-1}\times\mathbf{e}_i}{||\mathbf{e}_{i-1}\times\mathbf{e}_i||}\right)\cdot\left(\dfrac{\mathbf{e}_i\times\mathbf{e}_{i+1}}{||\mathbf{e}_i\times\mathbf{e}_{i+1}||}\right)\right), \\ \qquad\qquad if \quad \mathbf{e}_{i+1}\cdot(\mathbf{e}_{i-1}\times\mathbf{e}_i) \geq 0 \\ -\cos^{-1}\left(\left(\dfrac{\mathbf{e}_{i-1}\times\mathbf{e}_i}{||\mathbf{e}_{i-1}\times\mathbf{e}_i||}\right)\cdot\left(\dfrac{\mathbf{e}_i\times\mathbf{e}_{i+1}}{||\mathbf{e}_i\times\mathbf{e}_{i+1}||}\right)\right), \\ \qquad\qquad if \quad \mathbf{e}_{i+1}\cdot(\mathbf{e}_{i-1}\times\mathbf{e}_i) < 0 \end{cases} \end{cases} \tag{3}$$

In IC, θ_i is in the range of $[0,\pi]$, and its ordering in i runs as $i = 2,3...,N-1$. Also, φ_i is in the range of $[-\pi,\pi]$, and its ordering runs as $3,4,...,N-1$. For the convenience of further analysis, we lump the ICs into a set $\{\phi_\alpha: \theta_2,\varphi_3,\theta_3,\varphi_4,\theta_4...\varphi_{N-1},\theta_{N-1}\}$, where $\alpha = 1,2,...,2N-5$. The relation of index between CC and IC is

$$i = [\alpha/2] + 2 = \kappa(\alpha)$$
$$\alpha^{\varphi}(i) = 2i - 4(i \geq 3)$$
$$\alpha^{\theta}(i) = 2i - 3(i \geq 2) \tag{4}$$

Modified Elastic Potential Function. We defined a new potential function which has an extra term in addition to the potential function in the conventional eNMA.[1,5] The potential is

$$V = (\gamma/2)\sum_i\sum_j \sigma_{ij}(|\mathbf{r}_{ij}| - |\mathbf{r}_{ij}{}^0|)^2 + (\omega/2)\sum_\alpha(\phi_\alpha - \phi_\alpha{}^0)^2$$

$$\sigma_{ij} = \begin{cases} 1 & |\mathbf{r}_{ij}{}^0| \le r_c \\ 0 & |\mathbf{r}_{ij}{}^0| > r_c \end{cases}; \quad \omega = 3\min(H_{\alpha\alpha}{}^0) \tag{5}$$

where $H_{\alpha\alpha}{}^0$ is the diagonal element of the Hessian matrix of the conventional eNMA potential in IC. Note the summation in the second term goes over all the elements in the set of $\{\phi_\alpha\}$ (all of θ and φ). A scaling factor of 3 is chosen so that only the smallest diagonal term is dramatically changed.

This form of potential function also has its minimum at the current structure, thus no initial minimization is needed. The main purpose of the modified potential is to make the flexible degrees of freedom much stiffer so that the tip effect is eradicated.

Generalized Eigenvalue Problem in IC. The elements of the Hessian matrix in IC are related to those in CC as

$$\begin{cases} H_{\alpha\beta} = \dfrac{\partial^2 V}{\partial\phi_\alpha\partial\phi_\beta} = \sum \dfrac{\partial\mathbf{r}_i}{\partial\phi_\alpha}\cdot\mathbf{h}^{ij}\cdot\dfrac{\partial\mathbf{r}_j}{\partial\phi_\beta} \\ \dfrac{\partial\mathbf{r}_i}{\partial\phi_\alpha} = \dfrac{\partial\mathbf{r}_i}{\partial\theta_{\kappa(\alpha)}} = \begin{cases} (\mathbf{e}_{\kappa(\alpha)} \times \mathbf{e}_{\kappa(\alpha+1)}) \times \mathbf{r}_{\alpha i}, & i > \kappa(\alpha) \\ 0, & i \le \kappa(\alpha) \end{cases} & \phi_\alpha \in \theta, \\ \dfrac{\partial\mathbf{r}_i}{\partial\phi_\alpha} = \dfrac{\partial\mathbf{r}_i}{\partial\theta_{\kappa(\alpha)}} = \begin{cases} \mathbf{e}_{\kappa(\alpha)} \times \mathbf{r}_{\alpha i}, & i > \kappa(\alpha) \\ 0, & i \le \kappa(\alpha) \end{cases} & \phi_\alpha \in \varphi \end{cases} \tag{6}$$

where $\mathbf{r}_{\alpha i} = \mathbf{r}_i - \mathbf{r}_{\kappa(\alpha)}$, and $\mathbf{h}^{ij}$ is the submatrix of Hessian for the $i - j$ pair in CC. The calculation of the Hessian matrix in IC is accelerated from the order of n^4 to n^2. It uses a recursion relationship of matrix manipulation.[28]

$$H_{\alpha\beta} = \chi_\alpha^T \mathbf{R}^{i,j}\chi_\beta, \quad i = \kappa(\alpha), \quad j = \kappa(\beta)$$

$$\chi_\alpha^{(\theta)} = \begin{pmatrix} \mathbf{e}_{\kappa(\alpha)} \times \mathbf{e}_{\kappa(\alpha)+1} \\ (\mathbf{e}_{\kappa(\alpha)} \times \mathbf{e}_{\kappa(\alpha)+1}) \times \mathbf{r}_{\kappa(\alpha)} \end{pmatrix}$$

$$\chi_\alpha^{(\varphi)} = \begin{pmatrix} \mathbf{e}_{\kappa(\alpha)} \\ \mathbf{e}_{\kappa(\alpha)} \times \mathbf{r}_{\kappa(\alpha)} \end{pmatrix}$$

$$\mathbf{R}^{i,j} = \sum_{\substack{i'<i \\ j'\ge j}} \frac{\gamma}{r_{ij}^2}\begin{pmatrix} \mathbf{r}_i \times \mathbf{r}_j \\ \mathbf{r}_{ij} \end{pmatrix}(\mathbf{r}_i \times \mathbf{r}_j \quad \mathbf{r}_{ij}) = \sum_{\substack{i'<i \\ j'\ge j}} S_{i',j'} \tag{7}$$

and

$$\mathbf{R}^{i,j} = \mathbf{R}^{i,j+1} + \mathbf{R}^{i-1,j} - \mathbf{R}^{i-1,j+1} + S_{i-1,j}$$

$$\mathbf{R}^{1,j} = \mathbf{R}^{i,N+1} = 0 \tag{8}$$

recursion order: $(i,j) = (1,N),(1,N-1),...(1,1),$
$(2,N),...(2,2),(3,N),...,(N,N)$

where $\mathbf{R}^{i,j}$ is a 6×6 matrix, and $\chi_\alpha{}^{(\theta)}$, $\chi_\alpha{}^{(\varphi)}$ are 6×1 vectors. When analyzing dynamics, we should eliminate external motion by applying the Eckart Condition[29]

$$\sum_i m_i\mathbf{r}_i = 0; \quad \sum_i m_i\mathbf{r}_i \times \dot{\mathbf{r}}_i = 0 \tag{9}$$

First derivatives can also be calculated accordingly

$$\frac{\partial\mathbf{r}_i}{\partial\phi_\alpha} = \begin{cases} (\mathbf{P}_i\,\mathbf{1})\mathbf{W}_{\kappa(\alpha)}\chi_\alpha & i < \kappa(\alpha) \\ -(\mathbf{P}_i\,\mathbf{1})(\mathbf{1} - \mathbf{W}_{\kappa(\alpha)})\chi_\alpha & i \ge \kappa(\alpha) \end{cases}$$

$$\mathbf{W}_i = \begin{pmatrix} \mathbf{I}^{-1}\mathbf{I}_C & \mathbf{I}^{-1}\mathbf{P}_C^T \\ \mathbf{P}_C/M & M_C/M\mathbf{1} \end{pmatrix} \tag{10}$$

where

$$M_A = \sum_{i'\le i} m_{i'};\ \mathbf{P}_A = \sum_{i'\le i} m_{i'}\mathbf{P}_{i'};\ \mathbf{I}_A = \sum_{i'\le i} m_{i'}\mathbf{P}_{i'}^T\mathbf{P}_{i'}$$

$$\mathbf{I} = \sum_{i'} m_{i'}\mathbf{P}_{i'}^T\mathbf{P}_{i'};\ M = \sum_{i'} m_{i'}$$

$$\mathbf{I}_C = \mathbf{I} - \mathbf{I}_A;\ \mathbf{P}_C = -\mathbf{P}_A;\ M_C = M - M_A$$

$$\mathbf{P}_i = \mathbf{r}_i \times \quad = \begin{pmatrix} 0 & -z_i & y_i \\ z_i & 0 & -x_i \\ -y_i & x_i & 0 \end{pmatrix} \tag{11}$$

P is an operator in matrix form. Then, we have the kinetic matrix

$$T = \frac{1}{2}\sum_{\alpha,\beta}\mathbf{T}_{\alpha\beta}\dot{\phi}_\alpha\dot{\phi}_\beta$$

$$\mathbf{T}_{\alpha\beta} = \sum_i m_i \frac{\partial\mathbf{r}_i}{\partial\phi_\alpha}\cdot\frac{\partial\mathbf{r}_i}{\partial\phi_\beta}$$

$$= \chi_\alpha^T\mathbf{K}_{\alpha\beta}\chi_\beta$$

$$\mathbf{K}_{\alpha\beta} = \begin{pmatrix} \mathbf{P}_A^T \\ M_A\mathbf{1} \end{pmatrix}(\mathbf{P}_C\ M_C\mathbf{1})/M + \begin{pmatrix} \mathbf{I}_A \\ \mathbf{P}_A \end{pmatrix}\mathbf{I}^{-1}(\mathbf{I}_C\ \mathbf{P}_C^T) \tag{12}$$

We end up with solving a generalized eigenvalue problem

$$\mathbf{HV} = \Lambda\mathbf{TV} \tag{13}$$

After obtaining eigenvectors, we can get orthonormal vectors in CC.

$$\Delta\mathbf{r}_i^{(k)} = \sum_\alpha \frac{\partial\mathbf{r}_i}{\partial\phi_\alpha}\Delta\phi_\alpha^{(k)}$$

$$\sum_i \Delta\mathbf{r}_i^{(k)}\cdot\Delta\mathbf{r}_i^{(k')} = \delta_{k,k'} \tag{14}$$

Finally, we define a quantitative localization factor T for the "tip effect"

$$T = \sum_i \left(\frac{||\Delta\mathbf{r}_{i+1} - \Delta\mathbf{r}_i||}{||\mathbf{r}_{i+1} - \mathbf{r}_i||}\right)^3 \tag{15}$$

The larger the T, the more eminent the "tip effect".

Multichain Analysis. To generalize the above method to multichain molecules, we shall virtually connect the last Cα

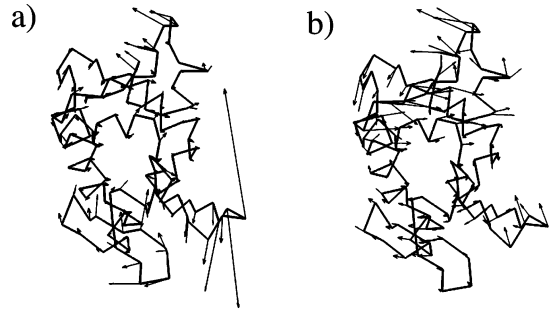


Figure 2. Motional patterns for the fourth mode of lysozyme (PDB code: 3lzt). (a) From the conventional eNMA, the lower-right portion has abnormal motions. (b) From the new eNMA, the motions for lower-right portion is much more realistic.

atom of the proceeding chain to the first Cα atom of the following chain. Hence, we will have 6 additional degrees of freedom for each additional chain. Among them, 5 are internal degrees of freedom and the sixth is for virtual bond length. It is noted that this virtual bond is the only flexible bond. We can make a new convention of order as

$$\{\phi_\alpha:\theta_2,\varphi_3,\theta_3,\varphi_4,\theta_4,...,\varphi_{N_1},\theta_{N_1},l_{N_1+1},\varphi_{N_1+1},\theta_{N_1+1},\varphi_{N_1+2},...\} \quad (16)$$

where N_1 is the number of Cα atoms in the first chain. For the virtual bond that connects the two chains, because

$$\frac{\partial \mathbf{r}_i}{\partial l_{\kappa(\alpha)}} = \mathbf{e}_{\kappa(\alpha)} \; (i > \kappa(\alpha)) \quad (17)$$

we only need to add an additional χ term in calculating matrices:

$$\chi_\alpha^{(l)} = \begin{pmatrix} 0 \\ -\mathbf{e}_{\kappa(\alpha)} \end{pmatrix} \quad (18)$$

Of course, with the additional degrees of freedom, the index order is also changed accordingly.

III. Results

Proteins with Single Polypeptide Chain. Here we first show an example of the tip effect. Lysozyme is a classic protein for NMA as its lowest-frequency mode represents a catalytically important hinge-bending motion.[30,31] Figure 2a shows the tip effect in one of the lowest-frequency modes, the fourth vibrational mode, calculated by the conventional eNMA.[5] The amplitude of motion of a flexible loop at the lower-right corner is abnormally large comparing with the rest of the system. In the new eNMA, the motion of the same region is much more realistic for this low-frequency mode (Figure 2b). Tests on various other proteins showed very similar results.

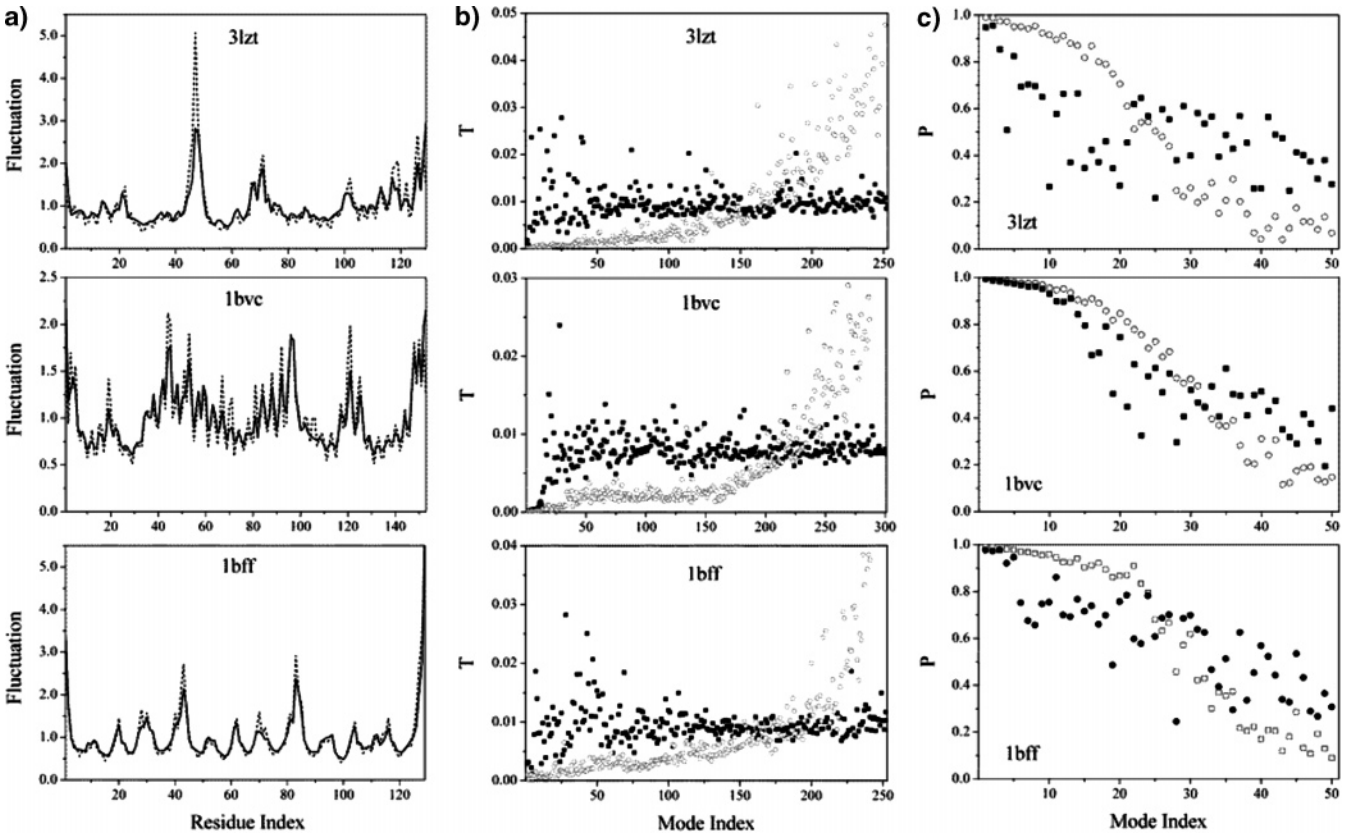


Figure 3. Test of the new eNMA on three different proteins (PDB codes: 3lzt, 1bvc, 1bff). (a) Fluctuation curves calculated by the conventional eNMA (dotted lines) and the new eNMA (solid lines). The curves by the new eNMA are smoother due to the absence of tip effect. (b) Comparison of localization factors (*T* values) for three proteins. The dark squares are for the conventional eNMA and the empty circles are for the new eNMA. The results of the new eNMA are much more rational in low-frequency regime. (c) Results of the projection of an individual mode by one eNMA method onto a subset of 50 low-frequency modes by another method. The empty circles are for case with one new eNMA mode projected onto 50 conventional eNMA modes. The dark squares are just the other way around.

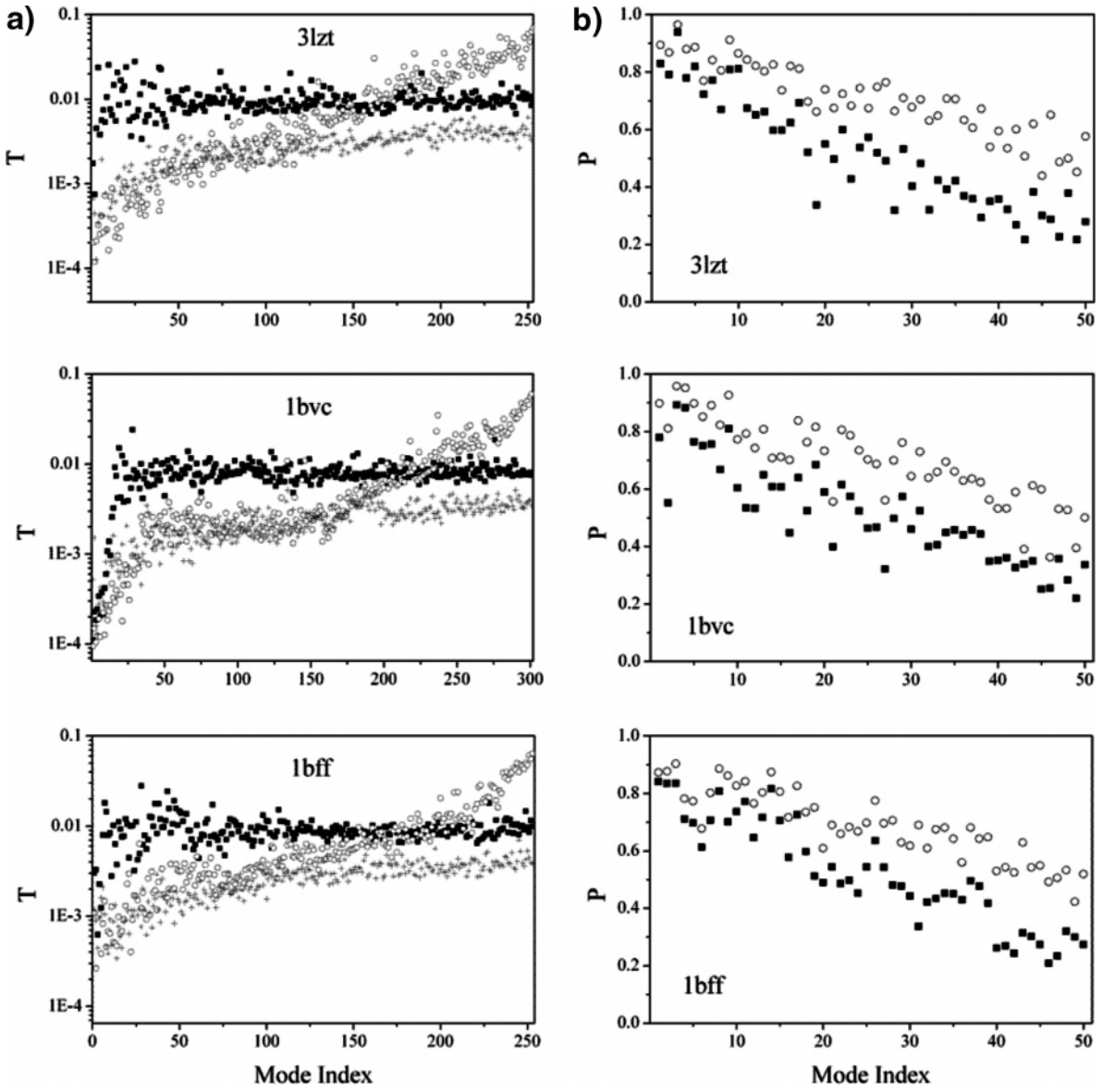


Figure 4. Comparison of modes from the conventional eNMA, the new eNMA, and CHARMM. (a) Localization factors for modes calculated by the conventional eNMA (dark squares), the new eNMA (empty circles), and CHARMM force field (cross). For all three proteins, in the low-frequency modes, the values of localization factor *T* are very much the same between the new eNMA and CHARMM but significantly worse in the conventional eNMA (note the vertical axes are in logarithmic scale). (b) The projection of an individual CHARMM mode (*x*-axis) onto 50 low-frequency mode subspace of the new eNMA (empty circles) and the conventional eNMA (dark squares). The projection coefficients for the new eNMA are systematically larger than those of the conventional eNMA, i.e., the modes from the new eNMA are closer to those of CHARMM modes.

We compared the atomic fluctuation curves computed for three different proteins by the conventional[5] and the new eNMAs in Figure 3a. The curves from the new eNMA (solid lines) are systematically smoother than those from the conventional eNMA as a result of eradication of the tip effect in the eigenvectors by the new eNMA. Figure 3b shows the comparison of the localization factor *T* of low-frequency modes (eq 15) for three proteins (the larger the value of *T* in low-frequency modes, the more localized the motion in that mode, i.e, more eminent the tip effect). In each case, the low-frequency modes from conventional eNMA has a much higher *T* value indicating the contamination of the tip effect. In fact, the *T* value from conventional eNMA stays more or less constant at higher frequency modes because of the divergence of eigenvectors. On the contrary, the *T* value for the modes from the new eNMA is small for low-frequency modes and increases progressively as the frequency increases. This is an expected behavior for normal modes, as the higher the frequency, the more localized motion in the modes. Figure 3c describes the projection of an individual mode by one method onto a subset of 50 low-frequency modes by another method. The purpose was to compare the overall similarity of modes between two methods. The empty circles are the results of using one mode from new eNMA protected onto a subset of 50 modes from conventional eNMA. It is clear that the individual mode of the new eNMA can be reasonably expressed by the subset of low-frequency modes of the conventional eNMA (the lower the frequency of the mode by new eNMA, the better in terms of linear expression). On the contrary, the other

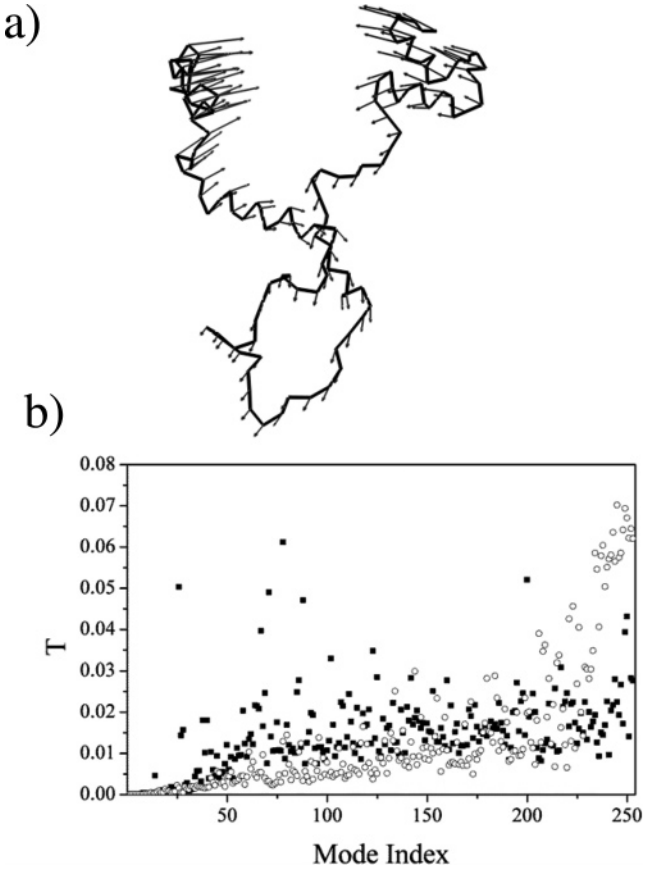


Figure 5. Results on lysozyme (PDB code: 3lzt) with a completed extended conformation. (a) Motional pattern of the first vibrational mode calculated by new eNMA. (b) Tip effect, the solid squares are for conventional eNMA, and the empty circles are for new eNMA.

way around gave much worse results (solid squares). This was because some of the modes of the conventional eNMA are contaminated by the tip effect so that they simply cannot be linearly expressed by a subset of modes of the new eNMA.

Figure 4a shows the comparison of T values for modes computed by the conventional eNMA, the new eNMA, and the CHARMM force field. Clearly the modes by the new eNMA are much closer to those of the CHARMM force field in terms of T values. The curve from the CHARMM force field rises slower than the new eNMA because the CHARMM force field has a much larger mode number. Figure 4b shows the projection of an individual CHARMM mode (x-axis) onto 50 low-frequency mode subspace of the new eNMA (empty circles) and the conventional eNMA (solid squares). The projection coefficients for the new eNMA are systematically larger than those of the conventional eNMA, indicating that the modes of the new eNMA are closer to those of CHARMM. Testing of the method on other proteins led to similar conclusions.

Proteins with a Completely Extended Conformation. Since the tip effect often occurs around less-packed regions in protein such as surface loops, we tested the new eNMA on protein conformations that were totally extended. The purpose of doing so was for cases in which one needs to conduct, for example, a Monte Carlo sampling of protein

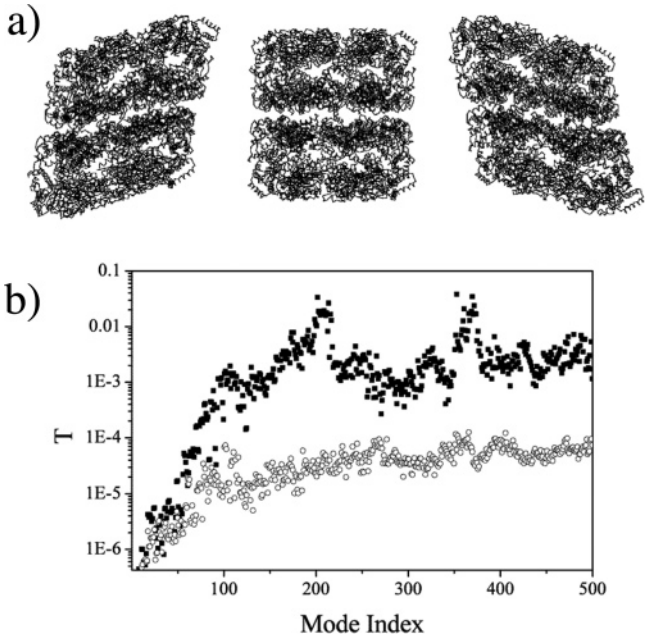


Figure 6. Results on multisubunit supramolecular complex, the molecular chaperonin GroEL. (a) Motional pattern of the second vibrational mode, which is a stretching mode along diagonal line of the molecule. (b) Tip effect, the solid squares are for conventional eNMA, and the empty circles are for new eNMA. Note the vertical axis is made in logarithmic scale.

conformation based on elastic normal modes.[32] Figure 5a shows the motional pattern of the first vibrational mode of the extended lysozyme chain calculated by the new eNMA. Figure 5b shows the T values, the black squares are for conventional eNMA, and the empty circles are for the new eNMA. The tip effect in the low-frequency modes by the conventional eNMA was much more severe.

Proteins with Multiple Polypeptide Chains. We also tested the new method on a multipolypeptide chain complex, the molecular chaperonin GroEL,[33] which is an ATP-driven molecular motor and a supramolecular complex. The molecule has a cylindrical shape formed by two back-to-back stacked 7-fold rings. It is known that GroEL has to undergo huge conformational changes in order to carry out its biological function, which is for facilitating correct folding of unfolded or misfolded polypeptide chains.[34] The system has been a subject of many computational studies.[14,35,36] As an example, Figure 6a shows the motional patterns of a second vibrational mode by the new eNMA. This mode is a stretching mode along a diagonal line of the molecule. The other low-frequency modes are also similar to what was previously observed.[14] For GroEL, however, the tip effect in the conventional eNMA (solid squares) is a few orders of magnitude worse than that in the new eNMA (empty circles). Figure 6b shows the T values for the first 500 vibrational modes; note, the plot was made in the logarithmic scale in the vertical axis.

IV. Concluding Discussion

In this paper, we have reported an improved method for coarse-grained elastic normal-mode analysis. The Hamiltonian used contains the regular term used in conventional

eNMA[1,5] and an additional term that involves pseudo bond angles and dihedrals along the Cα traces. Similar to the Hamiltonian of conventional eNMA, the new one has its minimum at the current structure so that no initial energy minimization is needed. The computational procedure was carried out in the internal coordinate system so that the softest link in elastic potential, the bond length stretching, was avoided. The essential idea of eliminating the tip effect was to increase the stiffness of degrees of freedom even when they are located in regions that are not very densely packed.

We tested our method on proteins with single polypeptide chains in their compact native conformations and extended denatured conformations. We also tested our method on multisubunit supramolecular complexes. Results demonstrated that the new method is capable of easing the tip effect in all cases. It was found that the improvement in the quality of modes was particularly substantial for supramolecular complexes. Therefore, we expect such a method would be a complementary tool to the existing methods. It should be particularly useful in certain special applications in which a set of continuous low-frequency modes are needed such as in structural refinement[37] and in normal mode based sampling.[24]

The scaling factor 3 in eq 5 affects the stiffness of modes. For larger complexes, the value of this factor may be adjusted. In our experience, the value between 3 and 15 is good.

The conventional eNMA was also successfully applied to cryo-EM density maps,[6,8] in which case one does not have the knowledge of chain connectivity. Extension of our current method to cryo-EM density maps requires certain efforts. It may be necessary to artificially construct a pseudochain among all the points used to represent the density maps using methods such as the Traveling Salesman Algorithm. We will report the results in a separate paper.

Acknowledgment. The authors gratefully thank support from the National Institutes of Health (R01-GM067801). B.P. is supported by a predoctoral fellowship from Houston Area Molecular Biophysics Program (HAMBP). M.L. is partially supported by a predoctoral fellowship from W. M. Keck Foundation of the Gulf Coast Consortia through the Keck Center for Computational and Structural Biology. Computer software is available upon request.

References

(1) Tirion, M. M. Large amplitude elastic motions in proteins from a single-parameter, atomic analysis. *Phys. Rev. Lett.* **1996**, *77*, 1905−1908.

(2) Haliloglu, T.; Bahar, I.; Erman, B. Gaussian Dynamics of Folded Proteins. *Phys. Rev. Lett.* **1997**, *79*, 3090−3093.

(3) Bahar, I.; Atilgan, A. R.; Erman, B. Direct evaluation of thermal fluctuations in proteins using a single- parameter harmonic potential. *Fold Des.* **1997**, *2*, 173−181.

(4) Doruker, P.; Jernigan, R. L.; Bahar, I. Dynamic of large proteins through hierarchical levels of coarse-grained structures. *J. Comput. Chem.* **2002**, *23*, 119−127.

(5) Atilgan, A. R.; Durell, S. R.; Jernigan, R. L.; Demirel, M. C.; Keskin, O.; Bahar, I. Anisotropy of fluctuation dynamics of proteins with an elastic network model. *Biophys. J.* **2001**, *80*, 505−515.

(6) Ming, D.; Kong, Y.; Lambert, M.; Huang, Z.; Ma, J. How to Describe Protein Motion Without Amino-acid Sequence and Atomic Coordinates. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 8620−8625.

(7) Ming, D.; Kong, Y.; Wu, Y.; Ma, J. Substructure Synthesis Method for Simulating Large Molecular Complexes. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 104−109.

(8) Tama, F.; Wriggers, W.; Brooks, C. L. Exploring Global Distortions of Biological Macromolecules and Assemblies from Low-resolution Structural Information and Elastic Network Theory. *J. Mol. Biol.* **2002**, *321*, 297−305.

(9) Hinsen, K. Analysis of domain motions by approximate normal mode calculations. *Proteins* **1998**, *33*, 417−429.

(10) Ming, D.; Kong, Y.; Wu, Y.; Ma, J. Simulation of F-actin Filaments of Several Microns. *Biophys. J.* **2003**, *85*, 27−35.

(11) Ming, D.; Kong, Y.; Wakil, S. J.; Brink, J.; Ma, J. Domain Movements in Human Fatty Acid Synthase by Quantized Elastic Deformational Model. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 7895−7899.

(12) Kong, Y.; Ming, D.; Wu, Y.; Stoops, J. K.; Zhou, Z. H.; Ma, J. Conformational Flexibility of Pyruvate Dehydrogenase Complexes: A Computational Analysis by Quantized Elastic Deformational Model. *J. Mol. Biol.* **2003**, *330*, 129−135.

(13) Beuron, F.; Flynn, T. C.; Ma, J.; Kondo, H.; Zhang, X.; Freemont, P. S. Motions and Negative Cooperativity between p97 Domains revealed by Cryoelectron Microscopy and Quantized Elastic Deformational Model. *J. Mol. Biol.* **2003**, *327*, 619−629.

(14) Keskin, O.; Bahar, I.; Flatow, D.; Covell, D. G.; Jernigan, R. L. Molecular Mechanisms of Chaperonin GroEL-GroES Function. *Biochemistry* **2002**, *41*, 491−501.

(15) Wang, Y.; Rader, A. J.; Bahar, I.; Jernigan, R. L. Global ribosome motions revealed with elastic network model. *J. Struct. Biol.* **2004**, *147*, 302−314.

(16) Xu, C.; Tobi, D.; Bahar, I. Allosteric changes in protein structure computed by a simple mechanical model: hemoglobin T ↔ R2 transition. *J. Mol. Biol.* **2003**, *333*, 153−168.

(17) Tama, F.; Valle, M.; Frank, J.; Brooks, C. L., III Dynamic reorganization of the functionally active ribosome explored by normal mode analysis and cryo-electron microscopy. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 9319−9323.

(18) Tama, F.; Brooks, C. L., III The Mechanism and Pathway of pH Induced Swelling in Cowpea Chlorotic Mottle Virus. *J. Mol. Biol.* **2002**, *318*, 733−747.

(19) Chacon, P.; Tama, F.; Wriggers, W. Mega-dalton biomolecular motion captured from electron microscopy reconstructions. *J. Mol. Biol.* **2003**, *326*, 485−492.

(20) Miyashita, O.; Onuchic, J. N.; Wolynes, P. G. Nonlinear elasticity, proteinquakes, and the energy landscapes of functional transitions in proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12570−12575.

(21) Ma, J. New advances in normal mode analysis of supermolecular complexes and applications to structural refinement. *Curr. Protein Pept. Sci.* **2004**, *5*, 119−123.

(22) Ma, J. Usefulness and Limitations of Normal Mode Analysis in Modeling Dynamics of Biomolecular Complexes. *Structure* **2005**, *13*, 373−380.

(23) Bahar, I.; Rader, A. J. Coarse-grained normal mode analysis in structural biology. *Curr. Opin. Struct. Biol.* **2005**, *15*, 586−592.

(24) Lu, M.; Ma, J. The Role of Shape in Determining Molecular Motions. *Biophys. J.* **2005**, *89*, 2395−2401.

(25) MacKerell, A. D.; Bashford, Jr., D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, I., W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem.* **1998**, *B102*, 3586−3616.

(26) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *J. Comput. Chem.* **1983**, *4*, 187−217.

(27) Kamiya, K.; Sugawara, Y.; Umeyama, H. Algorithm for normal-mode analysis with general internal coordinates. *J. Comput. Chem.* **2003**, *24*, 826−841.

(28) Noguti, T.; Go, N. A method of rapid calculation of a second derivative matrix of conformational energy for large molecules. *J. Phys. Soc. Jpn.* **1983**, *52*, 3685−3690.

(29) Eckart, C. Some studies concerning rotating axes and polyatomic molecules. *Phys. Rev.* **1935**, *47*, 552.

(30) McCammon, J. A.; Gelin, B. R.; Karplus, M.; Wolynes, P. G. The hinge-bending mode in lysozyme. *Nature* **1976**, *262*, 325−326.

(31) Brooks, B.; Karplus, M. Normal modes for specific motions of macromolecules: Application to the hinge-bending mode of lysozyme. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 4995−4999.

(32) Wu, Y.; Tian, X.; Lu, M.; Chen, M.; Wang, Q.; Ma, J. Folding of small helical proteins assisted by small-angle X-ray scattering profiles. *Structure (Camb.)* **2005**, *13*, 1587−1597.

(33) Xu, Z.; Sigler, P. B. GroEL/GroES: Structure and function of a two-stroke folding machine. *J. Struct. Biol.* **1998**, *124*, 129−141.

(34) Sigler, P. B.; Xu, Z.; Rye, H. S.; Burston, S. G.; Fenton, W. A.; Horwich, A. L. Structure and function in GroEL-mediated protein folding. *Anuu. Rev. Biochem.* **1998**, *67*, 581−608.

(35) Ma, J.; Karplus, M. The allosteric mechanism of the chaperonin GroEL: A dynamic analysis. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 8502−8507.

(36) Ma, J.; Sigler, P. B.; Xu, Z.; Karplus, M. A dynamic model for the allosteric mechanism of GroEL. *J. Mol. Biol.* **2000**, *302*, 303−313.

(37) Wu, Y.; Ma, J. Refinement of F-actin Model Against Fibre Diffraction Data by Long-range Normal Modes. *Biophys. J.* **2004**, *86*, 116−124.

CT050307U

JCTC Journal of Chemical Theory and Computation

Simulating Membrane Dynamics in Nonhomogeneous Hydrodynamic Environments

Lawrence C.-L. Lin† and Frank L. H. Brown*,‡

Department of Physics, University of California, Santa Barbara, California 93106-9530, and Department of Chemistry and Biochemistry, University of California, Santa Barbara, California 93106-9510

Received November 28, 2005

Abstract: Two previously introduced simulation algorithms for the dynamics of elastic membrane sheets embedded in a fluid medium are extended to account for inhomogeneous hydrodynamic environments. We calculate the height autocorrelation function for a lipid bilayer randomly pinned to a flat substrate and the influence of fluid confinement by the spectrin cytoskeleton on short wavelength membrane undulations of the human red blood cell. Altering the hydrodynamic environment of the membrane leads to significant changes in dynamics, and we discuss these effects in the context of recent experiments.

I. Introduction

Many interesting problems in membrane biophysics, cellular biology, and biochemistry involve length and time scales completely inaccessible to molecular dynamics simulation. Simple estimates[1] indicate that it will be several decades before fully atomic models will become viable tools for studying lipid bilayers over the micron and millisecond scales necessary to make connection with cellular scale behavior. Recent progress in developing coarse-grained lipid models[2−19] points to a more optimistic future for molecular simulation of biomembranes and related materials; however, these methods are still far too computationally intensive for practical use in studying length scales exceeding tens of nanometers and time scales exceeding hundreds of nanoseconds. Elastic models represent the only theoretical/computational means presently available for studying biomembranes and lipid bilayers over long length and time scales.

The traditional Canham-Helfrich elastic picture for membrane energetics[20−22] combined with overdamped dynamics in a hydrodynamic environment[23,24] has been successfully applied to numerous biophysical questions of interest. However such considerations are typically limited to problems simple enough to be treated analytically[23,25−27] or problems where dynamics are not explicitly considered.[28−31] Recently, we have proposed two simulation algorithms for studying the dynamics of membranes evolving under the influence of perturbations that do not allow for analytical treatment. The first of these methods is specific to harmonic potentials where it is possible to numerically (but impossible analytically) identify the normal modes of the system.[32] The second method is applicable to general anharmonic potentials.[33,34] We will refer to both of these methods as Fourier space Brownian dynamics (FSBD) in this work.

Both of the previously introduced FSBD algorithms have solved for the dynamics of elastic sheets residing in a fluid medium of infinite extent. More realistically, biological cells and the artificial systems developed to mimic biomembranes reside in finite environments, which can act to constrain the fluid flow proximal to the membrane surface. While this fact is appreciated in the theory literature,[23,35,36] we are unaware of prior simulations that have incorporated the effect of finite sized hydrodynamic environments. It turns out that our existing FSBD methodology may be extended to account for impermeable walls near a dynamically evolving membrane with little difficulty. The purpose of this article is to describe the necessary extensions to FSBD to model a dynamically fluctuating membrane near a wall impermeable to fluid flow. Similar considerations can be applied to semipermeable walls[38] as well.

The effect of altering the hydrodynamic environment around the membrane surface can be striking. Although the

* Corresponding author phone: (805)893-5494; fax: (805)893-4120; e-mail: flbrown@chem.ucsb.edu.
† Department of Physics.
‡ Department of Chemistry and Biochemistry.

10.1021/ct050293s CCC: $33.50
Published on Web 03/08/2006

thermal statistics of the membrane are not altered (as the energetics of the system do not change), the dynamics involved in sampling configurations can be modified by orders of magnitude. Physically, this effect is due to the difficulty in moving an incompressible fluid around inside a small volume and could be anticipated; however, the effect is surprisingly large and demonstrates the importance of including hydrodynamic effects properly in any simulation. For the purposes of illustration, we model the height−height correlation function of a membrane pinned to a solid support as seen in recent experiments[37,38] and show that the inclusion of an impermeable wall in the treatment slows dynamics by several orders of magnitude relative to treatments that ignore the presence of the solid support. Our results are in general agreement with experiment. We also consider the effect of hydrodynamic effects on the fluctuation dynamics of the red blood cell and related consequences for membrane protein diffusion. These results are compared to prior studies where such effects were neglected and significant influence of hydrodynamics is noted.

II. Background: Energetics and Dynamics of an Fluid Membrane Sheet

We begin by specifying the Hamiltonian for a thin membrane sheet. For the systems we study in this work, it is valid to assume that the bilayer is nearly flat and height fluctuations normal to the plane of the membrane are small. In this case, a convenient parametrization is the Monge gauge, which specifies the height of the membrane $h(\mathbf{r})$ in terms of the location $\mathbf{r} \equiv (x, y)$ in the xy plane. The total energy is therefore written as (Canham-Helfrich energetics)[20−22]

$$H = \int_A d\mathbf{r} \left\{ \frac{\kappa}{2}[\nabla^2 h(\mathbf{r})]^2 + \frac{\sigma}{2}[\nabla h(\mathbf{r})]^2 + \mathscr{H}_{\text{int}}[h(\mathbf{r})] \right\} \quad (1)$$

where κ is the bending modulus, σ is the surface tension, and $A = L^2$ is the projected area of a square patch of membrane. The height $h(\mathbf{r})$ is taken to be periodic with period L in both the x and y directions. The first two terms of H reflect bending energetics and energetics due to surface tension. These terms will be present even for an isolated membrane sheet suspended in solution. The last term allows for arbitrary functional dependence of H on $h(\mathbf{r})$ as dictated by any interactions between the membrane and its environment.

In the overdamped regime appropriate to cellular scale dynamics,[39] the velocity field of the membrane is expected to be linearly dependent upon the forces acting on the sheet.[40] The relationship is generally nonlocal in space due to hydrodynamic interactions between distant points on the membrane surface.[40] Assuming that height fluctuations away from $h(\mathbf{r}) = 0$ are small, the expression for membrane velocity is[24]

$$v_m(\mathbf{r}, t) = \frac{\partial h(\mathbf{r}, t)}{\partial t} \quad (2)$$

$$= \int_{-\infty}^{\infty} d\mathbf{r}' \, \Lambda(\mathbf{r} - \mathbf{r}')[F(\mathbf{r}', t) + \zeta(\mathbf{r}', t)]$$

where $F(\mathbf{r}, t) = -\delta H/\delta h(\mathbf{r}, t)$ is the force per area on the membrane resulting from the Hamiltonian H and $\Lambda(\mathbf{r} - \mathbf{r}')$ is a hydrodynamic kernel dependent upon the boundary conditions for the fluid medium in which the membrane resides. Additionally, we include a (Gaussian white) thermal random force $\zeta(\mathbf{r}', t)$ that satisfies the fluctuation−dissipation theorem[40]

$$\langle \zeta(\mathbf{r}, t) \rangle = 0$$

$$\langle \zeta(\mathbf{r}, t)\zeta(\mathbf{r}', t') \rangle = 2k_B T \Lambda^{-1}(\mathbf{r} - \mathbf{r}') \, \delta(t - t') \quad (3)$$

where the inverse of $\Lambda(\mathbf{r} - \mathbf{r}')$ is defined by

$$\int_{-\infty}^{\infty} d\mathbf{r}' \, \Lambda(\mathbf{r} - \mathbf{r}')\Lambda^{-1}(\mathbf{r}') = \delta(\mathbf{r}) \quad (4)$$

Although eq 2 is often quoted in the context of a membrane suspended in an infinite homogeneous hydrodynamic medium, the expression has generality beyond this case. It is this versatility that we exploit in this work. We shall present $\Lambda(\mathbf{r})$ for different boundary conditions corresponding to various physical situations in the next section.

The above equations for the membrane are most easily handled using Fourier modes defined by

$$h_{\mathbf{k}} = \int_A d\mathbf{r} \, h(\mathbf{r}) \, e^{-i\mathbf{k}\cdot\mathbf{r}} \quad (5)$$

$$h(\mathbf{r}) = \frac{1}{L^2} \sum_{\mathbf{k}} h_{\mathbf{k}} \, e^{i\mathbf{k}\cdot\mathbf{r}}$$

In Fourier space, the random force obeys[24]

$$\langle \zeta_{\mathbf{k}}(t) \rangle = 0$$

$$\langle \zeta_{\mathbf{k}}(t)\zeta_{\mathbf{k}'}(t') \rangle = 2k_B T L^2 \, \Lambda_{\mathbf{k}}^{-1} \, \delta_{\mathbf{k},-\mathbf{k}'} \, \delta(t - t') \quad (6)$$

The convenience of using the amplitudes $h_{\mathbf{k}}$ as dynamical variables is illustrated in the case where there are no additional interactions, $\mathscr{H}_{\text{int}} = 0$. All modes conveniently decouple, and eqs 1 and 2 simplify to

$$H = \frac{1}{2L^2} \sum_{\mathbf{k}} (\kappa k^4 + \sigma k^2) |h_{\mathbf{k}}|^2 \quad (7)$$

$$\frac{\partial h_{\mathbf{k}}(t)}{\partial t} = -\omega_{\mathbf{k}} h_{\mathbf{k}} + \Lambda_{\mathbf{k}} \zeta_{\mathbf{k}} \quad (8)$$

where we define the frequencies

$$\omega_{\mathbf{k}} \equiv \Lambda_{\mathbf{k}}(\kappa k^4 + \sigma k^2) \quad (9)$$

These equations describe an Ornstein-Uhlenbeck process[41] for each mode $h_{\mathbf{k}}$ whose known analytical solutions lead to the time correlation functions

$$\langle h_{\mathbf{k}}(t) h_{\mathbf{k}'}(0) \rangle = \frac{k_B T L^2}{\kappa k^4 + \sigma k^2} \, e^{-\omega_{\mathbf{k}} t} \, \delta_{\mathbf{k},-\mathbf{k}'} \quad (10)$$

$$\langle h(\mathbf{r}, t) h(\mathbf{r}, 0) \rangle = \frac{k_B T}{L^2} \sum_{\mathbf{k}} \frac{1}{\kappa k^4 + \sigma k^2} \, e^{-\omega_{\mathbf{k}} t} \quad (11)$$

In the general case where $\mathscr{H}_{\text{int}} \neq 0$, working in Fourier space is still useful for removing the convolution in eq 2

$$\frac{\partial h_{\mathbf{k}}(t)}{\partial t} = \Lambda_{\mathbf{k}}\{F_{\mathbf{k}}[h(\mathbf{r}, t)] + \zeta_{\mathbf{k}}(t)\} \tag{12}$$

The force $F_{\mathbf{k}}$ is a functional of the height field $h(\mathbf{r}, t)$ and, in general, depends on the entire set of amplitudes $\{h_{\mathbf{k}}\}$. In the case of harmonic interactions, it is possible to identify normal modes for the system, which still allows for analytical calculations and/or simplified simulations. Anharmonic potentials require a true simulation approach. We discuss both harmonic and anharmonic perturbations in the following sections. Before proceeding, however, the hydrodynamic kernel must be derived for the physical boundary conditions appropriate for the system. In the next section, we specify $\Lambda_{\mathbf{k}}$ for two different cases of interest.

III. Hydrodynamic Kernels

This section presents a brief motivation for the hydrodynamic kernels adopted in the following calculations (eqs 22, 24, and 26−29). Further elaboration on the calculation of these quantities may be found in the original papers by Seifert[35] and Gov et al.[36]

The effect of surrounding fluid flow on membrane motion is determined by the Navier−Stokes equations for an incompressible fluid. In cellular environments, the Reynolds number is small,[39] and inertia can be neglected so that (Stokes equations[42])

$$\eta\nabla^2\mathbf{v} = \nabla p \tag{13}$$

where $\mathbf{v}$ is the velocity of the fluid with viscosity η, and p is the pressure. Additionally, the incompressibility of the fluid requires that

$$\nabla\cdot\mathbf{v} = 0 \tag{14}$$

Boundary conditions for the system determine the solution for the velocity and pressure due to a given deformation of the membrane. In this work, the two cases of interest are that of a membrane in an infinite fluid and a membrane near an impermeable wall. Similar solutions for a membrane next to a permeable wall may also be derived[36] but will not be discussed here.

By performing the Fourier transform (in the x and y dimensions only as in eq 5) on the Navier−Stokes equations and the incompressibility condition, we derive the differential equation that $(v_z)_{\mathbf{k}}$ obeys

$$\partial_z^2[(-k^2 + \partial_z^2)(v_z)_{\mathbf{k}}] = k^2[(-k^2 + \partial_z^2)(v_z)_{\mathbf{k}}] \tag{15}$$

This equation can be reduced to

$$(-k^2 + \partial_z^2)(v_z)_{\mathbf{k}} = c_1' \sinh kz + c_2' \cosh kz \tag{16}$$

where the c_i' are arbitrary constants. This second order differential equation has the solution

$$(v_z)_{\mathbf{k}} = c_1 \sinh kz + c_2 \cosh kz + c_3 z \sinh kz + c_4 z \cosh kz \tag{17}$$

All of the c_i are arbitrary constants to be determined by the boundary conditions, which we specify below.

Taking the bilayer to be located at $z = 0$, the hydrodynamic kernel $\Lambda_{\mathbf{k}}$ can be derived from the requirement that the velocity of the fluid immediately adjacent to the membrane matches the velocity of the bilayer. This condition, expressed as $v_z^{\pm}(z = 0) = v_m$, combined with eq 12 leads to an equation involving the hydrodynamic kernel

$$[(v_z^{\pm})_{\mathbf{k}}]_{z=0} = \Lambda_{\mathbf{k}} F_{\mathbf{k}} \tag{18}$$

The superscript $\pm$ refers to the two fluid regions separated by the bilayer, the upper portion (+) and the lower portion (−). Defining the stress tensor for both upper and lower regions

$$T_{ij}^{\pm} \equiv \eta(\partial_i v_j^{\pm} + \partial_j v_i^{\pm}) - p^{\pm}\delta_{ij} \tag{19}$$

force balance between hydrodynamic stress and force on the bilayer implies

$$F_{\mathbf{k}} = [(T_{zz}^{-})_{\mathbf{k}} - (T_{zz}^{+})_{\mathbf{k}}]_{z=0} \tag{20}$$

The hydrodynamic kernel is therefore given by

$$\Lambda_{\mathbf{k}} = \left.\frac{(v_z)_{\mathbf{k}}}{(T_{zz}^{-})_{\mathbf{k}} - (T_{zz}^{+})_{\mathbf{k}}}\right|_{z=0} \tag{21}$$

where the velocity, the pressure, and the stress tensor are found by using eq 17 for a given set of boundary conditions.

We now specify the necessary boundary conditions for the cases considered in this work. For regions of fluid extending infinitely far away from the membrane, it must be the case that $\mathbf{v}(z) = 0$ at distances far from the membrane surface. The bilayer defines the other boundary for such a semi-infinite region. Continuity requires that the velocities of the fluid on either side of the membrane match at $z = 0$. This condition is expressed as $\mathbf{v}^+(z = 0) = \mathbf{v}^-(z = 0)$, where $\mathbf{v}^+$ and $\mathbf{v}^-$ are the solutions above and below the membrane, respectively. Additionally, a membrane with infinite regions of fluid on both sides is required by symmetry to respect $v_z^+(z) = v_z^-(-z)$. This condition leads to the solution

$$\Lambda_{\mathbf{k}}^{\text{inf}} = \frac{1}{4\eta k} \tag{22}$$

Our main interest is that of a membrane near an impermeable wall, which we take to be located at $z = -d$. The wall imposes the boundary condition $\mathbf{v}^-(z = -d) = 0$, which forces the velocity of the fluid to vanish there. To uniquely determine $\Lambda_{\mathbf{k}}$, an additional boundary condition must be specified at the membrane surface. Seifert[35] has argued for in-plane incompressibility of the membrane

$$[\partial_x v_x + \partial_y v_y]_{z=0} = 0 \tag{23}$$

which results in a solution

$$\Lambda_{\mathbf{k}}^{\text{S}} = \Lambda_{\mathbf{k}}^{\text{inf}} \frac{e^{2kd} + e^{-2kd} - 2[1 + 2(kd)^2]}{e^{2kd} - [1 + 2(kd)^2] + 2kd} \tag{24}$$

More recently, Gov et al.[36] have argued for boundary

conditions reflecting the fact that a fluid membrane is not expected to support shear stress. Mathematically, this translates to

$$[T_{xz}^{+} - T_{xz}^{-}]_{z=0} = 0 \tag{25}$$

(and similarly for T_{yz}) so that the shear stress is the same on both sides at the membrane. A different solution is found in this case

$$\Lambda_{\mathbf{k}}^{\mathrm{G}} = \Lambda_{\mathbf{k}}^{\mathrm{inf}}\, e^{-2kd}\{e^{2kd} - [\,1 + 2(kd\,)^2] - 2kd\} \tag{26}$$

Both solutions approach $\Lambda_{\mathbf{k}}^{\mathrm{inf}}$ as kd becomes large, as expected on physical grounds. The hydrodynamics should be unaffected by the wall if the wall is far enough away (the magnitude of "far enough" is seen to depend on the wavelength of the deformation). However, as kd becomes small, the limiting behavior of the two solutions differ by a factor of 4 such that $\Lambda_{\mathbf{k}}^{\mathrm{S}} \rightarrow k^2d^3/12\eta$ and $\Lambda_{\mathbf{k}}^{\mathrm{G}} \rightarrow k^2d^3/3\eta$. Using eq 9 for the noninteracting membrane as an approximate guide, we see that, in this limit, the relaxation times are longer for the scenario of Seifert.

We discuss the case of finding $\Lambda_{\mathbf{k}}$ when the viscosities above and below the membrane are different. The procedure for finding these solutions is similar to the one outlined already in this section. For notational convenience, we define $\bar{\eta} = (\eta^{+} + \eta^{-})/2$, $\xi^{\pm} = \eta^{\pm}/\bar{\eta}$, and $\Delta = (\eta^{+} - \eta^{-})/2\bar{\eta}$. The wall is located below the membrane where the viscosity is η^{-}. The solution for the membrane in the unbound fluid is

$$\Lambda_{\mathbf{k}}^{\mathrm{inf,general}} = \frac{1}{4\bar{\eta}k} \tag{27}$$

In the presence of the wall, the solution with Seifert's boundary condition is

$$\Lambda_{\mathbf{k}}^{\mathrm{S,general}} = \Lambda_{\mathbf{k}}^{\mathrm{inf,general}}\, \frac{e^{2kd} + e^{-2kd} - 2\,[1 + 2(kd)^2]}{e^{2kd} - [1 + 2(kd)^2]\xi^{+} + 2kd\xi^{-} + \Delta\, e^{-2kd}} \tag{28}$$

while the solution for Gov's boundary condition is

$$\Lambda_{\mathbf{k}}^{\mathrm{G,general}} = \Lambda_{\mathbf{k}}^{\mathrm{inf,general}}\, \frac{e^{-2kd}\,\{e^{2kd} - [1 + 2(kd)^2]\,\xi^{+} - 2kd\xi^{-}\} + \Delta\, e^{-4kd}}{1 - 2\Delta\, e^{-2kd}[1 + 2(kd)^2] + \Delta^2 e^{-4kd}} \tag{29}$$

In the case of $\eta^{+} = \eta^{-}$, $\xi^{\pm} = 1$ and $\Delta = 0$, and we recover the solutions in eqs 22, 24, and 26.

The general effect of the wall is to slow the relaxation of the membrane. The confinement of water between the membrane surface and the adjacent wall hinders relaxation of the membrane. For an estimate of the difference between the decay times of a bound and unbound fluid, we refer, again, to eq 9 for the noninteracting membrane. Since the relaxation time of the amplitude $h_{\mathbf{k}}$ is given by $1/\omega_{\mathbf{k}}$, we see that the presence of the wall slows the membrane by a factor of $\sim 1/(kd)^3$ in the limit of small kd. The dynamics of the membrane can be altered significantly if the wall is close enough to the bilayer or if the dominant modes have a sufficiently long wavelength.

IV. Independent Modes

Since we work with the amplitudes $h_{\mathbf{k}}$ as the dynamical variables, it is important to recognize that they are not all independent (even if $\mathscr{H}_{int} = 0$). Since $h(\mathbf{r})$ is a real quantity, the amplitudes must obey condition $h_{\mathbf{k}}^{*} = h_{-\mathbf{k}}$. In addition, we generally discretize $h(\mathbf{r})$ to avoid handling an infinite number of modes and to coarse-grain over microscopic details. Let the height field be discretized as an $N \times N$ matrix with a lattice spacing $\ell = L/N$ so that the allowed wave vectors are $\mathbf{k} = (m, n)2\pi/L$ with $-N/2 < m, n \leq N/2$. To proceed with any calculations or simulations, it is first necessary to find N^2 independent variables with which to work.

The discrete Fourier transform of a real $N \times N$ matrix has four explicitly real modes if N is even, and one explicitly real mode if N is odd. We consider only the case where N is even (the procedure for odd N follows in a straightforward fashion). The explicitly real modes are $(m, n) = (0, 0)$, $(N/2, 0)$, $(0, N/2)$, and $(N/2, N/2)$. Note that the center of mass h_{cm} is related to the $(m, n) = (0, 0)$ mode by the equation

$$h_{\mathrm{cm}} = \frac{h_{\mathbf{k}=0}}{L^2} \tag{30}$$

(In this work we always consider systems with a fixed center of mass and the $(m, n) = (0, 0)$ mode is not considered a dynamical variable.) We label the remaining three explicitly real modes as $\mathbf{q}_r$: $(m, n) = (N/2, 0)$, $(0, N/2)$, and $(N/2, N/2)$. The remaining independent modes are chosen based on the condition $h_{\mathbf{k}}^{*} = h_{-\mathbf{k}}$. We label these $N^2/2-2$ modes $\mathbf{q}_c$: (m, n) for $-N/2 < m < N/2$ and $0 < n < N/2$, $(m, 0)$ for $0 < m < N/2$, $(m, N/2)$ for $0 < m < N/2$, and $(N/2, n)$ for $0 < n < N/2$. These modes are illustrated graphically in Figure 1.

Defining the real and imaginary parts of the amplitudes to be $h_{\mathbf{k}} \equiv a_{\mathbf{k}} + ib_{\mathbf{k}}$, the set of N^2 independent modes can be listed as $(\{a_{\mathbf{q}_c}\},\{a_{\mathbf{q}_r}\},\{b_{\mathbf{q}_c}\})$. We define a vector of length $N^2 - 1$

$$\mathbf{q} \equiv (\{\mathbf{q}_c^{\mathrm{real}}\}, \{\mathbf{q}_r\}, \{\mathbf{q}_c^{\mathrm{imag}}\}) \tag{31}$$

as a way of indexing these independent modes. Although the sets $\{\mathbf{q}_c^{\mathrm{real}}\}$ and $\{\mathbf{q}_c^{\mathrm{imag}}\}$ both span the same values as $\{\mathbf{q}_c\}$, the superscript labels are used for clarity. Note that $\mathbf{k}$ is the Fourier index, while $\mathbf{q}$ is just a vector for the independent modes defined by eq 31.

All of the above also applies to $\zeta(\mathbf{r})$. Defining $\xi_{\mathbf{k}} \equiv f_{\mathbf{k}} + ig_{\mathbf{k}}$, the set of independent random forces is $(\{f_{\mathbf{q}_c}\},\{f_{\mathbf{q}_r}\},\{g_{\mathbf{q}_c}\})$. Using eq 6, the fluctuation−dissipation relations become

$$\langle\, f_{\mathbf{q}_c}(t)\rangle = \langle\, f_{\mathbf{q}_r}(t)\rangle = \langle g_{\mathbf{q}_c}(t)\rangle = 0$$

$$\langle\, f_{\mathbf{q}_c}(t) f_{\mathbf{q}_c}(t')\rangle = \langle g_{\mathbf{q}_c}(t) g_{\mathbf{q}_c}(t')\rangle = k_B T L^2 \Lambda_{\mathbf{q}_c}^{-1}\delta(t - t')$$

$$\langle\, f_{\mathbf{q}_r}(t) f_{\mathbf{q}_r}(t')\rangle = 2k_B T L^2 \Lambda_{\mathbf{q}_r}^{-1}\delta(t - t') \tag{32}$$

All cross correlations are equal to zero.

V. Harmonic Interactions

We focus on a physical situation for which a harmonic form of the interaction potential is appropriate. Specifically, we

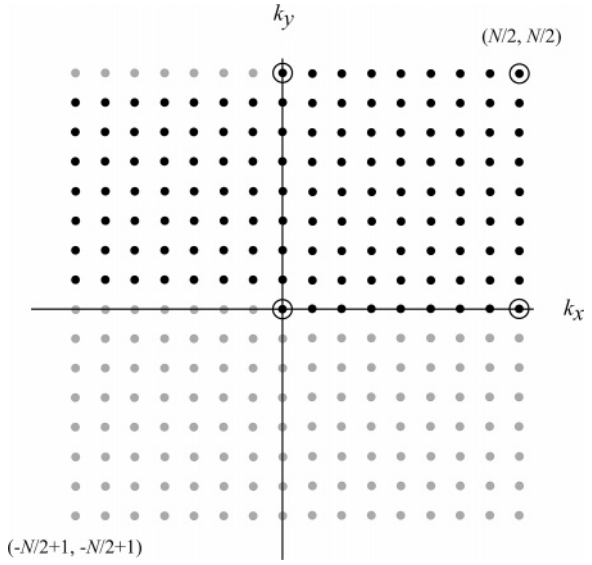


Figure 1. Plot of the N^2 modes (for N even) in **k** space. The four circled, black dots are the real modes consisting of the center of mass mode, $\mathbf{k} = 0$, and the other three modes $\{\mathbf{q}_r\}$. The remaining black dots are the independent, complex modes $\{\mathbf{q}_c\}$, while the gray dots are the dependent, complex conjugates of $\{\mathbf{q}_c\}$. Each $\mathbf{q}_r$ contributes one independent variable, while each $\mathbf{q}_c$ contributes two independent variables. The (m, n) values for the upper-right and lower-left modes are shown in the plot to define the boundaries in k_x and k_y space.

are interested in interactions that strongly bind the bilayer to a height of z_i at localized positions $\mathbf{R}_i$ in the plane of membrane. Such a pinning interaction is described by the equation

$$\mathcal{H}_{\text{pin}}[h(\mathbf{r})] = \frac{\gamma}{2}\sum_i [h(\mathbf{r}) - z_i]^2 \delta(\mathbf{r} - \mathbf{R}_i) \tag{33}$$

with a resulting force

$$F_{\text{pin}}(\mathbf{r}) = -\gamma \sum_i [h(\mathbf{r}) - z_i]\delta(\mathbf{r} - \mathbf{R}_i) \tag{34}$$

We discuss here and in the Appendix a method for obtaining averages and time correlation functions without the need for simulations. The procedure described here is an extension of previous work.[32] In Fourier space, the full set of equations can be written as

$$H = \frac{1}{2L^2}\left(\sum_{\mathbf{k},\mathbf{k}'} h_{\mathbf{k}}^* M'_{\mathbf{k}\mathbf{k}'} h_{\mathbf{k}'} + 2\sum_{\mathbf{k}} m'_{\mathbf{k}} h_{\mathbf{k}}\right)$$

$$\frac{\partial h_{\mathbf{k}}(t)}{\partial t} = \Lambda_{\mathbf{k}}\left(-\sum_{\mathbf{k}'} M'_{\mathbf{k}\mathbf{k}'} h_{\mathbf{k}'}(t) - m'_{\mathbf{k}} + \zeta_{\mathbf{k}}(t)\right) \tag{35}$$

where

$$M'_{\mathbf{k}\mathbf{k}'} = (\kappa k^4 + \sigma k^2)\delta_{\mathbf{k},\mathbf{k}'} + \frac{\gamma}{L^2}\sum_i e^{-i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{R}_i}$$

$$m'_{\mathbf{k}} = \gamma \sum_i (h_{\text{cm}} - z_i)\, e^{i\mathbf{k}\cdot\mathbf{R}_i} \tag{36}$$

and the sums exclude $\mathbf{k} = 0$. In this work, we are interested exclusively in systems with a fixed center of mass

$$h_{\text{cm}} = \frac{h_{\mathbf{k}=0}}{L^2} \tag{37}$$

which appears explicitly as a constant in $m'_{\mathbf{k}}$.

Since the modes in the above equations are mixed, we use matrix diagonalization to decouple the amplitudes $h_{\mathbf{k}}$. This procedure is somewhat complicated and is relegated to the Appendix. After diagonalization, the equations can be written in terms of the eigenmodes d_j, the eigenvalues ω_j, and the orthogonal transformation matrix **U**. The completely decoupled equations are

$$H = \frac{1}{L^2}\sum_j (\omega_j d_j^2 + 2n_j d_j) \tag{38}$$

$$\frac{\partial d_j(t)}{\partial t} = -\omega_j d_j(t) - n_j + s_j(t) \tag{39}$$

where n_j is a time-independent constant defined in the Appendix. The quantities $s_j(t)$ is related to the original random forces in eq 6 and obey the fluctuation−dissipation relations

$$\langle s_j(t)\rangle = 0$$

$$\langle s_j(t)s_j(t')\rangle = k_B T L^2 \delta(t - t') \tag{40}$$

The above equations define a set of independent Ornstein-Uhlenbeck processes[41] for the evolution of the harmonically pinned membrane surface. As such, it is a simple matter to simulate the stochastic behavior of the membrane by drawing random numbers from suitable Gaussian distributions. We do not pursue this type of simulation in the present work but instead calculate the average membrane shape, fluctuations, and dynamics analytically from the normal mode decomposition. Averages and time correlations for the eigenmodes d_j can be calculated in a way similar to the case of the noninteracting membrane. The results are

$$\langle d_j\rangle = -\frac{n_j}{\omega_j} \tag{41}$$

$$\langle d_j^2\rangle = \frac{k_B T L^2}{2\omega_j} \tag{42}$$

$$\langle d_j(\text{t}) d_j(0)\rangle - \langle d_j\rangle^2 = \sigma_j^2\, e^{-\omega_j t} \tag{43}$$

where $\sigma_j^2 \equiv \langle d_j^2\rangle - \langle d_j\rangle^2$. The equations lead to the average height and time correlation functions

$$\langle h(\mathbf{r})\rangle = h_{\text{cm}} + \sum_j \upsilon_j(\mathbf{r})\,\langle d_j\rangle$$

$$\langle h(\mathbf{r}, t)h(\mathbf{r}, 0)\rangle - \langle h^2(\mathbf{r})\rangle = \sum_j \upsilon_j^2(\mathbf{r})\,\sigma_j^2\, e^{-\omega_j t} \tag{44}$$

where $\upsilon_j(\mathbf{r})$ is defined in the Appendix.

VI. Nonharmonic Interactions

We now discuss the more general case in which the physical interaction cannot be written in terms of a harmonic potential.

In this situation, a simulation must, in general, be performed. We describe a Fourier Space Brownian Dynamics (FSBD) method[33,34] for time evolving the membrane based on standard Brownian dynamics.[43] The essential difference between the two methods is that, rather than using the position space variables, the amplitudes $h_{\mathbf{k}}$ are evolved instead. By making this choice, the computationally expensive convolution in eq 2 is avoided.

We first integrate eq 12 from t to $t + \Delta t$ for small Δt to get

$$h_{\mathbf{k}}(t + \Delta t) = h_{\mathbf{k}}(t) + \Lambda_{\mathbf{k}} F_{\mathbf{k}}(t)\Delta t + \Gamma_{\mathbf{k}}(\Delta t)$$

$$\Gamma_{\mathbf{k}}(\Delta t) \equiv \Lambda_{\mathbf{k}} \int_t^{t+\Delta t} dt'\, \zeta_{\mathbf{k}}(t') \tag{45}$$

Exactly N^2 independent random values must be chosen to properly construct $\Gamma_{\mathbf{k}}(\Delta t)$. We therefore restrict ourselves to the modes $\mathbf{q}$ defined eq 31. Using eq 32, it can be shown that the real and imaginary parts of $\Gamma_{\mathbf{q}_c}(\Delta t)$ are both drawn from a Gaussian distribution of mean zero and variance $k_B T L^2 \Lambda_{\mathbf{k}} \Delta t$, while $\Gamma_{\mathbf{q}_r}(\Delta t)$ is drawn from Gaussian distribution of mean zero and variance $2k_B T L^2 \Lambda_{\mathbf{k}} \Delta t$. After picking random values for $\Gamma_{\mathbf{q}}(\Delta t)$, the full matrix for $\Gamma_{\mathbf{k}}(\Delta t)$ can then be reconstructed.

The FSBD method is summarized below:

1. Evaluate the interaction part of the forces $F_{\text{int}}(\mathbf{r}) = -\delta \mathscr{H}_{\text{int}}/\delta h(\mathbf{r})$ in position space.

2. Compute the bending forces $-\kappa k^4 h_{\mathbf{k}}$ and surface tension forces $-\sigma k^2 h_{\mathbf{k}}$. Evaluate the interaction force $F_{\mathbf{k}}^{\text{int}}$ by Fourier transforming the result of the previous step.

3. As described above, draw random values from the appropriate Gaussian distributions for the independent modes only. Construct $\Gamma_{\mathbf{k}}(\Delta t)$ from these values.

4. Compute $h_{\mathbf{k}}(t + \Delta t)$ using eq 45. Inverse Fourier transform $h_{\mathbf{k}}$ to obtain $h(\mathbf{r})$ for use in the next iteration.

It is essential to choose a sufficiently small time step Δt such that the results have converged.

VII. Application of Harmonic Dynamics: Fluctuation Time Scales for Intermembrane Junctions

Lipid bilayers supported by solid substrates have become important tools for studying various membrane-related biophysical processes.[44] These bilayers are either directly fixed to the surface or are separated from the substrate by a thin layer of water or a polymer cushion. Membranes that are not directly attached to the surface retain their fluid properties,[45] making them useful as model systems for studying processes, such as the formation of the immunological synapse,[46,47] that involve the free diffusion of components in the bilayer.

In this section, we focus on a recent experimental study of supported intermembrane junctions by Kaizuka and Groves.[37] In these experiments, giant unilamellar vesicles were ruptured over a previously deposited planar supported membrane surface. The ruptured vesicles were observed to adhere to the underlying supported bilayer in two distinct fashions (see Figure 2). Either the two bilayers adhere to one another with a nearly uniform separation of a few nanometers ("Type 1" intermembrane junction), or the two bilayers form an irregular interface with average separation on the order of 50 nanometers held together by a sparse randomly distributed network of pinning sites where the two bilayers closely approach one another ("Type 2" intermembrane junction). In the Type 2 junctions, an appreciable volume of water is effectively trapped between upper and lower bilayers, making these systems a natural realization of hydrodynamic flow next to an impermeable wall. The reason for the formation of these distinct structures and the nature of the localized pinning in the Type 2 systems remain poorly understood. Kaizuka and Groves[37] have studied the average height and dynamic fluctuations of Type 2 junctions via fluorescence interference contrast (FLIC) microscopy,[48,49] and it is these data that we discuss here.

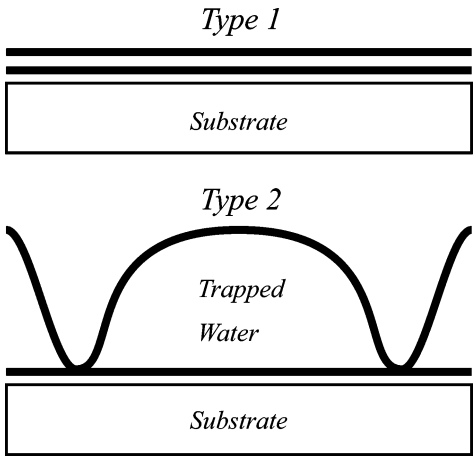


Figure 2. Illustration of the two types of junctions (adapted from ref 37). The thick line on the bottom and directly over the substrate represents the supported lipid bilayer. The giant unilamellar vesicle that has been ruptured over this planar bilayer is represented by another thick line above it. Type 1 junctions have intermembrane spacings of about 2−3 nm compared to Type 2 junctions which can have spacings of about 50 nm as well as localized adhesion sites.

We model this system by assuming that the upper membrane surface is locally pinned to the supported bilayer in a geometry directly inferred from experimental data—we make no attempt to account for the physical basis of this random adhesion. This localized pinning is achieved by using the potential in eq 33 with γ sufficiently large to fix the membrane at height $z_i = 0$ at the experimentally determined pinning sites $\mathbf{R}_i$ (i.e. we define the height of the upper bilayer to be zero ($h(\mathbf{R}_i) = 0$) where the pinning occurs). For the purposes of this simulation, the lower bilayer is an impermeable wall located at $z = 0$. The system size $L = N\ell \sim 5.4$ μm, with $N = 64$ and $\ell \sim 85$ nm correspond with the size and resolution obtained in experiment. We account for the trapped volume of water in the junction by setting h_{cm} to a positive constant value. This is equivalent to a constant volume condition for our periodic boundary condition simulations. The distance d of the membrane from the wall in the expression for $\Lambda_{\mathbf{k}}$ is therefore set equal to h_{cm}. We also assume that the interactions responsible for pinning the membrane to the supported bilayer dictate the slope of the membrane around the pin. To model this boundary condition

Table 1. Model Parameters for the Intermembrane Junction

parameter	description	value	reference
κ	bending modulus	7×10^{-12} ergs	*a*
σ	surface tension	0.08 dyn/cm	*b*
η	viscosity	0.01 poise	water
T	temperature	20 °C	*c*
L	system size	5.4 μm	*d*
l	lattice spacing	85 nm	*d*
s	slope around pinned site	0.29	*b*
h_{cm}	center of mass	46 cm	*b*
γ	pinning constant	10^8 ergs cm^{-2}	*e*

[a] Personal communication with Kaizuka and Groves.[37] [b] Fit to data of ref 37. [c] Reference 37. [d] From data provided by ref 37. [e] Sufficiently large to fix the membrane at the pinning site. Simulations with larger values of γ give identical results.

within our harmonic scheme, we harmonically bind the four sites immediately adjacent to the pin location at a height of $z_i = h_n$ using eq 33, which leads to a slope $s = h_n/\ell$. The average shape of the junction and magnitude of fluctuations can then be calculated using eq 44 assuming the experimental conditions of $T = 293$ K and $\eta = 0.01$ poise and implementing either of the two hydrodynamic kernels previously discussed.

The values chosen for the physical parameters h_{cm}, s, κ, and σ must be obtained by fitting to the experimental data, a process significantly complicated by the limited lateral resolution of the FLIC technique (i.e. the experimental data reflects a convolution in space over the point spread function of the laser). The procedure involved is beyond the scope of this work, so we simply quote the results for our parameters in Table 1. A detailed account of the agreement between experiment and theory for the average shape and fluctuations of the junction will be presented elsewhere.[50] The present work is concerned with the impact of different hydrodynamic kernels on relaxation time scales, and we focus on this question here. To give an idea of the shapes generated by these simple elastic considerations we do present the shape and fluctuations for a representative arrangement of pinning sites (Figure 3).

Time correlation functions for height fluctuations in this system have been measured experimentally with observed decay times in the range of hundreds of milliseconds to seconds[38] depending on the particular junction probed and the location of the observed fluctuation within the pinned geometry. Since the supported planar bilayer should act as an impermeable boundary to fluid flow, the hydrodynamic kernels in eqs 24 and/or 26 are expected to apply.

In Figure 4, we plot the predicted time correlation function at several locations distributed over the membrane surface and using both suggested hydrodynamic kernels. Our results show relaxation times in qualitative agreement with experimental results. The difference between results employing Seifert's and Gov's hydrodynamic boundary conditions are appreciable but not so striking as to rule one form out in favor of the other on the basis of available experimental data. What is clear is that the inclusion of the impermeable wall in some fashion is necessary to set relaxation time scales close to experimental values. Blindly applying the standard result in eq 22 for a membrane in an infinite fluid medium leads to relaxation times orders of magnitude too fast relative to experiment.

For comparison, we also plot the results without hydrodynamics by using a constant value $\Lambda_{\mathbf{k}} = 1/4\eta k_0$ for all $\mathbf{k}$, where $k_0 = 2\pi/L$ is the longest wavelength mode in the system. The choice of this constant value ensures that the longest wavelength mode evolves as it would with hydrodynamics. The other modes, however, evolve with a kernel that is k/k_0 times larger than the hydrodynamical value. For concreteness and simplicity, we first study the relaxation frequencies of the membrane without interactions in eq 9. The decay time for each mode is proportional to $1/\Lambda_{\mathbf{k}}$ so that the shorter wavelength modes have shorter decay times when the constant value of $1/4\eta k_0$ is used. This analysis should also be approximately true in the case where there are interactions. The effect of faster decay by the shorter wavelength modes can be seen over the shorter times scales in Figure 4.

The time scales seen in our figures are in general agreement with a more simplified approach to analyzing the

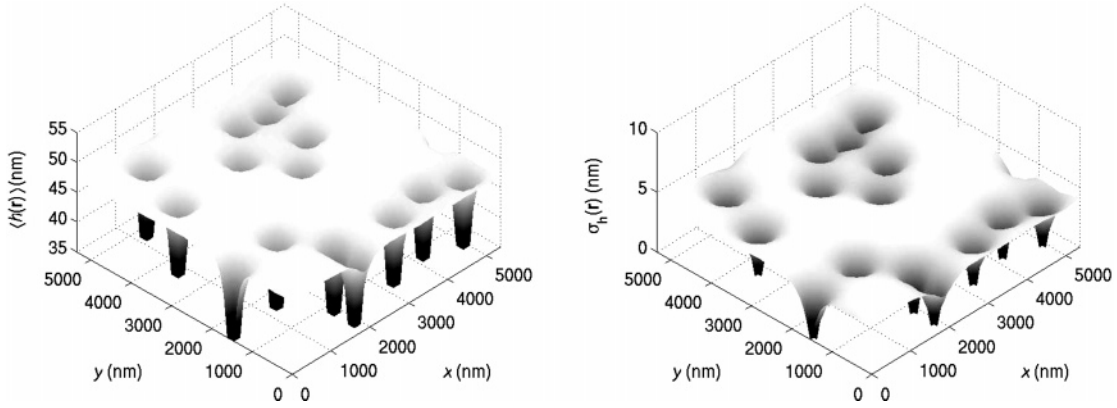


Figure 3. Average bilayer shape and fluctuations for a representative configuration of pinning sites in a Type 2 intermembrane junction. The results were obtained using the diagonalization procedure of section V with the physical constants listed in Table 1. The left panel shows the results for the average height $\langle h(\mathbf{r})\rangle$. Note that the bottom of the wells are pinned at a height of 0 (outside the range of the plot). The right panel shows the size of the height fluctuations $\sigma_h(\mathbf{r}) \equiv \sqrt{\langle h^2(\mathbf{r})\rangle - \langle h(\mathbf{r})\rangle^2}$. Both left and right panels display the data with z axis significantly expanded relative to x, y to show detail.

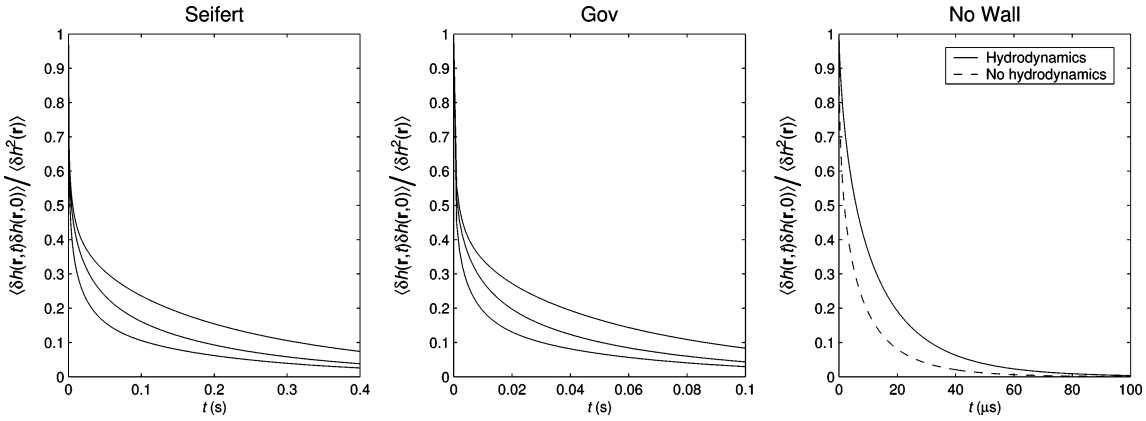


Figure 4. Plot of typical time correlations functions $\langle \delta h(\mathbf{r}, t)\delta h(\mathbf{r}, 0)\rangle$, where $\delta h(\mathbf{r}, t) \equiv h(\mathbf{r}, t) - \langle h(\mathbf{r})\rangle$, for different boundary conditions. The results are shown for Seifert's boundary condition in the left panel and Gov's boundary condition in the middle panel. The three curves indicate roughly the range of decay times for different locations on the membrane away from the pinning sites. In the right panel, a typical correlation function for a hydrodynamic kernel assuming an infinite fluid environment (no wall) is shown. For comparison, results without hydrodynamics are also plotted. In that case, we use a constant value $\Lambda_{\mathbf{k}} = 1/4\eta k_0$, where $k_0 = 2\pi/L$ is the longest wavelength mode in the system. The time scales for relaxation when there is no wall are orders of magnitude faster than the observed experimental decay times of hundreds of milliseconds to seconds.[38] Incorporation of the proper hydrodynamic kernel is necessary to model the experiment.

data suggested by Kaizuka and Groves.[38] In that work, Gov's form for the hydrodynamic kernel was employed, but an empirical Hamiltonian was introduced to constrain the upper membrane at a fixed height from the support. As far as the dynamics are concerned, the two methods produce similar results. The present study has the advantage of fitting all available experimental data to an elastic model that is physically motivated (albeit with an unexplained pinning interaction). We stress that the difference between the predicted relaxation rates for the $\Lambda_{\mathbf{k}}$ with and without the wall is several orders of magnitude. The presence of a wall slows the dynamics of the membrane significantly and must be included for proper modeling of experiment.

VIII. Application of Anharmonic Dynamics: Red Blood Cell Protein Mobility

In previous work,[33,34,51] we have studied the effect of thermal membrane undulations on the lateral diffusivity of band 3 protein over the surface of the red blood cell. Band 3 is known to diffuse freely over length scales less than ~100 nm, but its motion is hindered over larger distances.[52−56] The reason for the slower diffusion on longer length scales is the presence of the cell's cytoskeleton, which consists of roughly triangular corrals formed by spectrin filaments[57,58] attached to the membrane in a quasi-regular geometry. The microscopic diffusion constant,[59] $D = 0.53\ \mu\text{m}^2\ \text{s}^{-1}$, on scales smaller than the size of the corrals,[59] $L_c \sim 110$ nm, is in good agreement with estimates for protein diffusivity in a viscous quasi-two-dimensional environment.[60] However the macroscopic diffusion constant, relevant over longer length scales, takes on a value, $D_{\text{macro}} = 6.6 \times 10^{-3}\ \mu\text{m}^2\ \text{s}^{-1}$, 2 orders of magnitude smaller.[59,61−66]

Although the cytoskeleton hinders protein motion, band 3 does manage to escape confinement as evidenced by the finite macroscopic diffusion coefficient. One possible mechanism for this escape process is that thermal membrane fluctuations may help the protein to escape local confinement by lifting the cytoplasmic domain of the protein over the top of the cytoskeletal barriers.[33,34,51] Using methods similar to those of our prior studies, we calculate the expected D_{macro} resultant from fluctuations in bilayer shape. In contrast to previous work we include in this study a hydrodynamic kernel that approximately accounts for the hampered flow of cytoplasm (and hence slowed motion of the bilayer) due to the presence of spectrin filaments.

We model the interactions of the membrane with the cytoskeleton with a combination of two potentials. First, the bilayer is pinned to the membrane at discrete points using eq 33 to mimic the attachment of the membrane to the spectrin filaments. Second, the steric interaction between the membrane and the cytoskeleton is modeled with the repulsive hydration potential[67]

$$\mathscr{H}_{\text{rep}}[h(\mathbf{r})] = \epsilon\, e^{-h(\mathbf{r})/\lambda} \sum_i \exp\left\{-\left(\frac{a_i x + b_i y + c_i}{\ell/4}\right)^2\right\} \qquad (46)$$

where $a_i x + b_i y + c_i = 0$ specifies a particular linear segment of spectrin between pinning sites. The repulsive interaction is localized to this set of lines with a width of about $\ell = 7$ nm. The physical parameters used for this system are shown in Table 2.

Since the repulsive interaction in eq 46 is not harmonic, we must perform a simulation of the membrane to compute averages of interest, as detailed in section VI. In Figures 5 and 6, we show sample configurations of two different geometries generated using the FSBD simulation method. Since altering the hydrodynamics of the system only affects dynamics and not thermal statistics, these configurations are consistent with what was seen in our prior studies without the inclusion of hydrodynamic effects imposed by spectrin (see below).

We develop a model of protein mobility that incorporates thermal fluctuations of the lipid bilayer. Band 3 protrudes into the cellular interior a distance of $h_0 = 6$ nm[69] and

Table 2. Model Parameters for the Red Blood Cell

parameter	description	value	reference
κ	bending modulus	2×10^{-13} ergs	*a*
σ	surface tension	0	*b*
η^+	water viscosity	0.01 poise	water
η^-	cytoplasm viscosity	0.06 poise	*a*
T	temperature	37° C	body temp
h_0	depth of cytoplasmic domain of band 3	6 nm	*c*
D	band 3 diffusion constant	0.53 μm^2 s^{-1}	*d*
L_c	corral size	112 nm	*d*, *e*
l	lattice spacing	7 nm	*e*
t_D	random walk time step	23 μs	*e*
γ	pinning constant	100 ergs cm^{-2}	*e*
ϵ	repulsive potential energy scale	8.7×10^{-4} ergs cm^{-2}	*f*
λ	repulsive potential length scale	0.2 nm	*g*

[a] Reference 23. [b] Reference 68. [c] Reference 69. [d] Reference 59. [e] Reference 32. [f] References 33 and 34. [g] Reference 67.

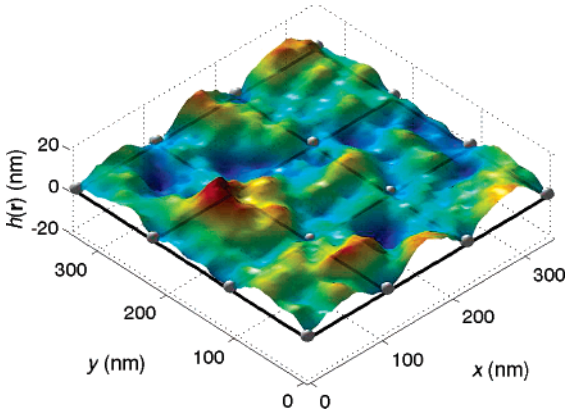


Figure 5. Sample configuration for a membrane with square pinning and cytoskeletal repulsion. The pinning sites are indicated by spheres and the repulsive interaction due to spectrin is localized along the black lines which connect between the pinning sites. The *z*-axis is expanded to help visualize fluctuations in the membrane.

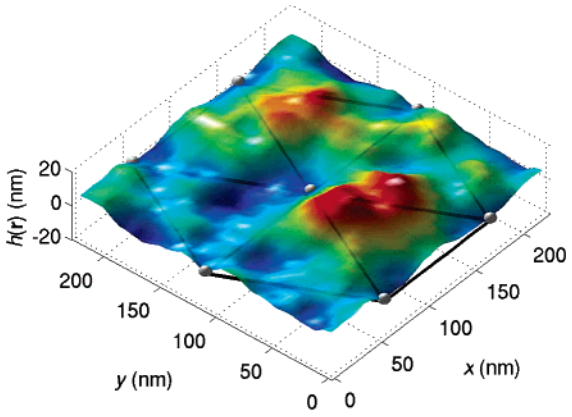


Figure 6. The same plot as in Figure 5, except for a triangular geometry.

interacts sterically with the cytoskeleton as it diffuses within the corral. The first requirement for band 3 escape from the corral is that the separation between spectrin and bilayer locally exceed h_0, so that there is a large enough gap for the protein to slip through. Second, the height fluctuation should persist long enough for the protein to diffuse a distance equal to the width of the spectrin filament (approximately $\ell = 7$ nm). On average, the time to diffuse this distance is $t_D = \ell^2/4D = 23\ \mu$s. An escape probability along the corral edge can be calculated by using FSBD to find correlated probability that $h(\mathbf{r}) > h_0$ at both times 0 and t_D. Knowledge of this escape probability leads to the macroscopic diffusion constant D_{macro} as detailed in our previous studies.[33,34]

In previous work,[32−34,51] we assumed that the red blood cell membrane was fluctuating in an unbound fluid characterized by the viscosity of the red cell cytoplasm. In recent work by Gov et al.,[26] it was proposed that the cytoskeleton acts as a source of confinement of hydrodynamic flow. The use of the hydrodynamic kernel with a wall in eq 26 provided improved fits to experimental dynamic membrane fluctuation data relative to models incorporating infinite hydrodynamic boundaries. Since we choose to use the value $d = 35$ nm determined using the hydrodynamic kernel proposed by Gov et al.,[26] we use that version of the kernel (eqs 26 and 29) in our simulations.

We illustrate the effect of the wall by studying the probability that the membrane height lies above h_0 at time t given that it was above h_0 at time 0 defined by

$$C(\mathbf{r}, t) = \frac{\langle \Theta(h(\mathbf{r}, t) - h_0)\Theta(h(\mathbf{r}, 0) - h_0)\rangle - P^2(\mathbf{r})}{P(\mathbf{r}) - P^2(\mathbf{r})} \qquad (47)$$

where $P(\mathbf{r})$ is the equilibrium probability that $h(\mathbf{r}) > h_0$ given by

$$P(\mathbf{r}) = \langle \Theta(h(\mathbf{r}) - h_0)\rangle \qquad (48)$$

and Θ is the Heaviside step function. The escape probability, and therefore D_{macro}, depends on the value of $P(\mathbf{r})C(\mathbf{r}, t_D)$. In Figure 7, we plot a comparison of $C(\mathbf{r}, t)$ for two different hydrodynamic kernels. As expected, the membrane relaxation is slowed in the presence of an impermeable wall.

Historically, we have calculated D_{macro} using $P(\mathbf{r})C(\mathbf{r}, t_D)$ because analytical results are available for harmonic potentials. Using FSBD however, it is possible to calculate the probability that the membrane remain above h_0 for the entire time interval t_D, defined to be

$$P_{open}(\mathbf{r}, t) = \langle \delta(t_D - \int_0^{t_D} ds\ \Theta\ (h(\mathbf{r}, s) - h_0))\rangle \qquad (49)$$

In our case, we require that $h(\mathbf{r}, t) > h_0$ at each time step Δt

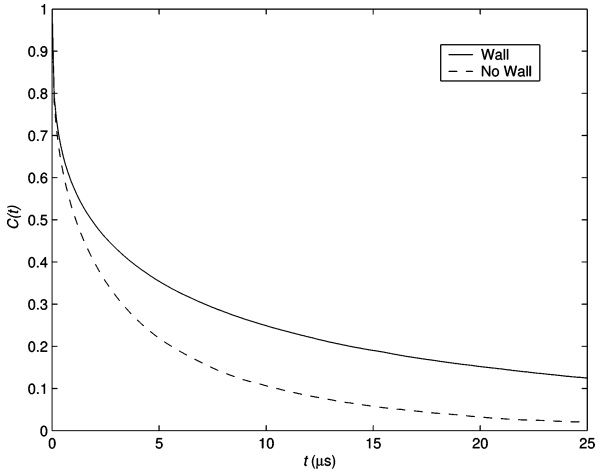


Figure 7. Plot of $C(\mathbf{r}, t)$ in eq 47, the correlated probability that the membrane height is greater than h_0 at time t given that it was greater than h_0 at time 0. The results shown are for the point $\mathbf{r} = (L_o/2, 0)$ in a square geometry. The relaxation of the membrane in the presence of the wall is slower than that of a membrane in an unbound fluid.

Table 3. Results for D_{macro} for Various Cases[a]

	$1000D_{\text{macro}}$ (μm^2 s^{-1}) using $P(\mathbf{r})C(\mathbf{r}, t_D)$	$1000D_{\text{macro}}$ (μm^2 s^{-1}) using $P_{\text{open}}(\mathbf{r}, t_D)$
experiment	6.6	
square, $\eta^+ = \eta^-$, no wall	37.8 ± 2.0	2.83 ± 0.25
triangle, $\eta^+ = \eta^-$, no wall	18.8 ± 5.3	2.02 ± 0.38
square, $\eta^+ = \eta^-$, wall	188 ± 3	20.3 ± 1.2
triangle, $\eta^+ = \eta^-$, wall	155 ± 3	18.0 ± 0.9
square, $\eta^+ \neq \eta^-$, no wall	7.40 ± 1.7	0.151 ± 0.048
triangle, $\eta^+ \neq \eta^-$, no wall	6.12 ± 0.17	0.0560 ± 0.0283
square, $\eta^+ \neq \eta^-$, wall	43.5 ± 1.9	0.859 ± 0.102
triangle, $\eta^+ \neq \eta^-$, wall	20.5 ± 2.2	0.490 ± 0.105

[a] The FSBD simulations were performed over a time of $t = 30$ ms using a time step of $\Delta t = 1$ ns. When the viscosities above and below the membrane are the same, we use the cytoplasmic viscosity $\eta = \eta^-$ (see Table 2) consistent with our prior work.[32−34,51]

in our simulation. It should be emphasized that, in the absence of detailed modeling of the interactions between the membrane, cytoskeleton, and protein, both $P(\mathbf{r})C(\mathbf{r}, t_D)$ or $P_{\text{open}}(\mathbf{r}, t)$ serve as approximations for the calculation of D_{macro}. It is not clear a priori which approximation should serve as a closer mimic to experiment.

The values of the macroscopic diffusion constant for various scenarios are shown in Table 3. We note several observations for the different scenarios listed. First, as a result of the increased temporal persistence of fluctuations due to the wall, the macroscopic diffusion constants are larger relative to the case of a membrane in an unbound fluid. Second, using the viscosity of water for η^+ so that $\eta^+ < \eta^-$ (a more realistic scenario for the case of a red blood cell in a biological environment) decreases the relaxation time for the membrane and therefore decreases D_{macro}. Finally, the use of $P_{\text{open}}(\mathbf{r}, t_D)$ is a more restrictive condition and decreases the macroscopic diffusion constant as compared to the value obtained using $P(\mathbf{r})C(\mathbf{r}, t_D)$ (in general $P_{\text{open}}(\mathbf{r}, t_D) < P(\mathbf{r})C(\mathbf{r}, t_D)$).

In previous studies, we have only obtained the results in the top two rows of Table 3. The last two rows represent the most accurate scenario assuming that the hydrodynamic flow is confined on the length scales in our system (the original study by Gov et al.[26] showed better fits to fluctuation data for length scales in the range of hundreds of nanometers to microns, a range larger than for our relevant length scale of $\sim L_c$). In the case where $P(\mathbf{r})C(\mathbf{r}, t_D)$ is used, the increase in decay times due to the wall is roughly canceled by a decrease due to the smaller viscosity of water. The final values for the last two rows is approximately the same as for the first two rows. For the case where we use $P_{\text{open}}(\mathbf{r}, t_D)$, the values in the last two rows are about 3 or 4 times smaller than the first two rows. This decrease results from the fact that using $\eta^+ \neq \eta^-$ has a more drastic effect than in the previous case and is not canceled by the presence of the wall. Short time fluctuations are important in the case where we require the membrane to be above h_0 for the entire interval, and therefore it is not surprising that decreasing the viscosity above the bilayer leads to a larger reduction in the value of the macroscopic diffusion constant. In the scenario of the last two rows, as for the results of the previous study, the experimental value, $D_{\text{macro}} = 6.6 \times 10^{-3}$ μm^2 s^{-1}, lies between the values obtained using the two different methods. In this sense, the contributions of the wall and unequal viscosities largely cancel one another.

IX. Conclusion

The computational limitations of molecularly based simulation schemes dictate that simpler models be developed to study phenomena over the length and time scales relevant to cellular biology and related physical systems. In this work we have extended the generality of the FSBD approach for modeling elastic membrane sheets by incorporating the possibility of noninfinite hydrodynamic environments. Although the resulting equations and algorithms are largely unaltered from the original FSBD methodology, the observation that generalized hydrodynamic kernels may be incorporated within this approach greatly enhances the practical utility of FSBD. We have presented two applications that clearly demonstrate the considerable influence of hydrodynamic conditions on membrane dynamics. Given the finite size and complex structures inherent to cellular systems, it is clear that noninfinite boundary conditions are a necessity to properly model most dynamics relevant to biology at the cellular scale.

Acknowledgment. This work was supported by the NSF (CHE-0349196, CHE-0321368). F.B. is an Alfred P. Sloan research fellow. We thank Jay Groves and Yoshihisa Kaizuka for providing us with their data. We also thank Rony Granek for helpful discussions.

Appendix

We describe in more detail the diagonalization procedure used in the Harmonic Interactions section. The first step is to write eq 35 in terms of the N^2 independent modes $(\{a_{\mathbf{q}_c}\},\{a_{\mathbf{q}_r}\},\{b_{\mathbf{q}_c}\})$ defined in the section Independent Modes. All quantities in the new equations are explicitly real. It is

then necessary to rescale the modes such that both the Hamiltonian and the equation of motion can be simultaneously diagonalized. The appropriate set of variables is

$$\mathbf{c}^{\mathrm{T}} \equiv \left(\left\{ \frac{a_{\mathbf{q}_c}}{\sqrt{\Lambda_{\mathbf{q}_c}}} \right\}, \left\{ \frac{a_{\mathbf{q}_r}}{\sqrt{2\Lambda_{\mathbf{q}_r}}} \right\}, \left\{ \frac{b_{\mathbf{q}_c}}{\sqrt{\Lambda_{\mathbf{q}_c}}} \right\} \right) \tag{50}$$

with a new random force is similarly defined as

$$\mathbf{p}^{\mathrm{T}} \equiv \left(\left\{ \frac{f_{\mathbf{q}_c}}{\sqrt{\Lambda_{\mathbf{q}_c}}} \right\}, \left\{ \frac{f_{\mathbf{q}_r}}{\sqrt{2\Lambda_{\mathbf{q}_r}}} \right\}, \left\{ \frac{g_{\mathbf{q}_c}}{\sqrt{\Lambda_{\mathbf{q}_c}}} \right\} \right) \tag{51}$$

where each component obeys

$$\langle p_{\mathbf{q}}(t) \rangle = 0$$

$$\langle p_{\mathbf{q}}(t) p_{\mathbf{q}}(t') \rangle = k_B T L^2 \, \delta(t - t') \tag{52}$$

In terms of these new variables, the equations become

$$H = \frac{1}{L^2} (\mathbf{c}^{\mathrm{T}} \mathbf{M} \mathbf{c} + 2\mathbf{m}^{\mathrm{T}} \mathbf{c})$$

$$\frac{\partial \mathbf{c}(t)}{\partial t} = -\mathbf{M}\mathbf{c}(t) - \mathbf{m} + \mathbf{p}(t) \tag{53}$$

where matrix multiplication is implied. The vector $\mathbf{m}$ is defined by

$$\mathbf{m}^{\mathrm{T}} = \frac{\gamma L^2}{2} \sum_i (h_{\mathrm{cm}} - z_i) \, \mathbf{w}(\mathbf{R}_i) \tag{54}$$

where

$$\mathbf{w}(\mathbf{r}) = \frac{2}{L^2} \left(\{ \sqrt{\Lambda_{\mathbf{q}_c}} \cos(\mathbf{q}_c \cdot \mathbf{r}) \}, \left\{ \sqrt{\frac{\Lambda_{\mathbf{q}_r}}{2}} \cos(\mathbf{q}_r \cdot \mathbf{r}) \right\}, \{ -\sqrt{\Lambda_{\mathbf{q}_c}} \sin(\mathbf{q}_c \cdot \mathbf{r}) \} \right) \tag{55}$$

The matrix $\mathbf{M}$ is defined to be

$$\mathbf{M} = \begin{pmatrix} \mathbf{P} + \mathbf{Q} & \sqrt{2}\mathbf{Q} & -\mathbf{R} + \mathbf{S} \\ \sqrt{2}\mathbf{Q} & \mathbf{P} & \sqrt{2}\mathbf{S} \\ \mathbf{R} + \mathbf{S} & \sqrt{2}\mathbf{S} & \mathbf{P} - \mathbf{Q} \end{pmatrix} \tag{56}$$

$$P_{\mathbf{q}\mathbf{q}'} = \sqrt{\Lambda_{\mathbf{q}} \Lambda_{\mathbf{q}'}} \left[(\kappa q^4 + \sigma q^2) \delta_{\mathbf{q},\mathbf{q}'} + \frac{\gamma}{L^2} \sum_i \cos[(\mathbf{q} - \mathbf{q}') \cdot \mathbf{R}_i] \right]$$

$$Q_{\mathbf{q}\mathbf{q}'} = \sqrt{\Lambda_{\mathbf{q}} \Lambda_{\mathbf{q}'}} \left(\frac{\gamma}{L^2} \sum_i \cos[(\mathbf{q} + \mathbf{q}') \cdot \mathbf{R}_i] \right)$$

$$R_{\mathbf{q}\mathbf{q}'} = \sqrt{\Lambda_{\mathbf{q}} \Lambda_{\mathbf{q}'}} \left(-\frac{\gamma}{L^2} \sum_i \sin[(\mathbf{q} - \mathbf{q}') \cdot \mathbf{R}_i] \right)$$

$$S_{\mathbf{q}\mathbf{q}'} = \sqrt{\Lambda_{\mathbf{q}} \Lambda_{\mathbf{q}'}} \left(-\frac{\gamma}{L^2} \sum_i \sin[(\mathbf{q} + \mathbf{q}') \cdot \mathbf{R}_i] \right) \tag{57}$$

where indices along each dimension are as defined in eq 31. For example, the block defined by $\mathbf{P} + \mathbf{Q}$ has dimensions of $(N^2 - 4)/2$ by $(N^2 - 4)/2$ and indices running through $\{\mathbf{q}_c^{\mathrm{real}}\}$ and $\{\mathbf{q}_c^{\mathrm{real}}\}$, the block defined by $-\mathbf{R} + \mathbf{S}$ has dimensions of $(N^2 - 4)/2$ by $(N^2 - 4)/2$ and indices running through $\{\mathbf{q}_c^{\mathrm{real}}\}$ and $\{\mathbf{q}_c^{\mathrm{imag}}\}$, the block defined by $\mathbf{P}$ has dimensions of 3×3 and indices running through $\{\mathbf{q}_r\}$ and $\{\mathbf{q}_r\}$, and the block defined by $\sqrt{2}\mathbf{P}$ has dimensions of $(N^2 - 4)/2 \times 3$ and indices running through $\{\mathbf{q}_c^{\mathrm{real}}\}$ and $\{\mathbf{q}_r\}$ (for the block in the first row, second column in $\mathbf{M}$). Note that the matrix $\mathbf{M}$ is a symmetric matrix resulting from the fact that $\mathbf{P}$, $\mathbf{Q}$, and $\mathbf{S}$ are symmetric and $\mathbf{R}$ is antisymmetric.

Diagonalization of $\mathbf{M}$ yields a transformation matrix $\mathbf{U}$, a set of eigenmodes $\mathbf{d} \equiv \mathbf{U}^{1-}\mathbf{c}$, and a set of eigenvalues ω. Rewriting the equations in terms of these quantities and defining $\mathbf{n} \equiv \mathbf{U}^{-1}\mathbf{m}$ and $\mathbf{s} \equiv \mathbf{U}^{-1}\mathbf{p}$, we derive eq 38. Finally, to compute averages, an expression for $h(\mathbf{r})$ in terms of the eigenmodes is required. By defining $\mathbf{v}(\mathbf{r}) = \mathbf{w}(\mathbf{r})\mathbf{U}$, we can write the height as $h(\mathbf{r}) = \mathbf{v}(\mathbf{r})\mathbf{d}$. This definition of $\mathbf{v}(\mathbf{r})$ is used in eq 44.

References

(1) Brannigan, G.; Lin, L. C.-L.; Brown, F. L. H. *Eur. Biophys. J.* **2006**, *35*, 104.

(2) Shelley, J. C.; Shelley, M. Y.; Reeder, R. C.; Bandyopadhyay, S.; Klein, M. L. *J. Phys. Chem B* **2001**, *105*, 4464.

(3) Marrink, S.; de Vries, A.; Mark, A. E. *J. Phys. Chem. B* **2004**, *108*, 750.

(4) Izvekov, S.; Voth, G. A. *J. Phys. Chem. B* **2005**, *109*, 2469.

(5) Smit, B.; Hilbers, P.; Esselink, K.; Rupert, L.; van Os, N.; Schlijper, A. G. *J. Phys. Chem* **1991**, *95*, 6361.

(6) Goetz, R.; Lipowsky, R. *J. Chem. Phys.* **1998**, *108*, 7397.

(7) Groot, R. D.; Rabone, K. L. *Biophys. J.* **2001**, *81*, 725.

(8) Ayton, G.; Bardenhagen, S. G.; McMurty, P.; Sulsky, D.; Voth, G. A. *J. Chem. Phys.* **2001**, *114*, 6913.

(9) Soddemann, T.; Dunweg, B.; Kremer, K. *Eur. Phys. J. E* **2001**, *6*, 409.

(10) Shillcock, J.; Lipowsky, R. *J. Chem. Phys.* **2002**, *117*, 5048.

(11) Yamamoto, S.; Maruyama, Y.; Hyodo, S.-A. *J. Chem. Phys.* **2002**, *116*, 5842.

(12) Rekvig, L.; Kranenburg, M.; Vreede, J.; Hafskjold, B.; Smit, B. *Langmuir* **2003**, *19*, 4897.

(13) Laradji, M.; Kumar, P. S. *Phys. Rev. Lett.* **2004**, *93*, 198105.

(14) Stevens, M. J. *J. Chem. Phys.* **2004**, ***121***, 11942.

(15) Farago, O. *J. Chem. Phys.* **2003**, *119*, 596.

(16) Brannigan, G.; Brown, F. L. H. *J. Chem. Phys.* **2004**, *120*, 1059.

(17) Brannigan, G.; Philips, P. F.; Brown, F. L. *Phys. Rev. E* **2005**, *72*, 011915.

(18) Cooke, I. R.; Kremer, K.; Deserno, M. *Phys. Rev. E* **2005**, *72*, 011506.

(19) Wang, Z.; Frenkel, D. *J. Chem. Phys.* **2005**, *135*, 234711.

(20) Canham, P. B. *J. Theor. Biol.* **1970**, *26*, 61.

(21) Helfrich, W. *Z. Naturforsch.* **1973**, *28c*, 693.

(22) Safran, S. A. *Statistical Thermodynamics of Surfaces, Interfaces and Membranes*; Westview Press: Boulder, CO, 1994.

(23) Brochard, F.; Lennon, J. F. *J. Phys. (Paris)* **1975**, *36*, 1035.

(24) Granek, R. *J. Phys. II (Paris)* **1997**, *7*, 1761.

(25) Zilman, A. G.; Granek, R. *Phys. Rev. Lett.* **1996**, *77*, 4788.

(26) Gov, N.; Zilman, A. G.; Safran, S. *Phys. Rev. Lett.* **2003**, *90*, 228101.

(27) Granek, R.; Klafter, J. *Europhys. Lett.* **2001**, *56*, 15.

(28) Gouliaev, N.; Nagle, J. F. *Phys. Rev. E* **1998**, *58*, 881.

(29) Gouliaev, N.; Nagle, J. F. *Phys. Rev. Lett.* **1998**, *81*, 2610.

(30) Lipowsky, R.; Zielenska, B. *Phys. Rev. Lett.* **1989**, *62*, 1572.

(31) Weikl, T. R.; Lipowsky, R. *Langmuir* **2000**, *16*, 9338.

(32) Lin, L. C.-L.; Brown, F. L. H. *Biophys. J.* **2004**, *86*, 764.

(33) Lin, L. C.-L.; Brown, F. L. H. *Phys. Rev. Lett.* **2004**, *93*, 256001.

(34) Lin, L. C.-L.; Brown, F. L. H. *Phys. Rev. E* **2005**, *72*, 011910.

(35) Seifert, U. *Phys. Rev. E* **1994**, *49*, 3124.

(36) Gov, N.; Zilman, A. G.; Safran, S. A. *Phys. Rev. E* **2004**, *70*, 011104.

(37) Kaizuka, Y.; Groves, J. T. *Biophys. J.* **2004**, *86*, 905.

(38) Kaizuka, Y.; Groves, J. T. *Phys. Rev. Lett.*, in press.

(39) Purcell, E. M. *Am. J. Phys.* **1977**, *45*, 3.

(40) Doi, M.; Edwards, S. F. *The Theory of Polymer Dynamics*; Clarendon Press: Oxford, 1986.

(41) van Kampen, N. G. *Stochastic Processes in Physics and Chemistry*; North-Holland: Amsterdam, 1992; pp 63, 83, 220−221.

(42) Happel, J.; Brenner, H. *Low Reynolds number hydrodynamics*; Kluwer: The Hague, The Netherlands, 1983.

(43) Ermak, D. L.; McCammon, J. A. *J. Chem. Phys.* **1978**, *69*, 1352.

(44) Sackmann, E. *Science* **1996**, *271*, 43.

(45) Sackmann, E.; Tanaka, M. *Trends Biotechnol.* **2000**, *18*, 58.

(46) Grakoui, A.; Bromley, S. K.; Sumen, C.; Davis, M. M.; Shaw, A. S.; Allen, P. M.; Dustin, M. L. *Science* **1999**, *285*, 221.

(47) Brian, A.; McConnell, H. M. *Proc. Natl. Acad. Sci.* **1984**, *81*, 6159.

(48) Lambacher, A.; Fromhertz, P. *Appl. Phys. A* **1996**, *63*, 207.

(49) Parthasarathy, R.; Groves, J. T. *Cell Biochem. Biophys.* **2004**, *41*, 391.

(50) Lin, L. C.-L.; Brown, F. L. H. Submitted for publication.

(51) Brown, F. L. H. *Biophys. J.* **2003**, *84*, 842.

(52) Cherry, R. J. *Biochim. Biophys. Acta* **1979**, *559*, 289.

(53) Schindler, M.; Koppel, D. E.; Sheetz, M. P. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 1457.

(54) Sheetz, M. P.; Schindler, M.; Koppel, D. E. *Nature* **1980**, *285*, 510.

(55) Koppel, D. E.; Sheetz, M. P.; Schindler, M. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 3576.

(56) Sheetz, M. P. *Semin. Hematol.* **1983**, *20*, 175.

(57) Byers, T. J.; Branton, D. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 6153.

(58) Liu, S.; Derick, L.; Palek, J. *J. Cell. Biol.* **1987**, *104*, 527.

(59) Tomishige, M.; Sako, Y.; Kusumi, A. *J. Cell Biol.* **1998**, *142*, 989.

(60) Saffman, P. G.; Delbruck, M. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *73*, 3111.

(61) Tsuji, A.; Ohnishi, S. *Biochemistry* **1986**, *25*, 6133.

(62) Tsuji, A.; Kawasaki, K.; Ohnishi, S.; Merkle, H.; Kusumi, A. *Biochemistry* **1988**, *27*, 7447.

(63) Edidin, M.; Kuo, S. C.; Sheetz, M. P. *Science* **1991**, *254*, 1379.

(64) Corbett, J. D.; Agre, P.; Palek, J.; Golan, D. E. *J. Clin. Invest.* **1994**, *94*, 683.

(65) Kusumi, A.; Sako, Y. *Curr. Opin. Cell Biol.* **1996**, *8*, 566.

(66) Tomishige, M. Ph.D. Thesis, The University of Tokyo, 1997.

(67) Podgornik, R.; Parsegian, V. A. *Langmuir* **1992**, *8*, 557.

(68) Fournier, J.-B.; Lacoste, D.; Raphael, E. *Phys. Rev. Lett.* **2004**, *92*, 018102.

(69) Zhang, D.; Klyatkin, A.; Bolin, J. T.; Low, P. S. *Blood* **2000**, *96*, 2925.

CT050293S

JCTC Journal of Chemical Theory and Computation

Calculation of Point-to-Point Short-Time and Rare Trajectories with Boundary Value Formulation

Dov Bai and Ron Elber*

Department of Computer Science, Upson Hall 4130, Cornell University, Ithaca, New York 14853

Received January 20, 2006

Abstract: Sampling rare, short-time, and reactive trajectories is of considerable interest in molecular simulations. These trajectories, which are also called "activated", hop between stable states separated by energy or entropy barriers. Simulations of activated trajectories with random sampling of initial conditions are inefficient, since most initial conditions lead to trajectories that do not pass the barrier in short times. A boundary value formulation is proposed that selects these rare trajectories, making the sampling of point-to-point reactive trajectories more effective. Earlier boundary value formulations by one of us focused on computations of approximate trajectories. In the proposed method, trajectories are accurate even when we employ a relatively large integration step (by a factor of about 100 compared to initial value methods). The boundary value solutions to short-time reactive trajectories tend to be unique and have significant statistical weights compared to other reactive trajectories of the microcanonical ensemble. Three numerical examples are considered: a transition in the Mueller potential, a conformational change in alanine dipeptide, and an isomerization in a Lennard-Jones cluster.

I. Introduction

Molecular dynamics simulations are powerful tools in the analysis of microscopic phenomena. Kinetic and thermodynamic properties of matter are studied on the computer by solving microscopic equations of motions. While highly successful in many cases, significant limitations remain. One of these limits is of sampling rare events. In numerous cases only a tiny fraction of the trajectories are "productive" after time t (reaching a desired final state). Because of the rarity of these reactive processes the overall rate of an ensemble of these trajectories (or the average time of reaction) can be very long even if the actual transition is rapid. Therefore, sampling of rare transitional events is computationally demanding. This paper focuses on this challenge and suggests a method for efficient sampling of rare, short-time, and reactive paths. We restrict the discussion to Newtonian's trajectories.

We describe the dynamics of the system with the coordinate vector $X(t)$—the (Cartesian) position of the system at a time instant t. We divide the coordinate (conformational) space into three states. The system can be in the reactant state (R), the product state (P), or in a transitional state (Tr) (Figure 1).

We determine the characteristics of trajectories that started in R at time zero and found in P after total time T.

The most straightforward way to compute reactive trajectories is to sample initial conditions $X(0),Q(0)$ from a starting probability density $\rho_0(X,Q)$ (Q is the momentum). The probability density we have in mind is either of the microcanonical ensemble $-(\rho_0(X,Q) = \delta(E(X,Q) - E_0))$ where E is the total energy and E_0 is a predetermined value of the energy) or of the canonical ensemble ($\rho_0(X,Q) = \exp[-\beta E(X,Q)/Z]$ where β is the inverse temperature and Z is the partition function ($Z = \int \exp(-\beta E(X,Q))dXdQ$)). With the initial conditions at hand we integrate the equations of motions up to the specified time T (m is the diagonal mass matrix and U is the potential energy).

$$\frac{dX}{dt} = m^{-1}Q \quad \text{and} \quad \frac{dQ}{dt} = -\frac{dU}{dX} \tag{1}$$

If the system is ergodic, and if it reaches equilibrium, then an ensemble average of a function $f(X,Q)$ - - $\langle f(X,Q)\rangle$ =

* Corresponding author e-mail: ron@cs.cornell.edu.

10.1021/ct060028m CCC: $33.50
Published on Web 03/28/2006

$\int \rho_0(X,Q) f(X,Q) dX dQ$ can be performed as an average over time and a single long trajectory $\langle f(X,Q)\rangle = 1/T \int_0^T f(X(t),Q(t))dt$. Here we do not require ergodicity or equilibrium, which means that averages must be done with respect to the initial distribution $\rho_0(X,Q)$. A trajectory with a final coordinate vector, $X(T) \in P$, is reactive. If the end configuration is not in P, then this trial failed and produced a nonreactive trajectory (nevertheless, the calculations of nonreactive trajectories is not a complete loss, since nonreactive trajectories can be used to estimate the fraction of reactive versus nonreactive paths). It is obvious that if only a tiny fraction of the trajectories initiated according to $\rho_0(X,Q)$ reacts, then a large fraction of straightforward trial trajectories would be a miss. The actual fraction can be so small that straightforward sampling of reactive trajectories may not be practical.

In light of the difficulties in straightforward sampling of reactive trajectories it is desirable to use a less straightforward approach and bias the sampling of the initial conditions to produce trajectories that spend less time in the reactant state and cross the barrier more readily than typical trajectories. This is in the spirit of the "biased sampling" of individual configurations[1] in computational statistical mechanics. Instead of configurations this time we bias the sampling of trajectories. Such biases are used at different levels:

(i) Generation of approximate or average trajectories from R to P to study qualitative reaction mechanisms.[2] Examples of subsets of "approximate" trajectories include minimum energy or free energy paths[3−8] and trajectories with filtered high-frequency modes[9−12] (of course, free energy paths have more uses than qualitative studies of mechanisms).

(ii) Generating exact trajectories from R to P but without accurate knowledge of their statistical weights. These Newtonian trajectories model better kinetic energy and inertial effects compared to option (i). The probability of the reactive trajectories that are sampled this way is assumed to be significant (see also Figure 1). However, because of the lack of exact weight, it is difficult to compute the reactive flux and the rate. Therefore, additional calculations are necessary to compute the rate in this way.

(iii) The third option computes reactive trajectories and their corresponding statistical weights. The weights enable, for example, the calculations of the rate of the reaction. Option (iii) is the most demanding computationally.

Significant progress was made in the direct application of option (iii) for large molecules in the condensed phase.[13,14] Of considerable theoretical interest is the transition path sampling approach.[13] Nevertheless, rate calculations are limited to systems in equilibrium and to reactants and products that are strong attractors. Trajectories are initiated in the transition region, and the equations of motion are integrated in the backward and forward directions to R and P, respectively. With the additional restriction of equilibrium the weights of the "patched" trajectories can be computed. If R and P are only weak attractors, trajectories initiated at Tr are unlikely to terminate at the desired reactants and products. For the last case this approach is not necessarily better than the straightforward method in which the trajectories are initiated simply at R.

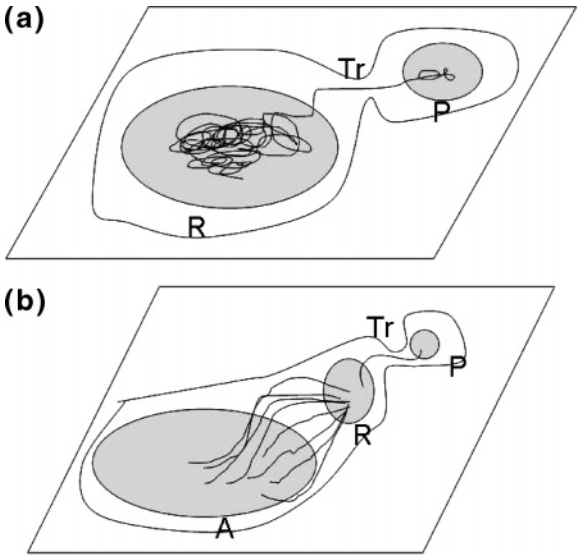


Figure 1. (a) A schematic drawing of an activated trajectory going from R to P. The system is initially trapped in a minimum R and is undergoing a rapid transition to minimum P, making the transition a rare event. Here we are primarily interested in the (short) time in which the trajectory is not in R or in P. (b) A schematic drawing of a rare transition due to geometrical factors. A nonequilibrium distribution is prepared at the state R. This state which is of high energy is relaxed rapidly to P and to A. Since the channel to A is much broader than the channel to P, most trajectories end at A. The transition to P is therefore rare (but fast).

Another important group of rare, short-time, and reactive paths is of nonequilibrium trajectories. The distribution $\rho_0(X,Q)$ can be a product of an external perturbation (absorption of a photon, temperature jump, etc.) that prepares the system in an initial nonequilibrium state and then is turned off. Typically, we follow the relaxation of the system to a new state either computationally or experimentally. For this class of problems equilibrium considerations apply only asymptotically at the long time limit. It is necessary to directly initiate trajectories according to the above distribution. If our interest is focused on a small fraction of these nonequilibrium trajectories that end at a known desired product (e.g. Figure 1), then the selection of reactive trajectories is difficult.

Here we propose an alternative approach for the sampling of reactive trajectories that is not based on initial value formulation. Instead we solve a boundary value problem and compute a whole trajectory from reactant to product as a single minimization task. The proposed approach can be classified as an intermediate between methods (ii) and (iii). The computed trajectories are exact, but the precise calculation of the weight can be expensive. In section VI we consider the computations of relative rates.

Studies of activated trajectories with boundary value formulation have a number of advantages compared to initial value solvers. Most importantly, boundary value formulation selects only reactive trajectories that end at the desired products, while initial value solvers are likely to have many misses. Nevertheless, there are also some negative aspects to the use of boundary value formulation. First, the optimiza-

tion of the whole trajectory (many time slices) is demanding in computer memory and expensive in terms of CPU time. These difficulties are to be contrasted with the calculation of a single frame at a time in initial value formulation that requires less memory and CPU per step. For a fixed time step and system size the boundary value formulation is considerably more expensive computationally. In practice boundary value calculations of exact classical trajectories are limited to short-time trajectories for which the cost of the calculation is not high. An estimate of the computational efficiency of boundary value calculations compared to initial value studies is provided in section VII.3

The limitation to short-time trajectories is not a significant restriction for the problem at hand since both of the processes that we are interested in: (a) passage over a large energy barrier, and (b) rapid relaxation to rare products, are quick (if we remove from trajectories of type (a) the "incubation" period, which is the (typically long) time spent at the reactant state prior to crossing the barrier).

Another nontrivial problem with the boundary value formulation is of uniqueness. In contrast to initial value formulation that generates a single unique trajectory (when the initial velocities and coordinates are given), the boundary value formulation (the initial and final coordinates and the total time are fixed) can have more than one solution. In section III.1 we illustrate numerically that the usual boundary formulation, based on Lagrangian mechanics, is ill posed. Numerous (and different) trajectories solve the same boundary value problem. The multiplicity of solutions needs to be addressed to ensure the proper convergence of the calculations to a unique solution and in order to estimate the weight of a particular boundary value. We introduce in section VI an algorithm to compute relative rates in which we sample trajectories, compute averages, and avoid the absolute weight problem.

The multiplicity of boundary value solutions should not come as a surprise. Consider a long time ergodic trajectory that reaches the "statistical-mechanics" limit in which a single trajectory samples extensively the coordinate space. In this case the probability of "hitting" the end point is independent of the initial conditions or the total time of the trajectory. This means that for a sufficiently large number of trials and long times more than one solution for the prespecified end point is likely to be observed.

II. Minimal Time, Point-to-Point Calculation of Classical Trajectories

The problem is formulated as follows: Given the two states R and P and two coordinate sets $X_i \in R$ and $X_f \in P$ of a reactive trajectory, compute a classical mechanics trajectory with a minimal time T and a bias to low energies. More specifically, we seek the stationary path of the classical action,[15] $S = \int_{0,X_i}^{T,X_f} L dt$, for which T is a minimum. The bias to low energies is added to a Monte Carlo procedure that searches for the classical path. The minimal time requirement removes a significant incubation period and, as we demonstrate below, makes the problem better posed.

Boundary value formulations (without the minimal time condition) were proposed by Gillilan and Wilson[16] in the context of a rubber band model of trajectories, by Cho, Doll, and Freeman[17] using Fourier space, and by Olender and Elber[9] in the context of the Onsager-Machlup action.[18] Recently, Passerone et al. studied activated trajectories using another boundary value formulation.[19]

As mentioned in the Introduction the application of the boundary value approach to compute *long time and exact* dynamics is difficult because of the following: (a) The optimization of an accurate long trajectory (including incubation time) is demanding in computer memory and CPU time compared to the calculation of a single frame at a time in initial value formulation. (b) We illustrate below that the solution for a long time boundary value problem is not unique. The above two problems are addressed by the minimal time condition that selects trajectories that are (i) nearly unique and (ii) more direct and therefore easier to compute. Hence besides focusing on the most interesting component of the transitions, these trajectories are more accessible to boundary value calculations.

For computational purposes it is convenient to write a discrete version of the classical action. The action becomes a function of the set of intermediate coordinates

$$S = \sum_{i=1}^{N} L_i \Delta t = \sum_{i=1}^{N} \left(\frac{m}{2} \left(\frac{X_i - X_{i-1}}{\Delta t} \right)^2 - U(X_i) \right) \Delta t \qquad (2)$$

The discrete coordinate sets provide a "basis-set" for the trajectory. X_i is the vector of (Cartesian) coordinates of time slice i. The condition that the action is stationary provides a set of equations for the intermediate coordinates, while X_0 and X_N are kept fixed.

$$\frac{1}{\Delta t} \frac{\partial S}{\partial X_i} = m \frac{X_{i+1} + X_{i-1} - 2X_i}{\Delta t^2} + \frac{\mathrm{d}U(X_i)}{\mathrm{d}X_i} = 0 \; i = 1,...,N-1 \qquad (3)$$

These equations are solved iteratively as follows:

1. Use a guess to interpolate between the two minima. In the examples below a straight line interpolation was used. Pick an initial time for the trajectory that should be small based on some knowledge of the system properties (e.g. a transition time in peptides is less than a picosecond). Use a small number of time slices (e.g. 20−40) to describe the trajectory.

2. Solve the second-order system of Verlet's equations (eq 3) with fixed boundaries. From (3) the residual vectors for the internal points $i = 1,...,N-1$ are $r_i = X_{i+1} + X_{i-1} - 2X_i + m^{-1}\Delta t^2 \, \partial U/\partial X_i$. The goal of the solution is to reduce all r_i to 0. This is accomplished by point-by-point iterations either by simulated annealing (SA) that minimize the sum of $r_i^t r_i$ or by Kaczmarz iterations[20] (the simple Gauss-Seidel iterations[21] do not converge). In Kaczmarz iterations, in contrast to Gauss-Seidel's, changes are made not only to a single point (e.g. point i) but also to all its neighboring points as well. The neighboring points are all those points appearing in the equation for the ith point. The coefficients of the neighboring points in the ith point equation determine the weights of the changes of these points. While iterating on the ith equation, the positions X_i, X_{i-1}, and X_{i+1} are updated so that r_i vanishes. A global minimizer (like simulated

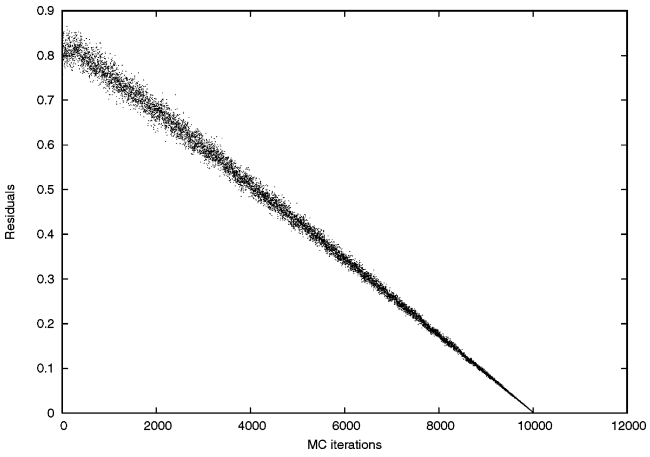


Figure 2. The convergence of the target function *R*2, which is the sum of the square of the residuals as a function of the simulated annealing step. Note the initial increase in the value of the target function before the calculations settle into approximate linear decay of the residuals. The example is from a calculation for the Lennard-Jones cluster (see section V for a more detailed discussion).

annealing) is necessary since elimination of local minima is required. The observation that the residuals sometimes increase before decreasing suggests that the system was trapped in a local minimum. A global minimizer also enables the search for multiple solutions. Convergence is assumed when the norm of residuals $||r|| = 1/N \sum_{i=1}^{N-1} |r_i|$ falls below a cutoff value (e.g. for the Mueller potential the cutoff is 10^{-9}) and the energy is conserved up to a few percents. A typical convergence plot for a simulated annealing run is shown in Figure 2.

3. Gradually decrease the total time T and use the spatial coordinates obtained with the larger T of the previous time as an initial approximation for the smaller T. Solve the resulting system as described in step 2.

4. When T becomes small enough in step 3, either convergence fails because of a sudden and significant change in the shape of the trajectory or some other significant changes occur in the system, such as an abrupt rise in the energy. The process stops at that T, and the solution of the previous total time is kept. We locate a *local* minimum of time in the neighborhood of the initial guess (of the minimal time) and in the neighborhood of a given energy. We do not search for a global minimum (which is a straight line and of high energy). A search bias toward lower energies can be added in the annealing process by a penalty function E_{penalty} which, at the time slice i, is equal to 0 if the energy E_i is below a given value E and to $\lambda(E_i - E)$ when $E_i > E$. λ is a positive constant that determines the relative weight of the penalty. In the examples presented in this work such a bias was added in the calculations of alanine dipeptide trajectories and in the trajectory calculations of the Lennard-Jones cluster.

III. Numerical Experiments on the Mueller Potential

To examine a number of alternative solutions that satisfy the same boundary value conditions, we perform an extensive sampling of trajectories using initial value formulation on the Mueller potential[22]

$$U(x,y) = \sum_{i=1}^{4} A_i \exp[a_i(x - x_i)^2 + b_i(x - x_i)(y - y_i) + c_i(y - y_i)^2]$$

$$A = (-200,-100,-170,15) \quad a = (-1,-1,-6.5,0.7)$$
$$b = (0,0,11,0.6)$$
$$c = (-10,-10,-6.5,0.7) \quad x = (1,0,-0.5,-1)$$
$$y = (0,0.5,1.5,1) \quad (4)$$

We choose the Mueller potential since it is simple and accessible to detailed analysis. It was used extensively in the past to study reaction paths, and many of its properties are well understood. Below we describe both initial and boundary value calculations of trajectories on the Mueller potential.

III.1. Initial Value Calculations To Solve the Boundary Value Problem. The Verlet integrator[23] is used with a time step of 0.0001. The mass was set to 1. The starting coordinates were at the lowest energy minimum (−0.5,1.5), the initial velocities were sampled with a random direction, and the total energy was fixed at −40.5 which is just above the barrier height. If the coordinate vector at any time slice during the trajectory was found within a distance of 0.01 from another minimum (0.5,0) we consider the computed trajectory a solution of the boundary value problem, and the time T of the trajectory was recorded (the end point can be determined only approximately within the neighborhood of the second minimum). We have repeated the calculation using two radii to check the convergence of our calculations and obtained similar results. Below we show a histogram plot (Figure 3) of the probability of finding a trajectory with the prespecified end points as a function of the total time of the trajectory.

The most striking feature of this plot is that trajectories obtained by the boundary value formulation are far from unique (at least in the numerical sense). The trajectories so computed are even "overdetermined" since the energy (in addition to total time and end points) was also fixed. In practice, depending on the calculation history, different trajectories may be obtained. The abrupt decay of the probability at time ≈ 400 is a reflection of the limited sample that we used rather than a property of the true distribution.

Since the directions of the velocities were sampled at random, it is possible that some of the sampled trajectories considered in Figure 3 are very similar. A question of interest is how different are computed trajectories that made it from reactant to product at the same total time and energy.

In Figure 4 we measure the similarity of 32 alternative trajectories with the same end points, energy, and total time T. The coordinates are sampled at three time slices: $\tau = T/4, T/2, 3T/4$ of each trajectory. At each time slice the variance $\sigma^2(\tau) = \langle (X_i - \langle X_i \rangle)^t (X_i - \langle X_i \rangle) \rangle$ of the coordinates X_i, $i = 1,...,32$, is computed from the trajectories. The shortest time trajectories converge to a very small variance (about zero), while at somewhat longer times a rapid increase in the variance (suggesting distinct trajectories) is evident.

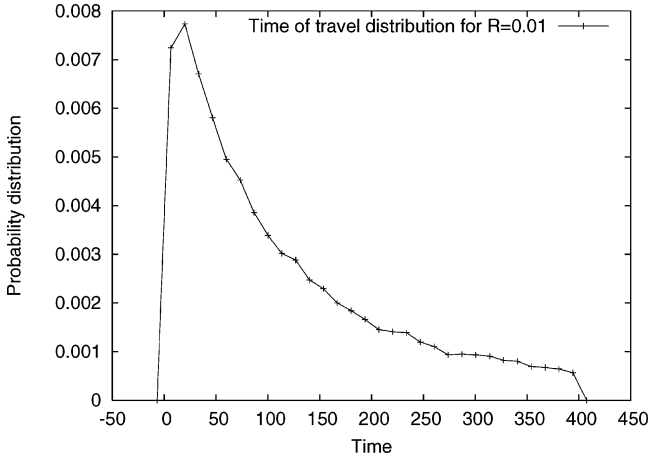


Figure 3. The distribution of the total time of the trajectories that are satisfying spatial boundary condition: The trajectories start at the lowest energy minimum (−0.5,1.5) and arrive at the second lowest minimum at (0.5,0). The total energy was fixed at −40.5, and the direction of the velocity was sampled at random. The drop of the probability density near 400 is due to limited statistics. The radius at the product was taken to be 0.01.

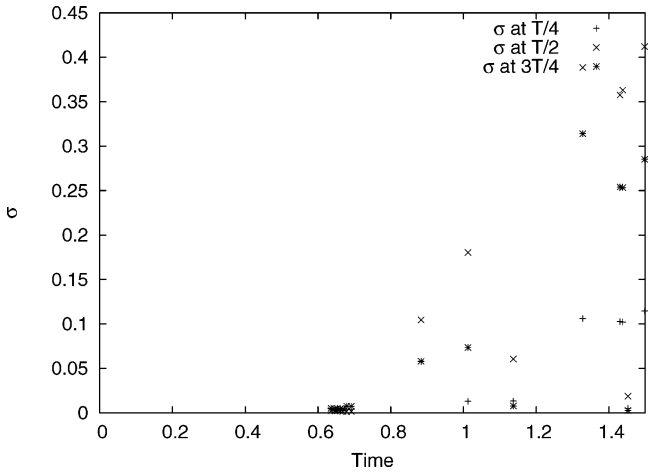


Figure 4. The variance of the alternative trajectories that solve the same boundary condition. As in Figure 3 the trajectories started from the same energy minima, and the total energy was the same. In contrast to Figure 3 we consider trajectories with the same total time and measured the variance of these trajectories as described in the text. It is evident that the variance is small for the shortest time trajectories, while the variation between the trajectories increases rapidly as a function of the trajectory time.

Figure 5a shows three trajectories satisfying the same boundary conditions that are clearly different. Figure 5b shows three trajectories at the short time limit that (within our numerical accuracy) correspond to essentially the same path.

III.2. Boundary Value Calculations. Here we used our proposed boundary value algorithm (section II). Trajectories were computed between the two energy minima of the Mueller potential that were studied with initial value formulation (see the previous section). Since the time step used for the discretization of the functional (eq 2) is significantly larger than the time step of the initial value

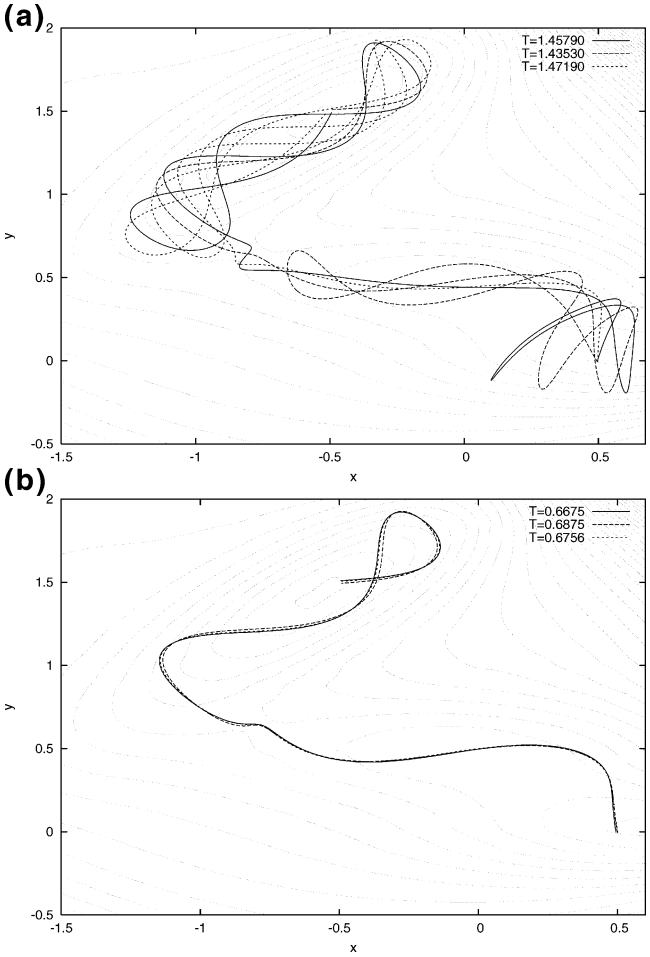


Figure 5. Sample of trajectories that satisfy the same boundary conditions as described in the text and the legends of Figures 3 and 4. (a) Three solutions of the same boundary value problem that are clearly different. (b) Three solutions of the same boundary value problem that seem to converge to the same solution.

formulation (0.01 versus 0.0001), it is desirable to use a refinement procedure. The refinement should convert the coarse grained boundary value trajectories to trajectories with the same time step as the initial value formulation, enabling a meaningful comparison of algorithm accuracies. We note however that for numerous other applications (like the calculation of the relative rate in section VI) the refinement is not necessary.

The proposed refinement is using iterations; i.e., given a coarse grained trajectory with time step ΔT, every iteration generates a new finer trajectory with half the time step ($\Delta T/2$). The boundaries of the finer trajectory are fixed to the same values as the coarse trajectory. Intermediate points are added using linear interpolation between points of the coarser trajectory as an initial approximation and followed by several Newton−Raphson steps that minimize the residuals of the equations of motion (eq 3). At each of the Newton−Raphson steps the unknown corrections δx_i and δy_i to the coordinates are calculated by solving simultaneously for all internal points $i = 1,N - 1$ the system of equations consisting of $\delta x_{i+1} + \delta x_{i-1} - (2 - \Delta t^2\ \partial^2 U/\partial x_i^2)\delta x_i + \Delta t^2\ \partial^2 U/\partial x_i \partial y_i\ \delta y_i = -r_{x_i}$ and their equivalent in the y direction.

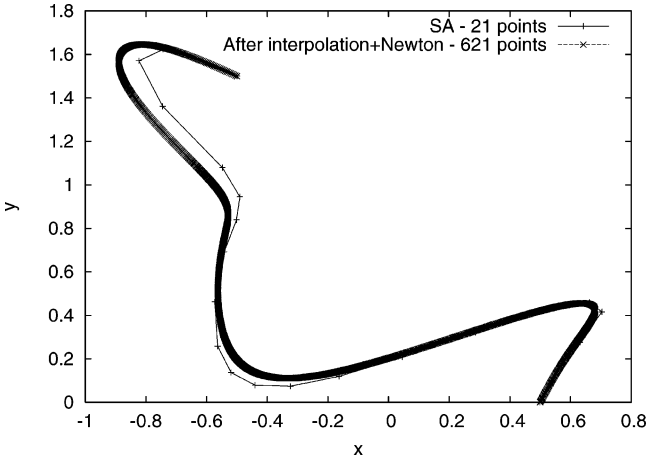


Figure 6. Coarse grained and highly refined boundary value trajectories on the Mueller potential. The refined trajectory is essentially identical to the short-time solution of the Verlet algorithm. The coarse grained trajectory (using a time step larger by a factor of 30) is quite accurate.

In Figure 6 we show a coarse grained and a highly refined trajectory. The coarse grained trajectory, which includes 21 points, is a minimal time trajectory obtained by the procedure described in section II. The initial guess was a straight line between the two end points. The initial total time was 0.5, and Kaczmarz iterations[20] were used until the residuals dropped to below 10^{-9}. The total time was gradually decreased until convergence stopped at total time 0.375. The coarse grained trajectory is the one obtained by the end of this process. The refined trajectory is very close to the exact Verlet trajectory. Even the coarse grained trajectory with a time step 30 times larger is quite accurate. This example demonstrates the benefit of working with short-time trajectories. The particle does not spend much time in oscillatory motion near the minima and focuses instead on executing the more interesting type of motion (the transition). We argue below that these trajectories have significant statistical weight in the microcanonical ensemble. Therefore they are worthy of investigation, even when the overall time scale is slow. Of course it is easier to compute exact trajectories if their total time is short, and if they are less curved and more direct.

IV. Molecular Example I: Transitional Trajectories in Alanine Dipeptide

We have repeated the boundary value computation also for the considerably more complex system of a conformational transition in alanine dipeptide. We consider the conformational transition from a helix to an extended chain. The coarse grained description of the trajectory includes 20 points. The time step was 0.00875 ps, making the total trajectory time 0.175 ps (which is the minimal time required for the transition as we confirmed using the procedure described in section II). This minimal time was also observed by extensive sampling of initial value solutions and recording the time of arrival to the product. The initial guess for the path was a straight-line interpolation, and a simulated annealing algorithm was used to minimize the norm of the residuals, defined identically to the Mueller problem. The starting residual norm was 10^5, and the cooling continued until the residuals dropped to about 10^{-5} where changes in space coordinates for the Metropolis algorithm were about 10^{-7} Å. In Figure 7 we show boundary value trajectories for several total times. All these trajectories are similar to the minimal time trajectory found by extensive sampling of initial value trajectories. The sampling of initial value trajectories was performed in a similar spirit to the calculations described for the Mueller potential. The alanine dipeptide system is considerably more complex than the Mueller potential, and it is possible to find new converged trajectories while making the total time shorter and shorter. However, as we decrease the total time, we obtain an abrupt rise in the total energy, shifting to a different class of trajectories. We terminate the process as described in section II when such a sharp change in the energy is observed.

V. Molecular Example II: An Isomerization Process in a Lennard-Jones Cluster

A more complex molecular example is the isomerization of a Lennard-Jones cluster. The potential energy of the system is given by $U = 4\sum_{j>k}[[1/q_{jk})^{12} - (1/q_{jk})^6]$ where q_{jk} is the distance between any two particles. The masses were set to one. A 38 atom LJ cluster has two minima with nearly the same lowest energy: the first is at the MacKay icosahedra and the other at a face-centered-cubic truncated octahedron[24] (Figure 8). In the following calculations the two boundaries are set to minima in the basins of the two distinct conformations.

To provide a rough estimate of the minimal time of the trajectory the period of oscillation of a Lennard-Jones term is computed by harmonic analysis as ~0.85. The expectation is that the minimal transition time will correlate with the vibrational time. The total time of the transition in the calculation was therefore set to $T = 2.0$. The number of slices was 33, and the uniform time step was set accordingly to $\Delta T = T/32$. The time was decreased in the optimization following section II until it was impossible to conserve the energy. The energy conservation was used as a stopping criterion.

The solution of the boundary value problem was done by Monte Carlo simulated annealing, minimizing the functional $R2 = \sum_{i=1}^{N-1} r_i^t \cdot r_i$. The term on the right-hand side is an inner product of the residual vector at each nonboundary time slice i. The slices are numbered between 0 and N. The initial approximation of the coordinates at nonboundary slices is obtained by linear interpolation from the boundaries.

The simulated annealing starts at a temperature $T_{\text{init}} = 5.0\text{e-}4$. The cooling protocol is linear, decreasing the temperature by a constant ΔT at each cooling step. The number of temperatures was $N_T = 1000$ and $\Delta T = T/N_T = 5.0\text{e-}7$. The process ends when the temperature reaches its final value, ΔT. To avoid trapping in a local minima the values of all space coordinates were discretized with a small mesh size $H_{\text{SA}} = 1.0\text{e-}4$. The initial coordinates are rounded to the nearest value which is an integral multiple of H_{SA}. The changes to coordinates in the Metropolis algorithm are always in integral multiples of H_{SA}. For each temperature, 200 Monte Carlo sweeps were done. A single sweep consists of a Metropolis step for each coordinate in the internal time

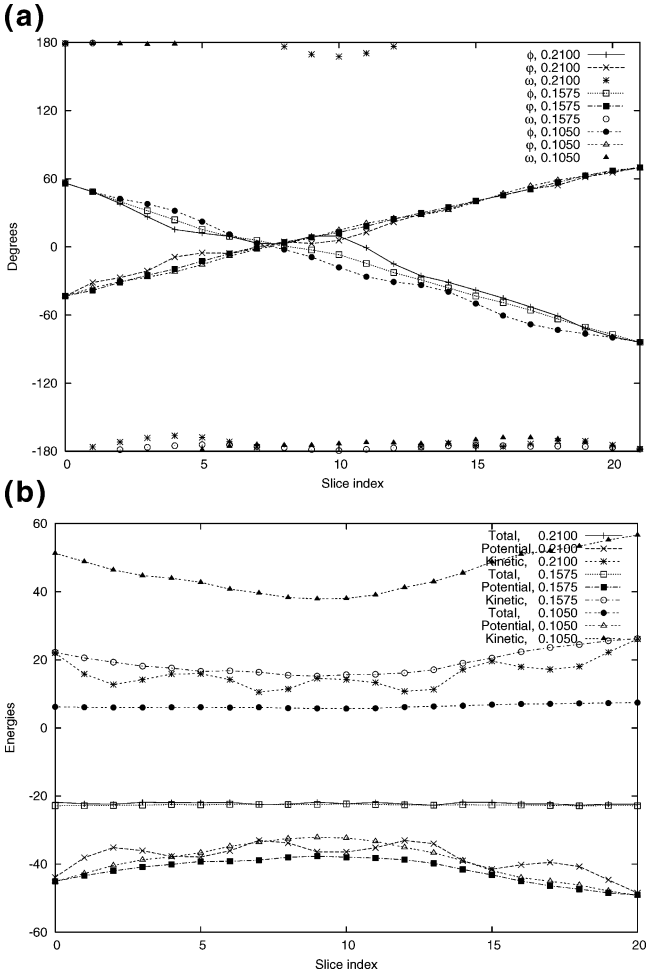


Figure 7. (a) A short-time trajectory for a conformational transition in alanine dipeptide. The calculations include all the 36 degrees of freedom of this molecular system. However, the projection onto the two soft degrees of freedom (the torsion angles φ and ψ) and one stiff degree of freedom (ω) is shown. A sequence of three times is shown to demonstrate our convergence criterion (shown in part b). (b) The energetic of the conformational transition in alanine dipeptide for three different trajectories of short time. Note the significant jump in the total energy and the kinetic energy when the time of trajectory is shortened from 0.1575 to 0.1050. We use this as an indicator that the minimal time trajectory (of low energy) converged at 0.1575.

slices. The value of the functional $R2$ was monitored to ensure that it decreases linearly with the temperature (Figure 2). At the initial temperature the maximum step size in the Metropolis algorithm was 400 H_{SA}, and the attempted changes to coordinates in the Metropolis algorithm were distributed uniformly between $-400H_{SA}$ and $400H_{SA}$. At later temperatures, the maximum allowable step size was decreased by a factor of 0.8 if the acceptance rate of the Metropolis algorithm fell below 0.4 in the previous temperature. The residuals decrease in the annealing process from a starting value of about 0.8 to a final value 0.002.

Figure 9 displays the total, potential, and kinetic energies at convergence as a function of time. The maximum of the potential energy is near -150 which is consistent with low kinetic energy when passing transition states. The total energy is nearly conserved, and the small fluctuations are likely to be due to the discretization.

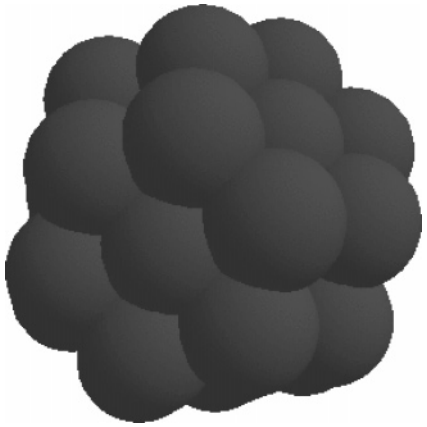


Figure 8. The initial structure of the conformational transition in the Lennard-Jones cluster (MacKay icosahedra). For the complete transition check http://www.cs.cornell.edu/ron/movies/lj_img.gif.

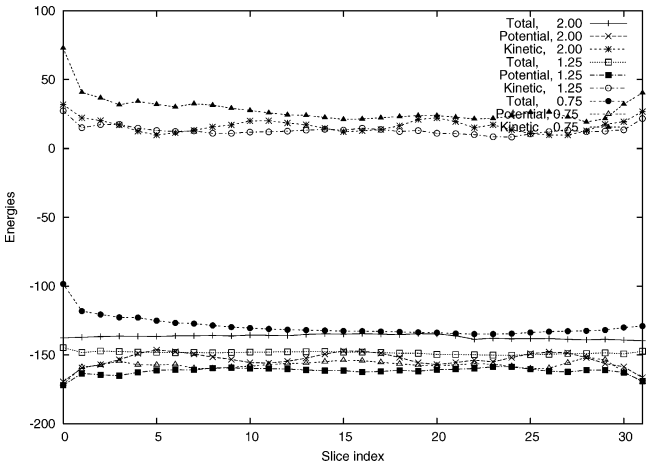


Figure 9. The energetic of a transition in 38 atoms of a Lennard-Jones cluster. Note that the total energy is clearly a constant for a total time of 2.00 but is fluctuating significantly for shorter times.

VI. Calculations of Relative Rates

Here we show how the short-time trajectories can be used in the calculation of rates and relative rates at short times. The restriction for short times is of particular interest for downhill but rare events and for activated processes (Figure 1).

In initial value formulation the reaction probability from R to P between time T and $T + dT$ is given by

$$p(T)dT = \left[\int_{\substack{X\in R \\ X(T)\in P}} Y_{RP}(X,Q,X(T),Q(T))\rho_0(X,Q)dXdQdX(T)\right]dT$$

where $Y_{RP}(X,Q,X(T),Q(T))$ is the probability density (in phase space and time) that a trajectory that starts in the phase space point (X,Q) in R will make a transition to P at time T. We used the shortcut (X,Q) to denote $(X(0),Q(0))$. For the (deterministic) Newton's equations of motion, $Y_{RP}(X,Q,X$-

$(T),Q(T))$ is either one or zero. Estimating the above integral there are two functions that require accurate and efficient sampling - - Y_{RP} and ρ_0. In initial value formulation we sample ρ_0 exactly, but the calculation of Y_{RP} can be difficult. If the process is rare, Y_{RP} will be zero most of the time. In the boundary value formulation Y_{RP} is always one, but the sampling is not done according to ρ_0 which complicated the calculation of the weight of the trajectories.

In the point-to-point approach we do not have the initial momentum as input. We sample configurations at both the reactants and the products sites $X(0) \in R$ and $X(T) \in P$ and compute trajectories of fixed time length T between any pair of sample points. The initial momentum is an output that we can compute by a finite difference formula from the calculated trajectory (i.e. $Q(0) \cong m(X_1 - X_0)/\Delta t$), once it is available.

Here we show how the relative rates of two different systems can be estimated, relying on a set of reactive trajectories only. Calculation of relative rates of similar systems is widely used in computational chemistry and biophysics. For example, we may consider changes of rates following the replacement of a hydrogen by a deuterium atom or a single residue mutation of a reacting protein. The time itself can also be a perturbation (e.g. by changing the size of Δt keeping the number of time slices fixed).

A concrete case is of a reaction at constant temperature. In the corresponding canonical ensemble, we have (E is the total energy)

$$\rho_0(X,Q) = w/Z$$

$$w = \exp(-\beta U(X)) \cdot \exp(-\beta Q^2/2m) \equiv \exp(-\beta E)$$

$$Z = \int \exp(-\beta E) dX dQ$$

Consider two systems with reactants and products R,P and $\bar{R},\bar{P}$, respectively, the ratio of the reaction probability densities is

$$\frac{\bar{p}(T)}{p(T)} = \frac{Z}{\bar{Z}} \cdot \frac{\int_{X,Q\in\bar{R}} \bar{w}(X,Q)\bar{Y}_{RP}(X,X(T))dXdQdX(T)dQ(0)}{\int_{X,Q\in R} \bar{w}(X,Q)Y_{RP}(X,X(T))dXdQdX(T)dQ(0)}$$

The ratio of the partition functions $(Z/\bar{Z})$ can be computed with standard methods in equilibrium statistical mechanics such as the umbrella sampling, free energy perturbation, dynamic integration, and more (for a recent discussion on alternative ways to compute free energy see ref 25).

$$\frac{\bar{Z}}{Z} = \frac{\int_{X\in R} \exp(-\beta(\bar{E} - E)) \exp(-\beta E) dX}{\int_{X\in R} \exp(-\beta E) dX} \equiv \langle \exp(-\beta(\bar{E} - E)) \rangle_{\exp(-\beta E)}$$

The above integration is restricted to the reactant state R.

Assume that a trajectory is found by the point-to-point approach. The weight of the initial structure is $\exp(-\beta\bar{E})$. This trajectory is perturbed to a new trajectory with a new energy, E. Since the *total energy* is not expected to vary violently from a trajectory to a trajectory, the average below should not be too difficult to perform.

The second ratio of integrals is called RR for "Relative Rates".

$$\mathrm{RR} = \frac{\int \exp(-\beta(E - \bar{E})) \frac{Y_{RP}(X,X(T))}{\bar{Y}_{RP}(X,X(T))} \exp(-\beta\bar{E}) \bar{Y}_{RP}(X(T) - X_P) dX dX(T)}{\int \bar{Y}_{RP}(X,X(T)) \exp(-\beta\bar{E}) dX_R dX_P}$$

$$= \left\langle \exp(-\beta(E - \bar{E})) \frac{Y_{RP}(X,X(T))}{\bar{Y}_{RP}(X,X(T))} \right\rangle_{\bar{Y}_{RP}(X,X(T))\exp(-\beta\bar{E})}$$

The existence of a corresponding point-to-point trajectory with a slightly different Hamiltonian is tested first (in our experience we were always able to find a nearby perturbed trajectory with a different energy), and the average of the energy differences is performed explicitly. This is an important difference from the initial value formulation in which the ratio of the *Ys* can be zero most of the time.

The algorithm to compute the RR is therefore as follows:

(i) Sample start and end configurations in R and P.

(ii) Compute a trajectory according to the "sampling" Hamiltonian using total fixed time T that is short but not minimal with the point-to-point approach,

(iii) Compute a trajectory for the perturbed Hamiltonian starting from the trajectory in (ii) as an initial guess. If convergence fails add zero to the average and go to (v). It is also possible to compute the change in energy by perturbation theory with no need to reoptimize the reference trajectory.

(iv) Estimate the difference in energy of the two trajectories and add to the RR.

(v) Check for termination conditions (convergence of the average or maximum number of allowed trajectories, etc.). Return to (i) if not terminating.

An example of a calculation of relative rates on the Mueller energy surface is sketched below. The original Mueller potential was used for sampling, and the perturbation was linear in the coordinate y, i.e., the perturbed potential is $U + 5y$, where U is the Mueller potential of eq 4. The reactant and product were defined with a radius of 0.2 around the location of the corresponding minima. The reactant was the minimum at the upper left corner of the Mueller potential. The total time was set to 0.38, slightly longer than the minimal time of 0.375 observed earlier. We first computed initial value trajectories with the Verlet algorithm (the time step was 0.001) and estimate the absolute rate for the two Hamiltonians (the absolute reaction probability of the unperturbed Hamiltonian was 0.0012). We then use a coarse grained trajectory with 38 points and estimated the perturbation. The rate difference was 10% in both the boundary and initial value approaches.

VII. Discussion

VII.1 Why Is It Possible To Use a Larger Time Step in Boundary Value Formulation? Equation 3, which is a discrete version of the Newton's equations of motion, is similar in accuracy to initial value algorithms. Rearranging eq 3 we can recover an initial value algorithm:

$$X_{i+1} = 2X_i - X_{i-1} - m^{-1}\Delta t^2 \frac{dU(X_i)}{dX} \quad (5)$$

If the coordinates at X_0 and X_1 are given, we can use the above formula to propagate the solution to its final configuration at X_N. A solution of the discrete boundary value problem is also a solution of the discrete initial value problem. This relationship is particularly intriguing in light of the observation on stability and accuracy made in the present study. We stated that the boundary value formulation is accurate and considerably more stable numerically compared to initial value formulation. Given that a solution of the boundary value formulation is identical to a solution obtained with initial value formulation, as discussed above, the statement seems false.

Therefore we have performed the reverse numerical experiment in which we start with initial value description. We sample different initial values (X_0 and X_1) and use eq 5 to generate trajectories of the same characteristics as in the boundary value formulation. If we employ the large time step that was used in the boundary value formulation, most of the initial value trajectories lose stability and "explode" in accord with the usual observations. Nevertheless, it is possible to find rare initial value trajectories with large time steps that maintain their stability (and accuracy). These are the trajectories that are picked by the boundary value formulation. Since the end structure in the boundary value formulation is sound (by construction) all the "exploding" unstable trajectories are eliminated at the start. Only the stable trajectories remain. This explains the observed stability of the approach proposed here. The enhanced stability of the boundary value formulation also supports the Stochastic Difference Equation approach[10−12] which focuses on generating approximate long-time (or length) trajectories that qualitatively describe reaction mechanisms. The present approach is centered on an accurate calculation of short-time trajectories.

VII.2. Are Minimal-Time Activated Trajectories Significant? We have invested considerable effort to show that the minimal time trajectories are nearly unique and as such can be computed accurately and efficiently with boundary value formulation. However, typical calculations of physical observations (such as calculations of rate) rely on an ensemble of trajectories and not on a unique path. Therefore the significance of minimal time trajectories and their contribution to the ensemble of reactive paths require examination.

A first impression suggests that these trajectories are not significant. Since all initial values (coordinates and momenta) in the microcanonical ensemble have the same statistical weight, a single trajectory cannot be representative (in this section we consider trajectories with fixed energy and time).

A more subtle argument focuses on the transitional component of the trajectories. Consider reactive trajectories of total time T. We will divide this long-time trajectory into two components: (i) an incubation part of length $(T - \tau)$ in which the system spends some time either in R or in P and (ii) a transient component (length τ) in which the system is in Tr. Since both R and P are attractive states we expect that $\tau \ll T$. A transitional trajectory is initiated at the interface between R and Tr, called RTr, and is terminated at the interface between Tr and P called TrP. A transitional trajectory spends all its time (τ) in Tr without "touching" RTr nor TrP at any time after initiation or before termination. A sample of such boundary value transitional trajectories makes it possible to compute, $P(\tau)$, the probability of observing a transitional trajectory of time−length τ. The details of the distribution will depend on the underlining energy surface, and we cannot be more specific. However we expect (due to the assumed transient nature of the trajectory) that $P(\tau)$ is a rapidly decreasing function of τ. Moreover, at very short times $P(\tau)$ must go to zero. Therefore it has at least one maximum.

Experiments rarely measure properties of transient times and trajectories. More often the overall time t (including the incubation time of component (i) of the trajectory) and rate are available. We therefore need to extend our argument to trajectories of total length t and transient time τ (communicated to us by Cristopher Dellago).

We now consider a discrete (numerical) representation of the trajectory. The numerical trajectory of total fixed time is described by a sequence of coordinates separated by a time step of Δt: $X_1,X_2,...,X_n,X_{n+1},...,X_m,X_{m+1},...,X_k$. Similarly to the continuous trajectory the discrete path is divided into three parts, $X_1,...,X_n$—a set of configurations that reside in the reactants, R, $X_{n+1},...,X_m$—a set of configurations that reside in the transition state, Tr, and $X_{m+1},...,X_k$—a set of configurations in the product, P. The transitional trajectory that we have in mind is the middle piece (from X_{n+1} to X_m, or to make the transition complete perhaps from X_n to X_{m+1}, we ignore this small difference). Assuming that we use eq 5, then with the initial pair of coordinates $-X_1$ and X_2, we can generate the whole trajectory from X_1 to X_k. Note however that we are not interested in the whole trajectory. We are only interested in the transitional component from X_{n+1} to X_m. The statistical weight of this trajectory is the number of initial conditions in the reactant space that results in the same *transitional* trajectory. Similarly to the reptation mechanism of polymer transport we now remove the first coordinate set and create a new coordinate set at the end of this trajectory (creating a trajectory from X_2 to X_{k+1}, X_{k+1} is uniquely defined since X_k and X_{k-1} are known). This is a trajectory of the same total length and exactly the same transitional trajectory (for $n > 1$). However, it has different initial conditions in the reactant state (they are now X_2 and X_3). We can repeat this process and move the first coordinate to the end of the trajectory, keeping the same transitional trajectory, n times. Since we consider a trajectory of total fixed length of time, a larger n means a smaller $m-n$. The shorter is the transitional trajectory the larger is n, and more initial conditions provide the same transitional trajectory.

Typically all initial conditions will have the same weight if sampled from the microcanonical ensemble; however, the above argument, restricted to an ensemble of trajectories with fixed energy and total time, suggests that shorter transitional trajectories have larger statistical weights that skew the original $P(\tau)$ to short times. The degree of skewing will depend on the particular energy surface.

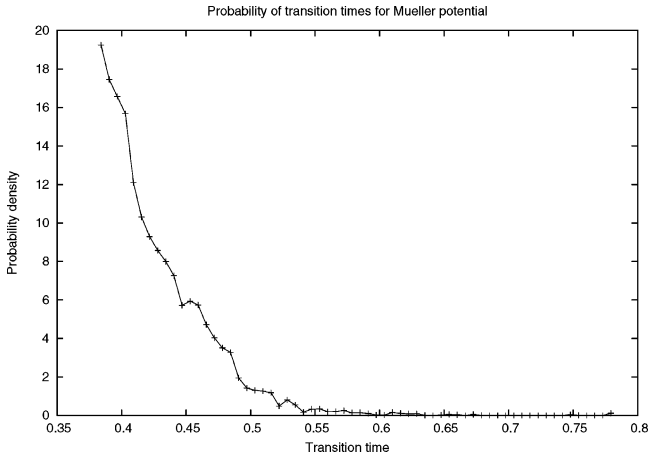


Figure 10. Distribution of transition times from two states in the Mueller potential. The reactant state is a circle with a radius 0.35 centered at (−0.5,1.5), and the product is another circle with a radius 0.3 centered at (0.5,0). Transitional trajectories were sampled by the Verlet algorithm for a total energy of −38 and a total time of 50. The time in which the trajectory is not in the reactant or the product is the transition time. Note that in the canonical ensemble (that prefers low energies and slower transition times) the peak shifts to the right.

To illustrate the above argument we computed the distribution of the transitional times for the Mueller potential. The probability density of transitional times for the Mueller potential is highly peaked at the shortest time (shown in Figure 10).

Note that for more diffusive transitions, and trajectories that cross slowly and/or many times the barrier, the argument above is not valid.

VII.3. Efficiency of the Calculation. The boundary value calculations of this manuscript are reasonably efficient. The basic unit for cost evaluation is the cost of energy calculation which is roughly proportional to the number of particles in the system, N. The cost of a single action evaluation is proportional to $L \times N$ where L is the number of time slices. The number of relaxation steps depends on the system but (in practice) is no worse than another factor of $L \times N$. Nevertheless, this factor is perhaps the major source of uncertainty in the formula provided below, and in some cases it can be larger. Accepting this estimate, the total cost is proportional to $(L \times N)^2$. This is similar to earlier boundary value calculations[9−12] in much larger systems, producing approximate trajectories.

For comparison, the computational cost for a single initial value trajectory is proportional to $\bar{L} \times N$, where $\bar{L}$ is the number of time slices we require in an initial value solver. We denote the probability of an initial value trajectory to be reactive by p. The computational cost for a reactive trajectory is therefore proportional to $p^{-1}(\bar{L} \times N)$. The ratio of costs of the two approaches is

$$\text{ratio of costs} = a\frac{(L \times N)^2}{p^{-1}(\bar{L} \times N)} = a\frac{L^2 \times N}{p^{-1} \times \bar{L}} \qquad (6)$$

where a is a numerical constant which is estimated to be of order of 10. In the most complex example of this manuscript (the Lennard-Jones cluster) L is 33 and $\bar{L}$ is about 100 times larger. The number of degrees of freedom, N, is 38×3. We can estimate what is the reaction probability in which the cost of computing a reactive trajectory using an initial value or the boundary value approach is roughly the same (ratio of costs equal one). $1 = 33^2 \times 38 \times 3 \times 10/p^{-1} \times 3300\ p \approx 0.001$, which is a process that is only marginally activated (i.e. not so rare). The above formula also suggests that as the system becomes larger the reaction probability better be smaller to make the boundary value calculations of the type described here more effective. Indeed the reaction probabilities of complex systems at short times are significantly smaller than 10^{-3}.

Parallelization does not change the above conclusions in an obvious way. Let the number of processors be n. Since the time slices can be distributed between the processors with limited communication overhead we can replace one of the L by L/n in eq 5 (the required number of minimization steps is reduced from the original cost of $L \times N$, but the extent is not clear). Initial value trajectories can be parallelized in a trivial way (launching independent trajectories on separate processors), and we therefore replace p^{-1} by $(np)^{-1}$. The cost ratio for n processors is therefore unchanged or perhaps slightly favors the boundary value formulation.

Finally we note that the calculation of the perturbed trajectory for the study of relative rate takes little effort within the boundary value formulation. In initial value formulation it takes the same effort as the generation of the unperturbed trajectory.

Acknowledgment. This research is supported by NIH Grant GM59796 to R.E. We thank Eric vanden Eijnden and David Shalloway for useful discussions. The initial and final coordinate sets of the Lennard-Jones cluster were kindly provided by David Wales.

Note Added after ASAP Publication. This article was released ASAP on March 28, 2006, with the incorrect Received Date. The correct version was posted on April 12, 2006.

References

(1) Torrie, G. M.; Valleau, J. P. Nonphysical sampling distributions in Monte-Carlo free energy estimation - umbrella sampling. *J. Comput. Phys.* **1977**, *23*, 187−199.

(2) Elber, R. a chapter 2, Reaction path studies of biomolecules. In *Recent developments in theoretical studies of proteins*; World Scientific: Singapore, 1996.

(3) Ulitsky, A.; Elber, R. A new technique to calculate steepest descent paths in flexible polyatomic systems. *J. Chem. Phys.* **1990**, *92*, 1510−1511.

(4) Jonsson, H.; Mills, G.; Jacobson, K. W. Nudge Elastic Band Method for Finding Minimum Energy Paths of Transitions. In *Classical and Quantum Dynamics in Condensed Phase Simulations;* Berne, B. J.. Ciccotti, G.. Coker, D. F., Eds.; World Scientific: Singapore, 1998; p 385.

(5) Weinan, E.; Weiqing, R.; Vanden-Eijnden R. String method for the study of rare events. *Phys. Rev. B* **2002**, *66*, Art. No. 52301, 1−4.

(6) Ren, W.; Vanden-Eijnden, E.; Maragakis, P.; Weinen, E. Transition pathways in complex systems: Applications of the finite-temperature string method to the alanine dipeptide. *J. Chem. Phys.* **2005**, *123*, Art No. 134109, 1−12.

(7) Micheletti, C.; Laio, A.; Parrinello, M. Reconstructing the density of states by history-dependent metadynamics. *Phys. Rev. Lett.* **2004**, *92*, Art No. 170601, 1−4.

(8) Noe, F.; Ille, F.; Smith, J. C.; Fischer, S. Automated computation of low energy pathways for complex rearrangements in proteins: Application to the conformational switch of ras p21. *Proteins: Struct., Funct., Bioinformatics* **2005**, *59*, 534−544.

(9) Olender, R.; Elber, R. Calculation of classical trajectories with a very large time step: Formalism and numerical examples. *J. Chem. Phys.* **1996**, *105*, 9299−9315.

(10) Elber, R.; Ghosh, A.; Cardenas, A. Long time dynamics of complex systems. *Acc. Chem. Res.* **2002**, *35*, 396−403.

(11) Cardenas, A. E.; Elber, R. The kinetics of cytochrome C folding: Atomisitically detailed simulation. *Proteins: Struct., Funct., Genet.* **2003**, *51*, 254−257, 2003.

(12) Cardenas, A. E.; Elber, R. Atomically detailed simulations of helix formation with the stochastic difference equation. *Biophys. J.* **2003**, *85*, 2919−2939.

(13) Dellago, C.; Bolhuis, P. G.; Geissler, P. L. Transition Path Sampling. *Adv. Chem. Phys.* **2002**, *123*, 1−78.

(14) Snow, C. D.; Sorin, E. J.; Rhee, Y. M.; Pande, V. S. How well can simulation predict protein folding kinetics and thermodynamics. *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 43−69. Voter, A. F. Hyperdynamics: Accelerated Molecular Dynamics of Infrequent Events. *Phys. Rev. Lett.* **1997**, *78*, 3908−3911.

(15) Landau, L. D.; Lifshitz, E. M. *Mechanics*; Butterworth-Heinenann: Oxford, 1993; Chapter 1.

(16) Gillilan, R. E.; Wilson, K. R. Shadowing, rare events and rubber bands—a variational Verlet algorithm for molecular dynamics. *J. Chem. Phys.* **1992**, *97*, 1757−1772.

(17) Cho, A. E.; Doll, J. D.; Freeman, D. L. The construction of double ended classical trajectories. *Chem. Phys. Lett.* **1994**, *229*, 218−224.

(18) Machlup, S.; Onsager, L. Fluctuations and irreversible processes 2. Systems with kinetic energy. *Phys. Rev.* **1953**, *91*, 1512−1515.

(19) Passerone, D.; Parrinello, M. Action-derived molecular dynamics in the study of rare events. *Phys. Rev. Lett.* **2001**, *87*, art. No. 108302, 1−4.

(20) Kaczmarz, S. Approximate solution of systems of linear equations. *Bull. Acad. Polonaise Sci. Lett. A* **1937**, *5*, 355−377.

(21) Golub, G. H.; Van Loan, C. F. *Matrix Computations;* John Hopkins University Press: 1996.

(22) Mueller, K. Reaction paths on multidimensional energy hypersurfaces. *Angew. Chem.* **1980**, *19*, 1−13.

(23) Verlet, L. Computer experiments on classical fluids. I. Thermodynamical properties of Lennard Jones molecules. *Phys. Rev.* **1967**, *159*, 98−103.

(24) Doye, J. P. K.; Wales, J. D. Thermodynamics of global optimization. *Phys. Rev. Lett.* **1998**, *80*, 1357−1360.

(25) Rodriguez-Gomez, D.; Darve, E.; Pohorille, A. Assessing the efficiency of free energy calculation methods. *J. Chem. Phys.* **2004**, *120*, 3563−3578.

CT060028M

JCTC Journal of Chemical Theory and Computation

Lattice-Boltzmann Simulations of Ionic Current Modulation by DNA Translocation

Sylvain Reboux,[†] Fabrizio Capuani,[‡] Nélido González-Segredo, and Daan Frenkel*

FOM Institute for Atomic and Molecular Physics, Kruislaan 407, 1098 SJ Amsterdam, The Netherlands

Received December 31, 2005

Abstract: We present a numerical study of the effect of DNA translocation on the ionic current through a nanopore. We use a coarse-grained model to solve the electrokinetic equations at the Poisson−Boltzmann level for the microions, coupled to a lattice-Boltzmann equation for the solvent hydrodynamics. In most cases, translocation leads to a reduction in the ionic current. However, at low salt concentrations (large screening lengths) we find ionic current enhancement due to translocation. In an unstructured pore, translocation of the helical charge distribution of the DNA has no effect on the ionic current. However, if a localized charge probe is placed on the wall of the nanopore, we observe ionic current modulations that, though weak, should be experimentally observable.

1. Introduction

The subtle and often counterintuitive effects underlying many electrokinetic phenomena, and their relevance in the behavior of many synthetic and biological complex fluids, make their study one of both practical and fundamental interest.[1,2] Electrokinetic effects such as electro-osmosis and electrophoresis can be used to control the transport of small molecules and colloids;[3] as microfluidic devices become ever more prevalent, there is an increasing number of electrokinetic effects that can be exploited for selective molecular transport.[4] An important example is the translocation of highly charged biopolymers such as DNA and RNA through biological or synthetic nanoscale pores. This topic has excited much experimental and theoretical interest[5−9] not only because of its potential implications for DNA sequencing but also because of the importance of biopolymer translocation in living organisms. Biopolymer translocation is involved, for example, in the transport of transcribed RNA out of the cell nucleus and viral injection of DNA into a host cell. The idea behind DNA sequencing using translocation through nanopores is to use biological or synthetic pores such as nanosensors to characterize very locally the spatial extent and charge distribution of the translocating polymer. This is done by measuring electric current fluctuations under a constant voltage bias.[8]

Experiments[9] and atomistic simulations[10−12] have been used to gain insight in the factors that determine the relation between the local structure of the translocating biopolymer and the associated ionic current. However, the interpretation of the experimental data in terms of a microscopic model is often not unique. Fully atomistic simulations make it possible to relate macroscopic observables, such as current modulations, to microscopic properties of the system under study, but such simulations are computationally very expensive. This makes them less suited for a systematic study of the effect of the different parameters (such as pore size and ionic strength) that affect the translocation current. In virtually all cases of practical interest, electrokinetic phenomena occur in confined systems of a rather complex geometry; as a consequence, the analytical solution of continuum—in particular, electrokinetic—equations can usually not be obtained. In addition, the dynamics of translocation is determined by a number of distinct characteristic lengths and times that may differ by several orders of magnitude: ionic radius, Debye screening length, channel/pore size and polyelectrolyte size, time of formation of the electrical double layer (EDL), and the transient time for the onset of convective currents. The

* Corresponding author e-mail: frenkel@amolf.nl.
[†] Current address: Division of Applied Mathematics, University of Nottingham, Nottingham NG7 2RD, U.K.
[‡] Current address: FIRC Institute for Molecular Oncology, Via Adamello 16, 20139 Milan, Italy.

10.1021/ct050340g CCC: $33.50
Published on Web 03/22/2006

need to span these scales may render atomistic simulation unpractical.

Mesoscopic ('coarse-grained') models for fluid dynamics can offer an alternative to atomistic simulation in the study of electrokinetic phenomena. To be more precise, coarse-grained models are useful to study phenomena on length and time scales where the details of a fully atomistic description can be represented by a small number of parameters. There exist several coarse-grained, mesoscopic methods to model hydrodynamics. Most notable among these are Dissipative Particle Dynamics (DPD), Stochastic Rotation Dynamics (SRD), and Lattice-Boltzmann (LB) methods. Dissipative Particle Dynamics can be viewed as a coarse-grained version of Molecular Dynamics, whereby mesoscopic particles representing elements of solvent move under the influence of conservative, friction, and stochastic forces.[13−15] Stochastic Rotation Dynamics[16,17] models the solvent as a viscous gas that undergoes stochastic many-body collisions. The interactions among solutes and between solute and solvent are described by conservative forces. Finally, the Lattice-Boltzmann (LB) method employs a coarse-grained lattice-gas cellular automaton model for a fluid which approximately solves the (continuum) Boltzmann transport equation, which in turn describes hydrodynamics (for details, see e.g. refs 18−20).

The relative merits of the different coarse-grained approaches to mesoscopic fluid dynamics depend not only on the nature of the problem but also on the state of development of the various techniques. At present, there are versions of the Lattice-Boltzmann method that can be used to model electrokinetic effects. For the study reported in the present manuscript, we employ the approach described in ref 21. The LB scheme of ref 21 integrates a Boltzmann transport equation for the solvent to a local diffusion equation for the charge dynamics. In ref 21 it is shown that this approach provides an excellent description of the sedimentation of charged colloids.

There also exists a version of DPD that takes charge transport into account.[15] However, unlike the LB method, DPD introduces an additional length scale into the description of electrolyte solutions, namely the size of the charged mesoparticles. These mesoions are larger than simple ions yet must be much smaller than the Debye screening length in order for charge modulations not to become obscured by spurious structural correlations. This makes it nontrivial to achieve a proper separation of length scales in DPD simulations of electrolytes. The LB method also has advantages when compared with more macroscopic approaches, such as a direct solution of the Navier−Stokes equation: the LB scheme tends to be numerically more stable, and it is straightforward to impose complicated boundary conditions. To our knowledge, there exists no viable version of SRD for electrolyte solutions.

The LB method that we use in the present paper treats electrostatics at the level of the nonlinear Poisson−Boltzmann equation, thus allowing us to model electrokinetic effects at high surface-charge densities. It should be noted that the Poisson−Boltzmann description does not account for correlation effects associated with the finite sizes of the salt ions nor for explicit charge correlations. Hence, physical phenomena that are sensitive to the discrete nature of the individual ions, such as like-charge attraction[33] or ion adsorption to highly charged polyelectrolytes (see e.g. ref 34), will not be captured by this formalism. However, our aim is to model electrokinetics at the simplest possible level in order to explore to what extent a purely continuum description can be used to describe the rather complex electrokinetic effects associated with DNA translocation. In this respect, our calculations can be considered as a "baseline" for more detailed (atomistic) simulations.

Explicit description of the individual ions is of course possible in more atomistic simulations (see e.g. ref 22), but the latter approach is computationally more expensive. Of course, the description of the fluid motion by hydrodynamics is also an approximation. However, one of the surprises in the study of the dynamics of simple liquids is that continuum hydrodynamics provides a good description of motions in fluids, even at the nanoscale. This is illustrated, for example, by the success of the Stokes−Einstein relation that connects the diffusion coefficient of microscopic particles to the hydrodynamic friction coefficient. This relation works extremely well for small colloids and even (to within a factor two) for atoms.

The remainder of this paper is organized as follows. In section II we present our model for a translocating DNA oligomer, and we discuss the equations that govern charge and mass transport. In section III we validate the numerical solution of the electrokinetic equations for the given boundary conditions (translocating rod) by (a) comparing the computed equilibrium profiles for the electrical double layer with the corresponding analytical solutions, in the low-charge regime, and (b) by showing that, in the low-salt and high-charge regime, the model correctly reproduces Manning condensation. In section IV, we show that DNA translocation may lead to both reduction and enhancement of the ionic current. Current enhancement is a low-salt and high-charge effect. We give a simple analytical estimate for this effect that matches well with our numerical calculations. In section V we address the possibility of detecting a signature of the DNA charge distribution in the ionic current. We show that current modulation during translocation should be observable, provided that there is a localized charge placed just inside the wall of the pore. In that case, we find that upon moving both single-stranded and double-stranded DNA through a nanopore, the ionic current undergoes modulations as a result of the overlapping of the electrical double layer of the DNA with that of the "probe" charge.

II. The Model

The length of nanopores, be they synthetic or biological, is usually less than the persistence length of most biopolymers. Hence, for a local description of the pore plus translocating biopolymer, it is reasonable to approximate the biomolecule by a rigid rod. This approximation is certainly justified for double-stranded DNA, a semiflexible polymer with a persistence length of ca. 50−100 nm. In what follows, we simplify the description even further by modeling the DNA-pore system as a nondeformable cylinder that is concentric

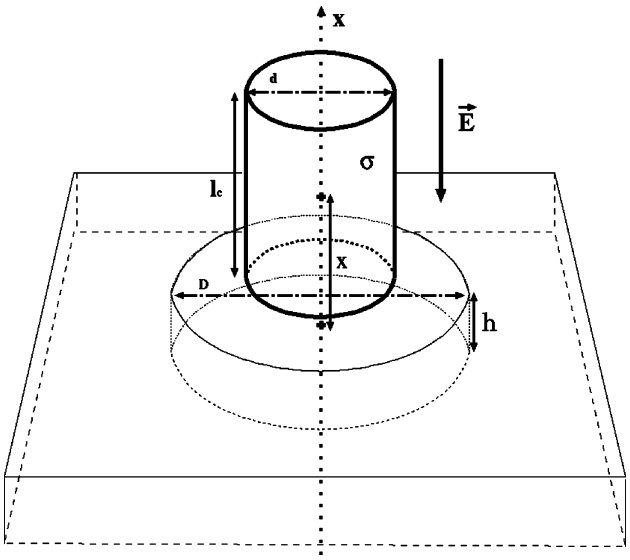


Figure 1. Sketch of the geometry of a rodlike polyelectrolyte of length *L* and diameter *d* translocating through a hole of height *h* and diameter *D*. The surface charge density of the polyelectrolyte is denoted by σ. The applied electric field **E** is parallel to the axis of the rodlike polymer.

with a cylindrical pore in an electrically neutral, rigid slab (see Figure 1). We studied two charge distributions on the surface of the cylinder: (a) homogeneous and (b) helical line charge on a single and a double helix.

In some of the experiments that probe the modulation of ionic current by DNA translocation, one end of a DNA oligomer is attached to a microbead that is trapped by an optical tweezer; the DNA chain is then straightened by the applied electric field. The position of the DNA in the pore can be varied at will. When the DNA is kept fixed, an applied electric field will first induce a transient current that then settles down to a steady-state value. In our simulations, we reproduce this scenario by positioning the charged cylinder at a fixed position close to or in the pore. We then let the surrounding electrolyte relax to steady state under the applied electric field.

The electrolyte solution in which the cylinder is immersed contains two oppositely charged microion species of concentration ρ^k, $k = +, -$, and negligible molecular size. The ions diffuse in a solvent of number density ρ. We assume that the flow of the solvent obeys the Navier−Stokes equation for incompressible fluids. The electrolyte can undergo two types of dynamics: (a) short-range, diffusive mass transport and (b) long-range, driven transport. To maintain charge neutrality of the total system (DNA + electrolyte solution), the electrolyte must contain an excess charge of counterions that matches the total charge on the cylinder.

Before applying an external uniform electric field **E** parallel to the longitudinal axis of the DNA, we allow a period of time for the formation of the electrical double layer, at which the electrolyte reaches electrostatic equilibrium. Once the external field is switched on, it exerts a force on those fluid elements that carry a net charge, thus creating an electro-osmotic current.

We considered several values of the cylinder diameter, salt concentration, and charge density. The uniform external electric field is kept constant at a value small enough to guarantee that the system is within the linear-response regime.

The Electrokinetic Equations. We utilize the hybrid LB scheme of Capuani et al.[21] to solve the electrokinetic equations on a lattice and with discrete time steps. We refer the reader to ref 21 for further details on this approach.

For the flows that we are considering, the Reynolds numbers are typically very small. Moreover, we can assume that local thermodynamic equilibrium holds, i.e., the microions become thermally equilibrated within the hydrodynamic time scale, $t_{hydr} \equiv L^2/\nu$, where ν is the kinematic viscosity of the solvent and L is a characteristic length over which the velocity of the solvent varies.

The electrostatic potential is computed by solving the Poisson equation

$$\nabla^2\phi = -4\pi l_B[\sum_{k=\pm} z^k(\rho^k + \rho^{k,W})], \quad k = +, - \qquad (1)$$

where $\phi \equiv \beta e\hat{\phi}$ is the electrostatic potential, $\hat{\phi}$, in units of k_BT/e, e is the fundamental charge, $\beta \equiv (k_BT)^{-1}$, $l_B \equiv e^2\beta/(4\pi\epsilon)$ is the Bjerrum length, and $z^k\rho^k$ and $z^k\rho^{k,W}$ are the charge densities of the microions (in solution) and of the embedded in solid objects, respectively. z^k is the valence of species k. We solved this equation in real space on the same lattice that was employed to solve the solvent hydrodynamics, using a successive over-relaxation method.[24] This approach allows great flexibility in the choice of the boundary conditions as long as they can be resolved on the lattice.

The time evolution of microion densities are governed by a convection-diffusion equation

$$(\partial_t + \mathbf{u}\cdot\nabla)\rho^k + \nabla\cdot\mathbf{j}^k = 0 \qquad (2)$$

where the ion fluxes (i.e. the number of ions crossing a unit area per unit time) are $\mathbf{j}^k \equiv -D^k\nabla\rho^k + \mu^k\rho^k\mathbf{F}^k$ (D^k and $\mu^k = \beta D^k$ are the relevant diffusion and mobility coefficients, respectively), and **u** is the solvent velocity. The first term corresponds to Fickian diffusion, the second to drift under the influence of a local electric force $\mathbf{F}^k \equiv -(ez^k)\nabla\hat{\phi}$. In our simulations, we impose strict local mass and charge conservation.[21] The Poisson−Boltzmann approximation prescribes that, in the absence of external forces, the ion densities relax to a canonical thermodynamic equilibrium, $\rho^{k,eq} = \rho_0 \exp(-z^k\phi)$, where ρ_0 is the salt concentration in the bulk. In fact, in our model, the stationarity implied by the equilibrium distribution is automatically satisfied as we impose that $\mathbf{j}^k$ vanishes for $\rho^k = \rho^{k,eq}$.

We assume that nonslip boundary condition for the flow field **u** on both the colloidal surfaces and the walls. The assumption of no slip is reasonable as all surfaces involved in our study are, in fact, rough on an atomic scale. As has been shown by Bocquet and Barrat,[32] a very small amount of roughness is enough to cause no-slip behavior on solid surfaces that are wetted by the fluid. For reasons of computational convenience, we choose a kinematic viscosity of $\nu = 1/6$ (in lattice units, l.u.). We choose the solvent density $\rho = 1$. The external (electric) field is chosen to be 10^{-6} (in reduced units). This value is well inside the linear-

response regime. The diffusivity of all microions is set to $D = 0.19$, a value for which spurious diffusion due to lattice advection (see ref 21) is negligible.

III. Tests for Weakly and Highly Charged Rods

Before discussing the transport properties of our model, we briefly describe a few tests that we carried out to verify that our numerical scheme correctly reproduces the equilibrium profile of the electrical double layer in the limit of both a weakly and a highly charged rod of finite length. For the weakly charged case, we can compare the profiles to the analytical solution of the linearized Poisson−Boltzmann equation around an infinite cylinder. For the highly charged case, we verify that our model reproduces Manning condensation.

For an infinite cylinder of radius a and linear charge density ξ, the linearized Poisson−Boltzmann equation $\nabla^2\phi = \kappa^2\phi$ can be solved analytically. Here $\kappa^{-1} \equiv \lambda_D \equiv (8\pi l_B z^2 \rho_0)^{-1/2}$ is the Debye length. In cylindrical coordinates, this equation reduces to $r^{-1}\partial_r(r\partial_r\phi) = \kappa^2\phi$, with the boundary condition $2\pi a\partial_r\phi|_{r=a} = \xi/\epsilon$. Its solution is $\phi(r) = (2\pi\epsilon)^{-1}Cf(r)$, where $C \equiv \xi/\kappa a$ and $f(r) \equiv K_0(\kappa r)/K_1(\kappa a)$, while K_0 and K_1 are the modified Bessel functions. Inserting this solution into the expression for the equilibrium densities $\rho^{\pm,eq}$ and expand to lowest order in $\phi \ll 1$ ("weak charging"), we obtain

$$\rho^{\pm,eq}(r) = \rho_0\left(1 \mp 2\frac{l_B}{e}Cf(r)\right) \tag{3}$$

This symmetric form around $\rho^{\pm,eq}(r) = \rho_0$ is a signature of the Debye−Hückel (linearized Poisson−Boltzmann) regime.

Figure 2 shows the computed density profiles of co- and counterions near a weakly charged cylinder ($\xi l_B/e = 0.07$). As can be seen from the figure, the numerical results are in excellent agreement with the results obtained of the linearized theory (eq 3). Of course, the accuracy of the lattice model improves as the lattice resolution is increased. However, the good match of Figure 2 was obtained with a fairly coarse-grained description (the diameter of the cylinder corresponded to 8 lattice units).

Manning Condensation. For higher line-charge densities (as is the case for DNA with a bare line-charge density of −2e/0.34 nm), the system enters a nonlinear regime where the deviations of ion profiles from their bulk values are no longer antisymmetric under charge reversal. As our simulations solve the nonlinear Poisson−Boltzmann equation, $\nabla^2\phi = 8\pi ez\rho_0\epsilon^{-1}\sinh(z\phi)$, we should expect to reproduce Manning condensation, an intrinsically nonlinear phenomenon that arises for highly charged cylinders in the limit of low salt concentrations, $\lambda_D \gg a$.[25−27]

The Manning parameter is defined as the number of charges in a length l_B on the rod, $q_0 \equiv \xi l_B/e$. There exists a critical value $q_c = 1$ above which the cylinder exerts such strong force on the counterions that these segregate into a condensed-phase layer close to the cylinder, plus a mobile, gaslike phase outside of this layer. The radius of this "Manning cloud" is $R_M = A(a\lambda_D)^{1/2}$, where A depends only weakly on q_0.[28] The effect of the counterions is to renormalize the bare charge of the rod to a universal value e/l_B.

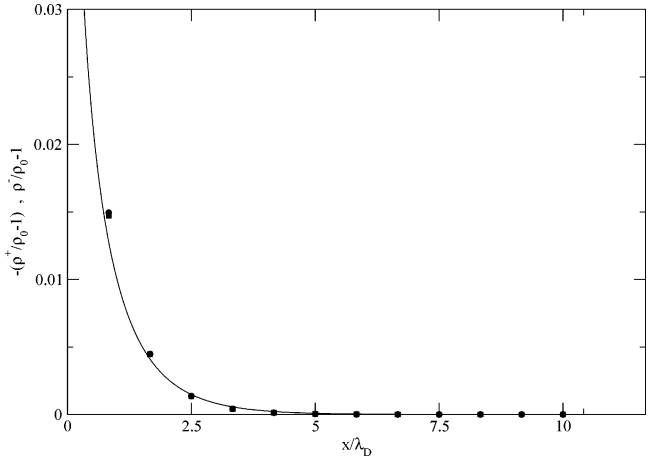


Figure 2. Ion density profiles near a uniformly charged cylinder of finite length, as a function of radial distance to the surface of the cylinder. The reduced charge density per unit length is $\xi l_B/e = 0.07$ (small positive charge density). Both charge density profiles are measured far away from the endpoints of the cylinder, and charge densities are normalized to their value in the bulk (salt concentration). Radial distances are expressed in units of the Debye screening length. The symbols denote simulation results ($\lambda_D/d = 0.15$) for counterions (spheres) and co-ions (squares); solid curves are solutions to the linearized Poisson−Boltzmann equation for an infinite cylinder, see eq 3.

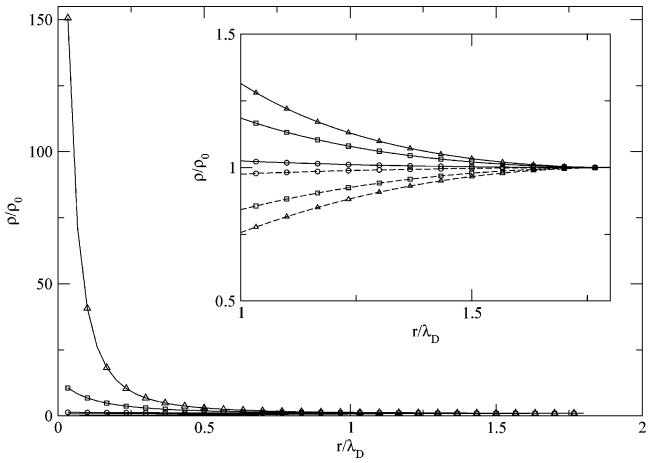


Figure 3. Ionic density profiles near a highly charged cylinder for different values of the Manning parameter q, namely 0.07 (circles), 0.7 (squares), and 2.8 (triangles). Curves are guides to the eye: solid curves represent counterions, dashed curves co-ions. Densities are normalized by their value in the bulk (salt concentration); the distance is normalized by the Debye screening length, $\lambda_D/d = 1.5$. The Manning radius is $R_M/\lambda_D \approx 1$. The inset shows the density profiles for $r > R_M$.

In Figure 3 we show the computed charge density profiles for different line charges: note the symmetry in the profiles for distances above $r/\lambda_D \approx 1$. In Figure 4 we report the values of the counterion concentration, both within and outside the Manning layer $r < R_M$, as a function of q. (Here, q is the Manning parameter resulting from redefining ξ as the linear charge density observed at a distance r from the rod.) Within the Manning layer, the counterion concentration increases roughly quadratically with q, in agreement with analytical

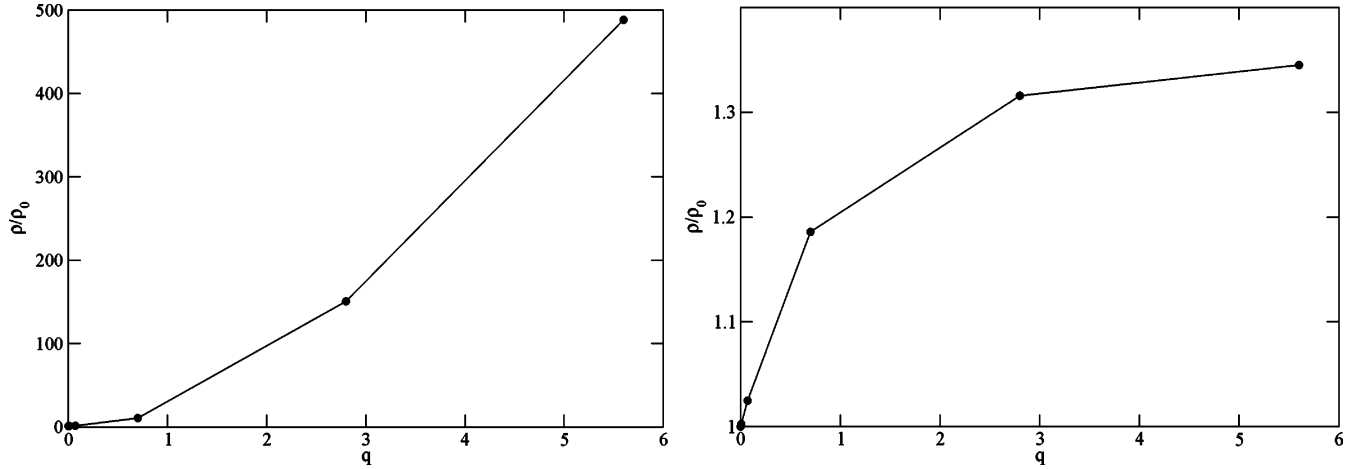


Figure 4. Counterion concentration for different values of the Manning parameter q. Left, within the Manning layer, $r < R_M$; right, outside the Manning layer, $r > R_M$. Symbols are simulation results; curves are guides to the eye. Densities are normalized by the value in the bulk. $\lambda_D/d = 1.5$ and $R_M/\lambda_D \approx 1$.

predictions, whereas outside the Manning layer the counterion concentration is less sentitive to charge increase for $q > 1$. This saturation effect is one of the signatures of Manning condensation. The fact that we do not reach a complete saturation in the counterion concentrations is consistent with the fact that our simulations are performed at small but finite salt concentrations, whereas Manning condensation is a low-salt effect.

IV. Current Modulation during Translocation: Symmetric Pore

Next, we consider the effect of the translocation of a charged rodlike polymer on the ionic currents through the pore. In our simulations, we placed a charged polymer at various positions in or near the translocation pore. We then first equilibrated the charge distributions around the charged polymer. Subsequently, we applied an electric field in a direction parallel to the polymer axis. Immediately after the field is switched on, we observe the onset of electro-osmotic currents. We let the system evolve until these current had reached their steady-state value. We then measure the ionic current, I, as the superposition of conductive plus convective contributions (in the remainder, we assume monovalent ions, $z = 1$, with the same mobility, μ, and we use the elementary charge as our unit of charge: $e = 1$)

$$I \equiv \int\int_{x=\text{const}} \mathrm{dydz}\,[-\mu(\rho^+ - \rho^-)\nabla\phi + (\rho^+ - \rho^-)u_x(y, z)] \tag{4}$$

where μ is the ion mobility and $u_x(y, z)$ is the component of the flow velocity parallel to the symmetry axis. Since mass and charge are conserved to machine precision, and since, in steady state, the total ionic current is a constant along axis Ox, the integral can be taken over any section perpendicular to Ox. To increase precision, we averaged over all planes Oxy. It is worth noting that, for the electric fields used in these simulations, the convective contribution to the current becomes negligible compared to the conductive term.

We measured I for several values of the rod's position, x/l_c, for two values of the screening length, $\lambda_D/d = 1$, 50, and four values of the surface charge densities, $\sigma/(e/d^2) = 0$, 0.1, 0.5, 1. We fixed the rod-to-pore diameter ratio at a value of 1/3. We chose the rodlike polymer to be twice as long as the pore. In Figure 5 we show how the ionic current varies with the position of the translocating polymer. Not surprisingly, we find that if the polymer is uncharged, I invariably decreases as the polymer enters the pore. In other words: the polymer simply blocks part of the current flow. The situation changes when we consider charged polymers. In the high-salt case, $\lambda_D/d = 1$, Figure 5 left, the presence of the polymer in the pore leads to a reduction of the ionic current, but the effect is smaller for highly charged polymers. However, in the low-salt case, $\lambda_D/d = 50$, Figure 5 right, the ionic current I is actually *increased* when a charged polymer is introduced into the pore.

This counterintuitive behavior has recently been observed in experiments[9,29] and large-scale atomistic simulations[10] of DNA 'unsteady' translocation, i.e., electrophoretic translocation of one free DNA oligomer in solution. The advantage of the current approach is that we can separate true steady-state effects from transients that are also present during rapid translocation. In fact, in the atomistic simulations of ref 10 the current enhancement is interpreted as a transient that occurs when DNA leaves the pore: it was attributed to the release of co- and counterions that had accumulated while the DNA was blocking the pore. However, even in ref 10 there is indirect evidence for the effects that we report here, as Figures 4 and 5 in ref 10 also show current enhancement just *before* double-stranded DNA enters the pore.

Figure 6 shows the computed electric current for a neutral polymer with different diameters. The figure shows that the current is proportional to the accessible cross section of the pore. As a consequence, the drop in the current intensity due to the cylinder entering in the pore (defined as the ratio of the currents measured as the cylinder is respectively far from and inside the pore) is proportional to $(d/D)^2$.

This rather trivial effect is nevertheless of practical importance as it is at the basis of experimental techniques

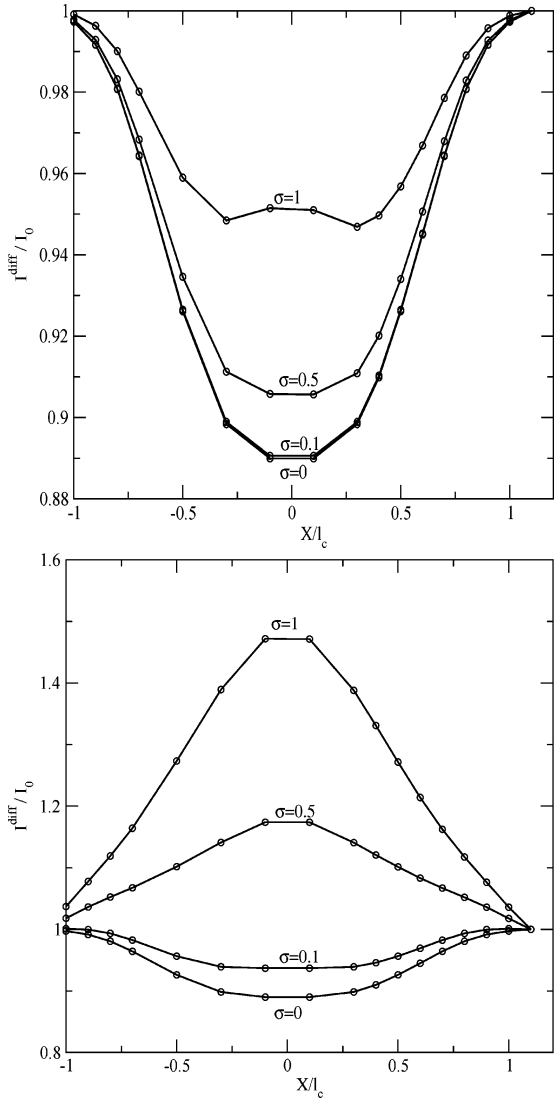


Figure 5. Diffusive ionic current measured at steady state as a function of the cylinder's position X from the pore. Intensities are normalized by the current measured when the cylinder is far from the pore. Results are shown for $\lambda_D/d = 1$ (left) and $\lambda_D/d = 50$ (right) and for different cylinder's surface charge σ. Symbols denote simulation results, while curves are guides to the eye.

to detect the translocation of rodlike macromolecules through a pore. This effect becomes more pronounced as the diameter of the pore becomes comparable to that of the cylinder, which is often the case of in the case of DNA translocation in vivo.

The behavior shown in Figure 5 is due to a competition between these two opposing effects. At low charge, the geometrical effect dominates. But when the cylinder is highly charged and the Debye screening length is large, the electrostatic effect is larger than that due to geometry, and the ionic current is enhanced with respect to the case where the nanopore is empty.

To understand the dependence of the ionic current I on the charge of the translocating polymer, we can use the following rough estimate. We assume that the electric field E and the charge densities ρ^+ and ρ^- are constant inside the pore. If the gap between the cylinder and the pore wall is small, we can approximate this gap by a flat slit. This allows

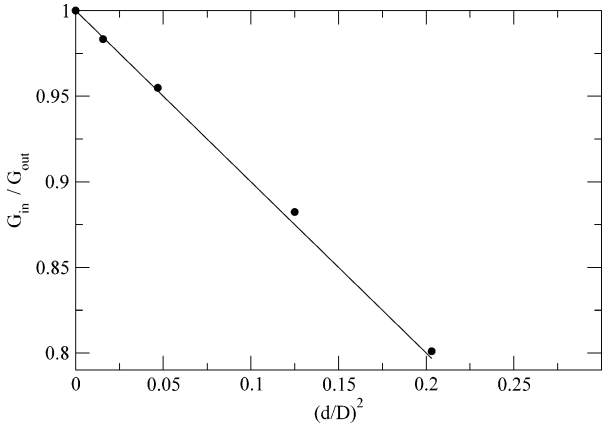


Figure 6. Drop of the conductivity of a nanopore as a function of the cylinder-to-pore cross-section ratio, for an uncharged cylinder. Symbols denote simulation results, and the line is described by the function $1 - (d/D)^2$. The conductivity is proportional to the cross-section of the gap between the pore and the rod.

us to use a simple analytical approximation[30] to estimate the total ionic concentration, $\rho^{+,W} + \rho^{-,W}$, at the pore wall

$$(\rho^{+,W} + \rho^{-,W}) = (\rho_0^+ + \rho_0^-) + \beta\frac{\sigma^2}{2\epsilon} = 2\rho_0\left(1 + \beta\frac{\sigma^2}{4\epsilon\rho_0}\right) \quad (5)$$

Here, $\rho_0^- = \rho_0^+ = \rho_0$ is the concentration of the two species at infinity, and ρ_0 is the salt concentration in the solution bulk.

Using this approximation, we can estimate the relative variation of the current intensity due to the surface charges, by using eq 5, and obtain

$$\frac{I(\sigma, \rho_0)}{I(\sigma = 0, \rho_0)} = 1 + \beta\frac{\sigma^2}{4\epsilon\rho_0}$$

Hence, recalling that $\lambda_D \equiv (8\pi l_B \rho_0)^{-1/2}$ and $l_B \equiv e^2\beta/(4\pi\epsilon)$, we obtain

$$\frac{\Delta I}{I_0} \equiv \frac{I(\sigma, \rho_0) - I(0, \rho_0)}{I(0, \rho_0)} = 8\pi^2\left(\frac{\sigma\lambda_D l_B}{e}\right)^2 \quad (6)$$

We can combine the effects of geometry and surface charge in one single (approximate) expression

$$\frac{\Delta I}{I_0} = \frac{I(\sigma, \rho_0) - I_0(0, \rho_0)}{I_0(0, \rho_0)} = \left(\frac{D-d}{D}\right)^2\left(1 + 8\pi^2\left(\frac{\sigma\lambda_D l_B}{e}\right)^2\right) - 1 \quad (7)$$

where I_0 denotes the current through the unobstructed pore. Note that sign of ΔI can be either positive or negative, depending on whether translocation enhances or decreases the current through the pore. Equation 7 shows that the ionic current increases with increasing Debye screening length. On the other hand, the partial blocking of the nanopore by the translocating polymer tends to decrease the electric current.

Figure 7 shows the relative variation of the current $\Delta I/I_0$ as a function of the dimensionless parameter $\sigma\lambda_D l_B/e$. As can

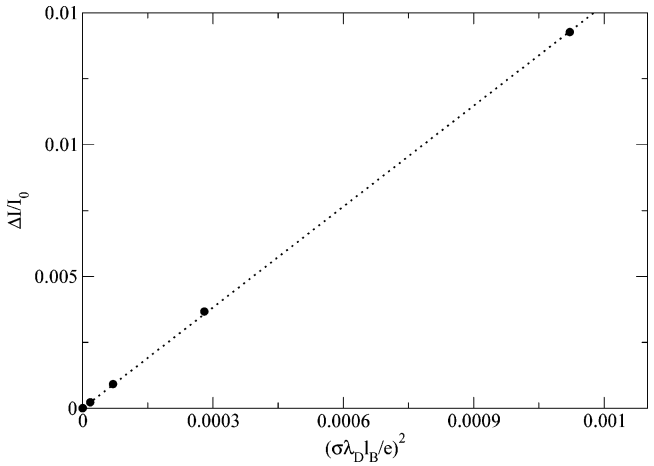


Figure 7. Relative variation of the ionic current, with respect to the ionic current measured for $\sigma = 0$, as a function of the surface charge squared on the cylinder. The surface charge σ is normalized by the length scales λ_D and l_B and the elementary charge e, resulting in a Manning parameter (see text). Symbols denote simulation results for $\lambda_D = 1.5$ and $D - d = 2$ (l.u). The dotted curve shows a quadratic fit to the numerical data, in good agreement with our analytic approximation (see text).

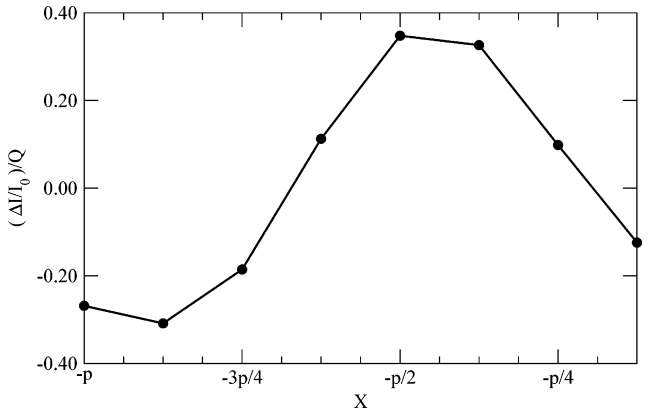


Figure 8. Modulation in the ionic current as ssDNA of radius 5 (l.u.) translocates through a pore of radius 10 (l.u.) at position X, shown as the relative variation $\Delta I/I_0$ of the ionic current with respect to the ionic current I_0, measured for $\sigma = 0$. We normalized this variation to the dimensionless parameter $Q \equiv (\sigma\lambda_D l_B/e)^2$ — this should correct for the explicit dependence of the current on the average charge density of the cylinder (eq 6). The diameters of the ssDNA and the pore are 10 and 20 l.u., respectively, and $p = 8$ l.u. is the helical pitch. Symbols denote simulation results, and the curves are guides to the eye.

be seen from the figure, the numerical data reproduce the quadratic dependence predicted by eq 6.

Of course, in experiments on polyelectrolyte translocation, the (bare) surface charge σ of the polyelectrolyte is dictated by its chemical composition and will usually not be varied. However, it is possible to vary the geometry of the nanopore and, more easily, the concentration of added salt. If we consider typical salt concentrations in the range from 10 mM to 1 M, then our simulations apply to nanopores with typical widths in the range $5 < D < 20$ nm and heights between $10 < h < 40$ nm.

V. Current Modulation during Translocation Past a Charged Probe

The holy grail of experiments on DNA translocation is to extract sequence information from the modulation of the ionic current. Although progress in this direction is certainly being made (see, e.g. ref 31), the generation and interpretation of a DNA translocation "signals" is not straightforward. To study current modulation during DNA translocation, we considered a model polyelectrolyte with constant thickness but a helical (rather than uniform) charge distribution. Translocation of such a helical polyelectrolyte through a symmetric nanopore does not give rise to any current modulation. The reason is that shifting the cylinder by an arbitrary distance h along its axis is equivalent to a rotation of angle $\theta = 2\pi h/p$, where p is the pitch of the helix—and simple rotations of a charged helix in a cylindrically symmetric pore do not change the total current. We therefore considered the interesting case of nanopore with a charged patch that acts as a probe, embedded in the tube wall.

Current nanofabrication techniques allow the production of synthetic pores of the diameter comparable to the thickness of the DNA oligomer (2 nm). Biological pores such as α-hemolysin have similar internal diameters. The design of a nanopore with a charged probe in the wall of the channel is therefore not beyond what is currently feasible. A charged patch acts as a probe because the overlap of the electrical double layers around the patch and the DNA affects the ionic current.

To obtain an estimate of the possible magnitude of such current modulations, we considered single- and double-helix charge distributions on the rod as models for ss- and dsDNA, respectively. As for ssDNA, we placed an infinite cylinder with a single helix charge distribution in the pore and measured the steady-state ionic current in response to a constant voltage for different displacements of the cylinder along its longitudinal axis. The surface charge of the patch was chosen to have the same sign and magnitude as that of the charges on the cylinder, but this assumption is not essential.

Figure 8 shows that, with such an asymmetrically placed probe, the total electric current is indeed modulated upon translocation (or, equivalently, rotation) of the ssDNA rod. In our (discretized) model, the helix pitch was equal to 8 lattice units (l.u.), and hence eight values for the current are reported in Figure 8.

In the case of ssDNA, one might argue that the fact that such molecules have a rather short persistence length (1.5−3 nm) would render the rigid rod model inadequate. However, flexibility would not affect the order of magnitude of the predicted modulations. Moreover, one might envisage situations where the ssDNA is prestretched by linking it to optically trapped colloidal beads on both sides of the pore.

This problem does not occur with the much more rigid double-stranded DNA. We model dsDNA as a cylinder of constant thickness covered by a double helical charge distribution (see Figure 9). We computed similar setups to those for the ssDNA, using the same charge patch inside the pore. To compare the effect of changing configuration

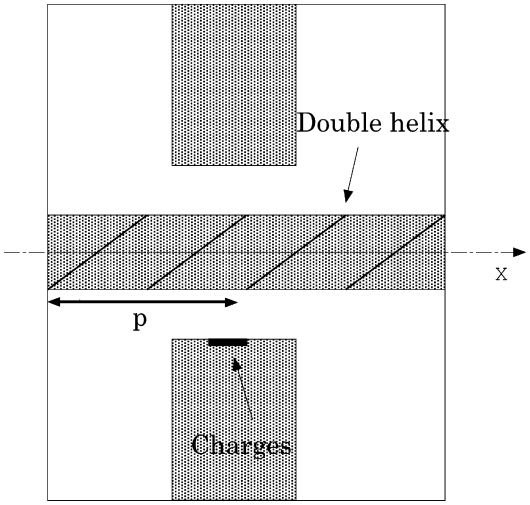


Figure 9. Schematic drawing of the translocation of dsDNA past a charged probe. On the dsDNA, two helical line charges form a double helix. The black area in the pore denotes the charged "probe". X denotes the position of the cylinder and *p* is the pitch of the helix. The ionic current is measured at steady state with a constant voltage bias. The overlap of the electrical double layer around the patch with that around the double helix locally increases the charge density, thus increasing the ionic current with periodicity *p*/2 as the cylinder translocates. A second patch on the opposite side of the pore would enhance the modulations in the current; however, this is only true for radially symmetric charge distributions, such as that of dsDNA.

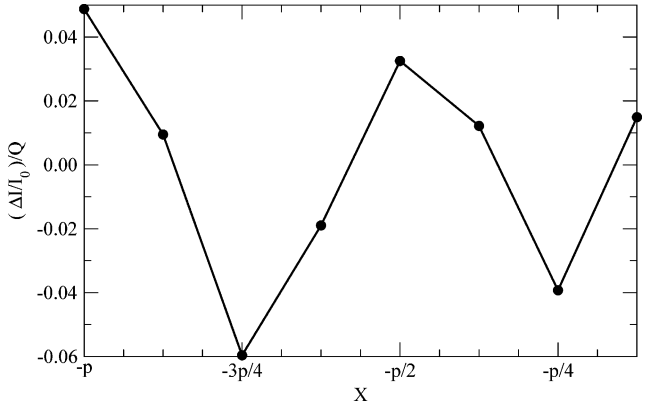


Figure 10. Modulation in the ionic current as dsDNA translocates through a pore at position X, shown as the relative variation of the ionic current with respect to the ionic current I_0, measured for $\sigma = 0$. We normalized this variation to the dimensionless parameter $Q \equiv (\sigma\lambda_D l_B/e)^2$ — this should correct for the explicit dependence of the current on the average charge density of the cylinder (eq 6). The diameters of the dsDNA and the pore are 10 and 20 l.u., respectively, while $p = 8$ l.u. is the helix pitch. The period of the modulations is half of that for ssDNA. Symbols denote the simulation results, and curves are guides to the eye. The actual modulation is much smaller (10^{-4}) than the average current. Hence, small discretization errors show up clearly in this figure: in the absence of such errors the modulation would be perfectly periodic with a period *p*/2.

(from single to double helix), we divide the current modulation by $Q \equiv (\sigma\lambda_D l_B/e)^2$, which should describe the effect of a smeared out charge density σ on the ionic current (eq 6). Our simulations again indicate that the ionic current is modulated upon translocation of the dsDNA (see Figure 10). The period of the modulations is now half the pitch of one helix. The screening length was chosen to be smaller than the pitch and comparable to the size of the charged patch. This choice maximizes the inhomogeneity of the electric potential. For dsDNA, these conditions can be achieved with a 1 M monovalent salt in water at room temperature, which is in the range of experimental parameters. This yields a screening-length of 0.3 nm for a pitch of 3 nm.

The amplitude of the scaled current modulation during translocation is larger (approximately seven times) for the single helix than the double helix. The value of this ratio suggests that the scaled current modulation is not simply proportional to the variance in the surface charge density. We have not attempted to quantify this dependence. Both for ssDNA and dsDNA, the unscaled current modulation is small (typically, between $1{:}10^2$ and $1{:}10^4$). In experiment, such a signal might still be detectable if the position of the DNA is modulated periodically.

We have not attempted to optimize the conditions for maximal modulation of the translocation current. Within the current model, there is certainly room for improvement. For instance, $\Delta I/I_0$ increases with increasing ratio of the diameter *d* of the DNA to the diameter *D* of the pore. In our simulations, the maximum value of *d*/*D* was determined by the lattice resolution. By increasing this resolution, larger *d*/*D* ratios can be studied. However, for DNA fitting snugly in the pore, it would not be justified to ignore the corrugation (grooves) of the DNA surface. Hence, we have not attempted to study this limit.

Conclusions

In summary, we have presented Lattice-Boltzmann simulations of the modulation in ionic current associated with the translocation of charged oligomers through nanopores under an applied electric field. Our first observation is that, depending on the conditions (charge of the polyelectrolyte, ionic strength of the solution), translocation may either decrease or, more surprisingly, increase the ionic current.

In addition, we have presented simulations that indicate that the modulation of the ionic current during translocation can be made sensitive to the microscopic charge distribution of the DNA by placing a localized charged "probe" on the wall of the nanopore. We argue that this effect should be experimentally observable, even though the predicted current modulations are small. Of course, thermal ("Nyquist") noise in the electrical current may partially obscure such measurements. However, unlike the thermal noise, the translocation-induced current modulation increases with increasing current. Moreover, if the DNA is connected to a colloidal bead that is moved periodically, then the signal-to-noise problem can be reduced by employing phase-sensitive detection.

Acknowledgment. The work of the FOM Institute is part of the research program of FOM and is made possible by financial support from The Netherlands Organization for Scientific Research (NWO).

Note Added after ASAP Publication. This article was released ASAP on March 22, 2006, with the incorrect Received Date. The correct version was posted on April 17, 2006.

References

(1) *Electrostatic effects in Soft Matter and Biophysics*; Holm, C., Kèkicheff, P., Podgornik, R., Eds.; NATO Sci. Series, Kluwer Academic Publisher: Dordrecht, 2001.

(2) Gelbart, W. M.; Bruinsma, R. F.; Pincus, P. A. DNA-inspired electrostatics. *Phys. Today* **2000**, *53*, 38−44.

(3) Probstein, R. F. *Physicochemical Hydrodynamics*; Butterworth: Boston, MA, 1989.

(4) Stroock, A. D.; Dertinger, S. K. W.; Ajdari, A.; Mezic, I.; Stone, H. A.; Whitesides, G. M. Chaotic mixer for microchannels. *Science* **2002**, *295*, 647−651.

(5) Kasianowicz, J. J.; Brandin, E.; Branton, D.; Deamer, D. W. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 13770−13773.

(6) Li, J.; Stein, D.; McMullan, C.; Branton, D.; Aziz, M. J.; Golovchenko, J. A. Ion-beam sculpting at nanometre length scales. *Nature* **2001**, *412*, 166−169.

(7) JLi, J. L.; Gershow, M.; Stein, D.; Brandin, E.; Golovchenko, J. A. *Nat. Mater.* **2003**, *2*, 611−615.

(8) Storm, A. J.; Chen, J. H.; Zandbergen, H. W.; Dekker, C. Translocation of double-strand DNA through a silicon oxide nanopore. *Phys. Rev. E* **2005**, *71*, 051903.

(9) Chang, H.; Kosari, F.; Andreadakis, G.; Alam, M. A.; Vasmatzis, G.; Bashir, R. DNA-mediated fluctuations in ionic current through silicon oxide nanopore channels. *Nano Lett.* **2004**, *4*, 1551−1556.

(10) Aksimentiev, A.; Heng, J. B.; Timp, G.; Schulten, K. Microscopic kinetics of DNA translocation through synthetic nanopores. *Biophys. J.* **2004**, *87*, 2086−2097.

(11) Heng, J. B.; Ho, C.; Kim, T.; Timp, R.; Aksimentiev, A.; Grinkova, Y. V.; Sligar, S.; Schulten, K.; Timp, G. Sizing DNA using a nanometer-diameter pore. *Biophys. J.* **2004**, *87*, 2905−2911.

(12) Aksimentiev, A.; Schulten, K. Imaging alpha-hemolysin with molecular dynamics: Ionic conductance, osmotic permeability, and the electrostatic potential map. *Biophys. J.* **2005**, *88*, 3745−3761.

(13) Hoogerbrugge, P. J.; Koelman, J. M. V. A. Simulating Microscopic Hydrodynamic Phenomena with Dissipative Particle Dynamics. *Europhys. Lett.* **1992**, *19*, 155−160.

(14) Groot, R. D.; Warren, P. B. Dissipative particle dynamics: Bridging the gap between atomistic and mesoscopic simulation. *J. Chem. Phys.* **1997**, *107*, 4423−4435.

(15) Groot, R. D. Electrostatic interactions in dissipative particle dynamics-simulation of polyelectrolytes and anionic surfactants. *J. Chem. Phys.* **2003**, *118*, 11265−11277.

(16) Malevanets, A.; Kapral, R. Mesoscopic model for solvent dynamics. *J. Chem. Phys.* **1999**, *110*, 8605−8613.

(17) Ihle, T.; Kroll, D. M. Stochastic rotation dynamics: A Galilean-invariant mesoscopic model for fluid flow. *Phys. Rev. E* **2001**, *63*, 020201.

(18) Succi, S. *The Lattice Boltzmann Equation for Fluid Dynamics and Beyond*; University Press: Oxford, 2001.

(19) Ladd, A. J. C. Numerical Simulations of Particulate Suspensions via a Discretized Boltzmann-Equation. 1. Theoretical Foundation. *J. Fluid Mech.* **1994**, *271*, 285−309.

(20) Gonzalez-Segredo, N.; Nekovee, M.; Coveney, P. V. Three-dimensional lattice-Boltzmann simulations of critical spinodal decomposition in binary immiscible fluids. *Phys. Rev. E* **2003**, *67*, 046304.

(21) Capuani, F.; Pagonabarraga, I.; Frenkel, D. Discrete solution of the electrokinetic equations. *J. Chem. Phys.* **2004**, *121*, 973−986.

(22) Thompson, A. P. Nonequilibrium molecular dynamics simulation of electro-osmotic flow in a charged nanopore. *J. Chem. Phys.* **2003**, *119*, 7503−7511.

(23) Alder, B. J.; Wainwright, T. E. Decay of the Velocity Autocorrelation Function. *Phys. Rev. A* **1970**, *1*, 18−21.

(24) Press, W. H.; Teukolski, S. A.; Vetterling, W. T. *Numerical Recipes*; Cambridge University Press: Cambridge, 1996.

(25) Oosawa, F. Interaction between parallel rodlike macroions. *Biopolymers* **1968**, *6*, 1633−1647.

(26) Manning, G. S. Limiting Laws and Counterion Condensation in Polyelectrolyte Solutions I. Colligative Properties. *J. Chem. Phys.* **1969**, *51*, 924−933.

(27) Manning, G. S.; Ray, J. Counterion condensation revisited. *J. Biomol. Struct. Dyn.* **1998**, *16*, 461−476.

(28) O'Shaughnessy, B.; Yang, Q. Manning-Oosawa counterion condensation. *Phys. Rev. Lett.* **2005**, *94*, 048302.

(29) Smeets, R.; Keyser, U.; Krapf, D.; Wu, M. Y.; Dekker, N.; Dekker, C. *Preprint*, 2005.

(30) Israelachvili, J. N. *Intermolecular and Surface Forces;* Academic Press: New York, 1991.

(31) Vercoutere, W. A.; Winters-Hilt, S.; DeGuzman, V. S.; Deamer, D.; Ridino, S. E.; Rodgers, J. T.; Olsen, H. E.; Marziali, A.; Akeson, M. Discrimination among individual Watson−Crick base pairs at the termini of single DNA hairpin molecules. *Nucleic Acids Res.* **2003**, *31*, 1311−1318.

(32) Bocquet, L.; Barrat, J. L. *Phys. Rev. Lett.* **1993**, *70*, 2726.

(33) Butler, J. C.; Angelini, T.; Tang, J. X.; Wong, G. C. L. *Phys. Rev. Lett.* **2003**, *91*, 028301.

(34) Bostrom, M.; Williams, D. R. M.; Ninham, B. W. *J. Phys. Chem. B* **2002**, *106*, 7908.

CT050340G

JCTC Journal of Chemical Theory and Computation

A Coarse-Grained Molecular Dynamics Study of Carbon Nanoparticle Aggregation

Sergei Izvekov[†] and Angela Violi*

Department of Mechanical Engineering, University of Michigan, Ann Arbor, Michigan 48109

Received January 21, 2006

Abstract: A multiscale coarse-graining procedure is used to study carbonaceous nanoparticle assembly. The computational methodology is applied to an ensemble of 10 000 nanoparticles (or effectively 2 million total carbon atoms) to simulate the agglomeration of carbonaceous nanoparticles using coarse-grained atomistic-scale information. In particular, with the coarse-graining approach, we are able to assess the influence of nanoparticle morphology and temperature on the agglomeration process. The coarse-graining of the interparticle force field is accomplished applying a force-matching procedure to data obtained from trajectories and forces from all-atom molecular dynamics simulation. The coarse-grained molecular dynamics results show rich and varied clustering behaviors for different particle morphologies. They are shown to reproduce accurately the structural properties of the nanoparticles systems studied, while allowing for molecular dynamics simulations of much larger self-assembled nanoparticles systems.

Introduction

In combustion environments, soot formation (particles with an average size of thousands of nanometers) is a fascinating multiscale problem of nanoparticle growth, both in length scale and in time scale. Figure 1 shows the chemical and physical processes for the formation of nanoparticle agglomerates starting from simple molecules, going through the formation of polycyclic aromatics and particle inception, and ending up with particle coagulation and agglomeration leading to primary particles (50 nm in diameter).[1] The processes involved in the formation of particles exhibit a wide range of time scales, spanning pico- or nanoseconds for intramolecular processes to milliseconds for intermolecular reactions. At the same time, the length scale also undergoes significant changes, going from a few angstroms for small polycyclic aromatic hydrocarbons to hundreds of nanometers for particle aggregates.

When the nanoparticles are able to reach a critical size (which is not always the case), they begin to aggregate on two separate length scales. The assembly on the smallest of these scales seems to be governed by the strongest interparticle interactions, so interesting patterns reminiscent of a spinodal decomposition process become evident. These smaller assemblies will, in turn, aggregate on an even larger

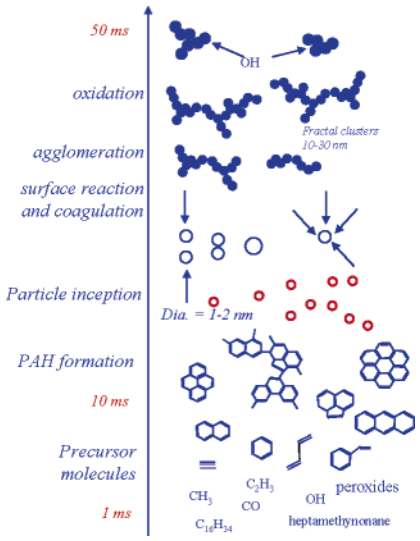


Figure 1. Particle formation in flames.

* Corresponding author phone: (734) 615-6448; fax: (734) 647-9379; e-mail: avioli@umich.edu.

[†] Permanent address: Department of Chemistry, University of Utah, Salt Lake City, UT 84112.

10.1021/ct060030d CCC: $33.50
Published on Web 04/26/2006

length scale through a process that resembles colloidal nucleation. Thus, the molecular science of the nanoparticles bridges with the statistical mechanics of nonequilibrium self-assembly and nucleation over two disparate length and time scales in a truly fascinating way.

These nanoparticles are quite robust and, in fact, stabilize a remarkably large number of radical sites on their surfaces. The high concentration of radical sites, along with the small size of the nanoparticles, leads to serious health implications when a person breathes these particles into their lungs.

This paper reports on a new multiscale coarse-graining simulation approach for bridging the time and length scales in the growth of the carbon nanocluster self-assembly. After a description of the methodology, results are shown for the assembly of carbonaceous nanoparticles of different morphologies in the temperature range 1100−1600 K.

Computational Methodology

Carbonaceous nanoparticle agglomeration is influenced by large length- and time-scale motions that extend to mesoscopic scales, that is, one micrometer or more in length and one microsecond or more in time. To increase the time and length scales accessible in simulations and to be able to simulate nanoparticle assembly, it is necessary to describe the particles on a more coarse-grained (CG) level. The strategy of these CG approaches is generally the same: to achieve a simpler description of the effective interactions in a given system while not losing the ability of the resulting models to predict the properties of interest. The existing approaches for obtaining effective CG potentials target the reproduction of a few average structural properties seen in atomistic simulations or experiments, for example, using an iterative adjustment of potential parameters starting from an approximation based on potentials of mean force,[2,3] or the Inverse Monte Carlo technique,[4] or they may be parametrized to match thermodynamic properties.[3−8] These approaches rapidly become computationally expensive and less reliable for systems with an "aggressive" coarse-graining, such as the nanoparticles systems. Recently, a novel approach for a reconstruction of CG potentials from an underlying explicit atomistic molecular dynamics (MD) simulation has been proposed.[9−11] The method, which is called "multiscale coarse-graining" (MS−CG), is built upon the force-matching (FM) technique developed in ref 9. The MS−CG method exploits the fact that an application of the FM procedure to the CG images from underlying atomistic trajectory/force data should produce the effective interaction between CG structural units as it seen in the atomistic dynamics. The most obvious way to map an atomic group into a CG site is to associate it with its center of mass (CM) because the force acting on the CM of the atomic group can be straightforwardly evaluated from the atomistic MD data.

The algorithmic development of the MS−CG method is presented elsewhere.[9,12−14] In the MS−CG framework, the site−site force $\mathbf{f}_i^p(r_{ij}) = f^p(r_{ij})\mathbf{n_{ij}}$ (acting from site j on site i), where r_{ij} is the modulus of the vector $\mathbf{r}_{ij} = \mathbf{r}_j - \mathbf{r}_i$ connecting two sites and $\mathbf{n}_{ij} = \mathbf{r}_{ij}/\mathbf{r}_{ij}$, was represented by the cubic splines connecting a set of points $\{r_k\}$ with $r_{k\max}$ equal to the cutoff, was used to represent the $f(r)$

$$f(r,\{r_k\},\{f_k\},\{f''_k\}) = A(r,\{r_k\})f_i + B(r,\{r_k\})f_{i+1} + C(r,\{r_k\})f''_i + D(r,\{r_k\})f''_{i+1} \quad (1)$$

with $r \in [r_i, r_{i+1}]$.

In eq 1, A, B, C, and D are known functions of r and $\{r_k\}$, and $\{f_k\}$ and $\{f''_k\}$ are tabulations of $f(r)$ and its second derivative on a radial mesh $\{r_k\}$. The parameters $\{f_k, f''_k\}$ are to be obtained from the fit. The choice of spline is advantageous for two reasons. First, it preserves the continuity of $f_i^p(r_{ij})$ and its first two derivatives and gives the flexibility of fitting arbitrary pairwise forces. Second, a spline is a linear form of its parameters $\{f_k, f''_k\}$ and therefore permits a reduction of the least-squares problem in the force-matching procedure[15] to an overdetermined system of linear equations.[11,16]

A linear dependence of the force on the fitting parameters permits a splitting of the least-squares fit into two phases, making the whole procedure computationally much less expensive. Practically, the MS−CG method is implemented as follows. All sets of atomic configurations recorded along the trajectories of the reference MD simulation are partitioned into blocks, each containing L configurations. For each block, the reference total force $\mathbf{F}^{\mathrm{ref}}_{\alpha il}$ which acts on the ith atom of kind α in the lth atomic configuration of the block and the same force predicted using the representation in eq 1 are required to be equal

$$-\sum_{\beta=1,K}\sum_{j=1,N_\beta}[f'(r_{\alpha il,\beta jl},\{r_{\alpha\beta,k}\},\{f_{\alpha\beta,k}\},\{f''_{\alpha\beta,k}\})]\mathbf{n}_{\alpha il,\beta jl} = \mathbf{F}^{\mathrm{ref}}_{\alpha il}$$
$$\alpha = 1, ..., \mathrm{K};\ i = 1, ..., N_\alpha;\ l = 1, ..., L \quad (2)$$

with respect to $\{f_{\alpha\beta,k}, f''_{\alpha\beta,k}\}$ which are force parameters. In eq 2, $r_{\alpha il,\beta jl}$ is the distance between atoms $\{\alpha i\}$ and $\{\beta i\}$ in the lth atomic configuration of the block; N_β and K are, respectively, the number of atoms of kind β and the total number of atomic kinds in the systems, and $\{r_{\alpha\beta,k}\}$ is a spline mesh which might be for different $\{\alpha\beta\}$. The size L of each block should be large enough to ensure that eq 2 overdetermines the force parameters. Standard equations which are linear with respect to $\{f_{\alpha\beta,k}, f''_{\alpha\beta,k}\}$ must be included into eq 2 for the $f(r)$'s first derivative, $f'(r)$, to be continuous across the boundary between two intervals.[17]

The solutions of eq 2 need to be averaged over all of the blocks to get the effective force field. More discussion of the MS−CG methodology is given in ref 9.

As a result of this new approach, the effective phase space of the system is significantly reduced in size, as is the number of costly force calculations. In the MS−CG approach, the CG parameters are therefore developed "on the fly" from the actual atomistic-level forces coming from a full MD "presimulation". Hence, the term "multiscale" is used for the MS−CG method, meaning that changes at the atomistic level of the force field will systematically propagate upward via the force-matching algorithm to the CG representation of the system. The MS−CG model will allow the simulations of the nanoparticle systems in this project to bridge upward in both length and time scale, so as to better access the properties influenced by those scales.

The capability of the MS−CG approach to develop the CG parameters "on the fly" from the actual atomistic-level forces coming from a full MD is very important in the study of particle formation in combustion. The newly nucleated particles grow by coagulation and coalescent collisions as well as surface growth. All these processes happen at the same time, and while the CG approach can describe particle coagulation, *the surface growth* needs to be carried out at the atomistic level. The changes that can occur on single nanoparticles because of surface growth at the level of the atomistic MD will propagate upward to the MS−CG effective force field via the force-matching algorithm in a multiscale fashion.

The MS−CG method has been successfully applied to several complex condensed-phase and biological systems.[9] The method proved to be workable for systems of carbonaceous nanoparticles.[18] This is an important result because the coarse-graining of nanoparticle systems presents a special challenge because of their shapes, that is, ranging from the highly symmetric C60 to more amorphous and irregular shapes. They also generally interact through the sum of many weaker interactions rather than via strong electrostatics in, for example, an aqueous environment.

The MS−CG method is computationally efficient, and this is an important feature of the approach because the coarse-graining of the system interactions may greatly reduce the transferability of the resulting models between different thermodynamic conditions. MS−CG allows fast parametrization of the CG force fields at the thermodynamic conditions of interest instead of using "universal" models that increase the reliability of the CG simulations.

Results

Molecular Models. The initial configurations of the carbonaceous nanoparticles chosen for this study are obtained using a combination of kinetic Monte Carlo and MD methodologies (**A**tomistic **M**odel for **P**article **I**nception, or AMPI code). The AMPI code has been developed to study the transformations that occur during the transition from the gas phase to particle inception,[19−21] providing information on both the chemical structure and the configuration of the nanoparticles. The capability of the AMPI code has been validated in different conditions, and nanoparticles have been characterized in terms of chemical structure and components, and relationships between the structure and pathways, structure and properties, and structure and reactivity population of active sites have been addressed. In particular, computed properties of nanoparticles have been compared with experimental data in terms of H/C trends, particle morphology, depolarization ratio, and free radical concentration.

Figure 2a,b shows two representative structures obtained with the AMPI code in different combustion environments. These results, in turn, correlate with the curvature of carbon layers seen in high-resolution transmission electron microscopy images of combustion-generated nanoparticles produced in flames, which show a different curvature of the carbon nanoparticles containing an amorphous, graphitic, or fullerenic nanostructure.[22−24]

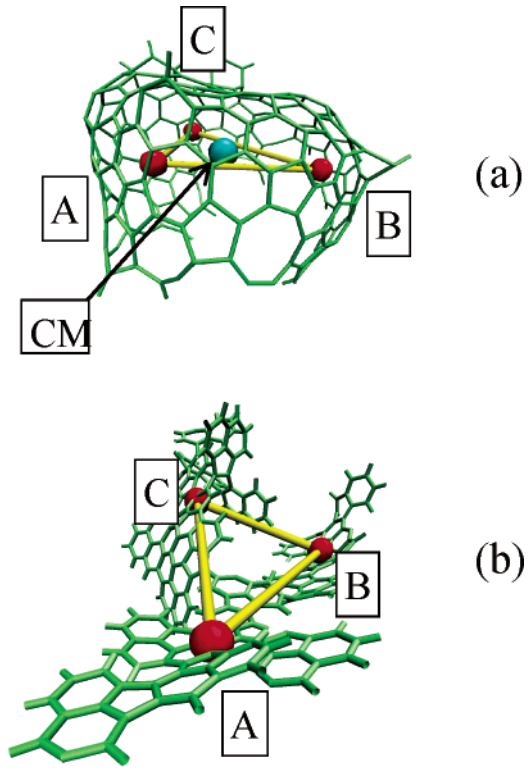


Figure 2. Atomistic and coarse-grained representations of two carbonaceous nanoparticles: (a) round (one-site and three-site representations) and (b) flat (three-site representation).

The particle in panel (a), named "round", has an aspect ratio close to unity with a formula of $C_{188}H_{53}$, while the compound in panel (b) (the "flat" particle) has a composition of $C_{202}H_{90}$. Both round and flat nanoparticles have a size of around 17 Å in the longest dimension. The round precursor consists of a network of 5−8-membered fused rings such as buckyballs, whereas the flat one has four planar aromatic moieties connected by freely rotating bonds.

CG Procedure. To construct and validate an effective force field for CG unit interactions (force-matching algorithm), the following steps are required:

(1) Atomistic MD simulations of the particle ensembles

(2) Partitioning of the systems in CG sites and a buildup of the CG interactions

(3) Comparison between the CG results and the fully atomistic data

1. Atomistic MD Simulation of the Particle Ensembles. To address the importance of particle morphology on the assembly behavior, three ensembles are considered in this study: (1) 32 round nanoparticles, (2) 32 flat nanoparticles, and (3) 16 round and 16 flat nanoparticles.

The CHARMM force-field parameters[25] are utilized to describe the interactions between nanoparticles in a constant NVT ensemble. Figure 3 reports the snapshots from MD simulations at 1100 K for the three ensembles. It is important to note that the radical chemistry is not included in these simulations, and all dangling bonds of the carbon atoms on the nanoparticles have been hydrogen-terminated.

The round particles tend to cluster, and they show a preferred orientation that is back-to-back (left panel). Some

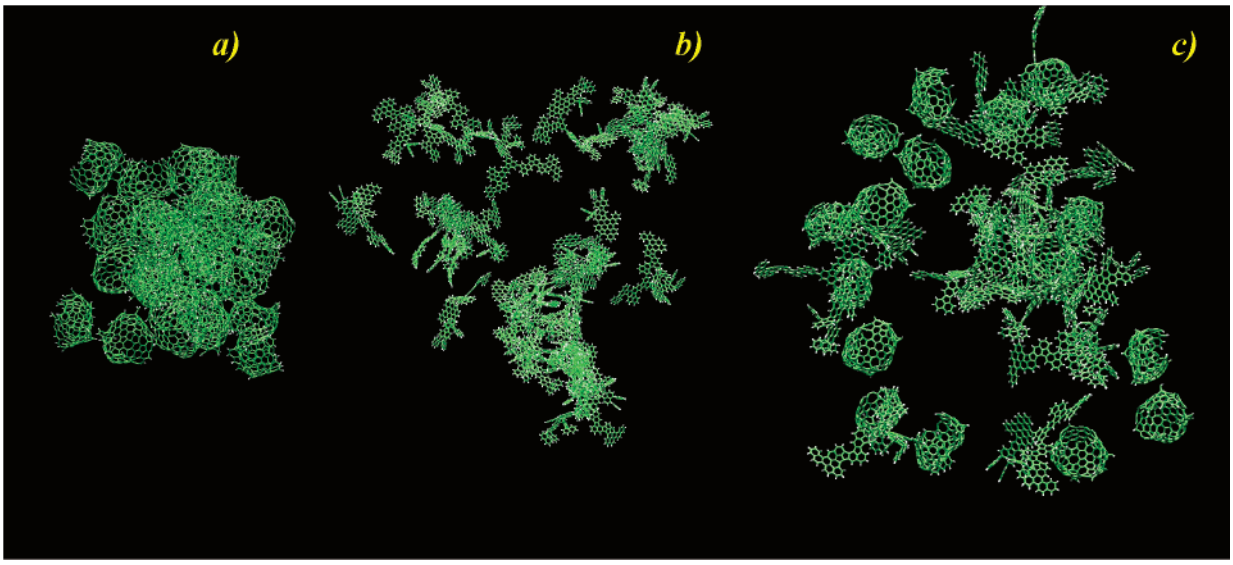


Figure 3. Snapshots from MD simulations of three ensembles of 32 nanoparticles (panel a, "round"; panel b, "flat"; and panel c, "round/flat") at 1100 K.

of the sheetlike particles do cluster in small agglomerates (middle panel), while the third panel shows a snapshot of the mixed systems in which some of the particles drift away from the ensemble and some form subclusters. In addition, the clustering behavior for flat−flat particles is quite different from that for the round−round ones: the orientation is more edge-to-edge or stacking, and the interacting particles are tightly constrained.

The current models present in the literature report a simplified notion of the coagulation process, in which two particles coagulate after collision if the kinetic energy of the particles is lower than their interaction potential, which depends on van der Waals attraction and Born repulsion forces.[26] The results presented in Figure 3 show a much more complex process in which particle morphologies play an important role. This information is necessary to build a realistic model for particle formation.

2. Partitioning of the System. Carbonaceous nanoparticles can be coarse-grained into a different number of sites. Generally, a larger number of interaction sites is advantageous because the CG simulation yields a better resolution. However, a higher computational cost and a lower accuracy of the MS−CG procedure as well as a more expensive CG simulation are major drawbacks of such models.

As reported in ref 18, the one- and three-site models of the particles perform very well in reproducing the structural properties of the carbonaceous nanoparticles obtained in an all-atom MD simulation. Figure 2 shows the CG sites for the two nanoparticles studied here. The three-site model accounts for orientational degrees of freedom of the particles themselves, which are very important when dealing with complex processes, such as self-assembly or structural transitions of nanoparticles. For the first two ensembles (32 particles of the same morphology), a three-site CG representation is used. The "round" particles were mapped into a three-site geometry by slicing each of them into three segments of approximately equal weight by planes perpendicular to the opening plane of the particle. The CG interaction sites are associated with the CM of the corresponding segments as shown in Figure 2 (sites A, B, and C). Similar to the round particle, the flat nanoparticle is divided into three segments in accordance with its three "wings". The CG sites were associated with the CM of the corresponding segments as shown in Figure 2 (sites A, B, and C).

For the third system composed of 32 mixed particles, the one-site model is used for the round particles, in which the CG site is associated with the CM of the whole particle, and the three-site model for the flat particles. This choice allows the reduction of the computational time without losing accuracy.[18] In particular, for the 32-particle system, the acceleration of the simulation is about 2000−5000 times.

2.1. Buildup of the CG Site Interactions. MS−CG models with such an "aggressive" coarse-graining as those of one and three sites are expected to have a low transferability among different phases (e.g., at different temperatures and different densities). As an example, a MS−CG model generated from the all-atom simulation at a temperature of 1600 K to study particle aggregation at a lower temperature could be questionable. Ideally, the CG model has to be parametrized at the same thermodynamic conditions as those at which a simulation of the phenomena of interest is intended.

To assess the issue of transferability, MS−CG force fields for ensembles 1 and 2 of round and flat particles are computed at two temperatures (1100 and 1600 K) and two particle densities. The sizes of the supercells used for the simulations are 10.5 nm ("high" density) and 12.0 nm ("low" density), respectively. For the mixed system—ensemble 3—the MS−CG force field is calculated at 1600 K and for a single density (size of the supercell of 10.0 nm).

Data from the all-atom simulations are collapsed into trajectories and forces of the CG sites, and the resulting CG trajectory and force data are used as input to the FM algorithm. The trajectory and the force data from the atomistic simulations are sampled at an interval of 0.1 ps.

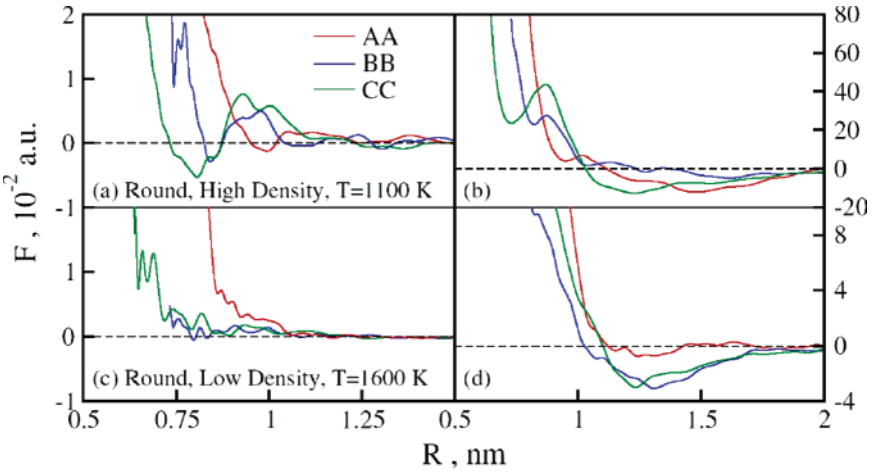


Figure 4. Effective pairwise forces (left) and potentials (right) for selected pairs of CG sites as a function of the intersite separation calculated by the force-matching method. Panels a and b: round particles at 1100 K, "high" density. Panels c and d: round particles at 1600 K, "low" density.

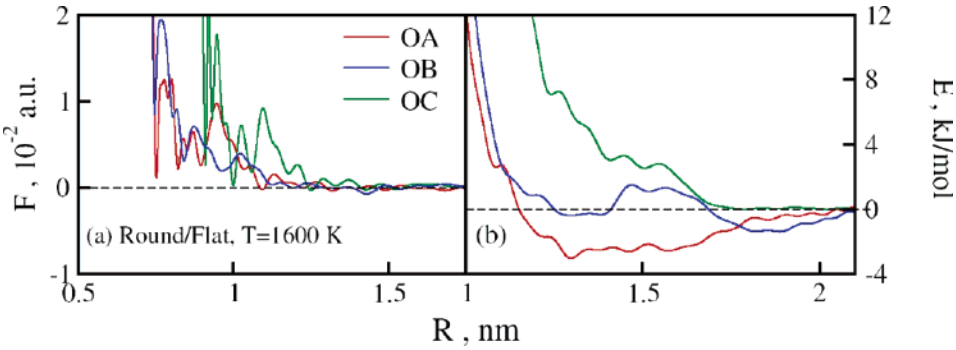


Figure 5. Effective pairwise forces (left) and potentials (right) for the MS−CG model of the mixed system of round (one-site representation − O) and flat (three-site representation − A, B, and C) nanoparticles.

For the simulation at 1600 K and low density, the overdetermined system of linear FM equations is solved repeatedly for 15 000 sets of atomic configurations from 3.0 ns MD trajectory data, with each set consisting of two configurations ($L = 2$ in eq 2) and then averaged over all of the sets. For the other cases, the all-atom simulation is increased to 6.0 ns with a proportional increase in the number of sampled all-atom configurations.

The force field is represented by a spline over a mesh with a grid spacing of approximately 0.01 nm. The potentials are calculated by integrating out the respective terms in the force fields and then shifting them to zero at the cutoff radius. The classical MD simulations with the new CG models were performed using the DL_POLY 2.9 computer code.[27]

The tabulated force fields and potentials for the systems investigated in this paper are available in a DL POLY code format.

Figure 4 reports the diagonal effective potential profiles for the system of round particles at two temperatures and different densities. The effective forces appear to be less sensitive to a change in the average particle density than to variations in the temperature. As a consequence, the difference between the two force fields can be attributed mainly to a change in the temperature.

At 1600 K (Figure 4c,d), the strongest interactions are between the A−A and C−C pairs of sites, and the potential shows a minimum at −3.1 kJ/mol that is smaller than k_BT = 13.3 kJ/mol at that temperature. At 1100 K, the binding energies of the A−A and C−C pairs dramatically increase to about −13 kJ/mol (k_BT at 1100 K = 9.1 kJ/mol), with the equilibrium distances shifted toward long separation distances suggesting a possibility of aggregation. The similarity of potential profiles for different pairs of sites is just a manifestation of the high symmetry of the round particles.

An analogous tendency is observed for the system of flat particles. However, the binding energies are smaller at both temperatures (1100 and 1600 K) than those of the system of round particles and at the same time are below the values of

Table 1. Properties of the Gas-Phase Dimer from Atomistic and CG Models of the Nanoparticles Parameterized at T = 1600 K and Low Density[a]

model	R_0 (nm)	Upot (kJ/mol)	θ (deg)
round atomistic	1.34	−105.4	50.3
round one-site	1.23	7.5	
round three-site	1.35	−16.74	44.5
flat atomistic	0.92	−156.4	21.1
flat three-site	0.97	28.2	23.7

[a] Shown are R_0, the distance between centers of mass of particles; Upot, the interaction energy; and θ, the angle between the planes through CG sites (for three-site models).

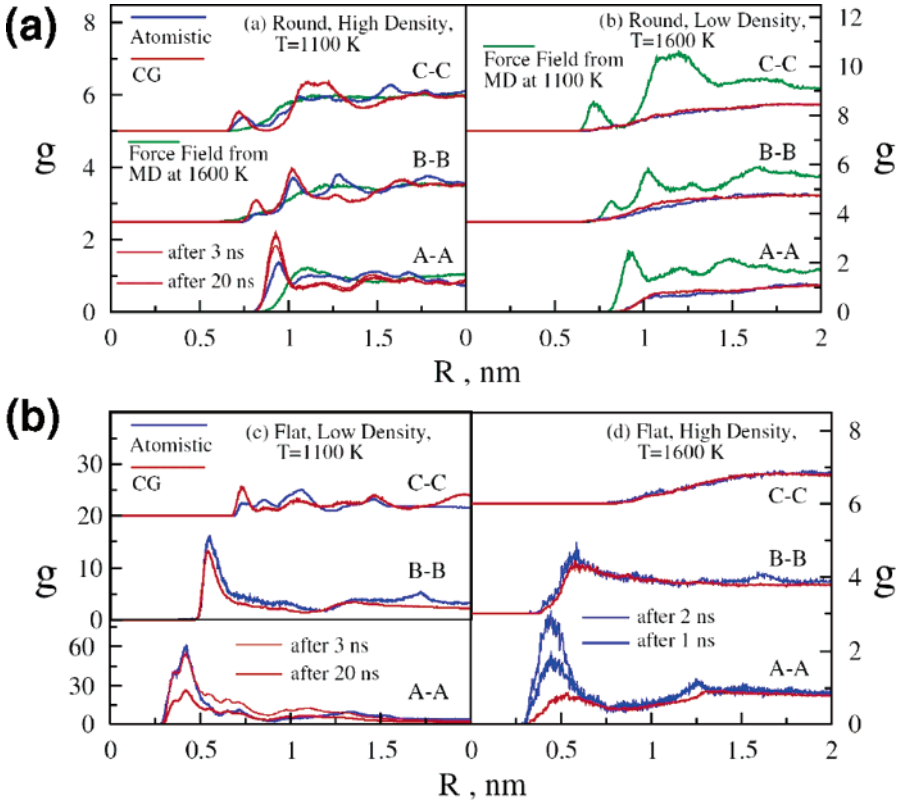


Figure 6. (a) Comparison between selected RDFs of the all-atom MD simulations (blue lines) and those obtained from the MS−CG MD simulation (red lines) for ensemble 1—round particles—at 1100 and 1600 K. The green line represents the RDFs obtained using the force field computed at 1600 K (left panel) and 1100 K (right panel). (b) Comparison between selected RDFs of the all-atom MD simulations (blue lines) and those obtained from the MS−CG MD simulation (red lines) for ensemble 2—flat particles—at 1100 and 1600 K.

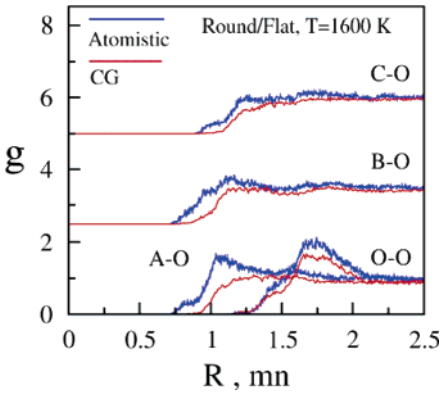


Figure 7. Comparison between RDFs of the all-atom MD simulations (blue lines) and those obtained from the MS−CG MD simulation (red lines) for ensemble 3—round and flat particles.

k_BT. These results indicate some difficulties for the MS−CG flat particle ensemble to aggregate even at low temperatures.

For the mixed system—ensemble 3—Figure 5 shows the effective interactions of one-site round particles with three-site flat particles at 1600 K. The round particle tends to bind strongly to the A site of the flat particle, and the binding energy is comparable to the one obtained for the system of round particles—ensemble 1 (see Figure 4)—suggesting a stronger ability of the mixed system to aggregate compared with the system of flat particles—ensemble 2.

Because the all-atom simulation of the mixed system yielded poor statistics for the MS−CG procedure because of the small number of sites contained in the reference system compared with the system of flat particles, for the simulations reported below, we replace the interactions between flat particles with those from ensemble 2 obtained at 1600 K and low density.

3. Validation of the MS−CG Models. To test the MS−CG model, we first examine the properties of nanoparticle dimers in a vacuum. The results for the MS−CG model parametrized at 1600 K and low density are summarized in Table 1. The computed CG geometry of the dimers matches very well the data obtained from the atomistic model.

For the many-particle systems (ensembles 1−3), the MS−CG simulations are carried out using the same geometries and thermodynamic conditions as those of the reference all-atom system.

To check the transferability of the MS−CG models between different thermodynamic conditions, the structural properties obtained through the MS−CG model at different temperatures were compared with the results obtained at the all-atom level. The MS−CG simulations were initiated using the configuration which was final in the reference all-atom MD simulations. In Figure 6, some comparisons for the systems of interest are presented.

The MS−CG models perform very well in reproducing the reference structural properties: some discrepancies observed for the 1100 K simulations and for the high density

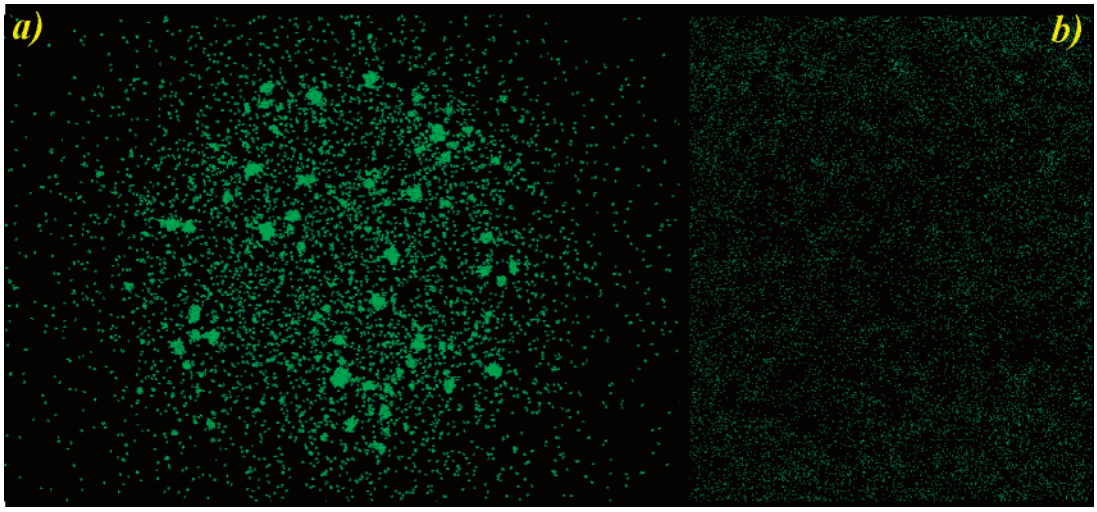


Figure 8. Snapshots from CG MD simulations for 10 000 round (panel a) and 10 000 flat (panel b) nanoparticles at 1100 K. The CG modes were fitted to all-atom MD simulations at the same temperature.

simulations at 1600 K are due to the limited simulation time that caused the radial distribution functions (RDFs) in the all-atom simulations not to be fully converged. This conjecture is supported by a comparison of RDFs from the MS−CG simulations at early stages of the dynamics. Figure 6a, panel a, reports the RDF from the MS−CG simulation of the system of round particles at 1100 K: a first peak at the AA RDF recorded after 3 ns shows a better agreement with the atomistic value than the peak recorded after 20 ps. This suggests that the all-atom RDFs were indeed not fully converged.

It should be noted that, because the dynamics in the MS−CG MD simulation are significantly accelerated, a fair comparison of time scales between the MS−CG and all-atom simulations is difficult. Similar conclusions can be drawn looking at the RDFs from different simulations of the system of flat particles shown in Figure 6b, panels c and d.

In Figure 6a—panel (a)—the RDFs computed using the force field generated at 1600 K are shown as green lines, and the structure is dramatically different from that of the MS−CG model generated at 1100 K. The RDFs from the 1600 K model bear a memory of the structure at that temperature which is very different form the low-temperature case.

A similar comparison is reported in Figure 6a—panel (b)—when the MS−CG model parametrized at 1100 K is used to compute the RDF at 1600 K.

The failure of the MS−CG models to correctly reproduce the structure at temperatures different from those at which they were parametrized unequivocally points toward a very poor transferability of the MS−CG models of the nanoparticles.

The RDFs from the 1100 K simulations as well as those from the simulation at 1600 K and high density exhibit a significant number of peaks, which indicates the formation of particle agglomerates. The peaks tend to amplify as the simulations proceed, thus, further supporting the aggregation behavior in these systems.

Figure 7 shows the RDFs from the MS−CG simulation of the mixed system. Some mismatch between atomistic and MS−CG RDFs is, partly, due to the use of the MS−CG model derived from the flat particle simulation, to describe interactions of flat particles with each other. Relatively well pronounced peaks in the AO and OO RDFs indicate a significant tendency to aggregate for these pairs of sites.

Behavior on Large Scales. To study the formation of bigger agglomerates (soot primary particles), the MS−CG MD simulations are then conducted on a system of 10 000 CG particles (or effectively 2 million atoms) for 500 ns. The density chosen for these runs is 5×10^{15} particles/cm^{-3}. Experimental data report a number concentration of around 1×10^{14} for slightly sooting flames of ethylene.[28] However, the measured value is an average, and there are pockets inside the region of interest with higher density. The simulations reported in this paper are relative to those areas because higher density will increase the collision frequency and reduce the computational time required to form the first nuclei. Three ensembles are considered: (1) 10 000 round particles; (2) 10 000 flat particles, and (3) 5000 flat and 5000 round particles.

Figure 8 shows snapshots from the end points of the simulations for ensembles 1 and 2. The system of round nanoparticles clearly exhibits the presence of clusters at 1100 K, which are observed to persist for the entire length of the simulation. The cluster size was identified to be around 15−20 nm. The system of flat particles shows a much lower degree of clustering for the same temperature and density. These results confirm what had been observed in a more limited fashion at the microscale level and described earlier.

A fusion of two clusters of round particles as well as a split of the cluster into several fragments (typically two) has been frequently observed during the simulations, as well as an absorption of the "gas-phase" particles and an evaporation of nanoparticles from clusters.

A snapshot from the simulation of ensemble 3 of round and flat particles at 1600 K is shown in Figure 9. This system shows a better ability to aggregate compared with the flat particle system, but worse than that in the pure round particle system. The agglomerates are of small size and contain either flat and round particles or only round particles. Flat particles tend to agglomerate through the A−A or A−B interactions,

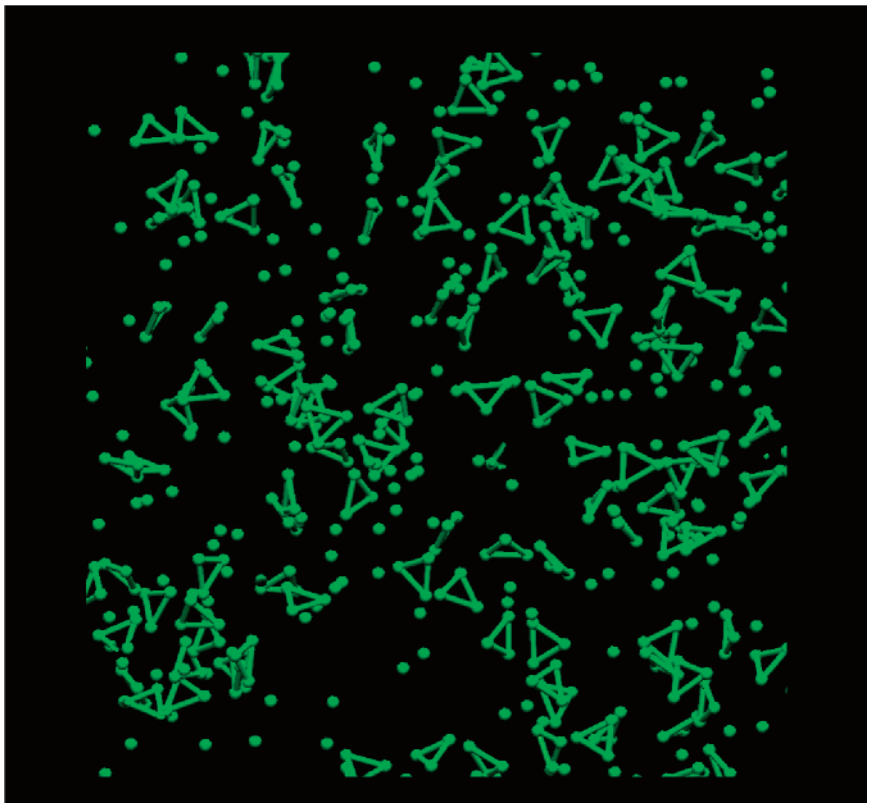


Figure 9. Snapshot from the CG MD simulation for a system of 5000 round and 5000 flat nanoparticles at 1100 K (one-site representation for the round particles and three-site representation for the flat particles).

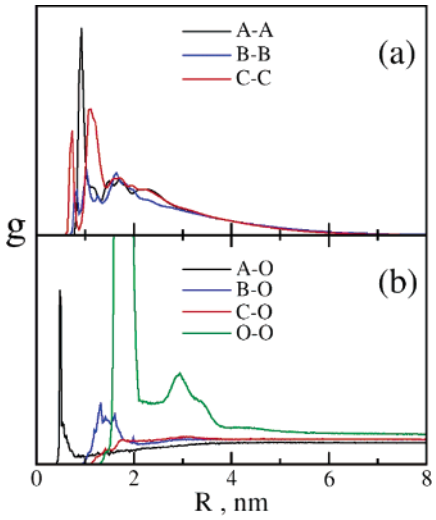


Figure 10. RDFs of the MS−CG MD simulations for ensemble 1—round particles (a) and for ensemble 3—round and flat particles (b). In the upper panel round particles have a three-site representation, A, B, and C and in the lower panel round particles have a one-site representation, O; flat nanoparticles have a three-site representation, A, B, and C.

which are the strongest. The round particles agglomerate primarily among each other and then to the A and B sites of the flat particles. Figure 10 shows the RDFs of the round system (a) reported in Figure 8a and the RDFs for the round/flat system (b) reported in Figure 9. From the width of the peaks, it is possible to derive information regarding the cluster sizes formed during the simulations. For the 10 000-round-particles system, the average cluster size is around 16 nm, and for the round/flat system, clusters of 6 nm are identified between the sites A, B, and C−O.

Conclusions

The recently developed MS−CG method for obtaining effective pairwise CG force fields from atomistic force and trajectory data has been shown here to be very successful in developing CG models for systems of nanoparticles. The methodology presented is quite general, and in this paper, we applied it to an important real world example, that is, particle formation from combustion sources.

This approach provides a connection between the various time and length scales in the nanoparticles self-assembly problem, together with an unprecedented opportunity for the understanding of the atomistic interactions underlying nanoparticles aggregation and self-assembly. The MS−CG methodology indeed provides an ideal multiscale route for bridging the explicit atomistic representation of the nanoparticles produced by the AMPI simulation method with a more coarse-grained representation necessary to study the agglomeration of the nanoparticles, which in the end can effectively involve millions of carbon atoms.

The agreement in the structural properties is an important result of this methodology because it reveals that even a simple MS−CG model with one or three interaction sites is able to accurately reproduce the structural properties for a system of particles having such a large size and complex geometries. The MS−CG methodology allows rapid parametrization of the CG force fields at the thermodynamic

conditions of interest instead of using "universal" models that increase the reliability of the CG simulations.

Future work will be conducted on mixed systems with a large variety of particles of different sizes and morphologies. Also, the results reported in this paper are relative to the early nucleation of the particles. Once the first nuclei are formed, the process becomes much faster than the one reported in the figures.

Acknowledgment. This research is supported in part by a National Science Foundation Nanoscale Interdisciplinary Research Team grant (EEC-0304433).

References

(1) Homann, K. H.; Wagner, H. G. Chemistry of Carbon Formation in Flames. *Proc. R. Soc. London. Ser. A* **1967**, *307* (1489), 141.

(2) Meyer, H.; Biermann, O.; Faller, R.; Reith, D.; Müller-Plathe, F. Coarse Graining of Nonbonded Interparticle Potentials Using Automatic Simplex Optimization to Fit Structural Properties. *J. Chem. Phys.* **2000**, *113*, 6264.

(3) Shelley, J. C.; Shelley, M. Y.; Reeder, R. C.; Bandyopadhyay, S.; Klein, M. L. A Coarse Grain Model for Phospholipid Simulations. *J. Phys. Chem. B* **2001**, *105* (19), 4464.

(4) Murtola, T.; Falck, E.; Patra, M.; Karttunen, M.; Vattulainena, I. Coarse-Grained Model for Phospholipid/Cholesterol Bilayer. *J. Chem. Phys.* **2004**, *121* (18), 9156.

(5) Marrink, S.; Mark, A. Molecular Dynamics Simulation of the Formation, Structure, and Dynamics of Small Phospholipid Vesicles. *J. Am. Chem. Soc.* **2003**, *125* (49), 15233.

(6) Marrink, S. J.; de Vries, A. H.; Mark, A. E. Coarse Grained Model for Semiquantitative Lipid Simulations. *J. Phys. Chem. B* **2004**, *108* (2), 750.

(7) Brannigan, G.; Brown, F. L. H. Solvent-Free Simulations of Fluid Membrane Bilayers. *J. Chem. Phys.* **2004**, *120* (2), 1059.

(8) Brannigan, G.; Philips, P. F.; Brown, F. L. H. Flexible Lipid Bilayers in Implicit Solvent. *Phys. Rev. E* **2005**, *72* (1), 011915.

(9) Izvekov, S.; Voth, G. A. A Multiscale Coarse-Graining Method for Biomolecular Systems. *J. Phys. Chem. B* **2005**, *109* (7), 2469.

(10) Izvekov, S.; Voth, G. A. Multiscale Coarse Graining of Liquid-State Systems. *J. Chem. Phys.* **2005**, *123* (13), 134105.

(11) Izvekov, S.; Violi, A.; Voth, G. A. Systematic Coarse-Graining of Nanoparticle Interactions in Molecular Dynamics Simulation. *J. Phys. Chem. B* **2005**, *109* (36), 17019.

(12) Izvekov, S.; Parrinello, M.; Burnham, C. J.; Voth, G. A. Effective Force Fields for Condensed Phase Systems from ab Initio Molecular Dynamics Simulation: A New Method for Force-Matching. *J. Chem. Phys.* **2004**, *120* (23), 10896.

(13) Hone, T. D.; Izvekov, S.; Voth, G. A. Fast Centroid Molecular Dynamics: A Force-Matching Approach for the Predetermination of the Effective Centroid Forces. *J. Chem. Phys.* **2005**, *122* (5), 54105.

(14) Izvekov, S.; Voth, G. A. Effective Force Field for Liquid Hydrogen Fluoride from ab Initio Molecular Dynamics Simulation Using the Force-Matching Method. *J. Phys. Chem. B* **2005**, *109* (14), 6573.

(15) Ercolessi, F.; Adams, J. B. Interatomic Potentials from 1st-Principles Calcualtions—The Force-Matching Method. *Europhys. Lett.* **1994**, *26* (8), 583.

(16) Lawson, C. L.; Hanson, R. J. *Solving Least Squares Problems*; Prentice-Hall: Englewood Cliffs, New Jersey, 1974.

(17) De Boor, C. *Practical Guide to Splines*; Springer-Verlag: New York, 1978.

(18) Violi, A.; Izvekov, S.; Voth, G. A. Nanoparticle Agglomeration. *Preprints of Symposia − American Chemical Society, Division of Fuel Chemistry* **2005**, *50* (1), 54.

(19) Violi, A. Modeling of Soot Particle Inception in Aromatic and Aliphatic Premixed Flames. *Combust. Flame* **2004**, *139*, 279.

(20) Violi, A.; Voth, G. A.; Sarofim, A. F. The Relative Roles of Acetylene and Aromatic Precursors during Soot Particle Inception. *Proc. Combust. Inst.* **2005**, *30*, 1343.

(21) Violi, A.; Sarofim, A. F.; Voth, G. A. Kinetic Monte Carlo Molecular Dynamics Approach To Model Soot Inception. *Combust. Sci. Technol.* **2004**, *176* (5−6), 991.

(22) Goel, A.; Hebgen, P.; Vander Sande, J. B.; Howard, J. B. Combustion Synthesis of Fullerenes and Fullerenic Nanostructures. *Carbon* **2002**, *40* (2), 177.

(23) Goel, A.; Howard, J. B.; Vander Sande, J. B. Size Analysis of Single Fullerene Molecules by Electron Microscopy. *Carbon* **2004**, *42*, 1907.

(24) Vander Wal, R. L.; Tomasek, A. J.; Street, K.; Hull, D. R.; Thompson, W. K. Carbon Nanostructure Examined by Lattice Fringe Analysis of High-Resolution Transmission Electron Microscopy Images. *Appl. Spectrosc.* **2004**, *58* (2), 230.

(25) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM—A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *J. Comput. Chem.* **1983**, *4* (2), 187.

(26) Narsimhan, G.; Ruckenstein, E. The Brownian Coagulation of Aerosols over the Entire Range of Knudsen Numbers—Connection between the Sticking Probability and the Interaction Forces. *J. Colloid Interface Sci.* **1985**, *104* (2), 344.

(27) Smith, W.; Forester, T. R. DL_POLY_2.0: A General-Purpose Parallel Molecular Dynamics Simulation Package. *J. Mol. Graphics* **1996**, *14*, 136.

(28) D'Alessio, A.; D'Anna, A.; Minutolo, P.; Sgro, L. A.; Violi, A. On the Relevance of Surface Growth in Soot Formation in Premixed Flames. *Proc. Combust. Inst.* **2000**, *28* (2), 2547.

CT060030D

JCTC Journal of Chemical Theory and Computation

Replica Exchange and Multicanonical Algorithms with the Coarse-Grained United-Residue (UNRES) Force Field

Marian Nanias,[†] Cezary Czaplewski,[†,‡] and Harold A. Scheraga*[,†]

Baker Laboratory of Chemistry and Chemical Biology, Cornell University, Ithaca, New York 14853-1301, and Faculty of Chemistry, University of Gdansk, Sobieskiego 18, 80-952 Gdansk, Poland

Received October 14, 2005

Abstract: Three algorithms, namely, a replica exchange method (REM), a replica exchange multicanonical method (REMUCA), and a replica exchange multicanonical method with replica exchange (REMUCAREM), were implemented with the coarse-grained united-residue force field (UNRES) in both Monte Carlo and molecular dynamics versions. The MD algorithms use the constant-temperature Berendsen thermostat, with the velocity Verlet algorithm and a variable time step. The algorithms were applied to one peptide (20 residues of alanine with free ends; ala_{20}) and two small proteins, namely, an α-helical protein of 46 residues (the B domain of the staphylococal protein A; 1BDD) and an $\alpha+\beta$ protein of 48 residues (the *Escherichia coli* MItd Lysm Domain; 1E0G). Calculated thermodynamic averages, such as canonical average energy and heat capacity, are in good agreement among all simulations for poly-L-alanine, showing that the algorithms were implemented correctly and that all three algorithms are equally effective for small systems. For protein A, all algorithms performed reasonably well, although some variability in the calculated results was observed, whereas for a more complicated $\alpha+\beta$ protein (1E0G), only replica exchange was capable of producing reliable statistics for calculating thermodynamic quantities. Finally, from the replica exchange molecular dynamics results, we calculated free-energy maps as functions of the RMSD and radius of gyration for different temperatures. The free-energy calculations show correct folding behavior for poly-L-alanine and protein A, while for 1E0G, the native structure had the lowest free energy only at very low temperatures. Hence, the entropy contribution for 1E0G is larger than that for protein A at the same temperature. A larger contribution from entropy means that there are more accessible conformations at a given temperature, making it more difficult to obtain an efficient coverage of conformational space to obtain reliable thermodynamic properties. At the same temperature, ala_{20} has the smallest entropy contribution, followed by protein A, and then by 1E0G.

1. Introduction

Efficient sampling algorithms have been an essential component of methods for studying protein structure and dynamics in structural biology and theoretical chemistry. A variety of sampling algorithms have been used in our laboratory, and depending on whether the goal is global optimization or folding simulations, they can be be categorized in the following way.

For successful prediction of the three-dimensional structure of a protein (based solely on its amino acid sequence), several classes of algorithms have been used. The first class includes modifications of the Metropolis Monte Carlo procedure,[1,2] such as Monte Carlo with minimization,[3,4] electrostatically

* Corresponding author tel: (607) 255-4034; fax: (607) 254-4700; e-mail: has5@cornell.edu.

[†] Cornell University.

[‡] University of Gdańsk.

10.1021/ct050253o CCC: $33.50
Published on Web 01/21/2006

driven Monte Carlo,[5,6] conformational family Monte Carlo,[7] and replica exchange Monte Carlo with minimization.[8] The second class includes deformation-based methods, such as the diffusion-equation method,[9] the distance-scaling method,[10] and the self-consistent basin-to-deformed-basin method.[11,12] The third class includes genetic algorithms such as the conformational space annealing (CSA) method.[13−15] For the study of protein-folding pathways, recently applied molecular dynamics (MD) with the united-residue (UNRES) force field[16−19] have been shown to be particularly effective. To evaluate thermodynamic properties, another class of sampling methods is necessary. This is because minimization-based methods violate the condition of microscopic reversibility required for producing Boltzmann statistics, and although methods such as molecular dynamics or Metropolis Monte Carlo can be used for estimating thermodynamic properties as well as for a global search, they easily become trapped for complex systems and, thus, are not the most effective methods for studying large systems.

The origins of one of the most popular advanced sampling methods, the replica exchange method (also known as exchange Monte Carlo[20] or parallel tempering[21]), can be traced back to the work carried out by Swendsen and Wang[22] for spin-glass systems, and the more familiar form of the algorithm was developed by Geyer[23] with his use of Metropolis-coupled Markov chain Monte Carlo. In the replica exchange method, several copies (replicas) of the system are simulated with standard Metropolis Monte Carlo[1,2] or molecular dynamics procedures (each replica differing from the others in a particular way, usually in temperature), while permitting an exchange among the replicas, and thus surmounting barriers in the rugged conformational energy landscapes. This method has been applied extensively in protein-folding simulations using both lattice[24−27] and off-lattice models.[28−32]

Recently, much attention has been paid to generalized ensemble algorithms whose advantage is efficient sampling of the conformational energy landscape. In this approach, efficient sampling does not mean locating the global minimum as quickly as possible but rather covering the landscape in such a way as to provide accurate statistics. Two well-known methods are the multicanonical algorithm[33,34] (also known as entropy sampling[35,36]) and simulated tempering[37] (also referred to as the method of expanded ensembles[38]). The multicanonical algorithm performs a one-dimensional random walk in energy space, while simulated tempering follows a random walk in temperature space, thereby inducing a random walk in the space of potential energy. Although these algorithms are generally too expensive for locating global minima,[39] they are useful for producing accurate statistics for thermodynamic averages of observed variables. However, the application of these algorithms is nontrivial and very tedious; in particular, the need to obtain the proper sampling weights often limits the use of generalized ensemble techniques.[40]

Due to the fact that the replica exchange method alleviates the problem of the tedious estimation of weight factors in the multicanonical algorithms, combinations of replica exchange with generalized ensemble methods have been developed, for example, REMUCAREM,[41] that is, replica exchange multicanonical algorithm with replica exchange; others include replica exchange simulated tempering or simulated tempering replica exchange.[42] Other modifications of replica exchange include replica exchange with solute tempering,[43] model hopping,[44] Hamiltonian replica exchange,[45] and the replica-exchange method using a generalized effective potential.[46]

Having demonstrated that the coarse-grained UNRES protein model is helpful in surmounting problems with all-atom models,[18,47] we apply the replica exchange method (REM), the replica exchange multicanonical method (REMUCA), and the REMUCAREM method, in both Monte Carlo (MC) and molecular dynamics versions, to the UNRES model in the present work. The advantage of replica exchange lies in its simplicity, and in contrast to other methods, it is not very sensitive to the few parameters involved therein (such as the cooling schedule in simulated tempering or the successful estimation of weight factors in multicanonical algorithms). The power of REMUCA lies in the effective estimate of the multicanonical weight factors from replica exchange simulations. REMUCAREM further exploits the idea of running several replicas of multicanonical simulations with different sets of multicanonical weights. The motivation behind the present work is to test the applicability of these algorithms to determine the thermodynamic properties of large systems. The ability to compute thermodynamic properties will thereby enable us to improve our UNRES model and, consequently, improve protein-folding simulations, that is, bring our simulated results closer to experimental ones.

2. Methods

2.1. UNRES Force Field. All the above-mentioned algorithms were implemented with the united-residue force field; hence, in this section, the UNRES model of polypeptide chains and the corresponding force field are described briefly. First, the UNRES model used with Monte Carlo procedures is described, followed by a description of the UNRES force field for molecular dynamics.

In the UNRES model,[48−58] a polypeptide chain is represented by a sequence of α-carbon (C^α) atoms linked by virtual bonds with attached united side chains (SC) and united peptide groups (p). Each united peptide group is located in the middle between two consecutive α-carbons. Only these united peptide groups and the united side chains serve as interaction sites, the α-carbons serving only to define the chain geometry. All virtual bond lengths (i.e., $C^\alpha\cdots C^\alpha$ and $C^\alpha\cdots$SC) are fixed; the distance between neighboring C^α's is 3.8 Å, corresponding to trans peptide groups, while the side-chain angles (α_{SC} and β_{SC}) and virtual-bond (θ) and dihedral (γ) angles can vary. The UNRES force field has been derived as a restricted free-energy (RFE) function of an all-atom polypeptide chain *plus the surrounding solvent*, where the all-atom energy function is averaged over the degrees of freedom that are lost when passing from the all-atom to the simplified system (i.e., the degrees of freedom of the solvent, the dihedral angles χ for rotation about the bonds in the side chains, and the torsional angles λ for

rotation of the peptide groups about the $C^{\alpha}\cdots C^{\alpha}$ virtual bonds).[52,53,59] The RFE is further decomposed into factors arising from interactions within and between a given number of united interaction sites.[53] Expansion of the factors into generalized Kubo cumulants[60] facilitated the derivation of approximate analytical expressions for the respective terms,[52,53] including the *multibody* or *correlation* terms, which are derived in other force fields from structural databases or on a heuristic basis.[61] The theoretical basis of the force field is described in detail in ref 53. The energy of the virtual-bond chain for Monte Carlo simulations is expressed by eq 1.

$$U_{\mathrm{MC}} = \sum_{i<j} U_{\mathrm{SC}_i\mathrm{SC}_j} + w_{\mathrm{SCp}} \sum_{i\neq j} U_{\mathrm{SC}_i\mathrm{p}_j} + w_{el} \sum_{i<j-1} U_{\mathrm{p}_i\mathrm{p}_j} + w_{\mathrm{tor}} \sum_i U_{\mathrm{tor}}(\gamma_i) + w_{\mathrm{tord}} \sum_i U_{\mathrm{tord}}(\gamma_i,\gamma_{i+1}) + w_{\mathrm{b}} \sum_i U_{\mathrm{b}}(\theta_i) + w_{\mathrm{rot}} \sum_i U_{\mathrm{rot}}(\alpha_{\mathrm{SC}_i},\beta_{\mathrm{SC}_i}) + w_{\mathrm{corr}}^{(3)} U_{\mathrm{corr}}^{(3)} + w_{\mathrm{corr}}^{(4)} U_{\mathrm{corr}}^{(4)} + w_{\mathrm{turn}}^{(3)} U_{\mathrm{turn}}^{(3)} + w_{\mathrm{turn}}^{(4)} U_{\mathrm{turn}}^{(4)} \quad (1)$$

The term $U_{\mathrm{SC}_i\mathrm{SC}_j}$ represents the mean free energy of the hydrophobic (hydrophilic) interactions between the side chains, which implicitly contains the contributions from the interactions of the side chain with the solvent. The term $U_{\mathrm{SC}_i\mathrm{p}_j}$ denotes the excluded-volume potential of the side-chain−peptide-group interactions. The peptide-group interaction potential ($U_{\mathrm{p}_i\mathrm{p}_j}$) accounts mainly for the electrostatic interactions (i.e., the tendency to form backbone hydrogen bonds) between peptide groups p_i and p_j. U_{tor}, U_{tord}, U_{b}, and U_{rot} are the virtual-bond dihedral angle torsional terms, virtual-bond dihedral angle double-torsional terms, virtual-bond angle bending terms, and side-chain rotamer terms, respectively; these terms account for the local propensities of the polypeptide chain. The terms $U_{\mathrm{corr}}^{(m)}$ represent *correlation* or *multibody* contributions from the coupling between backbone−local and backbone−electrostatic interactions, and the terms $U_{\mathrm{turn}}^{(m)}$ are correlation contributions involving *m* consecutive peptide groups; they are, therefore, termed turn contributions. The correlation contributions were derived[52,53] from a generalized-cumulant expansion[60] of the RFE of the system consisting of the polypeptide chain and the surrounding solvent. The multibody terms are indispensable for reproduction of regular α-helical and β-sheet structures.

The internal parameters of $U_{\mathrm{p}_i\mathrm{p}_j}$, U_{tor}, U_{tord}, $U_{\mathrm{corr}}^{(m)}$, and $U_{\mathrm{turn}}^{(m)}$ were derived by fitting the analytical expressions to the RFE surfaces of model systems computed by quantum mechanics at the MP2/6-31G** ab initio level,[57,58] while the parameters of $U_{\mathrm{SC}_i\mathrm{SC}_j}$, $U_{\mathrm{SC}_i\mathrm{p}_j}$, U_{b}, and U_{rot} were derived by fitting the calculated distribution functions to those determined from the PDB.[51] The *w*'s are the weights of the energy terms, and they were determined (together with the parameters within each cumulant term and the well depths of the side-chain pairwise interaction potential $U_{\mathrm{SC}_i\mathrm{SC}_j}$) by hierarchical optimization[62] of the potential-energy function.

Molecular dynamics with UNRES requires an extra degree of freedom, namely, the vibrations of the virtual bonds, which are treated with an additional harmonic potential. The complete UNRES potential-energy function for molecular dynamics is then expressed by the following equation:[18]

$$U_{\mathrm{MD}} = U_{\mathrm{MC}} + w_{\mathrm{vib}} \sum_i U_{\mathrm{vib}}(d_i) \quad (2)$$

where U_{MC} is the Monte Carlo UNRES potential energy described above (eq 1) and $U_{\mathrm{vib}}(d_i)$, d_i being the length of the *i*th virtual bond, are the simple harmonic potentials defined as $U_{\mathrm{vib}}(d_i) = (1/2)k_{d_i}(d_i - d^{\circ}_i)^2$, where k_{d_i} is the force constant of the *i*th virtual bond, currently set at 500 kcal/(mol Å^2), and d°_i is the average length (corresponding to that used in the fixed-bond UNRES potential) of the *i*th virtual bond; for example, $d^{\circ}_i = 3.8$ Å for a $C^{\alpha}\cdots C^{\alpha}$ virtual bond corresponding to a trans peptide group. As in previous work,[18] the weight w_{vib} was arbitrarily set at 1.

2.2. Replica Exchange Method (REM). The replica exchange method is an extension of the Metropolis Monte Carlo, or molecular dynamics, methods. The underlying idea is to run different copies (replicas) of the system at different levels of a certain property (such as temperature). Although different properties have been considered in published work,[45,46] the property of change across different replicas (i.e., how replicas differ from one another) in the present context is temperature. To summarize the method, a MC or MD simulation is carried out on each selected conformation at its assigned temperature for a determined number of MC or MD steps, after which, the neighboring replicas undergo an exchange with the acceptance criterion described below (in eq 4). Let

$$\Delta \equiv [(\beta_m - \beta_n)\{E(\mathrm{Y}) - E(\mathrm{X})\}] \quad (3)$$

where β_m is the inverse temperature defined as $1/(k_{\mathrm{B}}T_m)$ and $E(\mathrm{X})$ is the energy of conformation X. If one adopts the Metropolis method, the replica-exchange transition probability can be expressed as

$$W(\mathrm{X},\beta_m|\mathrm{Y},\beta_n) = \begin{cases} 1 & \text{for } \Delta \leq 0 \\ \exp(-\Delta) & \text{for } \Delta > 0 \end{cases} \quad (4)$$

That is, if Δ is less than or equal to 0, the exchange is performed (since the probability is 1); otherwise, a random number between 0 and 1 is generated and compared to the factor $\exp(-\Delta)$. If the value of this factor is smaller, the exchange is performed; otherwise, the exchange is rejected.

To evaluate thermodynamic quantities at any temperature, it is essential to extract maximum information from all replicas. For this purpose, a multihistogram reweighting technique[63,64] can be used. For a replica exchange simulation with *M* replicas at *M* distinct temperatures, a set of *M* energy histograms $N_m(E)$ is obtained. The densities of states [$n(E)$] are then obtained self-consistently from the following WHAM[63,64] equations:

$$n(E) = \frac{\sum_{m=1}^{M} g_m^{-1} N_m(E)}{\sum_{m=1}^{M} g_m^{-1} n_m \exp(f_m - \beta_m E)} \quad (5)$$

and

$$\exp(-f_m) \equiv \sum_E n(E) \exp(-\beta_m E) \tag{6}$$

where $N_m(E)$ is the histogram at temperature T_m, $\beta_m = 1/(k_B T_m)$ is the inverse temperature, n_m is the total number of samples in the mth replica, $g_m = 1 + 2\tau_m$, and τ_m is the integrated autocorrelation time at temperature T_m. In biomolecular systems, g_m is approximately constant[64] and, therefore, can be canceled in eq 5. The WHAM eqs 5 and 6 are evaluated self-consistently, and the resulting densities of states are used to evaluate the expectation value of any observable A in eq 7:

$$\langle A \rangle_T = \frac{\sum_E A(E)\, n(E) \exp(-\beta E)}{\sum_E n(E) \exp(-\beta E)} \tag{7}$$

2.3. Multicanonical Algorithm (MUCA). A single canonical simulation (MC or MD) by definition samples a very restricted energy region. Furthermore, when sampling the conformations of the protein in low-energy regions, the multiple-minima problem is usually encountered and the simulation can be trapped in a particular local energy minimum, making it difficult to obtain a reliable estimate of the density of states of proteins. In determining the density of states of a large system by simulation procedures, a clear criterion is needed about the stage of simulations at which all of the conformational space of the protein has been sampled sufficiently. Traditional MC or MD procedures do not provide such a convergence criterion. For these reasons, a multicanonical algorithm[33,34] (also known as entropy sampling[35,36]) has been used for protein studies. In Section 2.4, we show why MUCA is combined with REM to produce REMUCA, whose efficiency is explored in the present work. For this purpose, we first outline MUCA. In the next paragraph, we present the background of entropy sampling and tie it together with the multicanonical algorithm notation.

In the present work, we use the term "conformation" to indicate a particular structure and the term "state" to denote all the conformations that either have a given energy or are within a small energy interval. The probability of occurrence of a conformation x with energy E, denoted as $P(\mathrm{x})$, and the probability of occurrence of a state with energy E, denoted as $P(E)$, are related to each other in a canonical ensemble by the following relations, with E being written for $E(\mathrm{x})$:

$$P(\mathrm{x}) \propto \exp(-\beta E) \tag{8}$$

$$P(E) \propto n(E) \exp(-\beta E) = \exp[S(E)/k_B - \beta E] \tag{9}$$

where k_B is the Boltzmann constant, $\beta = 1/k_B T$ with T being the temperature, $n(E)$ is the number of conformations with energy E (i.e., density of states), and $S(E) = k_B \ln[n(E)]$ is the entropy of the state with energy E.

The entropy sampling method is based on an artificial distribution of states, in which the probability of occurrence of a state with energy E is scaled by the exponential of the *negative* of the entropy of the state, $S(E)$. In entropy sampling, the probabilities of occurrence of a conformation x and a state with energy E, respectively, are defined as

$$P(\mathrm{x}) \propto \exp\{-S[E(\mathrm{x})]/k_B\} \tag{10}$$

$$P(E) \propto n(E) \exp[-S(E)/k_B] \tag{11}$$

where $n(E)$ and $S(E)$ have similar meanings as described above. Equations 10 and 11 can be related to eqs 8 and 9 by first setting $\beta = 0$ (i.e., temperature to infinity) in eqs 8 and 9 and then multiplying the resulting probabilities by the weight factor $\exp[-S(E)/k_B]$. The physical meaning of this modification is that the larger the conformational entropy of a state, the smaller is the weight given to the state. In this way, the probabilities of occurrence of all states with different energies are constant in the new distribution; that is, $P(E)$ of eq 11 is a constant, taken as 1.

To connect the entropy sampling formalism to the commonly used multicanonical algorithm, we can define a new variable, the multicanonical energy E_{mu}, in the following way

$$E_{mu}(E;T_0) = T_0 S(E) = k_B T_0 \ln[n(E)] \tag{12}$$

where T_0 is the reference temperature and $S(E)$ is the microcanonical entropy as above. The reference temperature is the temperature at which the MC or MD multicanonical simulation is carried out. It should be noted that the reference temperature theoretically plays no role in calculating thermodynamics, because the formula for obtaining thermodynamic quantities (eq 7) is independent of T_0; however, in practice, the value chosen for T_0 affects the sampling efficiency of numerical simulations. Equations 10 and 11 then become

$$P(\mathrm{x}) \propto \exp\{-E_{mu}[E(\mathrm{x});T_0]/T_0 k_B\} \tag{13}$$

and

$$P(E) \propto n(E) \exp[-E_{mu}(E;T_0)/T_0 k_B] \tag{14}$$

Consequently, the multicanonical Monte Carlo simulation is carried out with the following modified Metropolis acceptance criterion:

$$W(X|Y) = \begin{cases} 1 & \text{for } \Delta E_{mu} \le 0 \\ \exp(-\beta_0 \Delta E_{mu}) & \text{for } \Delta E_{mu} > 0 \end{cases} \tag{15}$$

where $\beta_0 = 1/k_B T_0$, T_0 being a reference temperature, and $\Delta E_{mu} \equiv E_{mu}[E(\mathrm{Y});T_0] - E_{mu}[E(\mathrm{X});T_0]$.

The multicanonical molecular dynamics simulation is carried out by integrating the following modified Newton equation;[65−67] see eq 21 of ref 65:

$$\dot{p}_k = -\frac{\partial E_{mu}(E;T_0)}{\partial q_k} = \frac{\partial E_{mu}(E;T_0)}{\partial E} f_k \tag{16}$$

where $\dot{p}_k$ is the momentum, q_k is the generalized coordinate of the kth atom, and f_k is the force on the kth atom. Specifically, the UNRES MD equation of motion (eq 32 of ref 16) is modified as

$$\ddot{q}(t) = -G^{-1} \frac{\partial E_{mu}(U;T_0)}{\partial U} \nabla_q U[q(t)] \tag{17}$$

where U [being $U(x)$] is the UNRES potential energy (U_{MD} of eq 2), $q(t)$ are the generalized coordinates at time t, and

G is the mass matrix (eq 26 of ref 16). In practice, one can use cubic splines to approximate $\partial E_{mu}(U;T_0)/\partial U$.

Because the density of states is usually not known a priori, the multicanonical weights are usually obtained by iterating short runs;[36,68−70] that is, E_{mu} is obtained such that eq 14 is constant for all energies E. For this purpose, one uses the single histogram reweighting technique to obtain a new estimate of the densities of states after each iteration:

$$n(E) = \frac{N_{mu}(E)}{\exp[-\beta_0 E_{mu}(E;T_0)]} \tag{18}$$

where N_{mu} is the histogram obtained from the multicanonical simulation (either MC or MD) and $\exp[-\beta_0 E_{mu}(E;T_0)] = 1/n(E)$ are the input multicanonical weights. The new estimates of the density of states are then used in eq 12 to obtain new values of E_{mu} and, hence, new input weights. This procedure is repeated until the histogram N_{mu} obtained from the multicanonical simulation is sufficiently flat (i.e., the probability of visiting any part of the energy space is constant). The resulting weights are then used for a long multicanonical simulation, from which thermodynamic quantities can be calculated.

To obtain expected averages from a multicanonical simulation, the single histogram reweighting technique (eq 18) is first used to obtain a new estimate of the densities of states. The new estimates of densities of states are then used in eq 7 to obtain the thermodynamic averages.

2.4. Replica Exchange Multicanonical Algorithm (REMUCA). MUCA without REM converges very slowly and consequently is inefficient.[71−73] Therefore, we have explored the use of REMUCA, which differs from MUCA in how the starting weights for the simulation are obtained. While MUCA requires short iterative multicanonical simulations, REMUCA obtains the starting weights from a short replica exchange simulation, by first obtaining the densities of states from REM, which are then used to estimate the multicanonical weights $\{\exp[-E_{mu}(E;T_0)/k_B T_0]\}$ with eq 12. In practice, the values for the multicanonical potential energy, $E_{mu}(E;T_0)$, obtained from replica exchange, are reliable only in the range of $\langle E\rangle_{T_{min}} \le E \le \langle E\rangle_{T_{max}}$, where T_{min} and T_{max} are the lowest and highest temperatures in REM, respectively, and $E_{min} = \langle E\rangle_{T_{min}}$ and $E_{max} = \langle E\rangle_{T_{max}}$ are the canonical expectation values at those temperatures; that is, we use multicanonical sampling only in the region between E_{min} and E_{max} and canonical sampling outside of this region. The reason the weights are reliable only between E_{min} and E_{max} is because T_{min} and T_{max} (which determine E_{min} and E_{max}) are chosen arbitrarily for the REM simulation, such that the region sampled by overlapping replicas between E_{min} and E_{max} contains both the native structure and the most probable non-native structures. Therefore, the best region sampled by REM is the one between E_{min} and E_{max}, which determines that the multicanonical input weights should be reliable only between E_{min} and E_{max}. In principle, any sampling can be used below E_{min} and above E_{max} as long as the simulation returns back to the multicanonical region which should contain both the native structure and the most probable non-native structures; in practice, this calculation has been carried out with canonical sampling.

The only reason to explore the canonical region is to force a random walk from the multicanonical region, which may have wandered out of the multicanonical region, to return to the multicanonical region. In essence, by sampling for thermodynamic data only in the multicanonical region, it is being assumed that the multicanonical region is large enough to encompass both the native structure and the more probable (i.e., lower-energy) parts of the ensemble of non-native structures. In addition, at the upper (E_{max}) and lower energy (E_{min}) boundaries between the multicanonical and canonical regions, the constant probability in the multicanonical region decreases in the canonical region.

The canonical sampling is carried out by extrapolating the multicanonical energies $[E_{mu}(E;T_0)]$ linearly.[71] It should be noted that only data from the multicanonical region (between E_{min} and E_{max}) are used for calculating thermodynamic properties. Hence, the energy space in REMUCA is divided into three regions as follows:

$$\epsilon^0_{mu}(E) \equiv \begin{cases} E_{mu}(E_{min};T_0) + \left.\dfrac{\partial E_{mu}(E;T_0)}{\partial E}\right|_{E_{min}} (E - E_{min}) & \text{for } E \le E_{min} \text{ (canonical)} \\ E_{mu}(E;T_0) & \text{for } E_{min} \le E \le E_{max} \text{ (multicanonical)} \\ E_{mu}(E_{max};T_0) + \left.\dfrac{\partial E_{mu}(E;T_0)}{\partial E}\right|_{E_{max}} (E - E_{max}) & \text{for } E \ge E_{max} \text{ (canonical)} \end{cases}$$

where $\epsilon^0_{mu}(E)$ is substituted for $E_{mu}(E;T_0)$ in eqs 15 (for MC) and 17 (for MD) and T_0 is the reference temperature for the Monte Carlo and molecular dynamics simulations (the temperature at which the MC or MD simulation is carried out). Again, the reference temperature bears no significance in the results of the thermodynamic quantities (because eq 7 is independent of T_0). The rest of the simulation for both MC and MD proceeds as in a traditional MUCA simulation (eq 15 for MC and eq 17 for MD) with ϵ^0_{mu} replacing E_{mu}.

2.5. Multicanonical Replica-Exchange Method (MUCAREM). We also explore the use of the REMUCAREM algorithm, whose core is the same as that of the MUCAREM algorithm. Therefore, we first present the theoretical background of MUCAREM and later extend the discussion to REMUCAREM. Just as REM consists of several replicas of canonical MC or MD simulations, MUCAREM consists of several replicas of multicanonical simulations. The difference between REM and MUCAREM is that the replicas in REM are associated with different temperatures whereas, in MUCAREM, the replicas are associated with different energy ranges over which multicanonical simulations are carried out. The advantage of the MUCAREM approach over the traditional REM is that the probability distributions of energies of different replicas are broader in MUCAREM than in REM; therefore, a smaller number of replicas is required to cover the entire energy range.

The starting weights are obtained by short iterations of MUCA simulations, as described earlier in Section 2.3. The following procedures are carried out in *each* cycle:

1. Select an energy range for each replica, for which the replica will carry out the MUCA simulation. This energy

range of a given replica should overlap the energy ranges of the neighboring replicas, and the combined energy range from all replicas should cover the whole energy space (i.e., the combined energy range should contain the native structure and the most probable non-native structures). Assign a different random protein conformation to each energy range.

2. A MUCA simulation with MC or MD is carried out on each selected conformation within its energy range for a determined number of MC or MD steps. The MC or MD simulations are carried out with eqs 15 or 17, respectively, where E_{mu} is replaced by ϵ_{mu}^{m} defined as follows:

$$\epsilon_{mu}{}^{m}(E) \equiv \begin{cases} E_{mu}(E_{\min}^{m};T_m) + \left.\dfrac{\partial E_{mu}(E;T_m)}{\partial E}\right|_{E_{\min}^{m}} (E - E_{\min}^{m}) & \text{for } E \le E_{\min}^{m} \text{ (canonical)} \\ E_{mu}(E;T_m) & \text{for } E_{\min}^{m} \le E \le E_{\max}^{m} \text{ (multicanonical)} \\ E_{mu}(E_{\max}^{m};T_m) + \left.\dfrac{\partial E_{mu}(E;T_m)}{\partial E}\right|_{E_{\max}^{m}} (E - E_{\max}^{m}) & \text{for } E \ge E_{\max}^{m} \text{ (canonical)} \end{cases}$$

where m is the replica index (m = min...max) and min and max are the lowest and highest temperature replicas. $E_{\min}^{m}$ is then the canonical expectation value of the energy of the mth replica at temperature $T_{\min}^{m}$ [$E_{\min}^{m} = \langle E \rangle_{T_{\min}^{m}}$], and similarly, $E_{\max}^{m}$ is the canonical expectation value of the energy of the mth replica at temperature $T_{\max}^{m}$ [$E_{\max}^{m} = \langle E \rangle_{T_{\max}^{m}}$] for the mth multicanonical replica. It should be noted that $T_{\min}^{m}$ and $T_{\max}^{m}$ are different for different replicas (for different m's) and, thus, determine a different multicanonical energy range $E_{\min}^{m}$ and $E_{\max}^{m}$ for different replicas. Therefore, the multicanonical simulation with each replica is carried out in a different energy range ($E_{\min}^{m}$ and $E_{\max}^{m}$).

3. After carrying out a selected number of MC or MD steps, stop the simulation of each replica and attempt an exchange of the whole conformations between neighboring replicas with the following transition probability:

$$W(\mathrm{Y}|\mathrm{X}) = \begin{cases} 1 & \text{for} \Delta \le 0 \\ \exp(-\Delta) & \text{for} \Delta > 0 \end{cases} \tag{19}$$

where $\Delta \equiv \beta_{m+1}\{\epsilon_{mu}^{m+1}[E(\mathrm{Y})] - \epsilon_{mu}^{m+1}[E(\mathrm{X})]\} - \beta_m\{\epsilon_{mu}^{m}[E(\mathrm{Y})] - \epsilon_{mu}^{m}[E(\mathrm{X})]\}$.

4. Continue the simulation with each newly formed conformation at each new energy range as in step 2.

5. Iterate points 3 and 4 until the system sufficiently covers the entire energy range.

As in REM, the densities of states are obtained from a self-consistent evaluation of the following modified WHAM equations:

$$n(E) = \frac{\sum_{m=1}^{M} g_m{}^{-1} N_m(E)}{\sum_{m=1}^{M} g_{\mathrm{m}}{}^{-1} n_m \exp[f_m - \beta_m \epsilon_{mu}^{m}(E)]} \tag{20}$$

and

$$\exp(-f_m) \equiv \sum_{E} n(E) \exp[-\beta_m \epsilon_{mu}^{m}(E)] \tag{21}$$

where $N_m(E)$ is the histogram at temperature T_m, $\beta_m = 1/(k_{\mathrm{B}}T_m)$ is the inverse temperature, n_m is the total number of samples in the mth replica, and g_m is defined as in Section 2.2. The resulting densities of states are then used to evaluate the expectation value of any observable in eq 7, with g_m canceling out, as in eq 5.

2.6. Replica Exchange Multicanonical with Replica-Exchange Method (REMUCAREM). MUCAREM without input weights from REM converges very slowly and, consequently, is inefficient.[71−73] Therefore, we have explored the use of REMUCAREM, which, as in REMUCA, obtains the starting weights from replica exchange simulations as opposed to iterative short MUCA simulations. Everything else proceeds in the same manner as in MUCAREM.

3. Implementation Details

All the simulations were carried out on one peptide (20 residues of alanine with free ends; ala_{20}) and two small proteins, namely, the B domain of staphylococal protein A (an α-protein; 46 residues; 1BDD)[74] and the *Escherichia coli* Mltd Lysm domain (an α+β protein; 48 residues; 1E0G).[75] The ala_{20} peptide was used to check whether the algorithms perform correctly, and the proteins were chosen so that basic α and α+β topologies were tested, and their size was reasonable with respect to the computational time. As in our previous work,[76] the length of protein 1BDD was shortened from the original 60 residues in the PDB to 46 residues. The set of UNRES energy parameters, designated as the 4P force field[76] and used in the present work, was derived by optimizing the parameters for four proteins simultaneously: 1E0L[77] (a β protein; 37 residues), 1E0G[75] (an α+β protein; 48 residues), 1IGD[78] (an α+β protein; 61 residues), and 1GAB[79] (an α protein; 53 residues).

The MC simulations with REM, REMUCA, and REMUCAREM were carried out as follows. All four UNRES angles in every residue of the protein were subjected to a perturbation. One MC sweep consisted of updating all of these angles for each residue in the sequence, with a Metropolis evaluation after each perturbation. The MD simulations with these same algorithms were carried out with the Berendsen thermostat,[80] using the velocity Verlet algorithm[81] with a variable time step to integrate the equations of motion. The variable time step was accomplished by scaling the time step δt by powers of 2.[16] The cutoff change of acceleration δa_{cut} for the scaling procedure was increased to $\delta a_{\mathrm{cut}} = 4$ Å/mtu,[16] to allow for the multiplication of the forces in the modified Newton equation (in eq 17, MUCA MD utilizes a factor that multiplies the forces, i.e., accelerations, which would cause the maximum change of acceleration $\delta a_{\max}$ to exceed the cutoff value δa_{cut}, and thus, the time step would be unnecessarily reduced). The time step was set at 4.89 fs to yield stable trajectories.[16] However, this is only a formal time step, and because of the reduction of the number of degrees of freedom in UNRES, the time step is several times larger compared with that of all-atom MD (see ref 16 for details). The coupling constant to the thermal bath was increased to

Table 1. Parameters Used in ala_{20} Simulations[a]

	MD			MC		
simulation	replicas	temp	steps	replicas	temp	sweeps
REM	16	400−2000	16,000,000	30	100−2000	2,000,000
REMUCA	1	100	10,000,000	1		1,000,000
REMUCAREM	2	100,101	20,000,000	2		1,000,000

[a] The replicas column shows the number of replicas used for each simulation. The temp shows the reference temperature (K) or range of temperatures for simulations (for REMUCA MC and REMUCAREM MC, the reference temperature cancels out in the equations; therefore, the corresponding fields are empty; this is because REMUCA and REMUCAREM depend only on the input weights which are independent of T, whereas in REM, the replicas differ from one another in temperature, and therefore, temperature does not cancel out). The step is the number of UNRES MD time steps, where the maximum time step was set to 4.9 fs in all MD simulations. A sweep is defined as perturbing all four angles at all the positions along the peptide sequence (for ala_{20}, one sweep is equal to 80 energy evaluations).

0.2445 ps to overcome the limitation of the Berendsen thermostat and produce a more Boltzmann-like distribution.[17] Replica exchange MD was carried out using multiplexing,[82] in which several replicas were simulated at each temperature. Since MC lacks the gradient and is consequently much less efficient at exploring the energy space than MD, the temperature range in the MC version of REM was lower than that of the REM MD simulations (so that the low-temperature replicas in REM MC would involve a sufficient number of moves to explore the low energy basins), and the number of replicas and the frequency of exchange in REM were much higher in MC. In all the simulations (both MC and MD), the system was equilibrated for 20% of the simulation length, and the last 80% of the simulation was used for the calculations. All Monte Carlo simulations were started from random conformations, and the starting point for all molecular dynamics simulations was an extended chain; because the system was equilibrated and because REM uses high-temperature replicas and both REMUCA and REMUCAREM perform a random walk in the energy space, the simulations were independent of the starting conditions.

4. Results and Discussion

4.1. Poly-L-alanine. First, to test the algorithms, a very simple poly-L-alanine system (20 residues) was chosen, and REM, REMUCA, and REMUCAREM simulations were carried out with both MC and MD. The parameters used in all simulations for ala_{20} are shown in Table 1. REM simulations were carried out first, from which the densities of states were obtained. It was found that the densities of states obtained from REM simulations were not precise enough for REMUCA, because REMUCA simulations did not perform a random walk (i.e., did not have flat energy histograms). Therefore, after the first iteration of REMUCA simulations, the densities of states were reweighted with eq 18, and with these weights, a second iteration of REMUCA simulations was carried out. The second set of weights used for REMUCA was also used for REMUCAREM simulations. The simulation weights for alanine are shown as a solid or dashed curve in Figure 1. The dashed line shows an example of the multicanonical energy function (eq 12), used in the modified Metropolis criterion in MC simulations (eq 15), while the solid line shows its derivative, a factor multiplying the force in the modified Newton equation (eq 17). The results are summarized in Figures 2 and 3.

Figure 2 consists of six plots. Three plots on the top correspond to MC simulations, whereas the three plots on the bottom correspond to MD simulations. The two plots in each column are for REM, REMUCA, and REMUCAREM simulations, respectively. Each plot depicts the logarithm of the probabilities ln[$P(E)$] as a function of energy (E) for the given simulation. By comparing the top row to the bottom row, it can be seen that MC simulations cover a smaller energy range than their MD counterparts. This is due to the fact that the MD energy function contains the extra vibration term (eq 2) adding to the energy range for MD simulations. It is evident from the plots that REMUCA MC and REMUCAREM MC are flatter {constant ln[$P(E)$]} than REMUCA MD and REMUCAREM MD. This discrepancy probably arises from the fact that the MD versions of multicanonical simulations utilize the derivative of the multicanonical energy function (eq 17), whereas the MC simulations use only the multicanonical energy function itself (eq 15; Figure 1). As mentioned in the Methods section, the derivatives are fitted using cubic splines, which can cause problems if the entropy function is not smooth (the derivative will be rough, which will cause numerical instabilities in the integration of eq 17).

By comparing the plots for REM MC and REMUCA MC,

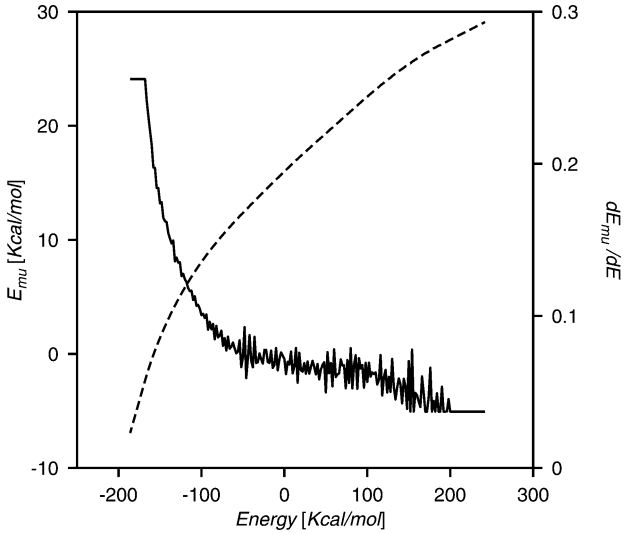


Figure 1. Parameters used for multicanonical simulations. The dashed line denotes the multicanonical energy function (eq 12), while the solid line denotes the derivative of this function fitted with cubic splines. The derivatives are used as a multiplicative factor [$\partial E_{mu}(U;T_0)/\partial U$] in the modified Newton equation (eq 17) in molecular dynamics. The flat regions of the derivative curve show where the multicanonical simulation changes to the canonical simulation.

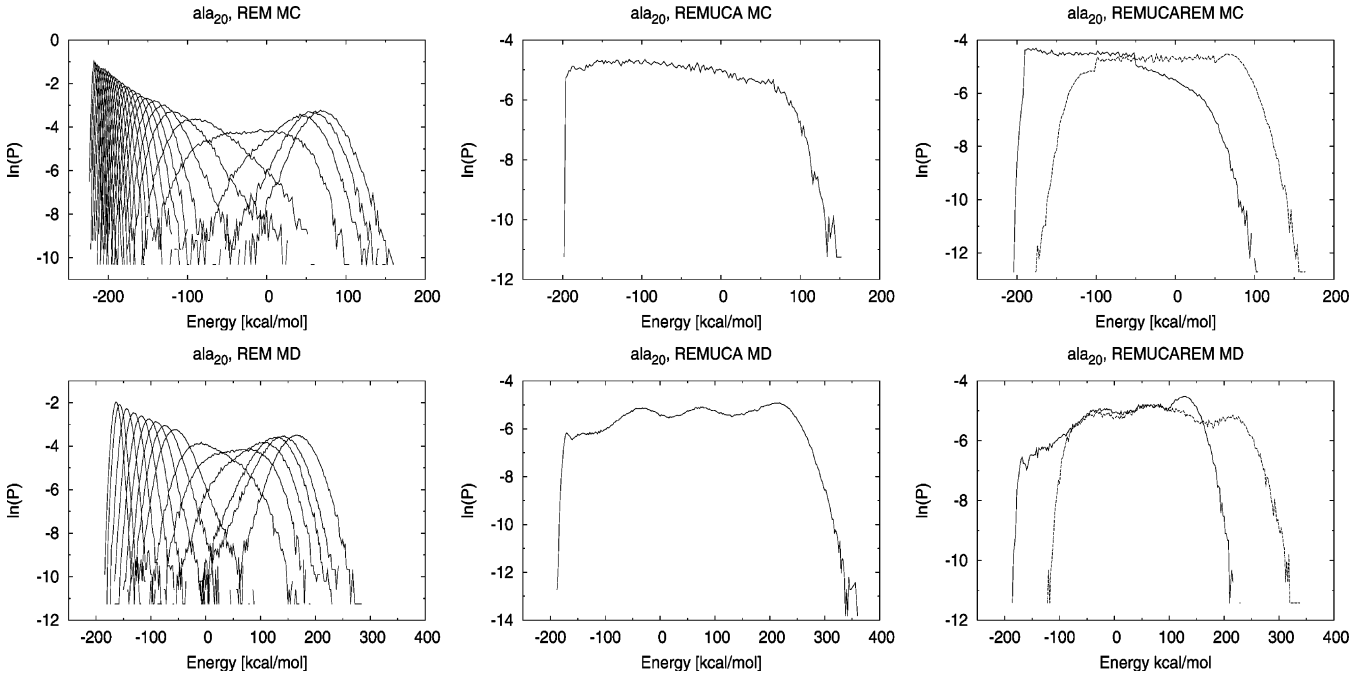


Figure 2. Histogram curves for simulations with alanine. The plots depict the logarithm of the probabilities ln[$P(E)$] as a function of energy (E). The top-row plots are from MC simulations (REM, REMUCA, and REMUCAREM, from left to right respectively). The bottom-row plots are from MD simulations. For REM and REMUCAREM (left and right columns, respectively), each curve corresponds to an individual replica at a different temperature (for REM) or different energy range (for REMUCAREM); see Table 1 for the number of such replicas.

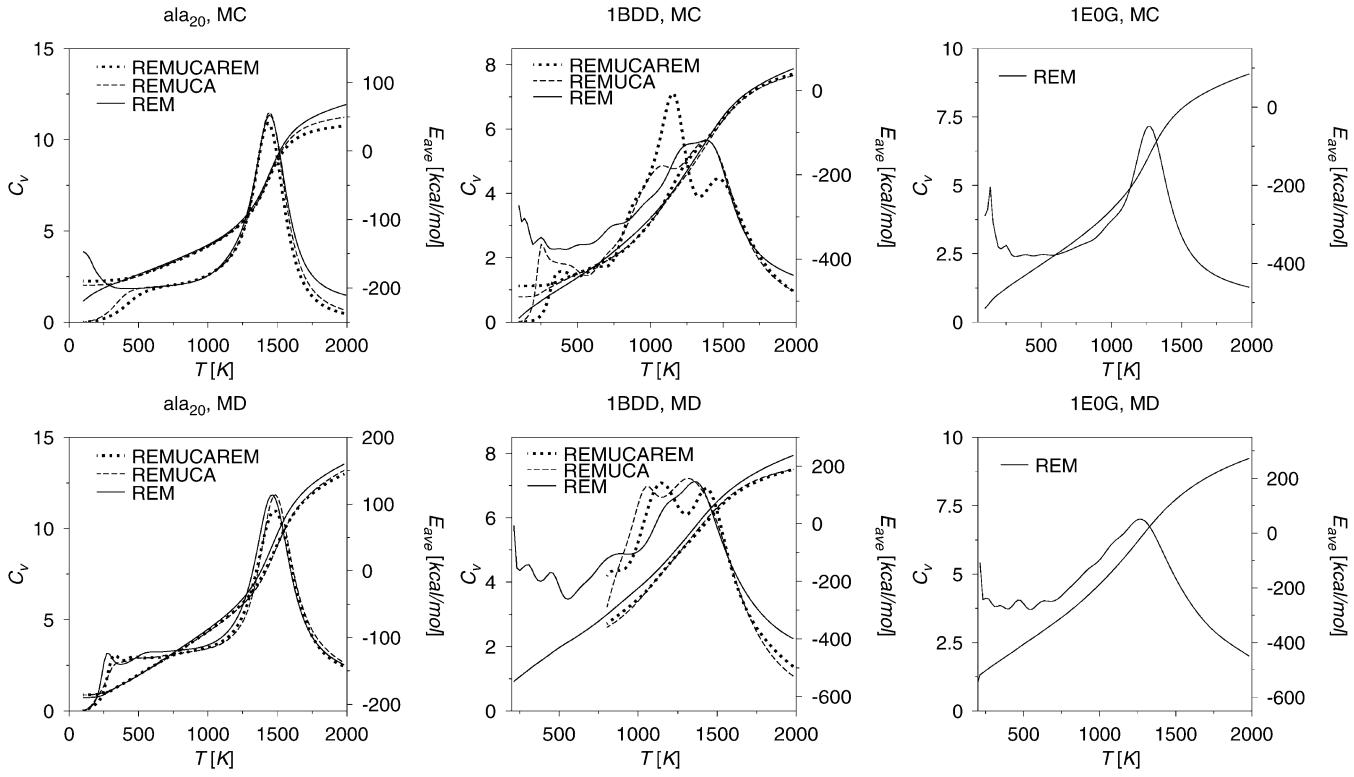


Figure 3. Heat capacity as well as average energy as a function of temperature for REM (solid line), REMUCA (dashed line), and REMUCAREM (dotted line) simulations with MC (top row) and MD (bottom row). The columns correspond to ala_{20}, 1BDD, and 1E0G, from left to right, respectively. The heat capacity curves are the ones with peaks at the folding temperatures. Good agreement for all three simulations for both MC and MD versions can be observed for ala_{20}; some overlap is observed for 1BDD, and only REM results (see text) are shown for 1E0G.

it can be seen that REMUCA MC does not cover the entire low-energy region but rather stops before −200 kcal/mol. This is because we shifted the low-energy boundary for multicanonical sampling up from the canonical average

Table 2. Parameters Used in 1BDD and 1E0G Simulations[a]

		MD			MC		
protein	simulation	replicas	temp	steps	replicas	temp	sweeps
1BDD	REM	30(×4)[b]	200−1800	240,000,000	50	50−1800	10,000,000
	REMUCA	1	50	30,000,000	1		1,000,000
	REMUCAREM	8	50−400	240,000,000	2		10,000,000
1E0G	REM	30(×4)[b]	200−1800	240,000,000	50	50−2000	10,000,000

[a] The replicas column shows the number of replicas used for each simulation. The temp shows the reference temperature (K) or range of temperatures for simulations (for REMUCA MC and REMUCAREM MC, the reference temperature cancels out in the equations; therefore, the corresponding fields are empty; this is because REMUCA and REMUCAREM depend only on the input weights which are independent of *T*, whereas in REM, the replicas differ from one another in temperature, and therefore, temperature does not cancel out). The step is the number of UNRES MD time steps, where the maximum time step was set to 4.9 fs in all MD simulations. A sweep is defined as 192 and 184 energy evaluations (four angles for each residue in the chain) for 1BDD and 1E0G, respectively. [b] Multiplexed replicas. 30(×4) means that four replicas for each temperature (with 30 temperatures) were simulated.

evaluated by the lowest temperature replica. The reason for doing this is that, when the boundary was lower in energy, the MC multicanonical simulations would walk in the entire energy range until they encountered the low-energy region, at which point the simulations would become trapped in deep local minima out of which they did not escape for the remainder of the simulation (data not shown). This issue was easily resolved for ala_{20} MC simulations by simply raising the low-energy boundary, but the issue reappears during both MC and MD simulations with 1BDD and 1E0G and is discussed further when describing the results for 1BDD and 1E0G.

Figure 3 also shows two rows of plots, one for MC and one for MD simulations. The first column corresponds to simulations with poly-L-alanine. Each plot consists of two graphs; one is the heat capacity, and the other is the average energy as a function of temperature. Each graph contains three curves, individually corresponding to REM, REMUCA, and REMUCAREM simulations. The average energy was calculated with eq 7, and the heat capacity was evaluated according to the following formula:

$$C_V = \beta^2 \frac{\langle E^2 \rangle_T - \langle E \rangle_T^2}{N} \tag{22}$$

For both MC and MD simulations with ala_{20}, all the curves overlap, suggesting that the simulations converged to the same distribution. The main peak of the specific heat curve indicates the temperature of the peptide collapse. For a simple system such as ala_{20}, the collapse occurs simultaneously with folding to the native α-helical state. This temperature appears to be 1400 K for MC and 1500 K for MD. It is important to note that the UNRES temperature has no relevance to the experimental temperature because UNRES is a coarse-grained potential in which the nonessential degrees of freedom have been averaged out, and energy parameter optimization was carried out with a hierarchical procedure[56] to provide the steepest decrease of energy with increasing native likeness[62] while ignoring the correspondence between the simulated and experimental thermodynamic characteristics of folding. Moreover, the decoy sets were generated using the CSA method which walks only in the space of local minima, thus violating the detailed balance condition. As mentioned further in the Conclusions section, we are currently revising our hierarchical force field optimization procedure,[62] to introduce entropy using methods applied in the present work and, consequently, to capture as much physics as possible.

4.2. 1BDD. We repeated the same procedure for 1BDD as for ala_{20}. The parameters used for the simulations with 1BDD are described in Table 2. Similarly, as for ala_{20}, the results for 1BDD are shown in Figure 4. First, since 1BDD has more degrees of freedom than ala_{20}, we used a larger number of replicas in both REM MC and REM MD algorithms, and in REM MD, we additionally multiplexed each replica to have more trajectories from which to sample. Although it might appear that, by using more replicas, REM would perform much better than both REMUCA and REMUCAREM, the advantage of REMUCAREM (as mentioned in Section 2.5) is that a smaller number of replicas is required to cover the entire energy range. To provide a fair comparison, we used the same number of steps for both REM and REMUCAREM (see Table 2); although many more steps were used in REMUCAREM than in REMUCA, the results with REMUCAREM are not substantially improved over those with REMUCA, as discussed later in this section. As for poly-L-alanine, the density of states from the replica exchange simulations was insufficient to carry out a random walk with REMUCA and REMUCAREM; therefore, the densities of states were reweighted. The multicanonical histogram curves in Figure 4 correspond to one iteration of reweighting. Additionally, we encountered a trapping problem in the low-energy region for both MC and MD simulations. As for ala_{20}, we increased the low multicanonical energy boundary to escape the trapping regions (Figure 4 shows that REMUCA and REMUCAREM MC and MD do not sample all the way to the lowest energy, i.e., not beyond −500 kcal/mol). To verify whether moving the multicanonical energy boundary is acceptable, we show the RMSD results in Figure 5. The left column shows the energy versus RMSD profile for replica exchange simulations. As can be seen from this column, both REM MC and REM MD cover a wide conformational space, which includes the native structure (centered ∼4.5 Å for REM MC and ∼4.0 Å for REM MD). The middle and the right columns show an RMSD trajectory for REMUCA and REMUCAREM simulations, respectively. It can be seen that the system folds and unfolds several times over the course of the run, that is, attains the low-RMSD region. Even though the multicanonical simulation should perform a random walk in the energy space, it is more important that the simulation fully samples

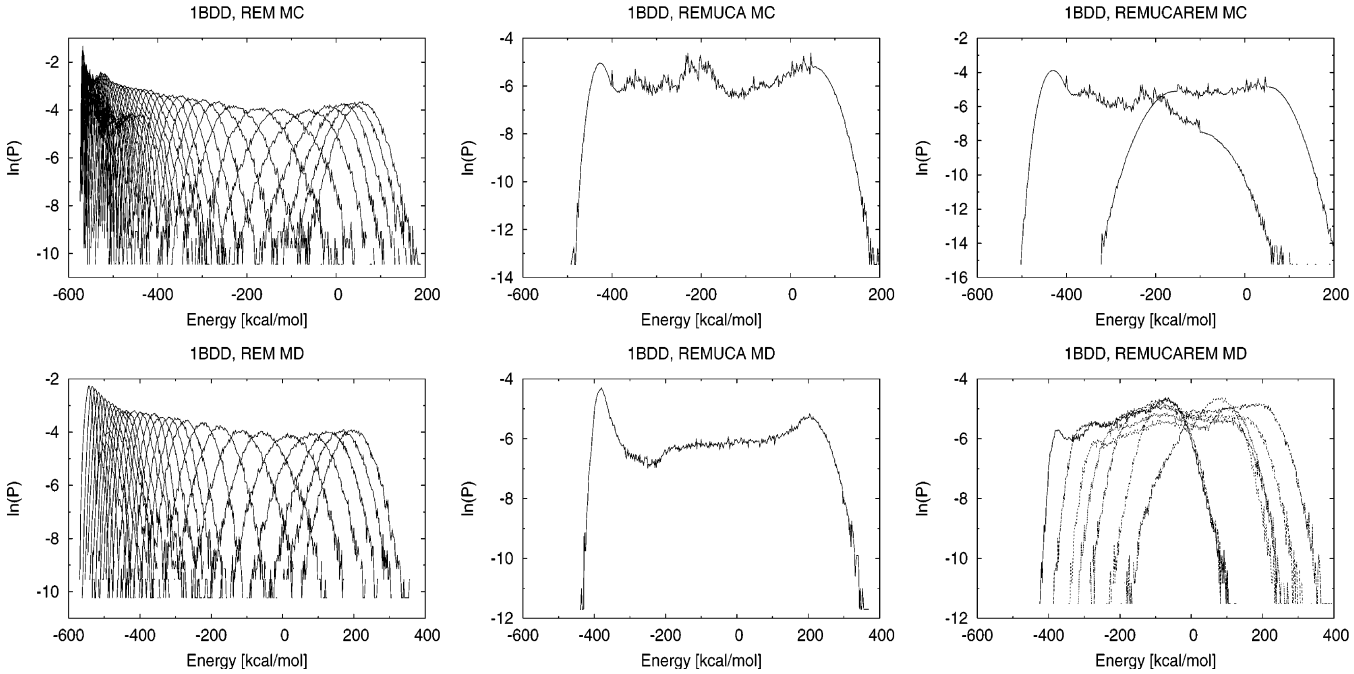


Figure 4. Histogram curves for simulations with 1BDD. The plots depict the logarithm of the probabilities as a function of energy. The top-row plots are from MC simulations (REM, REMUCA, REMUCAREM, from left to right respectively). The bottom-row plots are from MD simulations. For REM, and REMUCAREM (left, and right columns) each curve corresponds to an individual replica at a different temperature (for REM) or different energy range (for REMUCAREM); see Table 2 for the number of such replicas.

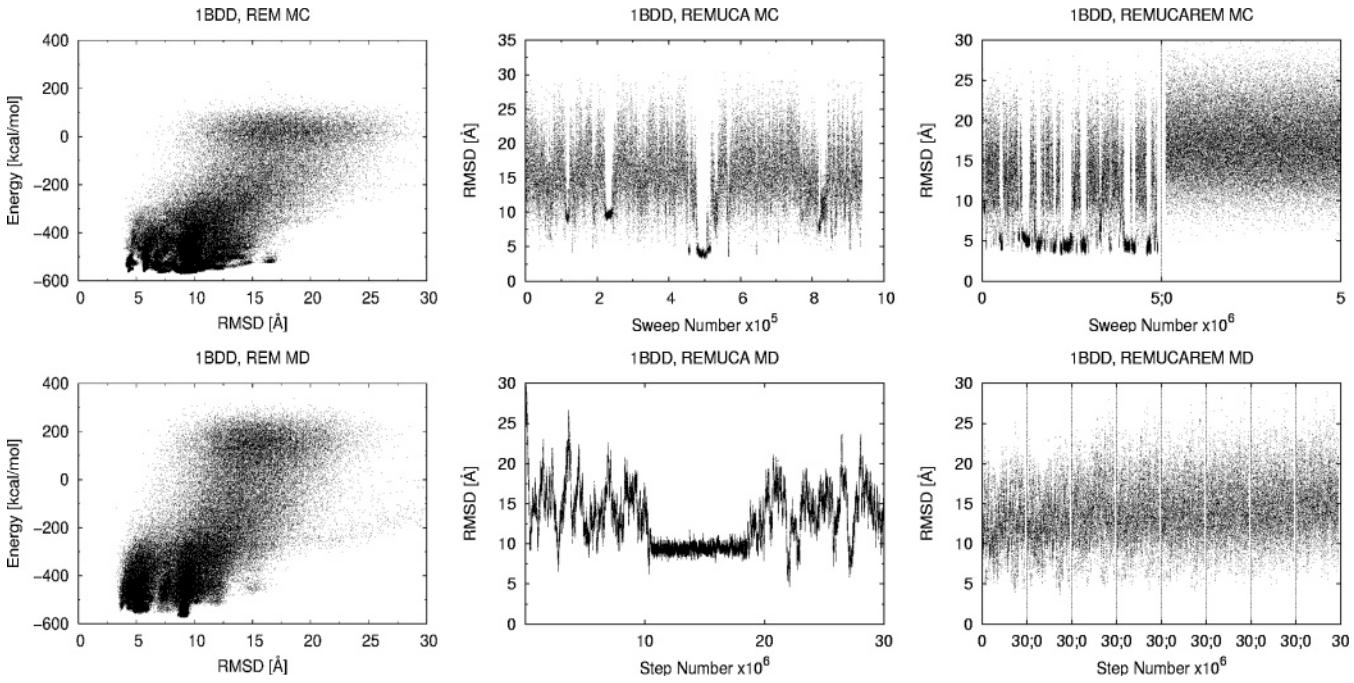


Figure 5. Simulation results for 1BDD. The top-row plots are from MC simulations (REM, REMUCA, and REMUCAREM, from left to right, respectively). The bottom-row plots are from MD simulations. The left column shows energy versus RMSD coverage of the energy space. The middle column shows the random walk of the REMUCA simulations, and the right column shows the random walk for all REMUCAREM replicas (one after another).

the conformational space, which can be observed in both the REMUCA and REMUCAREM RMSD trajectories.

The middle column of Figure 3 shows the calculated heat capacities and average energies for both MC and MD REM simulations with 1BDD. By contrast to the simulations with poly-L-alanine, 1BDD heat capacities have broad irregular peaks. The irregular peak is an overlap of two peaks, one corresponding to a collapse to a more compact state but without the final folding and one corresponding to a transition to the native state, as will be shown later. For 1BDD, REM, REMUCA, and REMUCAREM, peaks do not coincide as they do for poly-L-alanine. The fact that all simulations differ

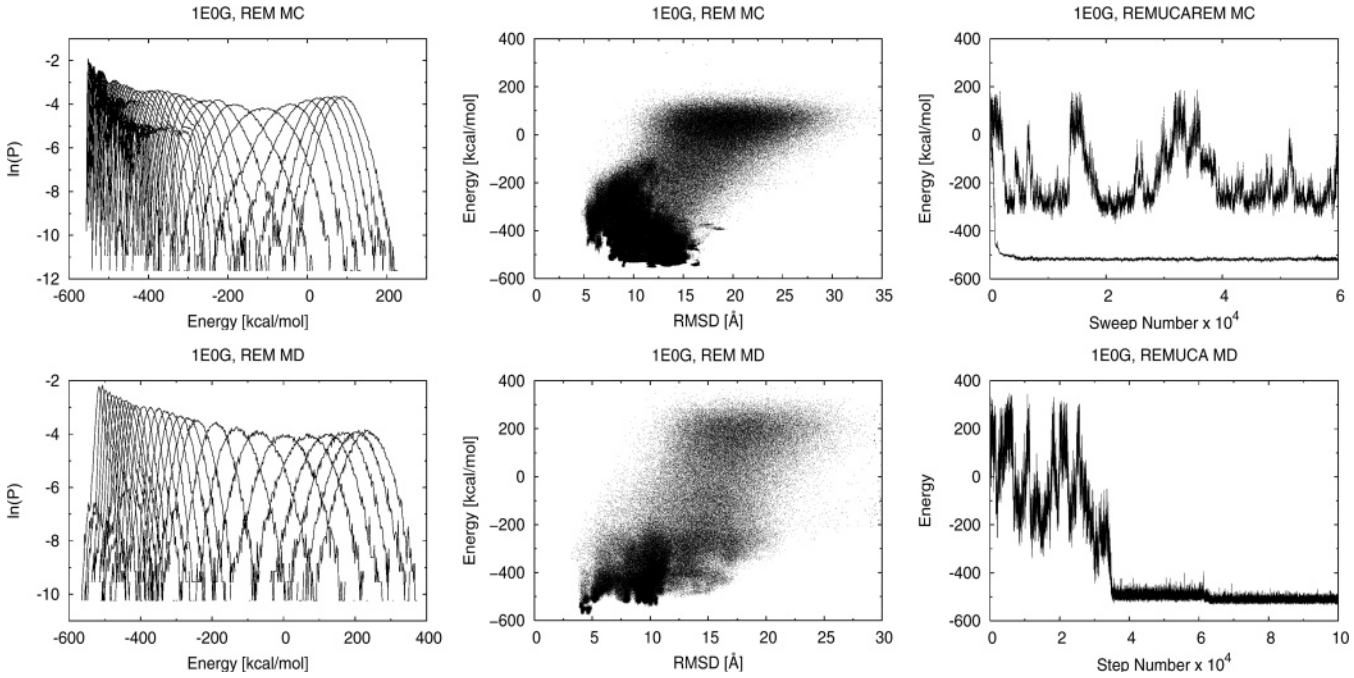


Figure 6. Simulation results for 1E0G. The top-row plots are from MC simulations, whereas the bottom-row plots are from MD simulations. The left-column shows the histogram curves for REM. Each curve corresponds to an individual replica at a different temperature. The middle column shows energy versus RMSD coverage of the energy space. The right-column shows energies at a series of steps of REMUCAREM for MC (top, with two replicas) and REMUCA for MD (bottom).

in the shape of their heat capacity curve suggests that all simulations have not converged to the same distribution. The reason the REMUCA and REMUCAREM curves do not cover the whole temperature range is that the multicanonical region was restricted to avoid trapping (i.e., the low multicanonical energy boundary was increased).

4.3. 1E0G. Finally, for 1E0G, replica exchange successfully sampled the energy space and produced reasonable statistics for thermodynamic quantities (Figure 6). The left column of Figure 6 shows the histograms for replica exchange simulations with both MC (top) and MD (bottom). The middle column depicts plots of energy as a function of the RMSD from the experimental structure, showing that the simulations cover an extended portion of the energy space. It can be seen that the REM MD simulation reaches the native state within an RMSD of around 4.5 Å and has low energy, whereas the REM MC simulation barely touches 5 Å RMSD, without reaching the low-energy region, which suggests incomplete N- and C-terminal β-strand contacts (correct β-strand packing provides a large contribution to decreasing the energy of the native structure and is necessary for the RMSD to be below 5 Å).

For multicanonical simulations (REMUCA and REMUCAREM), we were unable to obtain proper multicanonical weights, which would enable the system to carry out a random walk in the energy space. Even after several iterations of reweighting, the system would walk toward the low energy states, where it would stay for the remainder of the simulation. This behavior is shown in the right column of Figure 6, where a REMUCAREM simulation is shown for MC and a REMUCA simulation for MD. For REMUCAREM MC, it is evident that the lower-energy replica (replica 1) reaches low energies and remains trapped in a low-energy region, whereas the high-energy replica (replica 2) carries out a random walk. A similar behavior is observed for MD simulations (trapping of REMUCA MD is shown in Figure 6). This observation is similar to that from a study carried out by Bhattacharya and Sethna, who showed that, in the case of glassy systems, even multicanonical simulations have problems carrying out a random walk and instead become trapped in metastable states.[83] They implemented the entropy sampling version of the algorithm with Lennard-Jones glasses and observed that simulations that have dynamic updating of the microcanonical entropy function perform a random walk in the energy space, while the simulations with fixed weights (precomputed by iterative procedures) became trapped in metastable states. The dynamic updating of the weights (i.e., eq 5 of ref 36) is essentially a single histogram reweighting on the fly with the difference that not all regions might be visited, and typically the time between updates is much shorter. Dynamic updating ensures that the system does not remain in the same conformation for a long time. However, it also introduces discontinuities, and negative gradients into the E_{mu} function, which poses problems for the MD version of the REMUCA algorithm, with MD being more sensitive to the input weights because of its use of derivatives. The dynamic updating procedure pushes the system out of trapped states, but this violates the detailed balance condition and, thus, no longer guarantees convergence to the proper distribution or correct estimates of thermodynamic quantities. Because of the trapping problem, we did not calculate average energies and heat capacities from both REMUCA and REMUCAREM simulations for 1E0G (see Figure 3).

The third column of Figure 3 shows the calculated heat capacities and average energies for both MC and MD REM simulations with 1E0G. A sharp single peak for the heat capacity is observed for REM MC, whereas a broader peak is observed for REM MD simulations, and in both cases, it is centered at around 1270 K. As mentioned above (energy vs RMSD plot in Figure 6), the REM MC simulation does not quite sample the native region. This observation, and the fact that the heat capacity for REM MC has a sharper peak, suggests that REM MC predicts a collapse to a more compact state but without the final folding (i.e., there is no low-energy structure below a 5 Å RMSD as shown in the energy vs RMSD plot in Figure 6). On the other hand, the statistics from REM MD contain the native region (shown in the energy vs RMSD plot in Figure 6) and, thus, incorporates the contribution of the native region to the thermodynamic quantities. The collapse to a more compact structure and final folding do not seem to coincide (see the upcoming discussion about Figure 7), which broadens the heat capacity curve. For MC, the sharp peak is centered at 1270 K (Figure 3), which corresponds roughly to −130 kcal/mol of average energy. From the energy versus RMSD plot in Figure 6, it can be seen that the highest allowed energy for the collapsed structure (RMSD ∼ 5 Å) is also around −130 kcal/mol. Folding to the native state for MD occurs at lower energies, which broadens its heat capacity peak (see the discussion about Figure 7 in Section 4.4).

4.4. Free-Energy Diagrams. From our tests on ala_{20}, 1BDD, and 1E0G, we conclude that replica exchange molecular dynamics is the most efficient method for sampling and calculating thermodynamic quantities with a rugged energy landscape such as the 4P force field, applied to larger systems. Since the free energy is the most important quantity for the description of equilibrium properties of proteins, we used REM MD to calculate free-energy profiles for ala_{20}, 1BDD, and 1E0G. For this purpose, we used the densities of states obtained from the multihistogram analysis (eq 5). From the densities of states, we calculated the microcanonical entropy, $S(E_i) = k_B \ln[n(E_i)]$, for all conformations collected from the simulations and used it to compute the microcanonical free energies with the following expression: $F(E_i,T) = E_i - TS(E_i)$. To plot the restricted canonical free energy as a function of the RMSD (r) and radius of gyration (ρ), we calculated the restricted canonical free energy by evaluating the following expression for each grid point:

$$F(r,\rho,T) = -k_B T \ln \sum_{E_i \in N(r,\rho)} \exp\left(\frac{-F(E_i,T)}{k_B T}\right) \qquad (23)$$

where the index i enumerates conformations within the histogram bins, $N(r,\rho)$, for given ranges of the RMSD and radius of gyration.

Figure 7 shows the restricted canonical free-energy plots as a function of the RMSD and radius of gyration for various temperatures. Each column corresponds to simulations with ala_{20}, 1BDD, and 1E0G, from left to right, respectively. The temperatures are chosen so that the highest temperature is higher than that of the heat capacity peak (first row), within the peak (second row), below the peak (third row), and at zero K (fourth row) from top to bottom, respectively.

The highest-temperature free-energy plot for ala_{20} shows that, at this temperature, the peptide is preferentially completely unfolded, as indicated by the high RMSD (greater than 5 Å) and the high radius of gyration (greater than 9 Å), whereas at the heat capacity peak temperature (1460 K), the lowest free-energy region connects both the native and the non-native basins (RMSD between 2 and 5 Å). For 1000 K, the free-energy surface already appears very similar to the free-energy surface at 0 K, which represents the potential energy surface. The native state (RMSD lower than 2 Å) is the lowest free energy at this temperature, confirming our observation from the heat capacity curve. It should be noted that the range of energies observed in the potential energy plot is much larger than the range observed with nonzero temperatures, showing that the search for the native state is very much facilitated in the restricted canonical free-energy surface. In other words, the restricted canonical free-energy differences do not need to be very large in order to pass from the unfolded to the folded state, whereas large potential energy barriers must be crossed to pass from the unfolded to the folded state in the potential energy surface. For ala_{20}, we conclude that, even though the force field was optimized without any thermodynamics, we still observe a correct folding behavior.

For protein A (1BDD), the restricted canonical free-energy plots look similar to the plots for ala_{20}. At high temperatures, unpacked, open structures with a high RMSD and radius of gyration are observed. At 1000 K, the low free-energy region connects unfolded non-native states with compact states (both native and non-native). At a much lower temperature (600 K), the lowest free-energy regions belong to the native basin (centered around 5 Å RMSD) and to the mirror image (centered around 9 Å RMSD). It should be noted that, for ala_{20}, the native region had the lowest free energy at 1000 K whereas, for 1BDD, the temperature had to be lowered to 600 K for this to occur. Finally, the potential energy plot is again similar to the low-temperature free-energy plot but has a much larger energy range. It should be noted that, at 600 K, the free energy has well-defined regions of low free energy whereas, for the potential energy, the native state is more evenly connected with compact but non-native states, which has been observed previously in MD studies with protein A in our laboratory (all 10 simulations successfully folded protein A with the 4P force field at 800 K).[18]

For 1E0G, the high-temperature plot again shows a preference for unfolded structures. For 1000 K, the compact structures are not quite preferential in free energy. From previous MD work with 1E0G in our laboratory,[18] it was found that the successful folding trajectory starts with the formation of noninteracting helical structures, which then collapse to a native HTH motif (15 Å RMSD) and finally to one with a 3.9 Å RMSD from that of the experimental structure. The HTH motif structures appear to be preferable in terms of free energy at 1000 K, which is still within the broad peak of the heat capacity for 1E0G. For low temperatures, such as 600 K, the low free-energy region connects the HTH motif to compact nativelike structures without

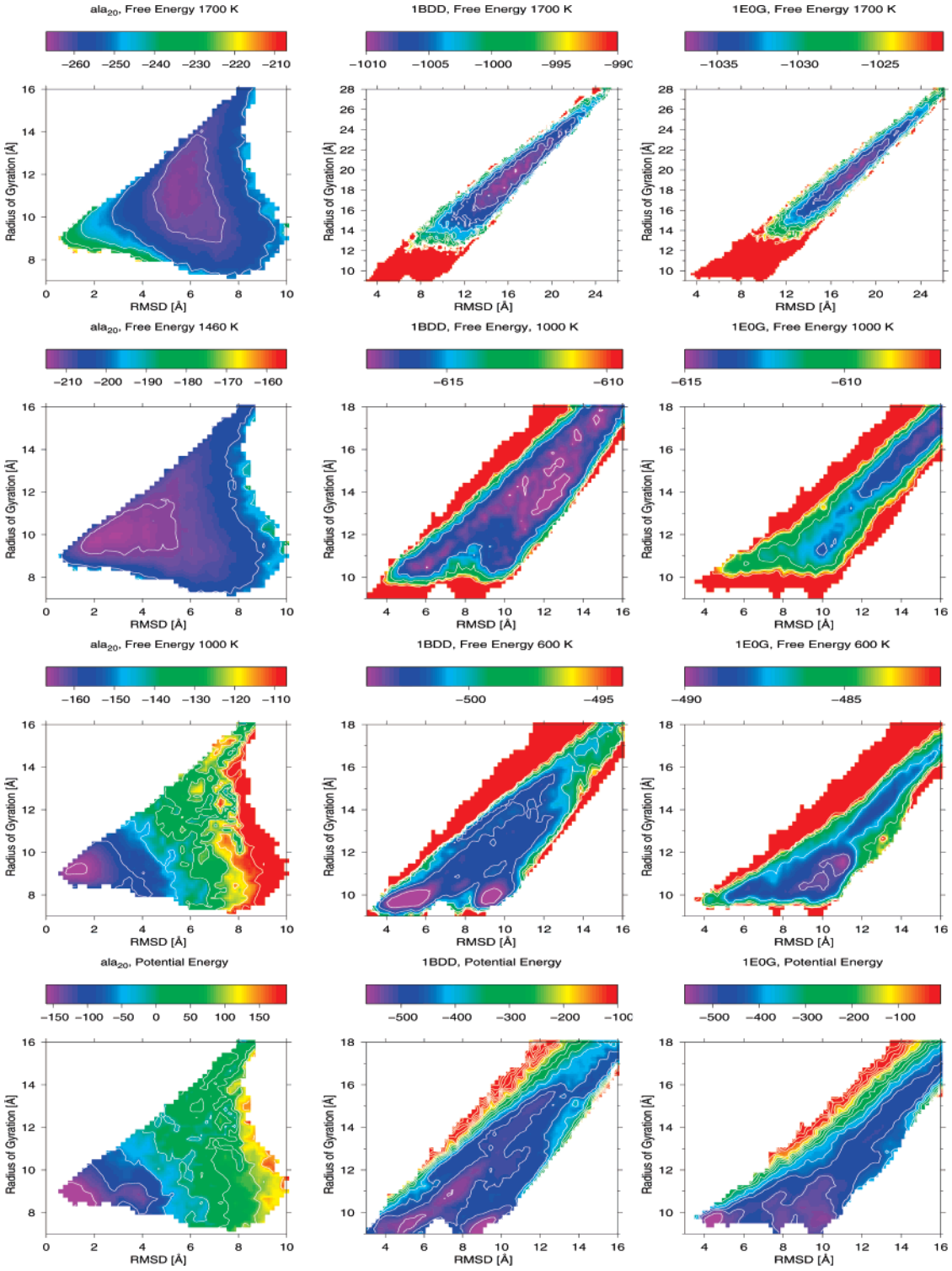


Figure 7. Restricted canonical free energy (in kcal/mol, indicated by the colored bars at the top of each graph) as a function of RMSD and radius of gyration for various temperatures. The free-energy surfaces were calculated from the REM MD simulations (see text). The columns correspond to simulations with ala_{20}, 1BDD, and 1E0G, from left to right, respectively. The temperatures are chosen so that the highest temperature is higher than that of the heat capacity peak (first row), within the peak (second row), below the peak (third row), and at 0 K (fourth row) for comparison.

β-strand contacts (around 6 Å RMSD). However, the fully formed native structure (centered at 4.5 Å RMSD) is at higher free energy, and it appears at the lowest free-energy region only at very low temperatures (where the free-energy plot is similar to the potential energy plot). Liwo et al. observed that only 6 out of 10 canonical MD simulations at 800 K yielded nativelike structures.[18] Our free-energy calculations show that the lowest free energy corresponds to non-native compact structures (i.e., with a low radius of gyration but a high RMSD); however, the native structures (with an RMSD less than 5 Å) have slightly higher free energy. Therefore, the non-native conformations are more

probable, but the native structures still have a finite probability to occur. Thus, our free-energy calculations agree with the results obtained by Liwo et al.

Since the temperature must be extremely low in order for the native state to be the global minimum of the free energy, the entropy contribution is much larger than that for the same temperature in protein A and ala_{20}. A larger contribution from entropy means more accessible conformations for a given temperature. Therefore, the multicanonical simulations have to sample a larger number of accessible conformations, which becomes difficult for 1E0G.

From Figure 7, it can be seen that, for a simple system such as ala_{20}, the collapse occurs simultaneously (at 1460 K) with folding to the native α-helical state (RMSD values and radii of gyration for low free-energy regions decrease simultaneously with temperature from 1700 to 1460 to 1000 K). For protein A and 1E0G, the low free-energy region at 1000 K extends all the way to the low radius of gyration and high RMSD values. For protein A, two low free-energy regions remain as the temperature is decreased to 600 K, one being the native and one being the mirror image. For 1E0G, the low free-energy region at 600 K with a low radius of gyration but a high RMSD appears first, and as the temperature is lowered (not shown here), the native region becomes the lowest free-energy basin. However, this occurs at very low temperatures, as described above. This explains why the heat capacity peaks for both protein A and 1E0G are broad and irregular. The two main events, collapse and folding to the native state, occur at different temperatures.

5. Conclusions

In the present work, we implemented REM, REMUCA, and REMUCAREM algorithms with the UNRES force field, utilizing Monte Carlo and molecular dynamics techniques. First, we tested all the algorithms on a simple poly-L-alanine system. For both the MC and MD algorithms, we obtained good agreement for heat capacity and average energy curves, which shows that all the simulations converged to the same distribution and that our implementation works as expected.

Next, we applied the simulations to two proteins, namely, to 1BDD and 1E0G. First, the 1BDD simulations performed reasonably well. The best performance was observed for the replica exchange algorithm in both the MC and MD simulations, since REM appeared to be much less sensitive to the input parameters (the only important parameter is the distribution of temperatures). To carry out a random walk, REMUCA and REMUCAREM depend on a proper estimation of the input weights and, as for ala_{20}, both REMUCA and REMUCAREM simulations had to be reweighted in order to obtain reasonably flat histograms. A trapping problem occurred at low energies, which was alleviated by raising the lower energy boundary for multicanonical simulations. However, by excluding a certain energy region from being sampled, the agreement among the heat capacity curves for all simulations was not so good.

Since 1E0G has a more complicated fold than 1BDD, multicanonical simulations broke down, and only replica exchange simulations were capable of exploring the energy region and computing the thermodynamic averages. This observation agrees with that from the study by Aleksenko et al.,[84] who concluded that the generalized ensemble approach is a useful study tool for proteins up to 30−40 residues with simple topology such as the α-helix. Furthermore, since the MD version of REMUCA and REMUCAREM use the derivative of the entropy function, MD multicanonical simulations are even more sensitive than their MC counterparts; therefore, they are more difficult to implement. Conversely, MD is much more capable of exploring the energy landscape than MC; hence, MD simulations are much more useful for larger systems.

Finally, we analyzed data from our REM MD simulations for all three test systems and calculated free-energy maps as a function of the RMSD and radius of gyration. The free-energy calculations show the correct folding behavior for poly-L-alanine and protein A, while for 1E0G, the native structure had the lowest free energy only at very low temperatures; hence, the entropy contribution is much larger than that for the same temperature in protein A and ala_{20}. The larger contribution from entropy means more accessible conformations for a given temperature. For the same temperature, ala_{20} has the smallest entropy contribution, followed by protein A, and then by 1E0G.

Although both REMUCA and REMUCAREM seem to have potential as sampling methods applied to smaller systems, replica exchange utilizing MD, coupled with multiplexing, appears to offer more insight into the behavior of protein folding for more complicated systems with a rough energy landscape. Moreover, since replica exchange is easy to implement and has few parameters to adjust, it is very suitable for implementation in the future revision of our hierarchical optimization procedure,[62] which is currently under development in our laboratory. The new optimization procedure is based on a hierarchical design of the potential-energy landscape such that the energy decrease follows the increase of native likeness[56] and utilizes MD as a sampling method to capture as much physics as possible. Preliminary tests (unpublished data) show that replica exchange together with umbrella sampling[85] (introduced when the native region is not sufficiently covered with the initial parameter set) covers a broader region of conformational space and, thus, produces better statistics for hierarchical optimization. Consequently, this will allow us to produce a coarse-grained force field suitable for molecular dynamics simulations, which will be capable of a more accurate evaluation of thermodynamic quantities.

Acknowledgment. We thank A. Liwo for valuable comments. This work was supported by grants from the National Institutes of Health (GM-14312), the National Science Foundation (MCB00-03722), the NIH Fogarty International Center (TW7913), and Grant 3 T09A 032 26 from the Polish Ministry of Scientific Research and Information Technology. This research was carried out by using the resources of (a) our 674-processor Beowulf cluster at the Baker Laboratory of Chemistry and Chemical Biology, Cornell University; (b) the National Science Foundation Terascale Computing System at the Pittsburgh Supercomputer Center; and (c) the National Center for Supercomputing

Applications System at the University of Illinois at Urbana−Champaign.

References

(1) Metropolis, N.; Ulam, S. *J. Am. Stat. Assoc.* **1949**, *44*, 335−341.

(2) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. *J. Chem. Phys.* **1953**, *21*, 1087−1092.

(3) Li, Z.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 6611−6615.

(4) Li, Z.; Scheraga, H. A. *THEOCHEM* **1988**, *179*, 333−352.

(5) Ripoll, D. R.; Scheraga, H. A. *Biopolymers* **1988**, *27*, 1283−1303.

(6) Ripoll, D. R.; Scheraga, H. A. *J. Protein Chem.* **1989**, *8*, 263−287.

(7) Pillardy, J.; Czaplewski, C.; Wedemeyer, W. J.; Scheraga, H. A. *Helv. Chim. Acta* **2000**, *83*, 2214−2230.

(8) Nanias, M.; Chinchio, M.; Oldziej, S.; Czaplewski, C.; Scheraga, H. *J. Comput. Chem.* **2005**, *26*, 1472−1486.

(9) Piela, L.; Kostrowicki, J.; Scheraga, H. A. *J. Phys. Chem.* **1989**, *93*, 3339−3346.

(10) Pillardy, J.; Olszewski, K. A.; Piela, L. *J. Phys. Chem.* **1992**, *96*, 4337−4341.

(11) Pillardy, J.; Liwo, A.; Groth, M.; Scheraga, H. A. *J. Phys. Chem. B* **1999**, *103*, 7353−7366.

(12) Pillardy, J.; Liwo, A.; Scheraga, H. A. *J. Phys. Chem. A* **1999**, *103*, 9370−9377.

(13) Lee, J.; Scheraga, H. A.; Rackovsky, S. *J. Comput. Chem.* **1997**, *18*, 1222−1232.

(14) Lee, J.; Scheraga, H. A. *Int. J. Quantum Chem.* **1999**, *75*, 255−265.

(15) Czaplewski, C.; Liwo, A.; Pillardy, J.; Oldziej, S.; Scheraga, H. A. *Polymer* **2004**, *45*, 677−686.

(16) Khalili, M.; Liwo, A.; Rakowski, F.; Grochowski, P.; Scheraga, H. A. *J. Phys. Chem. B* **2005**, *109*, 13785−13797.

(17) Khalili, M.; Liwo, A.; Jagielska, A.; Scheraga, H. A. *J. Phys. Chem. B* **2005**, *109*, 13798−13810.

(18) Liwo, A.; Khalili, M.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 2362−2367.

(19) Khalili, M.; Liwo, A.; Scheraga, H. A. *J. Mol. Biol.* **2006**, *355*, 536−547.

(20) Hukushima, K.; Nemoto, K. *J. Phys. Soc. Jpn.* **1996**, *65*, 1604−1608.

(21) Hansmann, U. H. E. *Chem. Phys. Lett.* **1997**, *281*, 140−150.

(22) Swendsen, R. H.; Wang, J. S. *Phys. Rev. Lett.* **1986**, *57*, 2607−2609.

(23) Geyer, C. *Computing Science and Statistics Proceedings of the 23rd Symposium on the Interface*; American Statistical Association: New York, 1991.

(24) Gront, D.; Kolinski, A.; Skolnick, J. *J. Chem. Phys.* **2001**, *115*, 1569−1574.

(25) Kolinski, A.; Gront, D.; Pokarowski, P.; Skolnick, J. *Biopolymers* **2003**, *69*, 399−405.

(26) Fenwick, M. K.; Escobedo, F. A. *J. Chem. Phys.* **2003**, *119*, 11998−12010.

(27) Romiszowski, P.; Sikorski, A. *Physica A* **2004**, *336*, 187−195.

(28) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141−151.

(29) Zhou, R.; Berne, B.; Germain, R. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 14931−14936.

(30) Sanbonmatsu, K.; Garcia, A. *Proteins* **2002**, *46*, 225−234.

(31) Garcia, A.; Onuchic, J. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13898−13903.

(32) Lin, C.-Y.; Hu, C.-K.; Hansmann, U. H. E. *Proteins: Struct., Funct., Genet.* **2003**, *52*, 436−445.

(33) Berg, B. A.; Neuhaus, T. *Phys. Lett. B* **1991**, *267*, 249−253.

(34) Berg, B. A.; Neuhaus, T. *Phys. Rev. Lett.* **1992**, *68*, 9−12.

(35) Lee, J. *Phys. Rev. Lett.* **1993**, *71*, 211−214.

(36) Hao, M.; Scheraga, H. A. *J. Phys. Chem.* **1994**, *98*, 4940−4948.

(37) Lyubartsev, A. P.; Martsinovski, A. A.; Shevkunov, S. V.; Vorontsov-Velyaminov, P. N. *J. Chem. Phys.* **1992**, *96*, 1776−1783.

(38) Marinari, E.; Parisi, G. *Europhys. Lett.* **1992**, *19*, 451−458.

(39) Gront, D.; Kolinski, A.; Skolnick, J. *J. Chem. Phys.* **2000**, *113*, 5065−5071.

(40) Hansmann, U. *Phys. Rev. E* **1997**, *56*, 6200−6203.

(41) Mitsutake, A.; Sugita, Y.; Okamoto, Y. *J. Chem. Phys.* **2003**, *118*, 6664−6675.

(42) Mitsutake, A.; Okamoto, Y. *J. Chem. Phys.* **2004**, *121*, 2491−2504.

(43) Liu, P.; Kim, B.; Friesner, R. A.; Berne, B. J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13749−13754.

(44) Kwak, W.; Hansmann, U. *Phys. Rev. Lett.* **2005**, *95*, 138102.

(45) Fukunishi, H.; Watanabe, O.; Takada, S. *J. Chem. Phys.* **2002**, *116*, 9058−9067.

(46) Jang, S.; Shin, S.; Pak, Y. *Phys. Rev. Lett.* **2002**, *91*, 058305.

(47) Oldziej, S.; Czaplewski, C.; Liwo, A.; Chinchio, M.; Nanias, M.; Vila, J.; Khalili, M.; Arnautova, Y.; Jagielska, A.; Schafroth, H. D.; Kazmierkiewicz, R.; Ripoll, D.; Pillardy, J.; Saunders, J.; Kang, Y.; Gibson, K.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 7547−7552.

(48) Liwo, A.; Pincus, M. R.; Wawak, R. J.; Rackovsky, S.; Scheraga, H. A. *Protein Sci.* **1993**, *2*, 1697−1714.

(49) Liwo, A.; Pincus, M. R.; Wawak, R. J.; Rackovsky, S.; Scheraga, H. A. *Protein Sci.* **1993**, *2*, 1715−1731.

(50) Liwo, A.; Ołdziej, S.; Pincus, M. R.; Wawak, R. J.; Rackovsky, S.; Scheraga, H. A. *J. Comput. Chem.* **1997**, *18*, 849−873.

(51) Liwo, A.; Pincus, M. R.; Wawak, R. J.; Rackovsky, S.; Ołdziej, S.; Scheraga, H. A. *J. Comput. Chem.* **1997**, *18*, 874−887.

(52) Liwo, A.; Kaźmierkiewicz, R.; Czaplewski, C.; Groth, M.; Ołdziej, S.; Wawak, R. J.; Rackovsky, S.; Pincus, M. R.; Scheraga, H. A. *J. Comput. Chem.* **1998**, *19*, 259−276.

(53) Liwo, A.; Czaplewski, C.; Pillardy, J.; Scheraga, H. A. *J. Chem. Phys.* **2001**, *115*, 2323−2347.

(54) Lee, J.; Ripoll, D. R.; Czaplewski, C.; Pillardy, J.; Wedemeyer, W. J.; Scheraga, H. A. *J. Phys. Chem. B* **2001**, *105*, 7291−7298.

(55) Pillardy, J.; Czaplewski, C.; Liwo, A.; Wedemeyer, W. J.; Lee, J.; Ripoll, D. R.; Arłukowicz, P.; Ołdziej, S.; Arnautova, Y. A.; Scheraga, H. A. *J. Phys. Chem. B* **2001**, *105*, 7299−7311.

(56) Liwo, A.; Arłukowicz, P.; Czaplewski, C.; Ołdziej, S.; Pillardy, J.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 1937−1942.

(57) Ołdziej, S.; Kozłowska, U.; Liwo, A.; Scheraga, H. A. *J. Phys. Chem. A* **2003**, *107*, 8035−8046.

(58) Liwo, A.; Ołdziej, S.; Czaplewski, C.; Kozłowska, U.; Scheraga, H. A. *J. Phys. Chem. B* **2004**, *108*, 9421−9438.

(59) Nishikawa, K.; Momany, F. A.; Scheraga, H. A. *Macromolecules* **1974**, *7*, 797−806.

(60) Kubo, R. *J. Phys. Soc. Jpn.* **1962**, *17*, 1100−1120.

(61) Kolinski, A.; Skolnick, J. *J. Chem. Phys.* **1992**, *97*, 9412−9426.

(62) Ołdziej, S.; Liwo, A.; Czaplewski, C.; Pillardy, J.; Scheraga, H. A. *J. Phys. Chem. B* **2004**, *108*, 16934−16949.

(63) Ferrenberg, A.; Swendsen, R. *Phys. Rev. Lett.* **1989**, *63*, 1195−1198.

(64) Kumar, S.; Bouzida, D.; Swendsen, R.; Kollman, P.; Rosenberg, J. *J. Comput. Chem.* **1992**, *13*, 1011−1021.

(65) Hansmann, U.; Okamoto, Y.; Eisenmenger, F. *Chem. Phys. Lett.* **1996**, *259*, 321−330.

(66) Nakajima, N.; Nakamura, H.; Kidera, A. *J. Phys. Chem. B* **1997**, *101*, 817−824.

(67) Bartels, C.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 865−880.

(68) Berg, B. *Int. J. Mod. Phys. C* **1992**, *3*, 1083−1098.

(69) Hansmann, U.; Okamoto, Y. *J. Phys. Soc. Jpn.* **1994**, *63*, 3945−3949.

(70) Hansmann, U.; Okamoto, Y. *Physica A* **1994**, *212*, 415−437.

(71) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **2000**, *329*, 261−270.

(72) Mitsutake, A.; Sugita, Y.; Okamoto, Y. *J. Chem. Phys.* **2003**, *118*, 6664−6675.

(73) Mitsutake, A.; Sugita, Y.; Okamoto, Y. *J. Chem. Phys.* **2003**, *118*, 6676−6688.

(74) Gouda, H.; Torigoe, H.; Saito, A.; Sato, M.; Arata, Y.; Shimada, I. *Biochemistry* **1992**, *31*, 9665−9672.

(75) Bateman, A.; Bycroft, M. *J. Mol. Biol.* **2000**, *299*, 1113−1119.

(76) Ołdziej, S.; Łagiewka, J.; Liwo, A.; Czaplewski, C.; Chinchio, M.; Nanias, M.; Scheraga, H. A. *J. Phys. Chem. B* **2004**, *108*, 16950−16959.

(77) Macias, M. J.; Gervais, V.; Civera, C.; Oschkinat, H. *Nat. Struct. Biol.* **2000**, *7*, 375−379.

(78) Derrick, J. P.; Wigley, D. B. *J. Mol. Biol.* **1994**, *243*, 906−918.

(79) Johansson, M. U.; de Chateau, M.; Wikstrom, M.; Forsen, S.; Drankenberg, T.; Bjorck, L. *J. Mol. Biol.* **1997**, *266*, 859−865.

(80) Berendsen, H.; Postma, J.; van Gunsteren, W.; DiNola, A.; Haak, J. *J. Chem. Phys.* **1984**, *81*, 3684−3690.

(81) Swope, W.; Andersen, H.; Berens, P.; Wilson, K. *J. Chem. Phys.* **1982**, *76*, 637−649.

(82) Rhee, Y.; Pande, V. *Biophys. J.* **2003**, *84*, 775−786.

(83) Bhattacharya, K.; Sethna, J. *Phys. Rev. E* **1998**, *57*, 2553−2562.

(84) Aleksenko, V.; Kwak, W.; Hansmann, U. *Physica A* **2005**, *350*, 28−37.

(85) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*; Academic Press: San Diego, CA, 2002.

CT050253O

JCTC Journal of Chemical Theory and Computation

YUP: A Molecular Simulation Program for Coarse-Grained and Multiscaled Models

Robert K. Z. Tan, Anton S. Petrov, and Stephen C. Harvey*

School of Biology, Georgia Institute of Technology, 310 Ferst Drive, Atlanta, Georgia 30332

Received December 16, 2005

Abstract: Coarse-grained models can be very different from all-atom models and are highly varied. Each class of model is assembled very differently, and some models need customized versions of the standard molecular mechanics methods. The most flexible way to meet these diverse needs is to provide access to internal data structures and a programming language to manipulate these structures. We have created YUP, a general-purpose program for coarse-grained and multiscaled models. YUP extends the Python programming language by adding new data types. We have then used the extended language to implement three classes of coarse-grained models. The coarse-grained RNA model type is an unusual nonlinear polymer, and the assembly was easily handled with a simple program. The molecular dynamics algorithm had to be extended for a coarse-grained DNA model so that it could detect a failure that is invisible to a standard implementation. A third model type took advantage of access to the force field to simulate the packing of DNA in viral capsids. We find that objects are easy to modify, extend, and redeploy. Thus, new classes of coarse-grained models can be implemented easily.

1. Introduction

All-atom molecular dynamics (MD) simulations have progressed from the treatment of small proteins in vacuo for a few picoseconds[1] to simulations in solution covering time scales as long as 1 μs.[2] One major challenge for MD simulations is the treatment of very large systems. Perhaps the largest system treated in all-atom detail is the ribosome,[3] for which about 10^6 CPU h were required for MD simulations totaling about 20 ns.

One way to reduce the computational burden for examining very large systems is to use reduced (coarse-grained) representations, in which pseudoatoms are used to represent pieces of the structure. These pseudoatoms may represent only a few atoms, e.g. methyl groups, or they can represent very large groups, depending on the resolution of the model. At very low resolution, a single pseudoatom might represent a protein, a nucleosome, or even larger structures.

The parametrization of all-atom models is based on the chemical properties of the atoms, and the transferability of parameters to a broad collection of molecules is a major goal. In contrast, coarse-grained models are generally developed for specific problems, so their parametrization tends to be empirical. The advent of multiscale models that contains different levels of resolution in different regions renders parametrization even more idiosyncratic.

We developed the Yammp molecular mechanics package with the specific purpose of facilitating coarse-grained modeling.[4] Although it can be used for all-atom modeling,[5] its principal applications have been aimed at tackling very large systems with reduced representations. Among these are supercoiled DNA,[6−8] the ribosome,[9−11] HIV,[12] and other viruses.[13,14] Yammp contains a number of unique force field terms, developed for different applications.

As coarse-grained and multiscale models become more popular, it would be desirable to have a single, integrated package that facilitates the development and application of such models. The revision of Yammp to Yammp 2 provides such a package. It extends the Python programming language, hence it is also known as YUP (Yammp Under Python). This paper describes the new package.

* Corresponding author e-mail: steve.harvey@biology.gatech.edu.

10.1021/ct050323r CCC: $33.50
Published on Web 03/18/2006

High-quality molecular simulation programs for all-atom models are widely available, and two such programs that influenced this work are AMBER[15,16] and CHARMM.[17,18]

A molecular simulation program has to provide at least two features. The first is a way for users to construct their model. Essentially, this is the process of converting the user's specification of the model (e.g. a base or amino acid sequence) to the terms of an energy function. This function expresses the interactions among the atoms: which atoms are involved in each interaction, the functional form of the interaction, and the specific parameters that are to be used to calculate the interaction. We call this process force field assembly, and ideally the procedure should be extensible to new classes of models. The second feature is the raison d'être for molecular simulation programs: routines for molecular mechanics calculations, and users should be able to extend existing methods and implement new ones.

Conventional, i.e., all-atom, modeling programs can be awkward tools for coarse-grained models. For example, most conventional programs include a mechanism to build only linear polymers with provisions for simple cross-links. This is sufficient to build proteins, nucleic acids, and most other molecules of computational interest. On the other hand, the best way to build a coarse-grained model may not necessarily be as a linear polymer, even if that is what the original is. In fact, a coarse-grained program will have to be much more versatile than a conventional program because the definition of a coarse-grained model is open-ended.

Version 2 of Yammp, or YUP (http://rumour.biology.gatech.edu/YammpWeb/), is designed for the easy implementation of coarse-grained models. This is to be achieved through two features: [1] most of the data structures of the program are user accessible and [2] a high-level language is available to manipulate the data. YUP is a component of the NIH Research Resource Center for the Development of Multiscale Modeling Tools for Structural Biology (http://mmtsb.scripps.edu/). YUP is available for two computer platforms, LINUX on Intel X86 and MacOS X on PowerPC, and can be downloaded without charge from the YammpWeb site.

Most simulation programs provide a scripting or programming language. Perhaps the most sophisticated example is the CNS[19] language, which can be used to implement many new algorithms. The predecessor of the CNS program package is XPLOR, which also influenced the development of Yammp.

2. Method

The Python[20] language (http://www.python.org/) is a good choice for user scripting because it has a simple and clear syntax that is close to natural languages and programs do not have to be explicitly compiled. Thus, it was chosen to be the scripting language in Yammp 1. The most computationally intensive parts of YUP have to be implemented in a low level language, and C was chosen for this purpose. However, high-level languages are much easier to use, and programmers are more productive with these languages. Thus, Python was chosen as a programming language for YUP.

MMTK[21] is a molecular simulation program that is similarly implemented in Python and C, but it is for all-atom models.

An alternative would be to use a bespoke language, as was the case with the CNS package. However, by using a standard language, we are freed from the burden of designing, implementing, and maintaining a language. Furthermore, YUP users will acquire skills in a standard language that can be used to solve other programming problems.

Limited Access to Internal Data. Python is an object-oriented language with a number of predefined data types such as floating-point numbers and strings of characters. YUP extends Python by adding new data types or objects. Programmers are not granted direct access to internal data structures. Instead, controlled access is provided through objects. Objects are bundles of data and code that define their properties and behaviors. The new data types or objects are designed for molecular simulations and are implemented either in the C language for speed or in Python for productivity. However, all the objects are to be used within the Python environment.

Object-Oriented Programming in Python. (This is a discussion of some of the concepts that will be used in this article.) A defining property of objects is *inheritance*: an object can be used to define another object and the derived object has the properties and behaviors of the parent object without having to be reimplemented. The new object may also extend or modify the properties and behaviors of the parent object. Objects or data types are defined in Python using the *class* statement. The class definition includes an initialization function, also known as the constructor. The class or data type definition provides a template for the construction of any number of objects each of which is an *instance* of the data type. The properties of a Python object are usually accessible from the *data attributes*, which are written using the dot notation, e.g. **a.b** is the data attribute named **b** of the object **a**. Some objects may support assignments to certain data attributes, e.g. **a.b=c** assigns the value **c** to the attribute **b** of the object **a**. Object behaviors are usually implemented as *methods*. These are functions written with the dot notation, e.g. **a.m(**...) applies the method **m** (with the arguments unspecified) to the target object **a**. Some objects can be mapped or subscripted, e.g. **a[k]** is a mapping or subscripting of **a** and **k** is the key or subscript. The key may be of any *immutable* type. A data type may also support assignments to mapping, e.g. **a[k]=v** assigns the value **v** to the mapping **a[k]**. Operators may also be defined for an object. These include the arithmetic and *bitwise* operators. For example, in the expression **a/b** the operator is /, and the expression is correct if the operator has been implemented for the objects **a** and **b**. The creator of an object can implement any action for an operator but would usually stay close to the established meaning. The order in which operators are applied (*precedence*) is never changed. One of the most useful data types in Python is the *tuple*, which is an immutable and heterogeneous list, i.e., a list that can contain any type of data but the *tuple* itself cannot be changed. A *tuple* is written as a comma-separated list

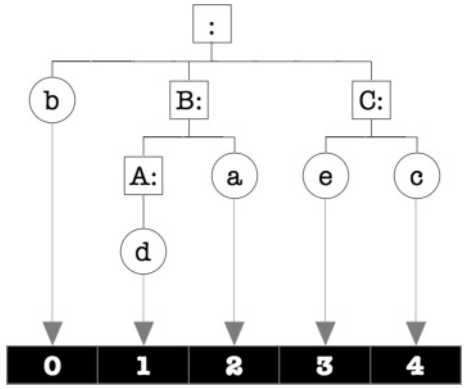


Figure 1. An *AtomMap* hierarchy and an associated *AtomVector* with group *AtomMaps* shown as squares, atom *AtomMaps* as circles, both enclosing the label, and *AtomVectors* are represented by black rectangles enclosing the index.

within parentheses. For example, **(3.142,(0,b),"a string")** is a *tuple* in which the second item is itself a *tuple*.

Types of Atoms and Interactions. YUP divides the potential energy terms into two categories: *Implicit* and *Explicit*. Interactions of closely bonded atoms such as bonds, angles, and torsions are always explicitly listed; we call such terms *Explicit*. Nonbond interactions involve all atom pairs except those that are explicitly excluded; we call such terms *Implicit*. YUP also provides *Explicit* versions of the nonbond interactions. Each atom can be assigned two atom types: an *inclusion* type and an *exclusion* type. Atom types are used as keys or indices to tables of force field parameters: *inclusion* atom types for parameters of *Explicit* interactions and *exclusion* atom types for parameters of *Implicit* interactions. Thus, parameters are specified separately for each of the two versions of nonbond terms. The parameters are keyed by as many atom types as the number of atoms in the interaction. The list of atom types is called an interaction type. For example, the parameters of the *Implicit* version of the van der Waals term (*VanderWaalsX*) are keyed by two *exclusion* atom types, and the parameters of the torsion term (*Torsions*) are keyed by four *inclusion* atom types.

Every model has to be associated with a specific parameter library. Each library defines a number of *inclusion* and *exclusion* atom types, some *Explicit* interaction types, and all *Implicit* interaction types. An atom can only be assigned an atom type that exists in the parameter library. The only exception is a nameless or *null* atom type. An interaction is assumed to have the *interaction* type that is formed from the atom types of the constituent atoms, but it is possible to assign a specific type to the interaction. In either case the interaction type must have been defined in the library. If at least one atom of an interaction has the *null* atom type, that interaction will not be evaluated. An atom that has a null *inclusion* type does not participate in any of the *Explicit* interactions. Atoms with null *exclusion* atom type are automatically excluded from all *Implicit* interactions.

The *AtomMap* Extension Data Type. This was implemented in the C language and compiled into a shared object (dynamic link library). This is closely associated with the *AtomVector* type that will be discussed later.

There are two varieties of *AtomMaps*. The "atom" variety is used to store atomic properties such as mass, charge, and so on. The "group" variety is used as a container of other *AtomMaps*. *AtomMaps* can be linked to form a hierarchy including the standard chain/residue/atom hierarchy. An *AtomMap* hierarchy resembles a file hierarchy in which the atom *AtomMaps* correspond to files and the group *AtomMaps* are like folders or directories. Thus, an *AtomMap* hierarchy is a filing system for atoms. Figure 1 shows a hierarchy containing five atoms and four groups.

The hierarchy shown in Figure 1 may be constructed as follows:

```
grR = AtomMap( datasource( "rrDNAv1" ) )

atb = grR.atom( "b", mass=79, … )

grB = grR.group( "B" )

grA = grB.group( "A" )

atd = grA.atom( "d", atomitype="cnt", … )

ata = grB.atom( "a", atomxtype="CNT", … )

grC = grR.group( "C" )

ate = grC.atom( "e", charge=-1.0, … )

atc = grC.atom( "c", mask=25, … )
```

A special group called the root always occupies the top of the hierarchy. It has no name, and it is created by a function (the constructor) instead of methods such as the rest of the hierarchy. The *AtomMap* constructor function requires the path to a library of force field parameters. The example makes use of the **datasource** function to construct the path to the standard parameter library named *rrDNAv1*.

The **atom** method adds a new atom *AtomMap* to an existing target that must be a group *AtomMap*. The expression a.**atom**("b") means to the group *AtomMap* a (the target), create and return a new atom *AtomMap* labeled "b". Additional arguments specify atomic properties such as mass, charge, *inclusion* type, and *exclusion* type. The **group** method is used to create and return a group *AtomMap* that is contained in the target *AtomMap*. In both cases, the returned *AtomMaps* will be printed if they are not assigned to variables. (These variables will be used to illustrate further examples.)

The name of the path between two *AtomMaps* is called the pathname of the destination *AtomMap*. It is formed by concatenating the names of the *AtomMaps* along the path and separating each name with a colon ":". *AtomMaps* support the mapping operation. The key must be a pathname, and the result of the mapping is the *AtomMap* of the destination. For example, the following are three references to the atom *AtomMap* labeled "d" in Figure 1:

The first line references the variable that was used to hold the atom *AtomMap* at its creation (in the earlier example). The second line shows an example of an absolute path, while the third line shows a relative path.

```
atd

grR[":B:A:d"]

grA["d"]
```

The *AtomVector* Extension Data Type. This is also implemented in the C language. *AtomVectors* contain arrays of floating-point numbers that can be manipulated in a limited number of ways, by using operators, methods, and functions. The operations include scalar and vector arithmetic, rotations and other coordinate transformations, and file operations. Generally, binary operations require conformal *AtomVectors*, i.e., the operands must have the same number of rows and columns. The contents of an *AtomVector* can be packed into or unpacked from a Python *tuple*. This allows for operations beyond those that have been defined for the *AtomVector* data type.

AtomVectors are used to represent such data as coordinates, velocities, gradients, atomic mass, atomic charges, and so on. An *AtomVector* can be associated with an *AtomMap* hierarchy.

The *AtomVector* shown in Figure 1 can be created as follows:

```
C = AtomVector( grR, numdimen = 3 )
```

This creates an *AtomVector* that is associated with the *AtomMap* hierarchy whose root is referenced by the variable *grR*. The *AtomVector* is assigned to the variable *C*, and it provides storage for five atoms and three dimensions. The number of dimensions is assumed to be three if it is not specified; thus the second argument in this example is redundant.

Any *AtomMap* from an *AtomMap* hierarchy can be used to map an *AtomVector* that is associated with this hierarchy. The result is another *AtomVector* that shares storage with the parent *AtomVector*. The derived *AtomVector* is called a slice or extract. Mapping always results in a single contiguous slice. For example

```
C[grB]
```

produces the slice with the indices 1 and 2 in Figure 1. This can also be accomplished with a function:

```
Extract( C, 1, 3 )
```

This is the only way to slice an *AtomVector* that is not associated with an *AtomMap* hierarchy.

Both the original and the extract are *AtomVectors*, and they have many properties and behaviors in common. (One of the few differences is as follows: an original *AtomVector* can be sliced many ways; extracted *AtomVectors* cannot be sliced at all.)

The molecular mechanics algorithms are implemented using *AtomVector* operations. The following is the statement that is at the heart of the steepest descent procedure for energy minimization

```
…
C.shift( G * stepsize )
…
```

where the ellipses represent the code that have been excised from the example: these include energy and gradient evaluations and such housekeeping codes as updating the nonbond list and printing intermediate results.

C is the *AtomVector* containing the coordinates, *G* is the *AtomVector* of the first derivatives, and *stepsize* is the negative step size, a scalar quantity. This example shows the use of an operator and a method.

For *AtomVectors*, the * operator is programmed to multiply the value from each cell of the left operand with the value from the corresponding cell of the right operand. Thus, the *AtomVectors* must be conformal, and the results are placed in a new *AtomVector* that is conformal with the operands. The right operand can be a scalar variable, in which case it is equivalent to having a conformal *AtomVector* that has the same value in every cell. The **shift** method adds the *AtomVector* specified in the argument to the target *AtomVector*, i.e., in-place addition. As with operators, the target and argument *AtomVectors* must be conformal, and the argument can also contain a scalar.

This code fragment applies equally to original *AtomVectors* or to their extracts, as long as the operands are conformal. All the current molecular mechanics methods are implemented in Python using *AtomVector* operations. Thus, these methods can be applied to an entire model or to contiguous slices of the model. Each method accepts a *limit* argument containing an *AtomMap*, and the calculations will be limited to the atoms contained in this *AtomMap*. The parts of the model that are not in the slice are not affected by these methods. If no limit is specified, then the calculations will be carried out on the entire model.

To take advantage of this capability, the *AtomMap* hierarchy has to be constructed so that slices of the model can be selected. The slices may overlap. This might be useful for piecewise refinement for example.

A YUP program runs under the control of a Python interpreter. The latter runs Python extension modules directly, while C extension modules are dynamically linked for execution. We will illustrate this with a simple example. The following defines a function to calculate and return the center of mass of an *AtomVector*.

```
def CenterOfMass( V ):

    from Yup.Taro.AtomVect import Reduction

    return Reduction( V, 0 ) / V.numatom
```

The **Reduction** function sums the rows or columns of the *AtomVector* specified in the first argument depending on the value specified in the second argument. The results are returned in a new *AtomVector*. The **Reduction** function is part of the *AtomVect* module, which was implemented in C and compiled into native machine code. The **CenterOfMass** function takes the *AtomVector* returned by **Reduction** and divides it by the number of atoms. The latter is available for any *AtomVector* as the *numatom* data attribute.

As an example of how the new function might be used, we will move the coordinates *R* so that the center of mass is at the origin:

```
R.shift( -CenterOfMass( R ) )
```

The Python interpreter will process the arguments first. *R* is a simple expression that needs no processing. Next, the **CenterOfMass** function will be evaluated. Like any extension module, *AtomVect* contains a table that maps each syntax element (such as the dot operator, the division operator, and the **Reduction** function in the definition of **CenterOfMass**) to an appropriate function. The value returned by **CenterOfMass** is an *AtomVector*. This is negated before it is passed to the **shift** method. These operations are also implemented as functions in the *AtomVect* module. In essence, the Python interpreter translates a YUP program into a list of functions.

The Energy Extension Data Types. These were implemented in the C language. These types are called *Bonds*, *Angles*, *Torsions*, and *VanderWaalsX*, and there are seven others. We will use the term represented by the *Angles* object for this discussion. This has the form

$$\sum_{i=1}^{N_\theta} k_{\theta,i}[\theta_i - \theta_{0,i}]^2$$

There are N_θ terms in total, and for the *i*th term, the force constant is $k_{\theta,i}$, the prevailing angle is θ_i while the equilibrium angle is $\theta_{0,i}$. To use any particular potential energy term in a model, the appropriate object must be created first:

```
Eangle = Angles( grR, "Angles" )
```

This creates an *instance* of the *Angles* data type. It is associated with the *AtomMap* hierarchy *grR*. The force field parameters will be obtained from the library that was named when *grR* was created. This object is given a label "Angles" to identify it; it can also be used to label printouts.

At this point, the *Angles* object is empty. In a moment, we will be adding the individual interactions.

The *Angles* object can be mapped with a string as key. For example

```
Eangle[":frnG:cntG:lftG:"]
```

The string is an *interaction* type, and it contains the names of three *inclusion* atom types, "frnG", "cntG", and "lftG", set off by colons ":". This mapping returns a *tuple* of numbers, the values (k_θ,θ_0) that have been defined for the *Angles* type ("frnG", "cntG", "lftG").

We can assign new values to an existing *Angles* type. For example, to preserve the force constant and change the equilibrium angle to 108°:

```
atype = ":frnG:cntG:lftG:"

ktheta, theta0 = Eangles[atype]

Eangles[atype] = ( ktheta, 108.0 )
```

(It is easier to assign the complicated *interaction* type string to a variable instead of having to type it twice.) This changes the parameters for this *interaction* type just for this *Angles* object.

The heart of force field assembly is the addition of the individual interactions. For example, to add an *Angles* interaction that involves the atoms in Figure 1, labeled "b", "e", and "d":

```
Eangle.add( ( atb, ate, atd ) )
```

This interaction is added to the *Angles* object *Eangle* using the **add** method. Note that the argument contains only one item, a *tuple* that happens to contain three values. The atoms are represented by their *AtomMaps*, which were assigned to variables in an earlier example.

When an interaction is added like the above, it is assumed that the interaction type is to be constructed from the atom types of the constituent atoms. The parameter library must contain data for this interaction type. If any one atom has a null atom type, then the interaction will not be evaluated even though it will remain in the *Angles* object. A specific interaction type can be specified as the second argument like this:

```
Eangle.add( ( atb, ate, atd ),
          ( "frnG", "cntG", "lftG" ) )
```

(Note that the *Angles* type is expressed differently from the key used in the earlier mapping.) If any one of the atom type names is replaced with a null string, the angle will still be registered but will not be evaluated.

When all the interactions have been added, the *Angles* object must be compiled before the energy (and gradients) can be evaluated.

```
Eangle.compile( C, G )
```

This registers the *AtomVector C* as atomic coordinates, and the *AtomVector G* will be used for the first derivatives. At this time, the *Angles* type is determined for each interaction: whether it is to be constructed from the atom types, or a specific type is to be used. Compilation fails if any interaction type cannot be found in the parameter library. Null interactions are not added to the list of interactions that are to be evaluated.

To evaluate the *Angles* energy

```
Eangle.evaluate()
```

The energy is evaluated for the conformation defined by the coordinates *AtomVector* registered earlier with the **compile** method. The energy is expressed in internal units.

The force field can be modified on the fly during a simulation, by applying the **modify** method, but the **compile** method must be applied to commit to the changes. For example, by writing

```
Eangle.modify( ( atb, ate, atd ) )

Eangle.compile()
```

This interaction will now be evaluated using the default *interaction* type instead of the specific type that was assigned earlier.

Table 1. Data Attributes of the *Model* Data Type

Name of Attribute	Data Type
Map[a]	*AtomMap*
Energy[a]	*Potential*
Coordinates[a]	*AtomVector*
Gradients[b]	*AtomVector*
Velocities[b]	*AtomVector*
InverseMass[b]	*AtomVector*

[a] The first three attributes must be added to a new *Model* object. [b] The remaining attributes will be created as the need arises.

It is possible to have more than one Energy object of the same kind. One reason to do this is to be able to isolate the interactions whose parameters are to be modified dynamically.

Other Energy objects are used in much the same way. Thus, for the *Bonds* object, each interaction requires two *AtomMaps* and two *inclusion* atom types. For the *Torsions* and *Impropers* types, four *AtomMaps* and four *inclusion* atom types are required. These are all *Explicit* Energy types; the interactions have to be added explicitly.

Some terms are also available in an *Implicit* version for example *VanderWaalsX*. An *Implicit* Energy object is completely full as soon as it is created. Exclusions have to be registered, by applying the **exclude** method on each atom pair that is to be excluded.

All nonbond Energy terms, whether *Implicit* or *Explicit*, are evaluated from a list of candidate pairs that have been explicitly added or have not been excluded or are not *null* interactions and are within a certain threshold distance. This list is updated periodically by applying the **update** method.

The *Potential* and *Models* Extension Data Type. These objects are implemented in Python. They are container objects, to hold lists of related objects so that they can be manipulated together. For instance, the *Potential* object holds a list of Energy objects. Thus, the **compile**, **evaluate**, and **update** methods can be applied to all the Energy objects without knowing what these objects are.

Defined for both data types are methods to add, delete, and find the component objects.

The *Model* object always consists of the root *AtomMap*, a *Potential* object, and four *AtomVectors*. These components define a complete model. The components are available as data attributes as listed in Table 1.

A *Model* data type may contain additional attributes. The *Draw* attribute specifies how a model is to be represented visually. If a *Model* object has this attribute, YUP can produce input files for the Kinemage[22] and POV-Ray (http://www.povray.org/) programs. This is especially useful for coarse-grained models, as most molecular graphics programs cannot handle idiosyncratic stereochemistry.

A *Model* type may also define attributes that are unique to a model type.

Thus, we have extended the Python language to be able to write programs to simulate coarse-grained models. Some simulation algorithms require users to define a target or objective function. The function will depend on the model and the problem, and therefore it is impractical to incorporate all possible functions into a program. For example, the CONTRA[23] algorithm to construct a minimally biased path between two conformations requires a user-defined function to provide a score for each conformation along the path. YUP is the ideal platform for such algorithms: many internal data structures are accessible to the user, and the user can use a powerful programming language to access the data and write a scoring function.

The YUP Package. YUP is a Python package, a collection of code modules organized in a file hierarchy. When a YUP program runs, the Python interpreter links each code module as it is needed. Code modules can be added to the package without disturbing existing code. These modules take care of memory management. Object size is theoretically limited only by the word size of the computer (typically 32 bits, increasingly 64 bits). The actual limits are determined by the amount of memory available on a system.The extension objects that were discussed earlier are in two subdirectories: *Yup/Energy/* for the Energy objects and *Yup/Taro/* for the remaining objects.

The molecular mechanics routines are in *Yup/Methods/*. There are routines for two energy minimization methods, a molecular dynamics method and three variants of the Monte Carlo method. The molecular mechanics algorithms are not implemented as functions but are objects derived from the *MolMech* parent object.

General-purpose utilities are kept in the *Yup/Tools/* directory. These are usually implemented as objects instead of procedures. Two of these modules are particularly important. The *FFA* module implements a template for a minimal force field assembler. The *Chains* module defines the *Chains* class, based on *FFA*, and it is a template for a force field assembler that links atoms into a simple unbranched chain. These two objects can be used as the basis for force field assemblers for more complex model types.

Three model types have been implemented so far, and they will be discussed shortly. Each model type requires a standard parameter library. The parameter library for a model type *ModelType* can be found in *Yup/Data/ModelType/*, and the code will be in *Yup/Models/ModelType/*. These models are implemented as objects that can be used to create future models. Programs that make use of the molecular simulation objects without reference to a specific model type can use the empty parameter library *Yup/Data/Null/*.

3. Results

The *rrRNAv1* Model Type. This is a coarse-grained model of RNA where each nucleotide is represented by one pseudoatom.[10] Helices are explicitly represented by semirigid structures in which hydrogen bonds between the strands are replaced by unbreakable interactions. The *rrRNAv1* model type cannot be easily assembled in conventional modeling programs, but a simple procedure has been developed for YUP.

A small RNA will be used as an example: the *rrRNAv1* model of the $tRNA^{Phe}$ molecule is shown in Figure 2.

This model can be viewed as a linear polymer of 76 monomers (each consisting of one pseudoatom), but the helices would have to be built separately (by cross-linking the monomers within each helix to give the proper three-

Table 2. YUP Package Directory[a]

Yup/					
Taro/	Energy/	Tools/	Methods/	Data/	Models/
AtomMap.so	Angles.so	Atoms.py	EnerMinim.py	rrDNAv1/	rrDNAv1/
AtomVect.so	Bonds.so	AutoGen.py	MolDynam.py	rrRNAv1/	rrRNAv1/
Model.py	ElectroI.so	Chains.py	MolMech.py	VirPack/	VirPack/
Potential.py	ElectroX.so	ChemNames.py	MonteCarlo.py	Null/	
	Impropers.so	FFA.py			
	Noens.so	Groups.py			
	SoftSphereI.so	LinSeq.py			
	SoftSphereX.so	MakeGraph.py			
	Torsions.so	MakeLine.py			
	VanderWaalsI.so	ParmTop.py			
	VanderWaalsX.so	TraceChain.py			
		Topology.py			
		misc.py			

[a] *.so files are shared objects or dynamic link libraries; *.py files are Python modules and directory names end in "/".

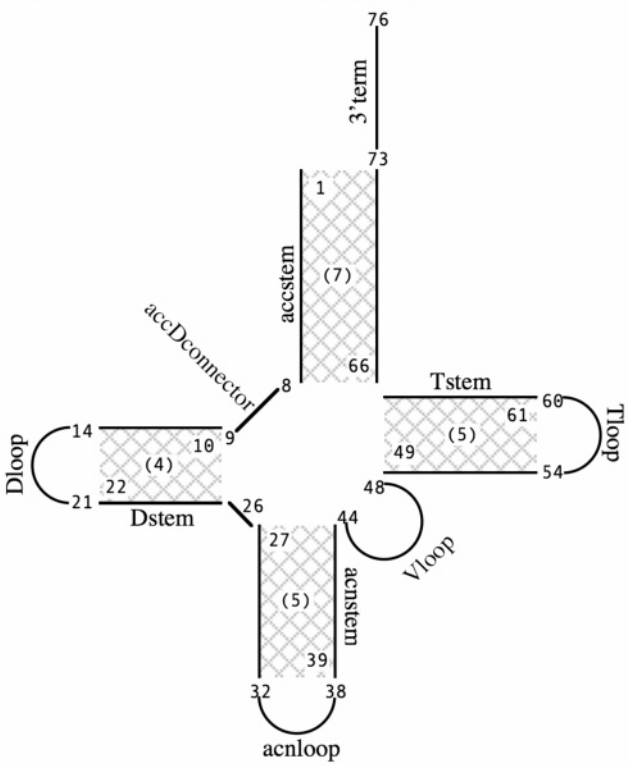


Figure 2. The *rrRNAv1* model of a $tRNA^{Phe}$ molecule: cross-hatched regions represent helices, the nucleotide indices are shown as unadorned numbers, numbers within parentheses indicate the lengths of helices, and each monomer is named except for nucleotide 26.

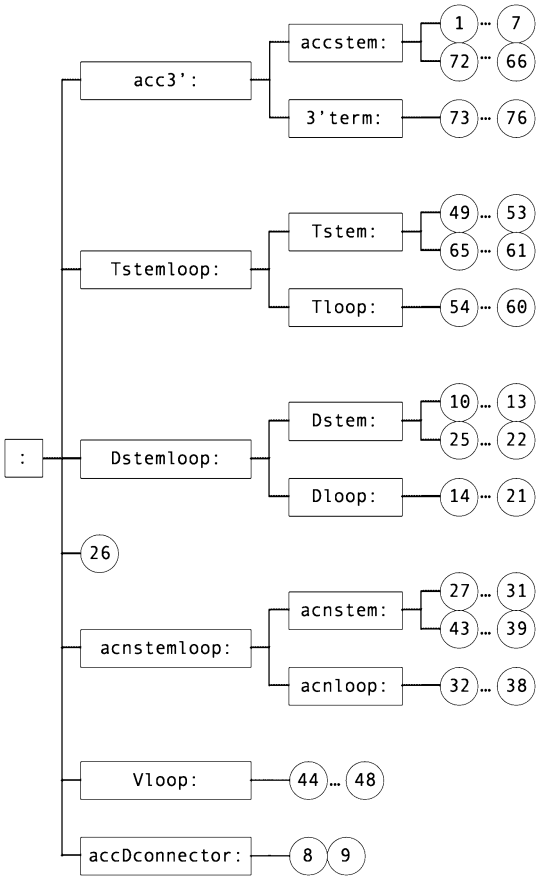


Figure 3. An *AtomMap* hierarchy of $tRNA^{Phe}$ with atom *AtomMaps* represented by circles, group *AtomMaps* by rectangles.

dimensional geometry dictated by base pairing), which would be a difficult task even for such a small RNA. A more systematic approach would be to treat the model as a nonlinear polymer of four helices and seven single-stranded tracts. The procedure in YUP relies on an *AtomMap* hierarchy in which each monomer is represented by a group *AtomMap*. Figure 3 shows one of many feasible hierarchies.

Note the additional levels in the hierarchy, e.g. *Dstem* and *Dloop* are placed in the *Dstemloop* group. Further groupings are possible, e.g. *Dstemloop*, nucleotide 26, and *acnstemloop* could all be placed in a new group; *acc3′* and *Tstemloop* could be placed in another. A hierarchy can be chosen to divide the model into possibly overlapping domains. Each domain can then be selected and treated as a rigid unit or as a focus of molecular mechanics calculations.

The desired hierarchy can be expressed as a *tuple* that contains components of three types. Each type is indicated by a keyword (*TRACT*, *HELIX*, or *DOMAIN*), followed by a label, and the specifications of a *TRACT* or *HELIX* or the contents of a *DOMAIN*. (The secondary structure may be read from a file, such as an RNAML[24] file, but the higher

levels of the hierarchy must be added manually.) The hierarchy of Figure 3 can be expressed as

```
( DOMAIN, "", (
     ( TRACT, "accDconnector", ( 8, 9 ) ),
     ( TRACT, "Vloop", ( 44, 48 ) ),
     ( DOMAIN, "acnstemloop", (
          ( HELIX, "acnstem", ( 27, 5, 39 ) ),
          ( TRACT, "acnloop", ( 32, 38 ) ) )),
     ( TRACT, "", ( 26, 26 ) ),
     ( DOMAIN, "Dstemloop", (
          ( HELIX, "Dstem", ( 10, 4, 22 ) ),
          ( TRACT, "Dloop", ( 14, 21 ) ) )),
     ( DOMAIN, "Tstemloop", (
          ( HELIX, "Tstem", ( 49, 5, 61 ) ),
          ( TRACT, "Tloop", ( 54, 60 ) ) )),
     ( DOMAIN, "acc3'", (
          ( HELIX, "accstem", ( 1, 7, 66 ) ),
          ( TRACT, "3'term", ( 73, 76 ) ) ))))
```

The contents of a *DOMAIN* are indented for clarity. This definition is constructed in one step, but larger RNAs will have to be defined in stages. The program to unravel an RNA definition, such as this, is very simple as the following code outline shows:

```
def adddomain( parent, parts ):
     G = makegroup( parent, parts[1] )
     if parts[0] == TRACT: addtract( G, parts[2] )
     elif parts[0] == HELIX: addhelix( G, parts[2] )
     elif parts[0] == DOMAIN: adddomain( G, parts[2] )
     else: raise RuntimeError, "unknown keyword"
```

The **adddomain** function constructs the *AtomMap* hierarchy, defined in the *parts* argument, underneath the group *AtomMap*, defined in the *parent* argument. The *parts* argument is a tuple containing three items: *parts[0]*, the keyword; *parts[1]*, the label; and *parts[2]*, the specification of *HELIX* or *TRACT* or the contents of *DOMAIN*. If the label is not a blank string, the **makegroup** function constructs a new group *AtomMap* under *parent* and returns the new group. Otherwise, the function returns the value of *parent*. Then, the keyword is parsed to build either a *HELIX* or a *TRACT* using the specification in the final part. The functions **addtract** and **addhelix** add the specifications for one and two tracts respectively to a buffer. The buffer is later used to add the interactions that are needed to join the tracts into a continuous chain. When the keyword is *DOMAIN*, the **adddomain** function is called to unravel *its* content. The program is short and easy to understand.

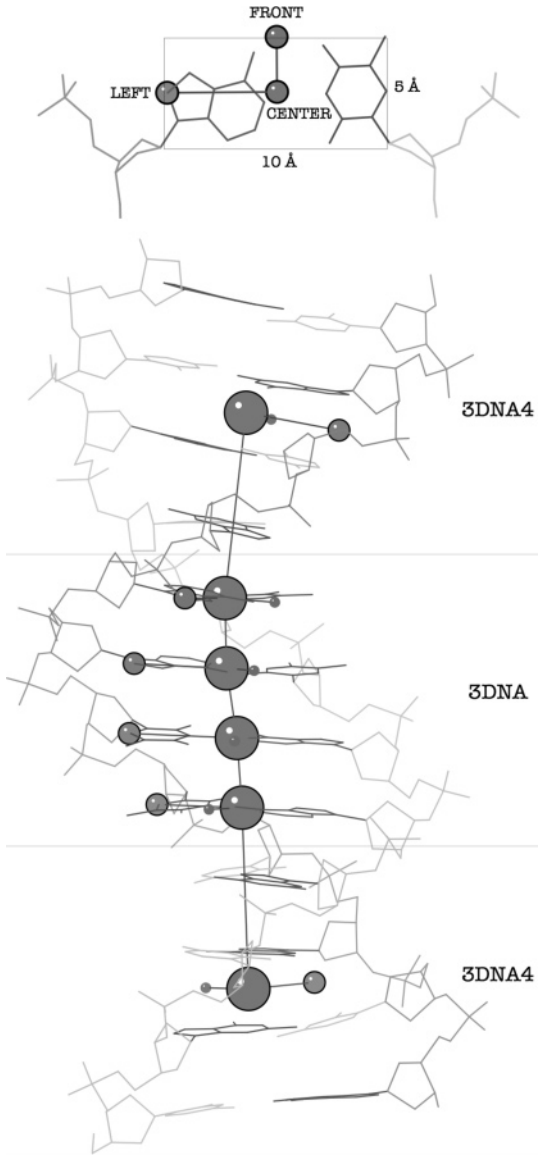


Figure 4. The conventional (all-atom) and *rrDNAv1* models of an isolated base pair (top) and a DNA duplex (bottom) are shown as wire-frames and spheres, respectively. The top panel shows how the pseudoatoms that make up the *rrDNAv1* base plane are related to the atoms of a base pair. The bottom panel shows the conventional and a multiscaled *rrDNAv1* model of a dodecamer. This figure was made using KiNG, a Kinemage[22] display program.

The *rrDNAv1* Model Type. This model type represents double-helical DNA at two levels of detail or scale: three pseudoatoms per base pair (*3DNA*)[6] and three pseudoatoms for every four base pairs (*3DNA4*). A model may contain both representations. The *rrDNAv1* model is used in the study of large pieces of DNA of hundreds to thousands of base pairs. Figure 4 shows a multiscaled *rrDNAv1* model of a dodecamer.

The *rrDNAv1* pseudoatoms, *CENTER*, *FRONT*, and *LEFT*, mark out a quadrant of a base plane. The latter is an idealized rectangle that lies in the average plane of the base pair. Idealized base pairs can be reconstituted from the *rrDNAv1* base plane. Note that the *3DNA4* base plane lies between the second and third base pairs of the tetramer that these

pseudoatoms represent. The base plane is also twisted equally between these base pairs. Only the *CENTER* pseudoatoms have bulk; they are connected to form a flexible rod.

The *rrDNAv1* force field assembler is based on the *Chains* template in *Yup/Tools/Chains.py*, which provides the code to assemble the *CENTER* atoms. The additional code is for the assembly of the *LEFT* and *FRONT* atoms. The *rrDNAv1* model is of a linear polymer (cyclized in closed circular DNA), and force field assembly is a straightforward process.

The nonbond interactions are somewhat unusual. As stated earlier, only the *CENTER* atoms have bulk, i.e., only one out of every three atoms takes part in nonbond interactions. Therefore, eight out of nine atom pairs never need to be considered. The list of interactions in the *Explicit* version of a nonbond term will be much shorter than the list of exclusions in the *Implicit* version. Thus, the *rrDNAv1* model type has historically used the *Explicit* version of the repulsion-only term (*SoftSphereI*). With support for *SoftSphereI* still in place, the *rrDNAv1* model now uses the *Implicit* version (*SoftSphereX*). Tests showed that *SoftSphereX* is significantly slower than *SoftSphereI* if a conventional exclusion list is used. However, if the *LEFT* and *FRONT* atom are assigned null *exclusion* atom types, these atoms will be automatically excluded from all *Implicit* interactions. Without the need to process a long exclusion list, *SoftSphereX* is now only slightly slower than *SoftSphereI*. (The switch to *SoftSphereX* allows *rrDNAv1* models to be used in *AMBER*, where all the nonbond terms are *Implicit*.)

The two strands of a circular DNA wind around each other, and we can count the number of times each strand crosses the other. This count is called the Linking Number.[25] Underwinding the DNA by N full turns ($360N°$) and closing the circle reduces the linking number by N. Overwinding increases the linking number.

The twist of two successive base pairs is modeled by an improper torsion angle between the equivalent *rrDNAv1* base planes. An underwound circular DNA can be constructed by setting all improper torsion angles below the equilibrium values. (The necessary offset at each base plane is 360° multiplied by the number of turns desired and divided by the number of base pairs in the model.) Given the chance, a linear model will relax, and the torsions will be returned to the equilibrium values. On the other hand, a circular model can only relieve the torsional stress by bending the helix. If the linking number is in sufficient deficit or excess, the DNA supercoils, i.e., the double helix winds around itself.

Supercoiling can occur only as long as the *CENTER* atoms form an impenetrable barrier to strand crossing. If permitted, strand crossing would relieve torsional stresses, and the linking number would revert to the equilibrium value. Further calculations would be unproductive since the model is no longer correct. We would like to detect failed calculations and terminate them quickly.

The problem is one of incorporating a nonstandard feature (calculation of the linking number) into a standard molecular mechanics method. The standard molecular dynamics class *Motors*, defined in *Yup/Methods/MolDynam.py*, applies the method **userafter** every *UserAfterInterval* steps of the procedure to solve the equations of motion. The solution is to customize the *Motors* class to have the properties and behaviors that are relevant to *rrDNAv1* models. Thus, we first derive a new class, *_motors*, from the *Motors* class:

```
class _motors( Motors ):

     def __init__( self, dnamodel ):

          Motors.__init__( self, dnamodel )

          self.UserAfterInterval = 10000

          self.__model = dnamodel

          self.__0link = CountLink( dnamodel )

          #… other settings

     def userafter( self ):

          if CountLink( self.__model ) != self.__0link:

               raise RuntimeError, "strands crossed, quitting…"
```

The new object *__motors* inherits the properties and behaviors of *Motors*, including the all-important methods to integrate the equations of motion. A function and a method are defined for the new class. The function **__init__** is special: it is called the constructor function, and it initializes a new instance of this class. Note that the parent class is initialized as part of the creation and initialization of the new class. The initialization routine sets a default value for the *UserAfterInterval* attribute (the user can set it to any value later on). The linking number is calculated (using the **CountLink** function), and the value is saved in the instance variable *__0link*. The **userafter** method definition overrides the method of the same name in the parent class. This method, called every *UserAfterInterval* steps, is now made to calculate the linking number and to compare it with the value saved by the initialization routine. If the numbers differ, then strand crossing has occurred, and the simulation is interrupted.

The *rrDNAv1* module can now use the *__motors* class, instead of *Motors*, to define molecular dynamics and simulated annealing (used for structural refinement) calculations. Thus, both structural refinement and simulations can benefit from the strand-crossing check, and failed calculations can be detected and terminated quickly.

The key feature of this implementation is the ability to store the initial linking number for the lifetime of the simulation. It would be much harder to do this if the simulation method were to be implemented as a traditional procedure.

The *VirPack* Model Type. This model type is used to study the packaging process in DNA viruses. The model represents a double-stranded DNA as a continuous chain containing one pseudoatom for every six base pairs, and the capsid is represented either as a sphere or a polyhedron. The shape and the size of the model depend on the virus that is being modeled. The *VirPack* model type makes use of many of the special features of YUP: access to internal data structures, null atom types, an unconventional *AtomMap* hierarchy, and molecular dynamics on slices of the model.

The *VirPack* force field assembler is derived from two classes: *Chains* to assemble the DNA and *Shapes* to assemble the capsid.

The class *Chains* is derived from the *Chains* template defined in *Yup/Tools/Chains.py*. The derived *Chains* class

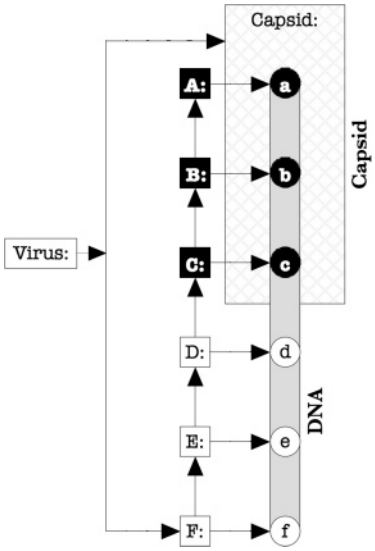


Figure 5. The *VirPack* model of a virus containing six DNA pseudoatoms (circles) linked to groups (rectangles) into a stepped hierarchy, showing the *AtomMaps* that are enclosed by the group labeled "C" (filled circles and rectangles).

links the DNA pseudoatoms into a stepped hierarchy, which allows the selection of an increasing length of the molecule. At this point in the assembly, the DNA pseudoatoms are assigned null atom types, i.e., these atoms contribute nothing to the force field. Figure 5 shows the stepped *AtomMap* hierarchy of a small system.

The *Shapes* class provides methods to assemble the capsid. The simplest type of capsid is a spherical cavity sized to match the volume of a virus, and it requires only one additional atom, the focal atom at the center of the sphere. The pseudoatoms of the DNA are subject to forces directed toward the focal atom for only as long as they are outside the capsid. The *Shapes* class is also capable of assembling more realistic capsids. Two polyhedra (icosahedron and a standard or elongated pentakisdodecahedron) of any size are available now. Barrier atoms are placed on the vertices, edges, and faces of the selected polyhedron. The barrier atoms are spaced as closely as it is required to prevent the passage of the DNA pseudoatoms. An entry hole through the capsid DNA (made by assigning null atom types to the barrier atoms in the hole) allows the passage of DNA. A more elaborate entry channel can also be constructed to mimic the portal proteins of the virus. The process of placing the barrier atoms on an icosahedron is illustrated in Figure 6.

As described earlier, the initial virus model contains a capsid part and a DNA component that is invisible to the force field. The packing procedure starts with the DNA pseudoatoms aligned on a straight line just outside the capsid. The coordinates of the exterior DNA are translated toward the capsid by half the length of the DNA−DNA bond. This pushes one DNA pseudoatom into the capsid, and this pseudoatom is now assigned the proper *inclusion* and *exclusion* atom types. The **compile** method is applied to all the Energy objects, and the newly injected DNA pseudoatom is now subject to the force field, as are all the pseudoatoms that are already in the capsid. All the interior pseudoatoms

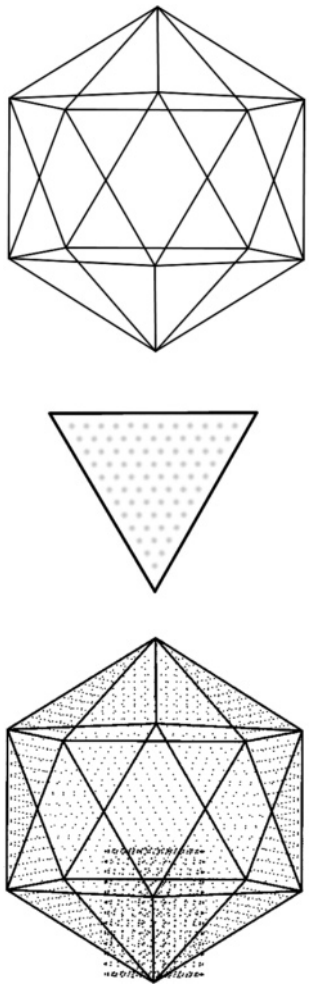


Figure 6. Construction of an icosahedral capsid: starting from the bare icosahedron (top), a template is triangulated to the necessary spacing, the grid is populated with barrier atoms (middle), and the template is copied and transformed into all the faces (bottom). The pseudoatoms defining the cylindrical core associated with the portal are also shown (bottom).

are selected for a molecular dynamics simulation. This relaxes the compressed bond near the entry as the DNA explores new conformations.

In Figure 5, the group labeled "A:" is shown linked to the DNA pseudoatom with the unique label "a". The next group up the hierarchy, labeled "B:", contains two pseudoatoms labeled "a" and "b". By going up the hierarchy, an increasing length of the DNA can be selected for molecular dynamics simulation. The group labeled "F:" encloses all the DNA pseudoatoms (for this simple system). The packing simulation is complete once the last DNA pseudoatom has been injected into the capsid and the last molecular dynamics simulation is finished. The capsid atoms, either the focal atom of a spherical capsid or the barrier atoms of a polyhedral capsid, are never subjected to molecular dynamics.

The results of a study of virus packing in spherical capsids are reported elsewhere.[26] Work on the polyhedral capsids is underway; these calculations are now taking two to four times longer to complete. Another elaboration of the model is to account for electrostatic interactions; these simulations are taking about three times the time needed for the earlier

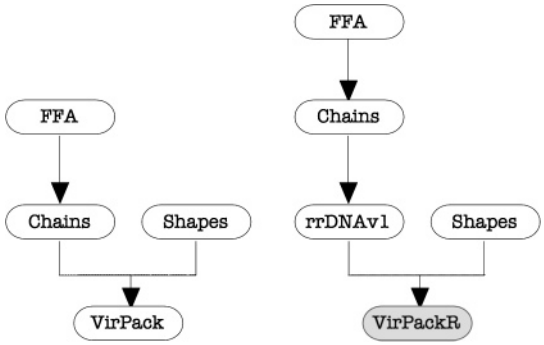


Figure 7. The DNA in the current *VirPack* model type is represented by *Chains* (left) or by *rrDNAv1* (right). The former is the simplest elastic model, with no torsional stiffness, while the latter is parametrized to match the torsional stiffness of DNA, thus mimicking the writhing and twisting of real DNA.

simulations (that lack electrostatics). Clearly, the virus model will have to remain coarse-grained even as it is enhanced.

4. Discussion

Access to internal data structures makes it possible to implement a diverse group of coarse-grained model types in YUP. Each model type has to be assembled in very different ways: *rrRNAv1* models as nonlinear polymers, *rrDNAv1* as conventional models but with very unusual nonbond interactions, and the atoms of *VirPack* models are organized in an unorthodox hierarchy. The data structures that represent the potential energy function were constructed in very different ways to suit each model. Access to internal data structures also makes it possible to customize the molecular simulation algorithms. For example, the *rrDNAv1* model can break down during a molecular dynamics simulation, but the failure is not detectable in a standard algorithm. The algorithm was extended to monitor strand crossing, and unproductive *rrDNAv1* simulations can now be avoided. The molecular dynamics algorithm is customized for the *VirPack* model to modify the force field on the fly in order to simulate the injection of DNA into a viral capsid.

The internal data structures are not directly accessible but are presented to the programmer as molecular simulation objects. This limits how the data are manipulated, but it also makes programming easier. The objects represent a higher level of abstraction than is possible with the underlying data structures. For example, the underlying data structure for the *AtomVector* object is an array of floating-point numbers. A large number of operations, including vector arithmetic, have been implemented for this object type. Thus, we can write more compact programs without the details that can obscure the algorithm.

The biggest benefit from the object-oriented approach is the ease by which objects can be extended, modified, and redeployed. Earlier, we showed how the molecular dynamics object was extended and modified for the *rrDNAv1* model type. As an example of object redeployment, consider the *VirPack* model type. The current model combines the DNA object (*Chains*) and the viral capsid object (*Shapes*). The DNA model lacks torsional stiffness, a defining feature of the molecule. This feature exists in *rrDNAv1*, a coarse-grained model of DNA. If we combine the *rrDNAv1* object with the *Shapes* object, we would then have a more realistic virus packing model (Figure 7).

Some code revisions will be required. A method has to be added to organize the *3DNA* pseudoatoms into the stepped hierarchy depicted in Figure 5. The customized packing procedure has to assign legal *rrDNAv1* atom types to each DNA unit (now consisting of three pseudoatoms) that is injected into the viral capsid. This is still a lot less effort that is required with a traditional approach.

Whenever possible, existing objects should be used to implement new model types. If it is necessary to write new code, it should be implemented as reusable objects. As the number of base objects increase, new models should be easier to implement.

Acknowledgment. This work is supported by NIH Grant P42 RR 12255 (C. L. Brooks III, P.I.).

References

(1) McCammon, J. A.; Gelin, B. R.; Karplus, M. Dynamics of folded proteins. *Nature* **1977**, *267*, 585−590.

(2) Duan, Y.; Kollman, P. A. Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science* **1998**, *282*, 642−643.

(3) Sanbonmatsu, K. Y.; Joseph, S.; Tung, C. S. Simulating movement of tRNA into the ribosome during decoding. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 15854−15859.

(4) Tan, R. K. Z.; Harvey, S. C. Yammp: Development of a Molecular Mechanics Program Using the Modular Programming Method. *J. Comput. Chem.* **1993**, *14*, 455−470.

(5) De Loof, H.; Harvey, S. C.; Segrest, J. P.; Pastor, R. W. Mean Field Stochastic Boundary Molecular Dynamics Simulation of a Phospholipid in a Membrane. *Biochemistry* **1991**, *30*, 2099−2113.

(6) Tan, R. K. Z.; Harvey, S. C. Molecular Mechanics Model of Supercoiled DNA. *J. Mol. Biol.* **1989**, *205*, 573−591.

(7) Sprous, D. Tan, R. K. Z.; Harvey, S. C. Molecular Modeling of Closed Circular DNA Thermodynamic Ensembles. *Biopolymers* **1996**, *39*, 243−258.

(8) Tan, R. K. Z.; Sprous, D.; Harvey, S. C. Molecular Dynamics Simulations of Small DNA Plasmids: Effects of Sequence and Supercoiling on Intramolecular Motions. *Biopolymers* **1996**, *39*, 259−278.

(9) Malhotra, A.; Harvey, S. C. A Quantitative Model of the *Escherichia coli* 16S RNA in the 30S Ribosomal Subunit. *J. Mol. Biol.* **1994**, *240*, 308−340.

(10) Malhotra, A.; Tan, R. K. Z.; Harvey, S. C. Modeling Large RNAs and Ribonucleoprotein Particles using Molecular Mechanics Techniques. *Biophys. J.* **1994**, *66*, 1777−1795.

(11) Valle, M.; Zavialov, A.; Li, W.; Stagg, S. M.; Sengupta, J.; Nielsen, R. C.; Nissen, P.; Harvey, S. C.; Ehrenberg, M.; Frank, J. Incorporation of aminoacyl-tRNA into the ribosome as seen by cryo-electron microscopy. *Nat. Struct. Biol*. **2003**, *10*, 899−906.

(12) Elgavish, T.; VanLoock, M. S.; Harvey, S. C. Exploring Three-Dimensional Structures of the HIV-1/tRNALys3 Initiation Complex. *J. Mol. Biol.* **1999**, *285*, 449−453.

(13) Arsuaga, J.; Tan, R. K. Z., Vazquez, M.; Sumners, de W.; Harvey, S. C. Investigation of viral DNA packaging using molecular mechanics models. *Biophys. Chem.* **2002**, *101*−*102*, 475−484.

(14) LaMarque J. C.; Le, T. V.; Harvey, S. C. Packaging double-helical DNA into viral capsids. *Biopolymers* **2004**, *73*, 348−355.

(15) Case, D. A.; Darden, T. A.; Cheatham, T. E. III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Wang, B.; Pearlman, D. A.; Crowley, M.; Brozell, S.; Tsui, V.; Gohlke, H.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Schafmeister, C.; Caldwell, J. W.; Ross, W. S.; Kollman, P. A. *AMBER 8*; University of California: San Francisco, CA, 2004.

(16) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E. III; DeBolt, S.; Ferguson, D.; Seibel, G.; and Kollman, P. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comput. Phys. Commun.* **1995**, *91*, 1−41.

(17) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *J. Comput. Chem.* **1983**, *4*, 187−217.

(18) MacKerell, A. D., Jr.; Brooks, B.; Brooks, C. L. III; Nilsson, L.; Roux, B.; Won, Y.; Karplus, M. In *The Encyclopedia of Computational Chemistry*; von Ragué Schleyer, P., Ed.; Wiley: Chichester, 1998; Vol. 1, CHARMM: The Energy Function and Its Parameterization with an Overview of the Program, pp 271−277.

(19) Brunger, A. T.; Paul, D.; Adams, G.; Clore, M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. Crystallography & NMR System: A New Software Suite for Macromolecular Structure Determination. *Acta Crystallogr.* **1998**, *D54*, 905−921.

(20) van Rossum, G.; Drake, F. L., Jr. *Python Reference Manual*; Python Software Foundation: 2005.

(21) Hinsen, K. The molecular modeling toolkit: A new approach to molecular simulations. *J. Comput. Chem.* **2000**, *21*, 79−85.

(22) Richardson, D. C.; Richardson, J. S. The kinemage: a tool for scientific communication. *Protein Sci.* **1992**, *1*, 3−9.

(23) Harvey, S. C.; Gabb, H. A. Conformational Transitions Using Molecular Dynamics with Minimum Biasing. *Biopolymers* **1993**, *33*, 1167−1172.

(24) Waugh, A.; Gendron, P.; Altman, R.; Brown, J. W.; Case, D.; Gautheret, D.; Harvey, S. C.; Leontis, N.; Westbrook, J.; Westhof, E.; Zuker, M.; Major, F. RNAML: a standard syntax for exchanging RNA information. *RNA* **2002**, *8*, 707−717.

(25) Bauer, W. R.; Crick, F. H.; White, J. H. Supercoiled DNA. *Sci. Am.* **1980**, *243*, 100−113.

(26) Locker, C. R.; Fuller, S. D.; Harvey, S. C. DNA organization and thermodynamics during viral packing. Submitted for publication.

CT050323R

JCTC Journal of Chemical Theory and Computation

An Analytical Electrostatic Model for Salt Screened Interactions between Multiple Proteins

Itay Lotan* and Teresa Head-Gordon*,†

Department of Bioengineering, University of California, Berkeley, Berkeley, California 94720

Received October 24, 2005

Abstract: We present a new general analytical solution for computing the screened electrostatic interaction between multiple macromolecules of arbitrarily complex charge distributions, assuming they are well described by spherical low dielectric cavities in a higher dielectric medium in the presence of a Debye−Hückel treatment of salt. The benefits to this approach are 3-fold. First, by exploiting multipole expansion theory for the screened Coulomb potential, we can describe direct charge−charge interactions and all significant higher-order cavity polarization effects between low dielectric spherical cavities containing their charges, while treating these higher order terms correctly at all separation distances. Second, our analytical solution is general to arbitrary numbers of macromolecules, is efficient to compute, and can therefore simultaneously provide on-the-fly updates to changes in charge distributions due to protein conformational changes. Third, we can change spatial resolutions of charge description as a function of separation distance without compromising the desired accuracy. While the current formulation describes solutions based on simple spherical geometries, it appears possible to reformulate these electrostatic expressions to smoothly increase spatial resolution back to greater molecular detail of the dielectric boundaries.

1. Introduction

Atomistic molecular dynamics simulations routinely and beneficially address materials problems in the condensed phase. However, there is another set of problems on the supramolecular scale where the limitations of size and time scales are reached, examples being the recognition events and subsequent complexation of multiple proteins in explicit solvent environments, or the study of phase behavior and interfacial properties of colloid systems. Atomistic modeling is too computationally demanding to evaluate for times long enough to measure the macromolecules' traversal over large spatial domains and to do this with enough statistical confidence to analyze complex phase behavior, association kinetics, or mechanism of assembly.

Fortunately, coarse-grained models may actually be the more sensible approach when large-scale spatial organization or dynamic events occurring over long time scales are operative. Spatial coarse-graining would involve, for example, removing explicit solvent molecules and ions as well as ignoring an individual macromolecule's internal motion for some period of time.

At large spatial separations between charged macromolecules in solution, electrostatic interactions will dominate, so that an appropriate coarse-grained model could describe them as complex charge distributions imbedded in a low dielectric medium surrounded by a high dielectric solvent continuum with salt screening defined by explicit microions or implicitly through a Debye−Hückel treatment. The electrostatic potential and forces and torques are found by solving the full Poisson−Boltzmann equation (PBE) or when expanding the exponential and keeping only linear terms, by solving the corresponding linearized PBE. Both linearized and the full PBE are typically solved numerically using either a finite-difference (FD) method, boundary-element (BE) meth-

* Corresponding authors e-mail: itayl@berkeley.edu (I.L.) and TLHead-Gordon@lbl.gov (T.H.-G.).

† Schlumberger Fellow, Chemistry Department, Cambridge University, Lensfield Road, Cambridge CB2 1EW, United Kingdom.

10.1021/ct050263p CCC: $33.50
Published on Web 03/16/2006

ods or based on more recent work with an adaptive multilevel finite element solution,[1] and the associated electrostatic fields are evaluated by a central difference approximation using the numerical potential. An additional feature to solving the full or linearized PBE numerically is that solutions are not restricted to simple geometries but can describe greater levels of molecular detail of the macromolecular cavity shape.

In this work, we have developed a new analytical solution for computing the electrostatic interaction between multiple macromolecules of arbitrarily complex charge distributions in aqueous salt solutions, assuming the macromolecules and their environment are well described by spherical low dielectric cavities in the presence of a Debye−Hückel treatment of salt.[2] By exploiting multipole expansion theory for the screened Coulomb potential, we are able to describe direct charge−charge interactions and all significant higher-order cavity polarization effects between low dielectric spherical cavities containing their charges, treating all higher order polarization terms correctly at all separation distances. Hitherto this level of completeness was only available through numerical solutions to the PBE, such as the grid-based FD solutions[3−6] and BE methods[7−9] (See ref 1 for a survey.). While the numerical methods are able to represent the shape of the molecule to high resolution, they are expensive to compute. They have an inherent tradeoff between spatial resolution (grid discretization for the FD methods and choice of boundary elements for the BE methods) and memory and time requirements, that limits them to system of few macromolecules. In contrast, our method is efficient and fast to compute even for many macromolecules, including frequent updates of changes in their charge distributions due to induced conformational changes. Unlike the effective charge approximation of Gabdoulline and Wade,[10] our formalism allows for automatically changing spatial resolutions of charge description on the fly as a function of separation distance so the desired accuracy is always achieved with no waste in computation.

Previous analytical solutions to the linearized PBE under spherical geometries have fallen short of providing complete and useable solutions for multiple macromolecules of arbitrarily complex charge distribution. When the system is comprised of only one macromolecule, a complete solution was offered by Kirkwood[11] more than 70 years ago, and a similar solution has recently been derived by Hoffman et al.[12] The interaction of two macromolecules has proven to be more difficult, and many different partial and approximate solutions have been proposed. Glendinning and Russel[13] proposed an analytical solution using multipole expansion for two equal spheres with uniform charge density. Assuming weak interaction between the two molecules, a superposition approximation was used, first by Verwey and Overbeek[14] for uniformly charged molecules and later by Sader and Lenhoff[15] for arbitrary charge distributions. Allen and Hansen[16] proposed a solution based on variational charge-density functional method for a system of two macromolecules, each with one charge on the line connecting their centers, and with no salt. Phillies[17] was the first to offer a complete analytical solution for two molecules with arbitrary charge distributions in solution with salt; however, his use of implicit re-expansion operators made the solution cumbersome and impractical. Later McClurg and Zukoski[18] improved on this solution by using, among other things, explicit re-expansion operators. While their solution is similar to ours, it is limited to the first few multipole orders and to only two molecules and thus is not as complete and general as the solution we provide here.

In devising this new formulation, much attention and care were directed at providing a computationally practical and reliable method. Our method is numerically stable even for large molecular radii and high multipole orders, which makes it usable even at short separation distances, independent of the Debye screening length. Our formulation also provides easy access to many quantities of interest, such as forces, torques, and interaction energies, and affords an intuitive physical understanding of the different factors that contribute to them. Our approach takes advantage of recent advances in the computation of Yukawa (screened Coulomb) potential[19,20] and the Helmholtz equation,[21,22] that makes the derivation and use of the re-expansion operator straightforward and computationally efficient. While the current analytical approach exhibits a reliance on simple spherical geometries, it appears possible to reformulate these electrostatic expressions to increase spatial resolution back to molecular level descriptions of cavity geometries.

2. Theory

2.1. Mathematical Preliminaries. In this section we will establish some definitions and recount some identities that will be used in developing the theory needed for computing the electrostatic interaction between charged spherical molecules. First, a definition of the *spherical harmonics*, following the work of Gumerov and Duraiswami[21]

$$Y_{n,m}(\theta,\phi) = (-1)^m \sqrt{\frac{(n-m)!}{(n+m)!}} P_{n,m}(\cos\theta) e^{im\phi} \qquad (1)$$

where $P_{n,m}(x)$ are the *Associated Legendre Polynomials*.[23] Note, this is not the standard definition of spherical harmonics, as it differs by a $(-1)^m$ factor. The resulting expressions, however, are simpler. This definition can be used for all $m \geq 0$. When m is negative we rely on the simple identity

$$Y_{n,-m}(\theta,\phi) = \overline{Y_{n,m}}(\theta,\phi) \qquad (2)$$

where $\overline{Y_{n,m}}$ is the complex conjugate of $Y_{n,m}$.

We also require the use of the *modified spherical Bessel functions* (MSBF) defined as

$$i_n(z) = \sqrt{\frac{\pi}{2z}} I_{n+1/2}(z)$$

$$k_n(z) = \sqrt{\frac{\pi}{2z}} K_{n+1/2}(z) \qquad (3)$$

where $I_n(z)$ and $K_n(z)$ are the *modified Bessel functions* of the first and second kinds, respectively.[23] We will make use of two addition theorems that use the MSBFs and the spherical harmonics.[23] For convenience we copy them below. Given two points in 3-D space specified by their spherical coordinates $\mathbf{p} = [\rho,\vartheta,\varphi]$ and $\mathbf{t} = [r,\theta,\phi]$, with $\rho < r$, and the Euclidean distance between them $R = |\mathbf{t} - \mathbf{p}|$, we have

$$\frac{1}{R} = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \frac{\rho^n}{r^{n+1}} \overline{Y_{n,m}}(\vartheta,\varphi) Y_{n,m}(\theta,\phi) \quad (4)$$

$$\frac{e^{-\kappa R}}{R} = \frac{2\kappa}{\pi} \sum_{n=0}^{\infty} \sum_{m=-n}^{n} (2n+1) i_n(\kappa\rho) k_n(\kappa r) \overline{Y_{n,m}}(\vartheta,\varphi) Y_{n,m}(\theta,\phi) \quad (5)$$

Note that while in eq 4 the exponentiated ratio of distances ρ^n/r^{n+1} appears explicitly; in eq 5 it is implicit in the MSBFs. Namely, $k_n(r) \propto 1/r^{n+1}$ and $i_n(r) \propto \rho^n$.

Kirkwood developed a general analytic solution to the linearized Poisson−Boltzmann equation (PBE) around a spherical solvent-excluding dielectric cavity

$$\Phi(\mathbf{t}) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \frac{A_{n,m}}{r^{n+1}} e^{-\kappa r} \hat{\mathrm{k}}_n(\kappa r) Y_{n,m}(\theta,\phi) \quad (6)$$

where the coefficients $A_{n,m}$ are determined by enforcing appropriate boundary conditions. His solution uses an adapted definition of the MSBF suitable for the conditions of the problem:

$$\hat{\mathrm{k}}_n(z) = \frac{2}{\pi} \frac{e^z z^{n+1}}{(2n-1)!!} k_n(z) \quad (7)$$

Based on this definition we can now rewrite the addition theorem in eq 5

$$\frac{e^{-\kappa R}}{R} = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \frac{\rho^n}{r^{n+1}} \hat{\mathrm{i}}_n(\kappa\rho) e^{-\kappa r} \hat{\mathrm{k}}_n(\kappa r) \overline{Y_{n,m}}(\vartheta,\varphi) Y_{n,m}(\theta,\phi) \quad (8)$$

where

$$\hat{\mathrm{i}}_n(z) = \frac{(2n+1)!!}{z^n} i_n(z) \quad (9)$$

Besides their usefulness in solving the Poisson−Boltzmann equation around spherical cavities, these *adapted* MSBFs are also more numerically stable than the *standard* MSBFs since the dependence on exponents of r and ρ that may grow rapidly with the order n of the functions has been taken out. A recursive method for directly computing $\hat{\mathrm{k}}_n$ and $\hat{\mathrm{i}}_n$ without first computing k_n and i_n is given in the Appendix.

2.2. Boundary-Value Problem. The system we are solving is comprised of N spherical molecules of radii a_i, whose centers are positioned at points $\mathbf{c}^{(i)}$. The molecules are immersed in a solvent, containing charged salt ions. The dielectric constant (permittivity) in the interior of each molecule is ϵ_p, and the dielectric constant of the solvent is ϵ_s. For simplicity we will assume all molecules have the same dielectric constant, although that is not a restriction of the solution. The inverse Debye screening length of the ions in the solution is κ.

The electrostatic potential of the system is governed by the PBE. For systems under physiological conditions the PBE can be safely linearized to give[2]

$$\nabla[\epsilon(\mathbf{r})\nabla\Phi(\mathbf{r})] - \epsilon(\mathbf{r})\kappa^2\Phi(\mathbf{r}) = 4\pi\rho(\mathbf{r}) \quad (10)$$

In the general case, when the dielectric boundary has arbitrary shape, only numerical solutions exist. For the case of spherical cavities, however, one can write general parametric expressions for the potential in different parts of the system. The values of the coefficients are then determined by enforcing a set of boundary conditions. These conditions stipulate the continuity of the electrostatic potential and the electrostatic field at the surface of each molecule. These conditions take the form of the following equations

$$\Phi_{out}^{(i)} = \Phi_{in}^{(i)}|_{r=a_i}$$
$$\epsilon_s \frac{\partial\Phi_{out}^{(i)}}{\partial r} = \epsilon_p \frac{\partial\Phi_{in}^{(i)}}{\partial r}\Big|_{r=a_i} \quad (11)$$

where $\Phi_{in}^{(i)}$ is the total potential inside molecule i and $\Phi_{out}^{(i)}$ is the total potential in the solvent, both expressed in the coordinate frame of molecule i.

The system at hand consists of two types of regions: the inside of each molecule and the outside solvent. The inside region is characterized by a low dielectric constant and an arbitrary distribution of free charges. In this region eq 10 reduces to $\epsilon_p\Delta\Phi(\mathbf{r}) = 4\pi\rho(\mathbf{r})$. The outside region is characterized by the high dielectric constant of the solvent and the presence of salt. In this region eq 10 reduces to $\Delta\Phi(\mathbf{r}) - \kappa^2\Phi(\mathbf{r}) = 0$. Following Kirkwood[11] we write general parametrized expressions for both regions. The electrostatic potential inside molecule i is described by

$$\Phi_{in}^{(i)}(\mathbf{t}) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \left(\frac{E_{n,m}^{(i)}}{\epsilon_p r^{n+1}} + B_{n,m}^{(i)} r^n \right) Y_{n,m}(\theta,\phi) \quad (12)$$

where $\mathbf{E}^{(i)}$ is the multipole expansion of the charges inside molecule i. It is defined as

$$E_{n,m}^{(i)} = \sum_{j=1}^{M_i} q_j^{(i)} (\rho_j^{(i)})^n Y_{n,m}(\vartheta_j^{(i)}, \varphi_j^{(i)}) \quad (13)$$

where M_i is the number of charges in molecule i, $q_j^{(i)}$ is the magnitude of the jth charge, and $\mathbf{p}_j^{(i)} = [\rho_j^{(i)}, \vartheta_j^{(i)}, \varphi_j^{(i)}]$ is its position in spherical coordinates. The expansion in eq 13 can be used in eq 12 only when $a_i \geq r > \max_j \rho_j^{(i)}$, a condition that is always met when computing the potential at the surface of the sphere. Its use facilitates the application of the boundary conditions and the subsequent computation of the coefficient vectors $\mathbf{B}^{(i)}$. If later on one desires to compute the potential at a point inside the molecule, the following variant of eq 12 can be used:

$$\Phi_{in}^{(i)}(\mathbf{t}) = \frac{1}{\epsilon_p} \sum_{j=1}^{M_i} \frac{q_j^{(i)}}{||\mathbf{p}_j^{(i)} - \mathbf{t}||} + \sum_{n=0}^{\infty} \sum_{m=-n}^{n} B_{n,m}^{(i)} r^n Y_{n,m}(\theta,\phi) \quad (14)$$

The general form of the potential outside all molecules (in a coordinate frame whose origin is the center of molecule i) is

$$\Phi_{out}^{(i)}(\mathbf{t}) = \frac{1}{\epsilon_s} \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \left(\frac{A_{n,m}^{(i)}}{r^{n+1}} e^{-\kappa r} \hat{\mathrm{k}}_n(\kappa r) + L_{n,m}^{(i)} r^n \hat{\mathrm{i}}_n(\kappa r) \right) Y_{n,m}(\theta,\phi) \quad (15)$$

The vectors $\mathbf{A}^{(i)}$ and $\mathbf{B}^{(i)}$ of eqs 12 and 15 are the free parameters (coefficients) that will be determined through the application of the boundary conditions of eq 11.

The coefficients $\mathbf{L}^{(i)}$ in eq 15 are a re-expansion of the external potential coefficients $\mathbf{A}^{(j)}$, $j \neq i$ of all other molecules in the system. It is defined as

$$\mathbf{L}^{(i)} = \sum_{\substack{j=1 \\ j\neq i}}^{N} \mathrm{T}^{(i,j)} \cdot \mathbf{A}^{(j)} \tag{16}$$

where $\mathrm{T}^{(i,j)}$ is the linear re-expansion operator that transforms a multipole expansion at $\mathbf{c}^{(j)}$ to a local (Taylor) expansion at $\mathbf{c}^{(i)}$. This operator is described in detail in the Appendix. The use of the $\mathrm{T}^{(i,j)}$ operators allows us to represent the potentials due to all molecules in the coordinate frame of a single molecule. This is crucial to our ability to analytically apply the boundary conditions and derive compact expressions for the free parameters of the formulation.

Applying the boundary conditions to eqs 12 and 15 yields the following expressions for the elements of the parameter vectors

$$A_{n,m}^{(i)} = \gamma_n^{(i)} \delta_n^{(i)} L_{n,m}^{(i)} + \gamma_n^{(i)} E_{n,m}^{(i)} \tag{17}$$

$$B_{n,m}^{(i)} = \frac{1}{\epsilon_s}\left(\frac{A_{n,m}^{(i)}}{a_i^{2n+1}} e^{-\kappa a_i} \hat{\mathrm{k}}_n(\kappa a_i) + L_{n,m}^{(i)} \hat{\mathrm{i}}_n(\kappa a_i)\right) - \frac{E_{n,m}^{(i)}}{\epsilon_p a_i^{2n+1}} \tag{18}$$

where the vectors of constants $\Gamma^{(i)}$ and $\Delta^{(i)}$ are

$$\gamma_n^{(i)} = \frac{(2n+1)e^{\kappa a_i}}{(2n+1)\hat{\mathrm{k}}_{n+1}(\kappa a_i) + n\hat{\mathrm{k}}_n(\kappa a_i)(\epsilon_p/\epsilon_s - 1)} \tag{19}$$

$$\delta_n^{(i)} = \frac{a_i^{2n+1}}{2n+1}\left[\kappa^2 a_i^2 \frac{\hat{\mathrm{i}}_{n+1}(\kappa a_i)}{2n+3} + n\hat{\mathrm{i}}_n(\kappa a_i)(1 - \epsilon_p/\epsilon_s)\right] \tag{20}$$

Given a solution to all the $\mathbf{A}^{(i)}$ vectors, the $\mathbf{B}^{(i)}$ vectors are easily calculated using eq 18. We therefore concentrate on solving for $\mathbf{A}^{(i)}$. Based on eq 17 we can write a system of linear equations that describes the electrostatics of the entire problem

$$\mathbf{A} = \Gamma \cdot (\Delta \cdot \mathrm{T} \cdot \mathbf{A} + \mathbf{E}) \tag{21}$$

where

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}^{(1)} \\ \mathbf{A}^{(2)} \\ \vdots \\ \mathbf{A}^{(N)} \end{bmatrix} \quad \Gamma = \begin{bmatrix} \Gamma^{(1)} & 0 & \cdots & 0 \\ 0 & \Gamma^{(2)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \Gamma^{(N)} \end{bmatrix}$$

$$\Delta = \begin{bmatrix} \Delta^{(1)} & 0 & \cdots & 0 \\ 0 & \Delta^{(2)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \Delta^{(N)} \end{bmatrix}$$

$$\mathrm{T} = \begin{bmatrix} 0 & \mathrm{T}^{(1,2)} & \cdots & \mathrm{T}^{(1,N)} \\ \mathrm{T}^{(2,1)} & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathrm{T}^{(N-1,N)} \\ \mathrm{T}^{(N,1)} & \cdots & \mathrm{T}^{(N,N-1)} & 0 \end{bmatrix} \quad \mathbf{E} = \begin{bmatrix} \mathbf{E}^{(1)} \\ \mathbf{E}^{(2)} \\ \vdots \\ \mathbf{E}^{(N)} \end{bmatrix} \tag{22}$$

and

$$\Gamma^{(i)} = \begin{bmatrix} \gamma_1^{(i)} & 0 & \cdots & 0 \\ 0 & \gamma_2^{(i)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \gamma_p^{(i)} \end{bmatrix} \quad \Delta^{(i)} = \begin{bmatrix} \delta_1^{(i)} & 0 & \cdots & 0 \\ 0 & \delta_2^{(i)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \delta_p^{(i)} \end{bmatrix} \tag{23}$$

Note that in order to write the system of equations we need to truncate all vectors $\mathbf{A}^{(i)}$ and concordantly all other coefficient vectors at a finite order (the maximal value for n), which we call p. The appropriate choice for p depends on practical considerations, which we will elaborate on in the next section.

We can give intuitive physical meaning to the matrices and vectors of eq 21. The vectors $\mathbf{A}^{(i)}$ represent the *effective* multipole expansion of the charge distributions of each molecule. The effective expansion represents an equivalent system, where the low dielectric cavities have been removed and the solvent (with the salt ions) is allowed to penetrate everywhere. Namely, one can use these effective expansions to represent the now missing dielectric boundary. This effective expansion is reminiscent of the effective charges of Gabdoulline and Wade[10] in the sense that it incorporates the effect of the dielectric boundary into the charge distribution. In contrast to Gabdoulline and Wade, our effective expansion also accounts for the polarization effects caused by the dielectric boundary. The diagonal matrix Δ can be understood as a cavity polarization operator. It transforms the coefficients of an external charge distribution (re-expanded around the center of the cavity) to yield the expansion coefficients of a polarization charge distribution that forms on the surface of the cavity (due to that external charge distribution). The diagonal matrix Γ can be taken to be a dielectric boundary crossing operator. It transforms the coefficients of a charge distribution expansion inside a dielectric cavity to yield the effective expansion as if the charges were in the solution. Both the Γ and Δ operators are functions of the parameters of the system being solved, namely the radii a_i, the dielectric constants ϵ_p and ϵ_s, and the inverse Debye screening length κ . Thus eq 21 can be understood to state the intuitive fact that the external potential field induced by a molecule is the sum of the contribution of its free charges and the contribution of polarization charges induced by other molecules, transformed by the effect of its dielectric boundary.

It is instructive to note that this formulation separates the solution into a free charge distributions (the vector $\mathbf{E}^{(i)}$) and operators that represent the configuration of the system (the matrices $\mathrm{T}^{(i,j)}$) and the geometric and physical conditions (the operators $\Gamma^{(i)}$ and $\Delta^{(i)}$). The charge distributions and the operators are independent except for the fact that they all need to be defined in terms of the same molecular centers. This observation may be exploited in devising approximate schemes that can deal with nonspherical molecules.

2.3. Interaction Energy. Plugging the expression in eq 18 into eq 12 and rearranging the terms yields

$$\Phi_{in}^{(i)}(\mathbf{t}) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n}\left[\frac{E_{n,m}^{(i)}}{\epsilon_p}\left(\frac{1}{r^{2n+1}} - \frac{1}{a_i^{2n+1}}\right) + \frac{\gamma_n^{(i)} E_{n,m}^{(i)}}{\epsilon_s a_i^{2n+1}} e^{-\kappa a_i}\hat{\mathrm{k}}_n(\kappa a_i)\right] r^n Y_{n,m}(\theta,\phi) + \frac{1}{\epsilon_s}\sum_{n=0}^{\infty}\sum_{m=-n}^{n}\gamma_n^{(i)} L_{n,m}^{(i)} r^n Y_{n,m}(\theta,\phi) \quad (24)$$

Note that the first double summation is the internal potential due to the charges inside the molecule and the effects of the external salt ions, while the second double summation is the potential due to the other charged molecules (external sources of charge), which we dub $\hat{\Phi}_{in}^{(i)}$:

$$\hat{\Phi}_{in}^{(i)}(\mathbf{t}) = \frac{1}{\epsilon_s}\sum_{n=0}^{\infty}\sum_{m=-n}^{n}\gamma_n^{(i)} L_{n,m}^{(i)} r^n Y_{n,m}(\theta,\phi) \quad (25)$$

Plugging eq 17 into eq 15 yields the expression for the potential anywhere in the solvent (outside all molecules) with the coordinate frame centered at molecule i

$$\Phi_{out}^{(i)}(\mathbf{t}) = \frac{1}{\epsilon_s}\sum_{n=0}^{\infty}\sum_{m=-n}^{n}\gamma_n^{(i)}(\delta_n^{(i)} L_{n,m}^{(i)} + E_{n,m}^{(i)})\frac{e^{-\kappa r}\hat{\mathrm{k}}_n(\kappa r)}{r^{n+1}} Y_{n,m}(\theta,\phi) + \frac{1}{\epsilon_s}\sum_{n=0}^{\infty}\sum_{m=-n}^{n} L_{n,m}^{(i)} r^n \hat{\mathrm{i}}_n(\kappa r) Y_{n,m}(\theta,\phi) \quad (26)$$

Note that by analogy to eq 6, the expansion of the total charge distribution of a molecule, dubbed $\mathbf{G}^{(i)}$, taking into account both its free charges and the charges due to the polarization of its dielectric cavity by external charges, is

$$G_{n,m}^{(i)} = \delta_n^{(i)} L_{n,m}^{(i)} + E_{n,m}^{(i)} \quad (27)$$

The interaction energy $\Omega^{(i)}$ of each molecule is the product of its charge distribution (both free and polarization charges) of eq 27 with the potential due to external sources of eq 25. We can compute this as the inner product of the coefficients of the two corresponding multipole expansions

$$\Omega^{(i)} = \frac{1}{\epsilon_s}\langle \Gamma^{(i)} \cdot \mathbf{L}^{(i)}, \mathbf{G}^{(i)} \rangle = \frac{1}{\epsilon_s}\langle \mathbf{L}^{(i)}, \mathbf{A}^{(i)} \rangle \quad (28)$$

where the inner product of two coefficient vectors is defined as

$$\langle \mathbf{U}, \mathbf{V} \rangle = \sum_{n=0}^{p}\sum_{m=-n}^{n} U_{n,m}\overline{V_{n,m}} \quad (29)$$

Note that using our formulation computing the interaction energy is very simple and straightforward.

2.4. Charge Distribution on the Surface of the Molecules. Because of the dielectric discontinuity, a charge distribution develops on the surface of each of the molecules in the system. These charges are a reaction to both the electrostatic field of the free charges of the molecule itself and the electrostatic field of other molecules in its vicinity. In the former case we shall call this charge distribution *self-polarization* charge, and in the latter case we shall call this charge distribution *external polarization* charge. The charged ions in the solution, whose distribution is governed by the potential field, also polarize the dielectric cavities. We will call this charge distribution the *salt polarization* charges. We can compute the surface charge density by using Gauss' Law on an infinitesimal patch on the surface of each molecule with zero thickness. Recall that Gauss' Law states that the net electrostatic field through the surface of a closed volume is proportional to the net charge inside. In the case of a small and very thin patch on the surface of a sphere, this translates to the following equality:

$$\left.\frac{\partial \Phi_{in}^{(i)}}{\partial r}\right|_{r=a_i} - \left.\frac{\partial \Phi_{out}^{(i)}}{\partial r}\right|_{r=a_i} = 4\pi\sigma^{(i)} \quad (30)$$

Plugging in the expression for $\Phi_{in}^{(i)}$ from eq 12 and for $\Phi_{out}^{(i)}$ from eq 15 we arrive at the following expression for the charge distribution on the surface:

$$\sigma^{(i)}(\theta,\phi) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n}\frac{(2n+1)}{4\pi a_i^{n+2}}\left[E_{n,m}^{(i)}\left(\frac{\gamma_n^{(i)}\hat{\mathrm{k}}_{n+1}(\kappa a_i)}{\epsilon_s e^{\kappa a_i}} - \frac{1}{\epsilon_p}\right) + \frac{nL_{n,m}^{(i)}\gamma_n^{(i)} a_i^{2n+1}}{\epsilon_s(2n+1)}\left(1 - \frac{\epsilon_p}{\epsilon_s}\right)\right] Y_{n,m}(\theta,\phi) \quad (31)$$

A careful look at the terms in eq 31 reveals a clear separation between self-polarization and external polarization. The first term in the brackets describes the self-polarization charge distribution. Note, however, that it also includes the salt polarization that depends on the potential of the free charges. We can extract the *pure* self-polarization term by setting $\kappa = 0$. We then get

$$\sigma_{self}^{(i)}(\theta,\phi) = \frac{\epsilon_p - \epsilon_s}{\epsilon_p}\sum_{n=0}^{\infty}\sum_{m=-n}^{n}\frac{(2n+1)}{4\pi a_i^{n+2}}\frac{(n+1)}{(n+1)\epsilon_s + n\epsilon_p} E_{n,m}^{(i)} Y_{n,m}(\theta,\phi) \quad (32)$$

Similarly the second term above describes the external polarization charge distribution (here too, the salt polarization governed by the external potential is included). Without the contribution of the salt-induced charges, the external polarization charge is

$$\sigma_{ext}^{(i)}(\theta,\phi) = \frac{\epsilon_s - \epsilon_p}{\epsilon_s}\sum_{n=0}^{\infty}\sum_{m=-n}^{n} a_i^{n-1}\frac{(2n+1)}{4\pi}\frac{n}{(n+1)\epsilon_s + n\epsilon_p} L_{n,m}^{(i)} Y_{n,m}(\theta,\phi) \quad (33)$$

Note that if among the external potential sources are also other dielectric cavities that are polarized by the potential of this molecule's charges, then there would be some (usually small) portion of the external polarization charges that actually depends on the free charges. Namely, in this case, removing the free charges will have some effect on the external polarization charges.

By integrating eq 31 over the surface of the molecule we can construct a multipole expansion representation of the polarization charge:

$$S_{n,m}^{(i)} = E_{n,m}^{(i)}\left(\frac{\gamma_n(i)\hat{\mathrm{k}}_{n+1}(\kappa a_i)}{\epsilon_s e^{\kappa a_i}} - \frac{1}{\epsilon_p}\right) + \frac{nL_{n,m}^{(i)}\gamma_n^{(i)}\, a_i^{2n+1}}{\epsilon_s(2n+1)}\left(1 - \frac{\epsilon_p}{\epsilon_s}\right) \tag{34}$$

We can now use this expansion to describe the potential outside molecule *i* (ignoring the direct contribution of other molecules) as if there was no dielectric boundary; namely, no difference in dielectric constant between the inside and the outside:

$$\Phi_{out}^{(i)}(\mathbf{t}) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n}(S_{n,m}^{(i)} + E_{n,m}^{(i)}/\epsilon_p)\frac{e^{\kappa a_i}}{\hat{\mathrm{k}}_{n+l}(\kappa a_i)}\frac{e^{-\kappa r}\hat{\mathrm{k}}_n(\kappa r)}{r^{n+1}}Y_{n,m}(\theta,\phi) \tag{35}$$

Here, the term $[e^{\kappa a_i}/\hat{\mathrm{k}}_{n+1}(\kappa a_i)]$ is a simplified version of the coefficients $\Gamma^{(i)}$ of eq 19, that result when there is no dielectric discontinuity, only a uniform dielectric, and salt cannot penetrate into the cavity.

Note that eq 35 is not equivalent to eq 15. Recall that molecule *i* not only caused a disruption in the dielectric medium of the solvent but also created a cavity devoid of salt ions. This displacement of salt affects the screening of the potential field of the external sources of charge by the salt and thus alters it. By examining the difference between the two equations we are able to extract another set of multipole expansion coefficients that accounts for this effect:

$$Z_{n,m}^{(i)} = \frac{1}{\epsilon_s}\frac{a_i^{2n+1}}{(2n+1)}\frac{\kappa^2 a_i^2\hat{\mathrm{i}}_{n+1}(\kappa a_i)}{(2n+3)}\frac{e^{\kappa a_i}}{\hat{\mathrm{k}}_{n+1}(\kappa a_i)}L_{n,m}^{(i)} \tag{36}$$

Figure 1 shows the surface charge on a Barstar molecule (see section 3 for details) interacting with a second Barstar molecule in close proximity (1Å separation distance between bounding spheres). The charge distribution is broken into its different components using the formulas we have developed above (eqs 31−33).

2.5. Forces and Torques. The force acting on a molecule due to other molecules in its vicinity is the gradient (in Cartesian coordinates) of the interaction energy with respect to the position of the center of the molecule, namely

$$F^{(i)} = \nabla_i\Omega^{(i)} = \frac{1}{\epsilon_s}[\langle\nabla_i\mathbf{L}^{(i)},\mathbf{A}^{(i)}\rangle + \langle\mathbf{L}^{(i)},\nabla_i\mathbf{A}^{(i)}\rangle] \tag{37}$$

where ∇_i is the gradient with respect to the coordinates of the center $\mathbf{c}^{(i)}$.

We compute the gradient of the coefficient vectors $\mathbf{A}^{(i)}$ by solving the linear system

$$\nabla_i\mathbf{A} = \Gamma\cdot\Delta\cdot[\nabla_i\mathrm{T}\cdot\mathbf{A} + \mathrm{T}\cdot\nabla_i\mathbf{A}] \tag{38}$$

which we arrive at by applying the gradient operator to both sides of eq 21. The gradient of the re-expansion operator, $\nabla_i\mathrm{T}^{(i,j)}$, can be computed analytically as described in the Appendix.

The torque on a molecule is the sum of the torques due to each of the charges in the molecule. By definition, the torque on a charge is the cross product of its position relative to the center of mass of the molecule (the moment arm) with the force it experiences. It is not difficult to convince oneself that the force acting on the external polarization charges generates no net torque. This is because rotating a spherical cavity cannot change the distribution of external polarization charges on its surface. The torque on molecule *i*, due to the charge $q_j^{(i)}$ inside it, is thus

$$\tau_j^{(i)} = \mathbf{p}_j^{(i)} \times \left[\frac{1}{-\epsilon_s}q_j^{(i)}\sum_{n=0}^{\infty}\sum_{m=-n}^{n}\gamma_n^{(i)}\,\nabla_i L_{n,m}^{(i)}(\rho_j^{(i)})^n Y_{n,m}(\vartheta_j^{(i)},\varphi_j^{(i)})\right] \tag{39}$$

where $\mathbf{p}_j^{(i)} = [x_j^{(i)}, y_j^{(i)}, z_j^{(i)}]$ in Cartesian coordinates. Due to the linearity of the cross product we can rewrite this as

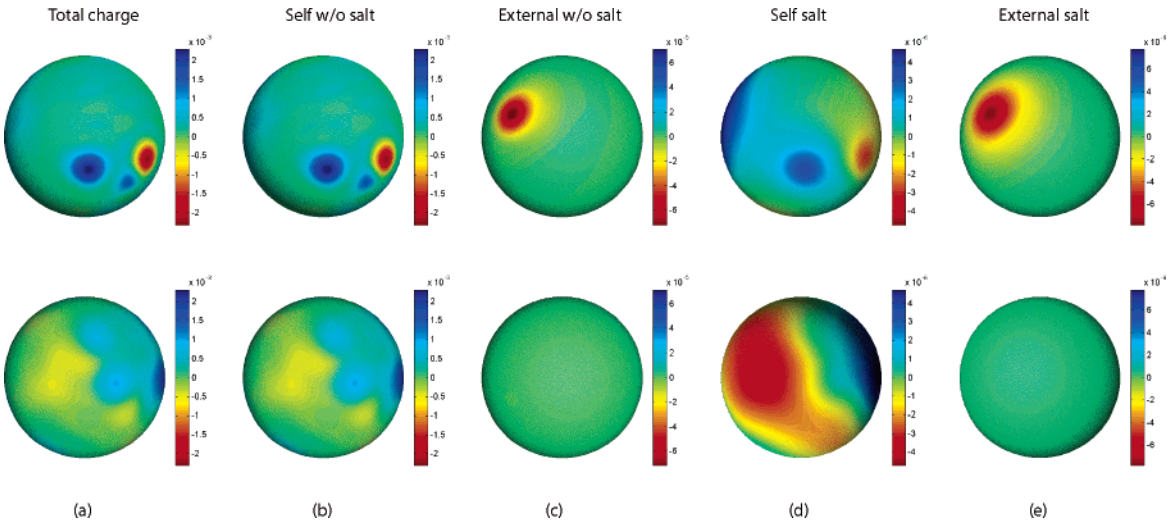


Figure 1. Surface charge distribution of a Barstar molecule broken up to its components: (a) the total charge distribution, (b) the self-polarization charge w/o the contribution of the salt ions, (c) the external-polarization charge distribution due to another molecule in close proximity w/o the contribution of the salt ions, (d) polarization charge distribution due to salt ions governed by the potential induced by the molecule's own charge, and (e) polarization charge distribution due to salt ions governed by the potential induced by external sources. Top row shows view from front and bottom row shows view from back.

$$\tau_j^{(i)} = \frac{1}{\epsilon_s} \sum_{n=0}^{\infty} \sum_{m=-n}^{n} [q_j^{(i)} \gamma_n^{(i)} (\rho_j^{(i)})^n Y_{n,m}(\vartheta_j^{(i)}, \varphi_j^{(i)}) \cdot \mathbf{p}_j^{(i)}] \times \nabla_i L_{n,m}^{(i)} \tag{40}$$

Summing over all charges of the molecule we get the following expression for the net torque

$$\tau^{(i)} = \frac{1}{\epsilon_s} [{}^x\mathbf{H}^{(i)}, {}^y\mathbf{H}^{(i)}, {}^z\mathbf{H}^{(i)}] \times [\nabla_i \mathbf{L}^{(i)}] \tag{41}$$

where

$$\begin{aligned} {}^xH_{n,m}{}^{(i)} &= \sum_{j=1}^{M_i} h_{j;n,m}^{(i)} x_j^{(i)} \\ {}^yH_{n,m}^{(i)} &= \sum_{j=1}^{M_i} h_{j;n,m}^{(i)} y_j^{(i)} \\ {}^zH_{n,m}^{(i)} &= \sum_{j=1}^{M_i} h_{j;n,m}^{(i)} z_j^{(i)} \end{aligned} \tag{42}$$

and

$$h_{j;n,m}^{(i)} = \gamma_n^{(i)} q_j^{(i)} (\rho_j^{(i)})^n Y_{n,m}(\vartheta_j^{(i)}, \varphi_j^{(i)}) \tag{43}$$

The cross product for two sets of three coefficient vectors, used in eq 41, is defined in analogy to the regular cross product for vectors in 3-D, where the complex inner product of two coefficient vectors defined in eq 29 replaces the regular multiplication of two vector elements. Note that the three coefficient vectors $[{}^x\mathbf{H}^{(i)}, {}^y\mathbf{H}^{(i)}, {}^z\mathbf{H}^{(i)}]$ do not change as long as the charges inside the molecule do not move and do not depend on the position of the molecule or other molecules.

3. Analysis of Electrostatic Potential

Two of the benefits of our formulation are the correct inclusion of cavity polarization to all orders, whose full contribution is not always included in partial analytical solutions and in numerical approaches such as Gabdoulline and Wade's effective charges[10] and desolvation terms,[24] as well as the ability to consider multiple proteins and thereby multibody effects through the cavity polarizations. To test the consequence of neglecting cavity polarization and multibody effects, we examined model configurations of multiple copies of both Barstar and Barnase,[25] which were chosen as representing typical macromolecules in terms of size and charge distribution. Barnase is an extracellular ribonuclease, and Barstar is its intracellular inhibitor. Barstar and Barnase are modeled as spheres of radii 21.8Å and 28Å, respectively. Their charge distribution is based on assignment of OPLS[26] partial charges centered on atomic coordinates taken from their PDB file (1BRS), after inclusion of hydrogens on heavy atoms using CHARMM.[27] Barstar has a relatively large net charge of −6e, and Barnase has a net charge of +2e. These two proteins were chosen somewhat arbitrarily, and we could have chosen for the sake of the computations we present here almost any other protein. The only assumption we make is that the charge distribution inside them is not significantly different from what is normally observed in proteins. The radii were chosen so that all charges are within a sphere centered at the geometric center of each molecule. It is possible to choose a different center and as a result have somewhat different radii; however, one must remember that the correctness of the method depends on the inclusion of all charges inside the cavity, and that if the chosen center differs from the center of mass, the computed torques must be translated to the true centers of mass. In all computations that follow we use $\epsilon_p = 4$, $\epsilon_s = 80$, and the salt concentration in the solution is 50 mM, which at room temperature yields $\kappa = 0.074$.

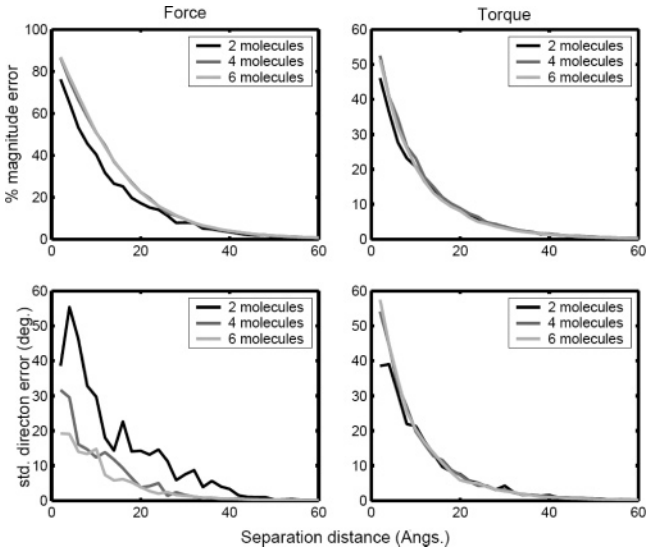


Figure 2. Higher order effect of the dielectric cavities (through polarization and salt displacement) on the forces and the torques in different configurations of Barstar molecules. The left column shows the effect on the forces acting on the different molecules, and the right column shows the effect on the torques. The top row shows the average relative error in magnitude, and the bottom row shows the standard deviation of the error in direction when ignoring the dielectric cavities. Results are shown for configurations of two, four, and six molecules, as a function of separation distance.

We considered configurations of two molecules, an equilateral tetrahedron with four molecules centered on each vertex, and six molecules configured as a 3-D cross, with the centers at the ends of the segments. In Figure 2 we show the relative error in magnitude and the standard deviation of the error in direction of both forces and torques, when cavity polarization is ignored, computed for different configurations at different separation distances for different numbers of Barstar molecules. The results for Barnase were similar and are not shown. We can see that magnitude errors as high as 85% in force and over 50% in the torque and is often as high as 60% in the direction of force or torque, at the point of closest protein contact. Although the neglect of molecular features of the interface will also be a significant error at these closest separations, at longer separations where the coarse-grained approximation is still a reasonable one, the cumulative effect of 10−30% error by neglecting cavity polarization could strongly influence, for example, a simulation of mechanism and rate of protein assembly.

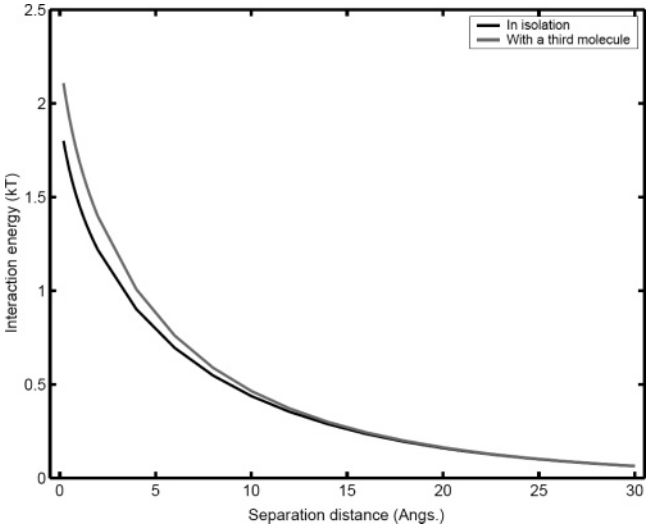


Figure 3. The effect of a third molecule on the interaction free energy of two Barstar molecules, as a function of separation distance. Results are averaged over all possible relative orientations of the molecules (by uniform sampling). In black is the interaction energy for the two molecules in isolation, and in gray the energy when a third molecule is placed an equal distance from both (forming an equilateral triangle).

In Figure 3 we compare the interaction energy of two Barstar molecules as a function of their relative separation (distance between surfaces of bounding spheres), to the effect on their potential of mean force (PMF) due to a third Barstar molecule. The three molecules are set up on the vertices of an equilateral triangle, and the PMF is calculated by uniformly sampling their relative orientations. The nonadditivity effects due to cavity polarization introduced by the third Barstar modulate the interactions of the two Barstar molecules so that there is an increased repulsion at closer distances. This effect would typically be missing in formulations of the electrostatic potential in which it is only feasible to consider pair interactions.

Figure 4 shows the power of our formulation in solving the electrostatics of very large systems. The interaction energy, force, and torque on a unit cell comprised of a single Barstar molecule in an infinite lattice is computed using neighborhoods as large as 3 surrounding layers or 7^3 unit cells (243 molecules participating in the calculation all together). As we can see the relative error going from two to three layers is smaller than 1% even at very close separations, which ensures us that a larger neighborhood is not required.

4. Implementation

4.1. Computing the Re-Expansion Operator and Its Derivatives. The literature describes a number of methods for computing the re-expansion operator $T^{(i,j)}$.[19−22,28,29] We based our implementation on the method of Gumerov and Duraiswami[21] for the solution of the Helmholtz equation because of its relative simplicity and efficiency. We adopt it to our formulation and extend it to the computation of partial derivatives. Note that in what follows we restrict our discussion to the practical aspects of using the re-expansion operator. The theory behind it can be found elsewhere.[21] The matrix $T^{(i,j)}$, which represents a re-expansion in an arbitrary direction, is decomposed (diagonalized) to a re-expansion along the *Z*-axis $S^{(i,j)}$. This is done using a rotation operator $R^{(i,j)}$ that orients the desired direction of re-expansion with the *Z*-axis and its inverse that rotates back to the original coordinate frame after the diagonalized re-expansion is performed. We thus have the following decomposition of $T^{(i,j)}$:

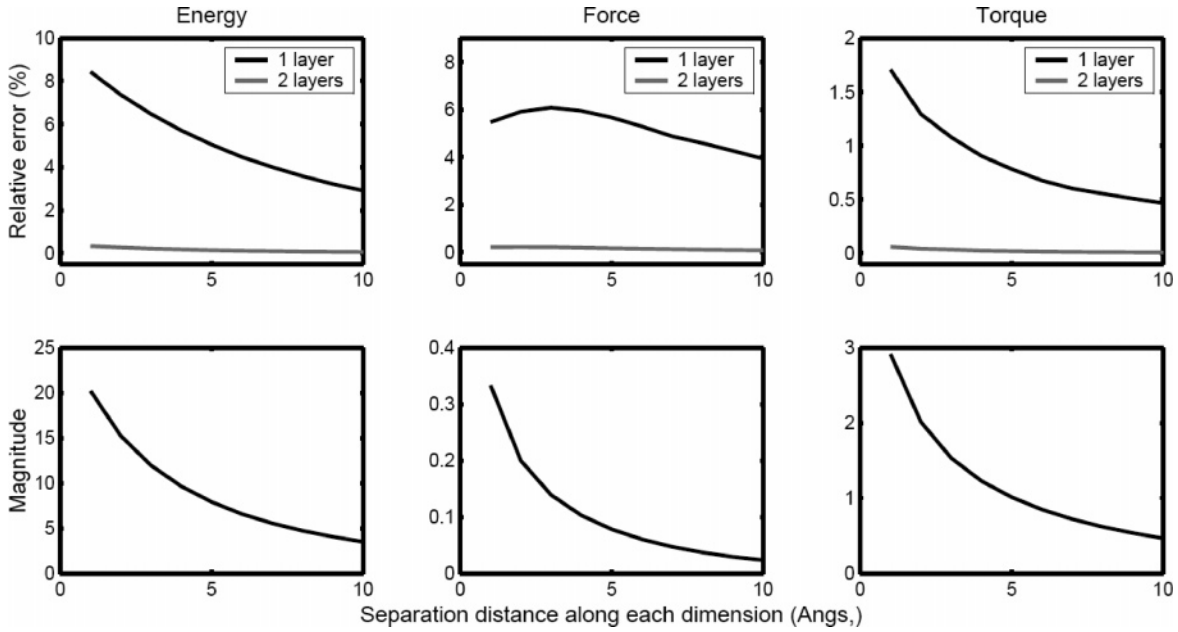


Figure 4. Computing the energy, force, and torque on a molecule in an infinite lattice. The top row shows the relative error in computing the energy, force, and torque of a unit cell comprised of a single Barstar molecule when taking into account interactions with only the first layer (a 3 × 3 × 3 neighborhood) and only the second layer (a 5 × 5 × 5 neighborhood). The second row shows the magnitude of the energy (kT), force (kT/Å), and torque (kT/rad) on a single Barstar molecule in the infinite lattice.

$$\mathrm{T}^{(i,j)} = (R^{(i,j)})^H \cdot S^{(i,j)} \cdot R^{(i,j)} \tag{44}$$

While $\mathrm{T}^{(i,j)}$ is a full matrix having possibly $O(p^4)$ independent elements, both $R^{(i,j)}$ and $S^{(i,j)}$ have only $O(p^3)$ nonzero elements, which can be computed recursively as described in Appendix A.1 in $O(p^3)$ time. Moreover, a simple relation exists between $\mathrm{T}^{(i,j)}$ and $\mathrm{T}^{(j,i)}$ namely, $\mathrm{T}^{(j,i)} = (\mathrm{T}^{(i,j)})^H$, which entails that

$$\mathrm{T}^{(j,i)} = (R^{(i,j)})^H \cdot (S^{(i,j)})^H \cdot R^{(i,j)} \tag{45}$$

Since both $R^{(i,j)}$ and $S^{(i,j)}$ are sparse matrices, whose nonzero elements are known in advance, we compute the re-expansion $\mathbf{Z} = \mathrm{T} \cdot \mathbf{X}$ in three steps:

$$X^1_{n,m} = \sum_{s=-n}^{n} R^{m,s}_n X_{n,s}$$

$$X^2_{n,m} = \sum_{l=|m|}^{p} S^m_{n,l} X^1_{l,m}$$

$$Z_{n,m} = \sum_{s=-n}^{n} \overline{R^{s,m}_n} X^2_{n,s} \tag{46}$$

The partial derivatives of the re-expansion matrix have a similar decomposition that can be computed by differentiating eq 44. For re-expansion along $\mathbf{v}^{(i,j)} = \mathbf{c}^{(i)} - \mathbf{c}^{(j)} = [r,\theta,\phi]$ we get

$$\frac{\partial \mathrm{T}^{(i,j)}}{\partial r} = (R^{(i,j)})^H \cdot \frac{\partial S^{(i,j)}}{\partial r} \cdot R^{(i,j)}$$

$$\frac{\partial \mathrm{T}^{(i,j)}}{\partial \theta} = \left(\frac{\partial R^{(i,j)}}{\partial \theta}\right)^H \cdot S^{(i,j)} \cdot R^{(i,j)} + (R^{(i,j)})^H \cdot S^{(i,j)} \cdot \frac{\partial R^{(i,j)}}{\partial \theta}$$

$$\frac{\partial \mathrm{T}^{(i,j)}}{\partial \phi} = \left(\frac{\partial R^{(i,j)}}{\partial \phi}\right)^H \cdot S^{(i,j)} \cdot R^{(i,j)} + (R^{(i,j)})^H \cdot S^{(i,j)} \cdot \frac{\partial R^{(i,j)}}{\partial \phi} \tag{47}$$

Note that $R^{(i,j)}$ is independent of r and $S^{(i,j)}$ is independent of θ and ϕ . To convert the derivatives with respect to the spherical coordinates to derivatives with respect to Cartesian coordinates the following 3×3 matrix is used:

$$\begin{bmatrix} \frac{\partial Z_{n,m}}{\partial x} \\ \frac{\partial Z_{n,m}}{\partial y} \\ \frac{\partial Z_{n,m}}{\partial z} \end{bmatrix} = \begin{bmatrix} \sin\theta\cos\phi & \frac{\cos\theta\cos\phi}{r} & \frac{-\sin\phi}{r\sin\theta} \\ \sin\theta\sin\phi & \frac{\cos\theta\sin\phi}{r} & \frac{\cos\phi}{r\sin\theta} \\ \cos\theta & \frac{-\sin\theta}{r} & 0 \end{bmatrix} \cdot \begin{bmatrix} \frac{\partial Z_{n,m}}{\partial r} \\ \frac{\partial Z_{n,m}}{\theta} \\ \frac{\partial Z_{n,m}}{\partial \phi} \end{bmatrix} \tag{48}$$

A careful examination of eqs 47 and 48 leads to the conclusion that $\nabla_j \mathrm{T}^{(i,j)} = -\nabla_i \mathrm{T}^{(i,j)}$. The derivatives with respect to θ and ϕ encounter a singularity when $\sin\theta \to 0$. In Appendix A.2 we discuss this in detail and explain how to handle the singularities.

4.2. Iterative Computation of Forces and Torques. A one-time exact solution of the forces for a given configuration of the molecular system can be computed by inverting eqs 21 and 38. We first compute the values of the $\mathbf{A}^{(i)}$ vectors by solving

$$\mathbf{A} = (\mathrm{I} - \Gamma \cdot \Delta \cdot \mathrm{T})^{-1} \cdot \Gamma \cdot \mathbf{E} \tag{49}$$

and then for every molecular center $\mathbf{c}^{(j)}$ we compute all $\nabla_j \mathbf{A}^{(i)}$ by solving

$$\nabla_i \mathbf{A} = (\mathrm{I} - \Gamma \cdot \Delta \cdot \mathrm{T})^{-1} \cdot \Gamma \cdot \Delta \cdot \nabla_i \mathrm{T} \cdot \mathbf{A} \tag{50}$$

which is an inversion of eq 38. Note that the expensive matrix inversion operation need only be performed once, since the same inverted matrix is used in both eqs 49 and 50. The values of $\mathbf{A}^{(i)}$ and $\nabla_i \mathbf{A}^{(i)}$ are then plugged into eq 37 to yield the force on molecule i and into eq 41 to yield the torque on molecule i.

As long as the distance between any pair of molecules is larger than the sum of their radii, which is a requirement of our theory, the inverted matrix of eq 49 is strictly diagonally dominant.[30] This guarantees that an iterative approach such as the Gauss-Seidel method[30] will converge quickly. Moreover, in a simulation setting, the position and orientation of each molecule changes little between time-steps, thus it makes sense to *fix* the solution computed at the previous time-step, instead of computing a new solution from scratch. Therefore instead of a costly matrix inversion we use the Gauss-Seidel method, first to solve eq 21 and then to solve eq 38 for each center $\mathbf{c}^{(j)}$. Casting eq 21 as an iteration step in a Gauss-Seidel solution of the linear system yields

$$\begin{aligned} \mathbf{A}^{(1)}_t &= \Gamma^{(1)} \cdot (\Delta^{(1)} \cdot \sum_{j=2}^{N} \mathrm{T}^{(1,j)} \cdot \mathbf{A}^{(j)}_{t-1} + \mathbf{E}^{(1)}) \\ &\vdots \\ \mathbf{A}^{(N)}_t &= \Gamma^{(N)} \cdot (\Delta^{(N)} \cdot \sum_{j=1}^{N-1} \mathrm{T}^{(N,j)} \cdot \mathbf{A}^{(j)}_t + \mathbf{E}^{(N)}) \end{aligned} \tag{51}$$

At iteration t each $\mathbf{A}^{(i)}_t$ is computed in turn, i going from 1 to N using the most recent values of all $\mathbf{A}^{(j)}$, $j \neq i$. The iterations stop when the relative change to $\mathbf{A}^{(i)}$ falls below the desired precision. The relative change is computed as

$$\text{change} = \frac{1}{2Np^2} \sum_{i=1}^{N} \sum_{k=1}^{p^2} \frac{|\mathbf{A}^{(i)}_{t,k} - \mathbf{A}^{(i)}_{t-1,k}|}{|\mathbf{A}^{(i)}_{t,k}| + |\mathbf{A}^{(i)}_{t-1,k}|} \tag{52}$$

In a simulation setting, $\mathbf{A}_0{}^{(i)} = \Gamma^{(i)} \cdot \mathbf{E}^{(i)}$ for the computation at the first step of the simulation, and for each subsequent step $\mathbf{A}_0{}^{(i)}$ is set to its final value in the previous step.

Once all $\mathbf{A}^{(i)}$ are computed, an analogous iterative form of eq 38 can be used to compute the $\nabla_j \mathbf{A}^{(i)}$:

$$\begin{aligned} \nabla_j \mathbf{A}^{(1)}_t &= \Gamma^{(1)} \cdot \Delta^{(1)} \cdot \sum_{k=2}^{N} [\nabla_j \mathrm{T}^{(1,k)} \cdot \mathbf{A}^{(k)} + \mathrm{T}^{(1,k)} \cdot \nabla_j \mathbf{A}^{(k)}_{t-1}] \\ &\vdots \\ \nabla_j \mathbf{A}^{(N)}_t &= \Gamma^{(N)} \cdot \Delta^{(N)} \cdot \sum_{k=1}^{N-1} [\nabla_j \mathrm{T}^{(N,k)} \cdot \mathbf{A}^{(k)} + \mathrm{T}^{(N,k)} \cdot \nabla_j \mathbf{A}^{(k)}_t] \end{aligned} \tag{53}$$

Here $\nabla_j \mathbf{A}_0{}^{(i)} = \mathbf{0}$ for the first step of a simulation. At each subsequent step the final value of the previous time-step is used.

4.3. Adaptive Control of Precision. Since we must truncate each vector of expansion coefficients at a finite order, there will always be an inherent error in the computation of the forces. However by carefully choosing the

truncation order we can guarantee the error will never exceed a preset value. In the interest of speed, we would like to use the smallest order required to give the desired precision. A theoretical upper bound on the error when computing the potential at a distance R using a truncated multipole expansion is $O(a/R)^{p+1}$, where a is the radius of the charges and $R > a$.[31] When computing the interaction between two sets of charges whose centers are a distance R apart, with radii a_i and a_j and $R > a_i + a_j$, an upper bound on the truncation error is

$$err_p = O\left(\frac{a_i}{R - a_j}\right)^{p+1} \tag{54}$$

Thus for each interaction between a pair of molecules a different truncation order should be used that depends on the respective separation and radii of the pair.

The theoretical bound of eq 54 is often not tight for proteins, because the constant hidden in the bound depends on $\sum_i |q_i|$, which for proteins is much larger than $\sum_i q_i$. Thus we use a heuristic bound that is much tighter in practice (although its correctness cannot be guaranteed). We approximate the truncation error as

$$err_p^{(i,j)} = \left| \frac{[(\mathbf{A}^{(i)})^H \cdot \mathrm{T}^{(i,j)} \cdot \mathbf{A}^{(j)}]_{p+1} - [(\mathbf{A}^{(i)})^H \cdot \mathrm{T}^{(i,j)} \cdot \mathbf{A}^{(j)}]_p}{[(\mathbf{A}^{(i)})^H \cdot \mathrm{T}^{(i,j)} \cdot \mathbf{A}^{(j)}]_p} \right| \tag{55}$$

Namely, the relative error in computing the interaction by truncating at order p is the difference between the value using maximal order $p + 1$ and the value using maximal order p, divided by the value using maximal order p. Since the values of $\mathbf{A}^{(i)}$ and $\mathbf{A}^{(j)}$ at the current time-step are not yet known, their values from the previous step are used. Assuming these values change very little between time-steps the error introduced by this approximation will be small. When $err_p^{(i,j)}$ is larger than the desired error, the truncation order is incremented and tested again. If $err_p^{(i,j)}$ falls below a second, lower threshold, the order we are using is too high and should be decremented and tested again. This heuristic bound starts to break as the ratio in eq 54 approaches 1; in this case it may be safer to also test truncation order $p + 2$ and even $p + 3$ in order to make sure the truncated terms indeed fall off and do not diverge.

4.4. Computational Complexity. The size of the computational problem at hand is governed by a number of parameters: N: the number of molecules, $M = \max_i(M_i)$: the number of partial charges in each molecule, and p: the maximal expansion order used in the calculations. To compute the forces and torques as described above, the following calculations are performed:

1. Compute the multipole expansion for each molecule. $O(Mp^2)$ per molecule, $O(NMp^2)$ in total.

2. Invert a matrix of size $Np^2 \times Np^2$ which takes $O(N^3p^6)$ time (eq 49).

3. Compute all $\mathbf{A}^{(i)}$ and all $\nabla_j\mathbf{A}^{(i)}$ (eqs 49 and 50). Involves multiplying Np^2 vectors by $Np^2 \times Np^2$ matrices which takes $O(N^2p^4)$ per molecule or $O(N^3p^4)$ in total.

4. Compute the forces and torques, which involves dot products of vectors of p^2 elements. $O(p^2)$ per molecule or $O(Np^2)$ in total.

The total computational complexity for a one time calculation is thus $O(NMp^2 + N^3p^6)$.

During a simulation, as mentioned above, the configuration of the system changes very little between time-steps, and an iterative method can be used to *fix* the forces and torques computed at the previous step. Also, as long as the charge distribution does not change, the multipole expansion of the charges of each molecule need only be rotated to represent the new orientation of the molecule, and there is no need to compute it from scratch. Thus for an update step during a simulation the calculations are as follows:

1. Rotate the multipole expansion of each molecule. $O(p^3)$ per molecule or $O(Np^3)$ in total.

2. Iteratively update all $\mathbf{A}^{(i)}$ and all $\nabla_j\mathbf{A}^{(i)}$. This requires $O(N^2p^3)$ per cycle through the lines of eq 51 as well as per cycle through the lines of eq 53. For all molecules, assuming a maximum of k cycles, the time is $O(kN^3p^3)$.

3. Compute the forces and torques. Same as above $O(Np^2)$ in total.

The total computational complexity for an update step is thus $O(kN^3p^3)$, which is independent of M, and assuming $k \leq p$ is asymptotically faster than a one time computation by at least a factor of p^2. Note that when the charges of a molecule are slightly perturbed, its multipole expansion may need to be recomputed. The update procedure, however, can still be used effectively, since the previous values of $\mathbf{A}^{(i)}$ and $\nabla_j\mathbf{A}^{(i)}$ still constitute good initial guesses.

4.5. Numerical Stability. When the use of high order poles is required to maintain the desired precision of the computations, we run the risk of exceeding the precision bounds of the machine we are using. A simple scaling scheme can be used to significantly alleviate this problem. We define a scaling factor λ to be the average molecular radius of the system we are solving. We then use scaled versions of the multipole expansions in the computation that will be numerically better behaved:

$$\check{\mathrm{E}}_{n,m}^{(i)} = \frac{1}{\lambda^n} E_{n,m}^{(i)}, \qquad \check{\mathrm{H}}_{n,m}^{(i)} = \frac{1}{\lambda^n} H_{n,m}^{(i)} \tag{56}$$

Note that the scaling should be done before the actual construction of these expansions, by multiplying the centered coordinates of each charge by λ^{-1}. Each of the parameters $\delta_n^{(i)}$, defined in eq 20, also needs to be scaled down to $\check{\delta}_n^{(i)} = \lambda^{-2n}\,\delta_n^{(i)}$. This too should be done during the construction of the $\delta_n^{(i)}$ to ensure numerical stability. Finally the translation coefficients $S^{(i,j)}$ need to be scaled up. This is explained in detail in the Appendix A.1.

4.6. Performance Evaluation. We ran a number of experiments using our implementation of the theory, to test how well it performs in realistic scenarios. First we looked at the performance of the Gauss-Seidel linear system solver. We expected the number of iterations to be low because the matrix we need to invert is diagonally dominant. We looked at the solution of configurations of 2, 4, and 6 Barstar molecules similar to those we used in section 3. The error defined in eq 52 was set to 1%. The number of iterations required for the solution of both the effective charge expansion vectors $\mathbf{A}^{(i)}$ and their gradients $\nabla_j\mathbf{A}^{(i)}$ is recorded

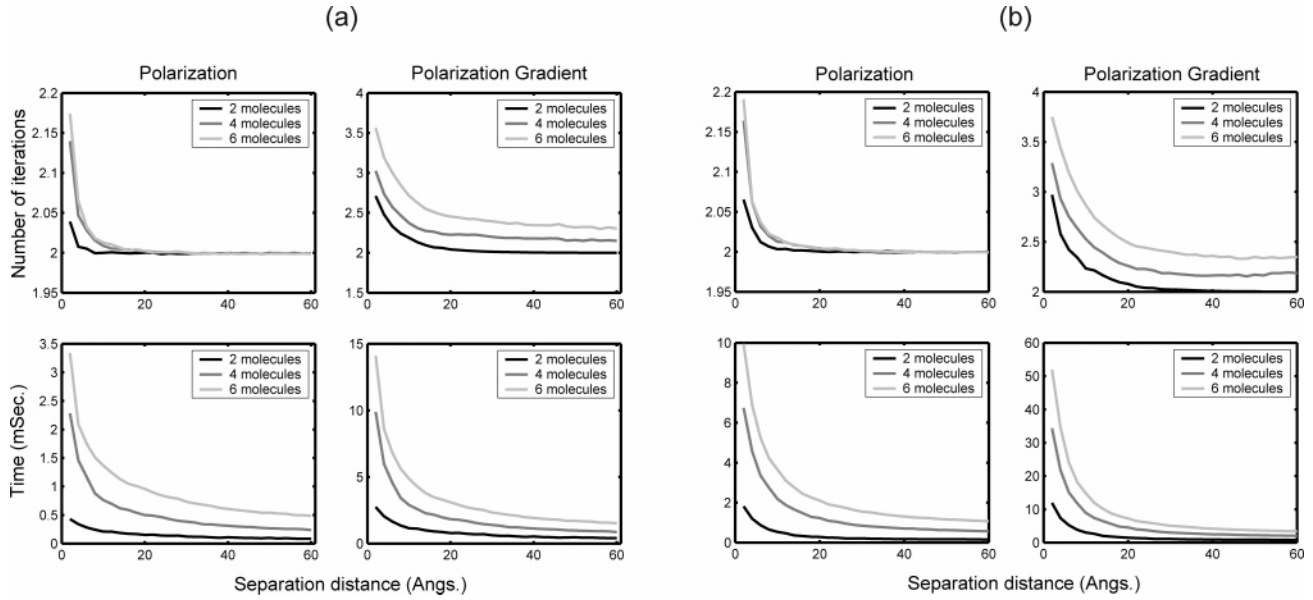


Figure 5. The computation time required to compute the energy and forces for different configurations of (a) Barstar and (b) Barnase molecules as a function of separation distance. The first column shows the number of iterations and time required to compute the polarization effects (solve the linear system of eq 51). The second column shows the number of iterations and time required to compute the polarization effects (solve the linear system of eq 51). The second column shows the number of iterations and time required to compute the partial derivatives of the polarization effects (solve the linear system of eq 53 for a single molecular center). The first row shows the number of Gauss-Seidel iterations, and the second row shows the computation time required to converge to within 1% of the correct solution.

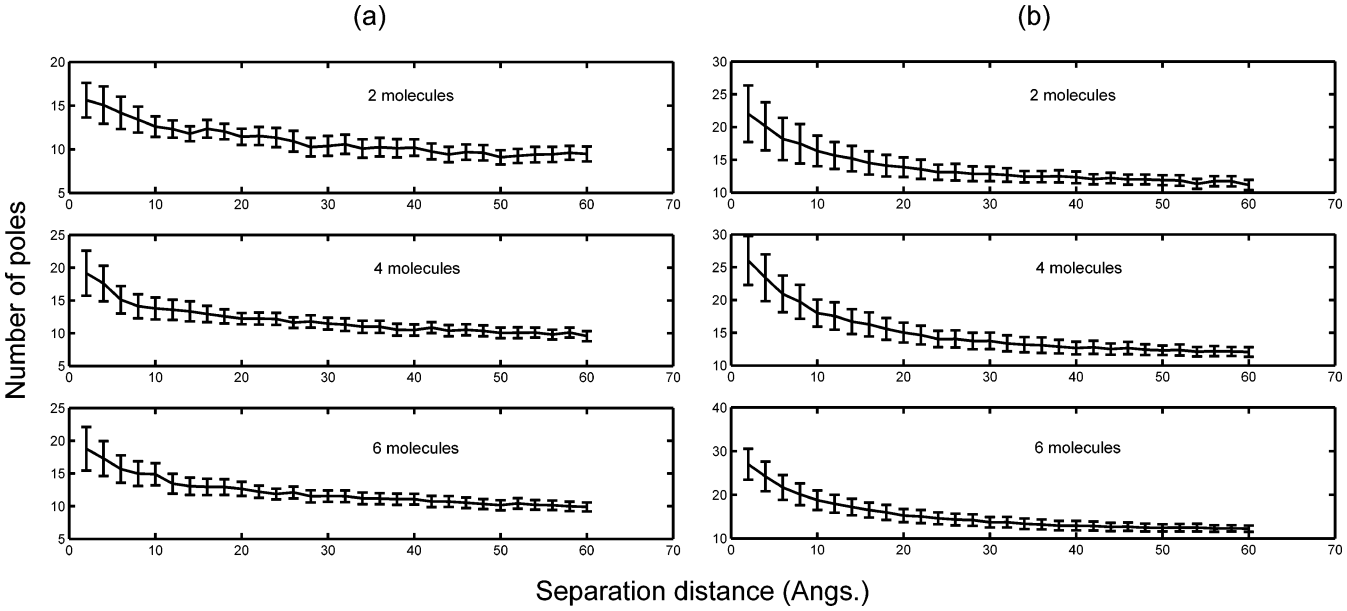


Figure 6. The average and standard deviation of the number of poles in a multipole expansion needed to compute the interaction between two molecules to within 0.1% error as a function of separation distance. Results are shown for three configurations of (a) Barstar and (b) Barnase molecules: two molecules at the top, four in the middle, and six at the bottom graph.

for 5000 different relative random orientations of the tested molecules. In Figure 5a we show the iterations and timing results for a range of separation distances for the molecule Barstar. The number of iterations to compute the $\mathbf{A}^{(i)}$ vectors is only slightly more than 2 and seems to vary as a function of size only for short separations. The computation of the gradient vectors $\nabla_j\mathbf{A}^{(i)}$ requires as many as 3.5 iterations for 6 molecules at short separations, and in general there is a small yet persistent difference between the different configuration sizes across separation distances, where the larger the configuration the more iterations that are required on average. Also included as a reference are the actual average running times on an apple G5 workstation. These running times have been normalized by the size of the configuration. Thus one needs to multiply the reported times for polarization by a factor of N (the size of the configuration) and the reported times for the polarization gradient by a factor of N^2 to get the total computation time for the configuration.

The iteration and timing results for Barnase are in Figure 5b. Comparing the two, one notices that the number of

iterations required for convergence is about the same for both molecules; however, the time per iteration is about 3 times larger for Barnase because higher order expansion was needed to get the same desired level of precision. The number of poles (the maximal expansion order) that are used on average to represent the $\mathbf{A}^{(i)}$ vectors determines the actual size of the computations that are performed. Recall that this number is chosen automatically by the algorithm to maintain the desired error of eq 55. In Figure 6a we show the number of poles required to reach a desired relative error of 0.001 (mean and standard deviation) for different numbers of Barstar molecules, and in Figure 6b we show the same for Barnase.

Computing these systems using numerical methods, such as DelPhi, would take many seconds, and we expect our method to be at least 2−3 orders of magnitude faster. Moreover, for the larger configurations (4 and 6 molecules) a very large grid would be needed to get high levels of accuracy. Numerical methods, in general, are designed to handle complex molecular surfaces and are unable to take advantage of simple spherical geometries, unlike our method. Thus, already from a theoretical perspective, it is always considerably better, in terms of both precision and time, to use our method when spherical geometry is assumed.

5. Conclusion

We have presented a new general analytical solution for computing the screened electrostatic interaction between multiple macromolecules of arbitrarily complex charge distributions, assuming they are well described by spherical low dielectric cavities in a higher dielectric medium in the presence of a Debye−Hückel treatment of salt. Our new formulation builds upon foundations of electrostatic theory of dielectric boundary problems laid down by Kirkwood,[11] Phillies,[17] Hansen and co-workers,[12,16] and McClurg and Zukoski[18] and unites it with the scientific computing community's recent developments in solving for the screened Coulomb (Yukawa) potential[19,20] and solution to the 3-D Helmholtz equation.[21,22] Our formulation offers the further advantage that it is general for arbitrary numbers of macromolecules with arbitrarily complex charge distributions, and the use of straightforward and computationally efficient multipole re-expansion operators to solve the screened Coulomb problem has made our formulation practical as well. In addition we present its algorithmic scaling costs and implementation (with code made available upon request to one of the corresponding authors).

Often a macromolecule in solution is represented by two concentric spheres, an inner sphere to represent the molecule and a concentric layer around it (a 3-D annulus) to represent a charge-free region due to the finite size of ions, a solvent-free region due to the finite size of solvent molecules, or an ionic Stern Layer. See, for example Kirkwood[11] and Hoffman et al.[12] When assuming that this layer has the same permittivity as the interior of the macromolecule our formulation can be used as is. The radius of the molecule is simply extended to include this outer layer as well, and any charges in it are treated as part of the molecule's charge distribution. If, however, a different permittivity is used, one would need to derive new expressions for the electrostatic potential, where now each cavity would have two boundaries, and expressions for the potential would be required for three different areas (inside, external layer, and outside). This has been done before for a single molecule.[11,12] Our approach would be applicable for this case as well; however, the expressions would be much longer and more cumbersome and thus we have chosen not to derive them. We could also model the Stern layer ions explicitly since our formulation allows for arbitrary numbers of spheres of arbitrary radii.

We plan to apply this generally formulated electrostatic model in the future to problems in protein complexation using Brownian dynamics. In typical implementations of protein−protein association using Brownian dynamics, the charge distribution of the protein in isolation is assumed to be fixed throughout the stochastic simulation of proteins in solution.[24,32,33] While this approximation is best for earlier stages of the diffusion reaction, the approximation degrades significantly as the relative separation of the two proteins become smaller, since the accumulated forces at shorter range acting on the charge sites should shift the conformational state of the two proteins, thereby altering their overall charge distribution and shape. This scenario can be further complicated by the effect of the presence of multiple proteins on the pair interactions. Our analytical solution is general to arbitrary numbers of macromolecules, is efficient to compute, and can therefore simultaneously provide on-the-fly updates to changes in charge distributions due to protein conformational changes. We can also change spatial resolutions of charge description as a function of separation distance without compromising the desired accuracy.

While for many applications the simple spherical geometry of the macromolecule is sufficient and reasonable, in some cases it falls short of capturing the desired level of atomic detail. In our future work we plan to explore a number of paths that promise to extend the usability of our formulation to cases where a more accurate description of the influence of shape on the electrostatics near the surface of proteins is required. As we discuss in section 2.2, our solution separates the description of the charge distribution from the description of the geometric and electrostatic properties (dielectric constants and salt screening length). The latter is restricted to the Γ operator that describes the behavior of the dielectric boundary between each molecule and the solution and the Δ operator that describes the polarization properties of the molecule. Thus, instead of using the analytic values for these operators computed for spherical geometry, it may be possible to compute generalized operators for more complex geometries such as unions of spheres, where more than a single expansion center is used to better approximate the electrostatic potential.

Acknowledgment. We would like to thank Rosalind Allen and Jean-Pierre Hansen for useful discussions. We would like to thank Nail Gumerov for help in the implementation of the re-expansion operator. We acknowledge financial support from DOE/LDRD and the National Institutes of Health Grant No. GM76730-01. T.H.G. gratefully acknowledges the support of a Schlumberger Fellowship while on sabbatical in the Theoretical Chemistry Sector at Cambridge University.

Appendix

A.1. Recursive Computation of Re-Expansion Coefficients. The re-expansion coefficients are computed using the decomposition of eq 34. Thus a set of rotation and translation coefficients needs to be computed. Gumerov and Duraiswami[21] propose a recursive method for the re-expansion coefficients for the Helmholtz equation, which we adopt here for the very similar Poisson−Boltzmann equation. Other formulations of the re-expansion operator also exist.[20,22,29] Below we lay out the recursive procedure for computing the rotation and translation coefficients for a re-expansion along the vector $v = [r,\theta,\phi]$.

Rotation Coefficients and Their Derivatives. The rotation coefficients $R_n^{m,s}$ $(0 \le n \le p-1,\ -n \le m,s \le n)$ are used as described in eq 46. The recursive procedure for computing $R_n^{s,m}$ is the following:

1. Set $R_n^{0,s} = Y_{n,-s}(\theta,\phi)$ for $0 \le n \le 2p-1$ and $-n \le s \le n$.

2. For each $0 \le m \le p-1$ starting at $m = 0$, fill all coefficients with $m+2 \le n \le 2p-m-1$ and $-n+1 \le s \le n-1$ using the following rule:

$$R_{n-1}^{m+1,s} = \frac{1}{b_n^m}\left[\frac{1}{2}e^{-i\phi}(1+\cos\theta)b_n^{s-1}R_n^{m,s-1} - \frac{1}{2}e^{i\phi}(1-\cos\theta)b_n^{-s-1}R_n^{m,s+1} + \sin\theta a_n^s R_n^{m,s}\right] \quad (1.1)$$

For negative values of m use $R_n^{m,s} = \overline{R_n^{-m,-s}}$.

The constants used in the recursion are

$$a_n^m = \sqrt{\frac{(n+m+1)(n-m+1)}{(2n+1)(2n+3)}}$$

$$b_n^m = \mathrm{sign}(m)\sqrt{\frac{(n-m-1)(n-m)}{(2n-1)(2n+1)}} \quad (1.2)$$

The partial derivatives of the rotation operator $\partial R_n^{m,s}/\partial\theta$ and $\partial R_n^{m,s}/\partial\phi$ can be computed in a similar fashion. For $\partial R_n^{m,s}/\partial\theta$ the following recursion should be used:

1. For $0 \le n \le 2p-1$ use the formula

$$\frac{\partial R_n^{0,-s}}{\partial\theta} = s\cot\theta Y_{n,s}(\theta,\phi) - \sqrt{(n-s)(n+s+1)}e^{-i\phi}Y_{n,s+1}(\theta,\phi) \quad (1.3)$$

when $s \ge 0$, and then use $\partial R_n^{0,-s}/\partial\theta = \overline{\partial R_n^{0,s}/\partial\theta}$ to complete the positive values of s.

2. For each value of $0 \le m \le p-1$ starting at $m = 0$ fill all coefficients with $m+2 \le n \le 2p-m-1$ and $-n+1 \le s \le n-1$ using the following rule:

$$\frac{\partial R_{n-1}^{m+1,s}}{\partial\theta} = -\frac{1}{b_n^m}\left[\frac{1}{2}e^{i\phi}b_n^{-s-1}\left(\sin\theta R_n^{m,s+1} + (1-\cos\theta)\frac{\partial R_n^{m,s+1}}{\partial\theta}\right) + \frac{1}{2}e^{-i\phi}b_n^{s-1}\left(\sin\theta R_n^{m,s-1} - (1+\cos\theta)\frac{\partial R_n^{m,s-1}}{\partial\theta}\right) - a_n^s\left(\sin\theta\frac{\partial R_n^{m,s}}{\partial\theta} + \cos\theta R_n^{m,s}\right)\right] \quad (1.4)$$

3. For negative values of m use $\partial R_n^{m,s}/\partial\theta = \overline{\partial R_n^{-m,-s}/\partial\theta}$. For $\partial R_n^{m,s}/\partial\phi$ the formula is

$$\frac{\partial R_n^{m,s}}{\partial\phi} = -isR_n^{m,s} \quad (1.5)$$

Translation Coefficients and Their Derivatives. The translation coefficients $S_{n,l}^m$ $(0 \le n,l \le p-1,\ -n \le m \le n)$ are used as in eq 46. The recursive procedure for computing $S_{n,l}^m(r)$ is the following:

1. Set $S_{0,l}^0 = (\lambda/r)^l(\hat{\mathrm{k}}_l(\kappa r)e^{-\kappa r}/r)$ for $0 \le l \le 2p-1$ and use $S_{n,0}^0 = (-1)^n S_{0,n}^0$ for $0 \le n \le 2p-1$.

2. Use the following recursion to compute $S_{n,l}^0$ for $0 \le n \le p-2$ and $n+1 \le l \le 2p-n-2$:

$$S_{n+1,l}^0 = -\frac{1}{\alpha_n^0}[\beta_{l-1}^0 S_{n,l-1}^0 + \beta_{n-1}^0 S_{n-1,l}^0 + \alpha_l^0 S_{n,l+1}^0] \quad (1.6)$$

3. For each $0 \le m \le p-2$ use the following two steps to compute all $S_{n,l}^{m+1}$:

I. For $m \le l \le 2p-2-m$ compute all $S_{m+1,l}^{m+1}$ using

$$S_{m+1,l}^{m+1} = -\frac{1}{\eta_{m+1}^{-m-1}}[\mu_l^{-m-1}S_{m,l-1}^m + \eta_{l+1}^m S_{m,l+1}^m] \quad (1.7)$$

II. For $m \le n \le p-2$ and $n+1 \le l \le 2p-n-2$ use

$$S_{n+1,l}^{m+1} = -\frac{1}{\alpha_n^{m+1}}[\beta_{l-1}^{m+1}S_{l-1}^{m+1} + \beta_{n-1}^{m+1}S_{n-1,l}^{m+1} + \alpha_l^{m+1}S_{n-1,l}^{m+1}] \quad (1.8)$$

4. To complete the set of coefficients two identities can be used:

I. For all $S_{n,l}^m$ where $n > l$ use $S_{n,l}^m = (-1)^{n+l}S_{l,n}^m$.

II. For all $S_{n,l}^m$ where $m < 0$ use $S_{n,l}^{-m} = S_{n,l}^m$.

The constants used in the recursion are the following:

$$\alpha_n^m = \sqrt{(n+m+1)(n-m+1)}$$

$$\beta_n^m = \frac{\lambda^2\kappa^2\alpha_n^m}{(2n+1)(2n+3)}$$

$$\eta_n^m = \mathrm{sign}(m)\sqrt{(n-m-1)(n-m)}$$

$$\mu_n^m = \frac{\lambda^2\kappa^2\eta_n^m}{(2n-1)(2n+1)} \quad (1.9)$$

Recall that λ is the uniform scaling factor that was defined in section 4.5.

The derivatives of the translation coefficients $\partial S_{n,l}^m/\partial r$ are computed using a similar procedure, which differs only in step #1, where the initial values are computed using

$$\frac{\partial S_{0,l}^0}{\partial r} = \left(\frac{\lambda}{r}\right)^l\frac{e^{-\kappa r}}{r^2}(l\hat{\mathrm{k}}_l(\kappa r) - (2l+1)\hat{\mathrm{k}}_{l+1}(\kappa r)) \quad (1.10)$$

A.2. Dealing with Singularities. The method we use to compute the derivatives of the re-expansion operator breaks

as the direction of the re-expansion vector $v = [r,\theta,\phi]$ approaches the *z*-axis (either the positive or negative directions). In this case the angle $\theta \rightarrow 0,\pi$ and $\sin\theta \rightarrow 0$. This is not due to a true singularity of the re-expansion operator but is rather due to the fact that the gradient of this operator is computed first in Spherical coordinates and then converted to Cartesian coordinates. This singularity affects the derivatives with respect to θ and ϕ . Note that when $\theta = 0,\pi$, the coordinate ϕ becomes meaningless and we can arbitrarily set it to 0.

For the case of $\partial R_n^{m,s}/\partial\theta$, when $\sin\theta \rightarrow 0$, eq 1.3 reduces to

$$\frac{\partial R_n^{0,-s}}{\partial\theta} = s\cot\theta Y_{n,s}(\theta,\phi) \tag{1.11}$$

which is nonzero only when $s = \pm 1$ because then the sin θ term in both $Y_{n,s}(\theta,\phi)$ and $\cot\theta$ cancels out. Examining carefully the recursion rule in eq 1.4 tells us that only the $\partial R_n^{m,s}/\partial\theta$ coefficients where $m = s \pm 1$ will be nonzero and thus the calculation and application of this operator in the singular case is significantly simplified (only $O(p^2)$ operations).

For the case of $\partial R_n^{m,s}/\partial\phi$ we also need to take into account the term $\partial\phi/\partial y = \cos\phi/(r\sin\theta)$ in the operator that converts derivatives with respect to spherical coordinates to derivatives with respect to Cartesian coordinates (see eq 48) because of its dependence on $1/\sin\theta$. Similarly to the case of the $\partial R_n^{m,s}/\partial\theta$ coefficients, here too, because we end up multiplying each $\partial R_n^{m,s}/\partial\phi$ by $\partial\phi/\partial y$ we get a cancellation of $\sin\theta$ in both. Recall that the $\partial R_n^{m,s}/\partial\phi$ coefficients have a simple dependence on the $R_n^{m,s}$ coefficients (see eq 1.5). We thus define in this case new coefficients $\partial R_n^{m,s}/\partial y = (\partial R_n^{m,s}/\partial\phi\ \partial\phi/\partial y)$, noting that when $\sin\theta = 0$, $\partial R_n^{m,s}/\partial y$ depends only on $\partial R_n^{m,s}/\partial\phi$. Here too we end up with only the $\partial R_n^{m,s}/\partial y$ coefficients, where $s = \pm 1$, being nonzero.

A.3. Computing the Adapted MSBFs. Following Kirkwood, a simple recursive formula allows for the generation of all $\hat{k}_n(z)$

$$\hat{k}_{n+1}(z) = \hat{k}_n(z) + \frac{z^2\hat{k}_{n-1}(z)}{(2n+1)(2n-1)} \tag{1.12}$$

where the starting values are $\hat{k}_0(z) = 1$ and $\hat{k}_1(z) = 1 + z$. The derivatives of this function, which are used in the derivations of section 2, can be generated using the relation

$$\frac{d\hat{k}_n(z)}{dz} = \frac{(2n+1+z)\hat{k}_n(z) - (2n+1)\hat{k}_{n+1}(z)}{z} \tag{1.13}$$

We have also derived a recursive formula for the generation of all $\hat{i}_n(z)$:

$$\hat{i}_{n+1}(z) = (2n+1)(2n+3)\frac{\hat{i}_{n-1}(z) - \hat{i}_n(z)}{z^2} \tag{1.14}$$

where the starting values are $\hat{i}_0(z) = \sinh(z)/z$ and $\hat{i}_1(z) = (3/z^2)(\cosh(z) - \hat{i}_0(z))$. While this formula is correct and useful in theoretical derivations, its usefulness in actual finite precision computation is limited. Instead we recommend the following formula

$$\hat{i}_n(z) = 1 + \sum_{j=1}^{L} t_j^n(z^2/2)$$

$$t_j^n(y) = \frac{1}{j}\, t_{j-1}^n(y)\frac{y}{2n+2j+3} \tag{1.15}$$

where the starting value is $t_0^n(y) = y/(2n+3)$. The summation in the formula should be continued until the change falls below the desired precision (we used $L = 20$). The derivatives can be generated using the relation

$$\frac{d\hat{i}_n(z)}{dz} = \frac{z\hat{i}_{n+1}(z)}{2n+3} \tag{1.16}$$

Finally a relation between the two kinds of the adapted MSBFs that is useful in simplifying complex expressions is

$$\hat{i}_n(z)\hat{k}_{n+1}(z) + \hat{k}_n(z)\hat{i}_{n+1}(z)\frac{z^2}{(2n+1)(2n+3)} = e^z \tag{1.17}$$

References

(1) Baker, N. A. In *Methods in Enzymology*; Brand, L., Johnson, M. L., Eds.; Academic Press: 2004; Vol. 383, pp 94−118.

(2) Honig, B.; Nicholls, A. *Science* **1995**, *268*, 1144−1149.

(3) Davis, M. E.; Madura, J. D.; Luty, B. A.; McCammon, J. A. *Comput. Phys. Commun.* **1991**, *62*, 187−197.

(4) Madura, J. D.; Briggs, J. M.; Wade, R. C.; Davis, M. E.; Luty, B. A.; Ilin, A.; Antosiewicz, J.; Gilson, M. K.; Bagheri, B.; Scott, L. R.; McCammon, J. A. *Comput. Phys. Commun.* **1995**, *91*, 57−95.

(5) Nicholls, A.; Honig, B. *J. Comput. Chem.* **1991**, *12*, 435−445.

(6) Rocchia, W.; Alexov, E.; Honig, B. *J. Phys. Chem.* **2001**, *105*, 6507−6514.

(7) Beard, D. A.; Schlick, T. *Biopolymers* **2001**, *58*, 106−115.

(8) Bordner, A. J.; Huber, G. A. *J. Comput. Chem.* **2003**, *24*, 353−367.

(9) Boschitsch, A. H.; Fenley, M. O.; Zhou, H. X. *J. Phys. Chem.* **2002**, *106*, 2741−2754.

(10) Gabdoulline, R. R.; Wade, R. C. *J. Phys. Chem.* **1996**, *100*, 3868−3878.

(11) Kirkwood, J. G. *J. Chem. Phys.* **1934**, *2*, 351−361.

(12) Hoffmann, N.; Likos, C. N.; Hansen, J. P. *Mol. Phys.* **2004**, *102*, 857−867.

(13) Glendinning, A. B.; Russel, W. B. *J. Colloid Interface Sci.* **1983**, *93*, 95−104.

(14) Verwey, E. J. W.; Overbeek, J. T. G. *Theory of the stability of lyophobic colloids*; Dover Publications: Mineola, NY, 1999.

(15) Sader, J. E.; Lenhoff, A. M. *J. Colloid Interface Sci.* **1998**, *201*, 233−243.

(16) Allen, R.; Hansen, J. P. *J. Phys.: Condens. Matter* **2002**, *14*, 11981−11997.

(17) Phillies, G. D. *J. Chem. Phys.* **1974**, *60*, 2721−2731.

(18) McClurg, R. B.; Zukoski, C. F. *J. Colloid Interface Sci.* **1998**, *208*, 529−542.

(19) Boschitsch, A. H.; Fenley, M. O.; Olson, W. K. *J. Comput. Phys.* **1999**, *151*, 212−241.

(20) Greengard, L. F.; Huang, J. F. *J. Comput. Phys.* **2002**, *180*, 642−658.

(21) Gumerov, N. A.; Duraiswami, R. *Siam J. Sci. Comput.* **2003**, *25*, 1344−1381.

(22) Rokhlin, V. *Appl. Comput. Harmonic Anal.* **1993**, *1*.

(23) Arfken, G. B.; Weber, H.-J. *Mathematical methods for physicists*, 4th ed.; Academic Press: San Diego, CA, 1995.

(24) Gabdoulline, R. R.; Wade, R. C. *J. Mol. Biol.* **2001**, *306*, 1139−1155.

(25) Buckle, A. M.; Schreiber, G.; Fersht, A. R. *Biochemistry* **1994**, *33*, 3, 8878−8889.

(26) Jorgensen, W. L.; Tiradorives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657−1666.

(27) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.

(28) Danos, M.; Maximon, L. C. *J. Math. Phys.* **1965**, *6*, 766−778.

(29) Epton, M. A.; Dembart, B. **1995**, *16*, 865−897.

(30) Golub, G. H.; Van Loan, C. F. *Matrix computations*, 3rd ed.; Johns Hopkins University Press: Baltimore, MD, 1996.

(31) Greengard, L.; Rokhlin, V. *J. Comput. Phys.* **1987**, *73*, 325−348.

(32) Gabdoulline, R. R.; Wade, R. C. *Methods* **1998**, *14*, 329−341.

(33) Gabdoulline, R. R.; Wade, R. C. *Curr. Opin. Struct. Biol.* **2002**, *12*, 204−213.

CT050263P

JCTC Journal of Chemical Theory and Computation

Parametrization and Validation of Coarse Grained Force-Fields Derived from ab Initio Calculations

Giacomo Prampolini*

Dipartimento di Chimica e Chimica Industriale, Università di Pisa, via Risorgimento 35, I-56126 Pisa, Italy

Received December 23, 2005

Abstract: A novel multisite interaction potential, suitable for computer simulations of complex materials as liquid crystals or polymers, is proposed and parametrized. Its validation is achieved through Monte Carlo numerical experiments at constant temperature and pressure, performed on the *p*-*n*-phenyls series and a typical mesogenic molecule (5CB). The model is constructed by connecting an array of anisotropic Gay-Berne sites and a collection of isotropic Lennard-Jones sites. The former mimics the rigid planar six-membered rings of the molecule, while the latter represents the flexible chain, if present. Such intermolecular potential, coupled with an intramolecular part to account for molecular flexibility, is parametrized from ab initio information only, obtained through the recently proposed Fragmentation-Reconstruction Method (FRM). Computer simulations are performed on all systems by exploring phase behavior at several temperatures and by comparing the resulting thermodynamic and structural properties with the relevant experimental data. Despite the simplicity of the present models, the good agreement with the experimental measures suggests the possibility of adopting such hybrid potentials for those systems with a large number of atoms, where high computational cost does not allow the use of more accurate atomistic potentials.

1. Introduction

In the past decades, computer simulations have considerably aided the study of advanced materials such as liquid crystals or polymers.[1,2] Unlike simple liquids, both the wide range of length and time scales that characterizes the dynamics of such complex substances, and the large dimensions of their forming molecules have suggested the adoption of simplified or coarse grained models.[2−5] In liquid crystal field, for instance, many anisotropic single site interaction models have been reported in the literature.[6−9] Among them, the most successfully employed is certainly the Gay-Berne (GB) potential,[3,9−11] which has much contributed to clarify the basic features responsible for mesogenic behavior. The main lack of these single-site models is the absence of molecular flexibility, which cannot be neglected if one aims to accurately reproduce experimental behavior. On the other hand, the use of atomistic force-fields becomes rapidly unfeasible with the growth of the molecular dimensions. Even standard united atom (UA) approaches, by which groups of methyl and methylene atoms are represented with a single Lennard-Jones (LJ) interaction site, might not solve the problem if one aims to sensibly increase the speed of the calculation or the quality of the statistics.

For these reasons, several models have been proposed in the past few years,[1,2,12−20] where a reasonable computational cost is reached through a reduction of the number of interaction sites, without losing the basic features of the molecular interactions. For instance, following the UA approach, LaPenna and co-workers[13] have first employed a coarse grained model potential, by coupling one GB anisotropic potential (representing the rigid molecular core of a typical mesogenic substance) with a collection of standard LJ sites describing the flexible aliphatic chain. Such a hybrid GB/LJ model has then been refined and successfully adopted in simulations of both polymers[21,22] and liquid crystals.[14,15,17,20] Despite their simplicity, in fact, hybrid models

* Corresponding author e-mail: giacomo@dcci.unipi.it.

10.1021/ct050328o CCC: $33.50
Published on Web 03/28/2006

have proven valuable to study the basic structure−property relationships subtending the bulk behavior. Nevertheless, it should be mentioned that most of these force-fields[13−15,21,22] were parametrized in a semiempirical manner, being their aim to represent the features of a general (liquid−crystalline or polymeric) bulk phase, rather than reproduce with good accuracy the experimental properties of a specific substance. Even in the most recent works,[17,20] the proposed hybrid GB/LJ model is parametrized on an empirical potential derived from the widely used OPLS[23] atomistic force-field. This route of parametrization might be inappropriate when the interest focuses on a specific substance, with a well defined molecular composition. To reproduce the thermodynamic, structural, and dynamic properties of a real advanced material, one needs to employ model potential functions able to accurately describe the interactions between its forming molecules, taking into account the chemical details that characterize the microscopic structure.

The Fragmentation Reconstruction Method (FRM), recently proposed by our group,[18,19,24−26] allows to compute, through accurate ab initio calculations, the interaction potential energy surface (PES) of dimers of large molecular dimensions. Once the FRM-PES has been computed, it can be used to parametrize the intermolecular potential model through a fitting procedure, at several possible degrees of accuracy and complexity. In this work, a model potential based on representing the phenyl groups in each molecule with single interaction sites is parametrized on the base of previously computed FRM data[18,19,24,25] and employed in MC computer simulations.

Two test cases have been chosen to validate the aforementioned model and its parametrization procedure: the *p*-*n*-phenyls series ($n = 2-5$) and 4-*n*-pentyl 4′-cyanobiphenyl (5CB), a typical mesogenic molecule. The series of *p*-polyphenyls certainly exhibits fascinating properties both in the field of polymers and liquid crystals. In the former, the applications of poly(*p*-phenylene) vary from ribbons and fibers to solid-state lubricants, and the polymer itself has been the object of several reviews.[27] In the latter, *p*-quinquephenyl and *p*-sexiphenyl show nematic and smectic phases, respectively. Moreover, the torsional potential between the phenyl rings of all the series has been recently modeled by our group[28] through accurate DFT calculations. The coupling of such an intramolecular term to the intermolecular hybrid potential should refine the results of preliminary runs already reported for these systems.[19] On the other hand, the 5CB molecule has been the object of several simulation studies,[12,16,17,20,26,29−34] being a sort of prototype nematogen. Empirical atomistic simulation models[26,30−34] have been proven capable to reproduce with a certain accuracy most of the static and dynamic properties of such liquid−crystal nematic phase. Moreover, an atomistic force-field, derived by our group through the FRM approach,[25] has been successfully employed to reproduce the bulk behavior of crystalline, nematic, and isotropic 5CB, yielding a phase diagram in good agreement with the experimental trends. However, since atomistic modeling cannot be extended to much larger molecules, a hybrid potential parametrization of 5CB PES is tempted in this work and compared with the results achieved by the atomistic model, so to evaluate the possibility of its use for computationally more expensive systems.

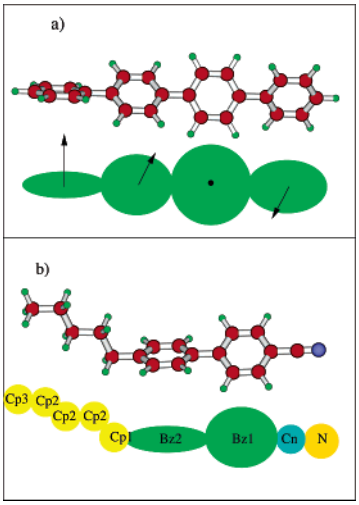


Figure 1. (a) The adopted model is shown for *p*-quaterphenyl. All the series members have been modeled in analogous manner. The black arrows on each disk represent the added quadrupoles. (b) The hybrid GB/LJ modeling is reported for 5CB. All UA interaction sites have been represented with a LJ potential, except the phenyl rings, which have been modeled with a generalized GB potential (green disks). For both models, all parameters are listed in Tables 1−2 and 4, respectively.

The paper is organized as follows. Section 2 gives details on the hybrid potential and on the employed MC technique. In the first part of section 3, the results of the parametrization of both *p*-*n*-phenyls and 5CB are discussed. In the second part, the simulation results are reported and compared with the relevant experimental data. A comparison with the results of the more detailed atomistic model[25] is also discussed. Finally, section 4 collects the main conclusions.

2. Computational Details

2.1. Hybrid Force-Fields and Fitting Procedure. *p*-Polyphenyls. All the considered *p*-*n*-phenyls ($n = 2-5$) have been modeled by representing the phenyl rings of each molecule with oblate Gay-Berne ellipsoids,[9,10] plus a linear quadrupole (Q) along the orientational axis of the disk (i.e. perpendicular to the phenyl plane). The upper panel of Figure 1 shows the adopted model for the third member of the series, namely *p*-quaterphenyl. The word "hybrid" here stands to indicate the use, in modeling one molecule, of several bonded GB sites, normally used as a single site potential.[3]

In this hybrid model, the intermolecular potential U_{AB}^{inter} between two interacting sites A and B is thus expressed by a sum of two terms, i.e.

$$U_{AB}^{inter} = U_{AB}^{gGB} + U_{AB}^{Q} \quad (1)$$

With regards to the first term of eq 1, a pair of nonequivalent GB interaction sites can be defined[9,10,19] by the specification of two parameter arrays, namely $\mathbf{P}_k = [\sigma_k^0,\sigma_k^{ee},\sigma_k^{ss},\epsilon_k^0,\epsilon_k^{ee},\epsilon_k^{ss},\mu_k,\nu_k,\xi_k]$, with $k =$ A, B.

Their interaction energy, the generalized Gay-Berne potential (U_{AB}^{gGB}),[10] depends on the distance vector $\hat{r}_{AB}$ and on the disks orientational axes $\hat{u}_A$ and $\hat{u}_B$

$$U_{AB}^{gGB} = E_{AB} \times [(S_{AB})^{12} - (S_{AB})^{6}] \tag{2}$$

where

$$S_{AB} = \frac{\sigma_{AB}^{0}\xi_{AB}}{r_{AB} - \sigma_{AB}(\hat{u}_A, \hat{u}_B, \hat{r}_{AB}) + \sigma_{AB}^{0}\xi_{AB}} \tag{3}$$

and

$$E_{AB} = 4\epsilon_{AB}^{0}[\epsilon_{AB}^{(1)}(\hat{u}_A, \hat{u}_B)]^{\nu_{AB}}[\epsilon_{AB}^{(2)}(\hat{u}_A, \hat{u}_B, \hat{r}_{AB})]^{\mu_{AB}} \tag{4}$$

In eqs 3 and 4, the functions $\sigma_{AB}(\hat{u}_A, \hat{u}_B, \hat{r}_{AB})$, $\epsilon_{AB}^{(1)}(\hat{u}_A, \hat{u}_B)$, and $\epsilon_{AB}^{(2)}(\hat{u}_A, \hat{u}_B, \hat{r}_{AB})$ are defined in terms of the quantities χ_{AB}, α_{AB}, χ'_{AB}, and α'_{AB} through eqs 5−7, i.e.

$$\sigma_{AB}(\hat{u}_A, \hat{u}_B, \hat{r}_{AB}) = \sigma^{0}_{AB}\left[1 - \chi_{AB}\frac{(\alpha_{AB}^{2}(\hat{r}_{AB}\cdot\hat{u}_A)^2 + \alpha^{-2}_{AB}(\hat{r}_{AB}\cdot\hat{u}_B)^2 - 2\chi_{AB}(\hat{r}_{AB}\cdot\hat{u}_A)(\hat{r}_{AB}\cdot\hat{u}_B)}{1 - \chi_{AB}^{2}(\hat{u}_A\cdot\hat{u}_B)^2}\right]^{-1/2} \tag{5}$$

$$\epsilon_{AB}^{(1)}(\hat{u}_A, \hat{u}_B) = [1 - \chi_{AB}^{2}(\hat{u}_A\cdot\hat{u}_B)^2]^{-1/2} \tag{6}$$

and

$$\epsilon_{AB}^{(2)}(\hat{u}_A, \hat{u}_B, \hat{r}_{AB}) = 1 - \chi'_{AB}\left[\frac{(\alpha'^{2}_{AB}(\hat{r}_{AB}\cdot\hat{u}_A)^2 + \alpha'^{-2}_{AB}(\hat{r}_{AB}\cdot\hat{u}_B)^2 - 2\chi'_{AB}(\hat{r}_{AB}\cdot\hat{u}_A)(\hat{r}_{AB}\cdot\hat{u}_B)}{1 - \chi'^{2}_{AB}(\hat{u}_A\cdot\hat{u}_B)^2}\right] \tag{7}$$

In ref 10, the authors suggest a mixing rule to calculate, from the aforementioned sets $\mathbf{P}_A$ and $\mathbf{P}_B$, the combinations of the quantities χ_{AB} and α_{AB}, as appearing in eqs 5−7, that is

$$\chi_{AB}\alpha_{AB}^{2} = \frac{(\sigma_A^{ee})^2 - (\sigma_A^{ss})^2}{(\sigma_A^{ee})^2 + (\sigma_B^{ss})^2}; \quad \chi_{AB}\alpha_{AB}^{-2} = \frac{(\sigma_B^{ee})^2 - (\sigma_B^{ss})^2}{(\sigma_B^{ee})^2 + (\sigma_A^{ss})^2} \tag{8}$$

$$\chi_{AB}^{2} = \frac{[(\sigma_A^{ee})^2 - (\sigma_A^{ss})^2][(\sigma_B^{ee})^2 - (\sigma_B^{ss})^2]}{[(\sigma_B^{ee})^2 + (\sigma_A^{ss})^2][(\sigma_A^{ee})^2 + (\sigma_B^{ss})^2]} \tag{9}$$

To obtain the other quantities (e.g. ϵ^{0}_{AB} or μ_{AB}), since no combination rule of this kind was found in the literature, a Lorentz−Berthelot approach has been followed where possible, thus defining

$$\epsilon^{0}_{AB} = (\epsilon^{0}_{A}\epsilon^{0}_{B})^{1/2}; \quad \sigma^{0}_{AB} = \frac{1}{2}(\sigma^{0}_{A} + \sigma^{0}_{B}) \tag{10}$$

$$\mu_{AB} = \frac{1}{2}(\mu_A + \mu_B); \quad \nu_{AB} = \frac{1}{2}(\nu_A + \nu_B); \quad \xi_{AB} = (\xi_A\xi_B)^{1/2} \tag{11}$$

The three remaining quantities, namely $\chi'_{AB}\alpha'^{2}_{AB}$, $\chi_{AB}\alpha'^{-2}_{AB}$, and χ'_{AB}, were here derived by comparing the standard GB expressions[9] for χ' with eqs 8 and 9, found in the literature[10] for $\chi\alpha^2$, $\chi\alpha^{-2}$, and χ^2. These considerations have lead to the definitions

$$\chi'_{AB}\alpha'^{2}_{AB} = \frac{(\epsilon_A^{ss})^{1/\mu_A} - (\epsilon_A^{ee})^{1/\mu_A}}{(\epsilon_A^{ss})^{1/\mu_A} + (\epsilon_B^{ee})^{1/\mu_B}}; \quad \chi'_{AB}\alpha'^{-2}_{AB} = \frac{(\epsilon_B^{ss})^{1/\mu_B} - (\epsilon_B^{ee})^{1/\mu_B}}{(\epsilon_B^{ss})^{1/\mu_B} + (\epsilon_A^{ee})^{1/\mu_A}}; \tag{12}$$

$$\chi'_{AB} = \frac{[(\epsilon_A^{ss})^{1/\mu_A} - (\epsilon_A^{ee})^{1/\mu_A}][(\epsilon_B^{ss})^{1/\mu_B} - (\epsilon_B^{ee})^{1/\mu_B}]}{[(\epsilon_B^{ss})^{1/\mu_B} + (\epsilon_A^{ee})^{1/\mu_A}][(\epsilon_A^{ss})^{1/\mu_A} + (\epsilon_B^{ee})^{1/\mu_B}]} \tag{13}$$

Equations 12 and 13 correctly reproduce the standard GB interaction in the limiting case A = B. Furthermore, they also can be used to model the interaction between an anisotropic and a spherical site, i.e. a gGB-LJ pair. In fact, this has been done in the 5CB hybrid model, which will be further discussed.

The second term of eq 1, U_{AB}^{Q}, is a quadrupole−quadrupole contribution,[3,35] whose asymptotic expression *i*

$$U_{AB}^{Q} = \frac{3}{4}Q_AQ_B[1 + 2(\hat{u}_A\cdot\hat{u}_B)^2 - 5(\hat{u}_A\cdot\hat{r}_{AB})^2 - 5(\hat{u}_B\cdot\hat{r}_{AB})^2 - 20(\hat{u}_A\cdot\hat{u}_B)(\hat{u}_A\cdot\hat{r}_{AB})(\hat{u}_B\cdot\hat{r}_{AB}) + 35(\hat{u}_A\cdot\hat{r}_{AB})^2(\hat{u}_B\cdot\hat{r}_{AB})^2]/r_{AB}^{5} \tag{14}$$

where the orientational $\hat{u}_i$ (i = A, B) and intermolecular vectors $\hat{r}_{AB}$ have the usual meaning, and Q_i (i = A, B) is the quadrupole value of the ith interacting site. The model intermolecular potential E^{inter}, of a couple of molecules i and j, can be expressed by

$$E^{inter} = \sum_{A}^{N_i}\sum_{B}^{N_j}U_{AB}^{inter} \tag{15}$$

where U_{AB}^{inter} is given in eq 1, and N_i and N_j are the number of interaction sites of molecules i and j, respectively.

The gGB parameter arrays **P** as well as the molecular quadrupoles Q of all interaction sites were obtained by a least-squares fitting procedure, using as reference the dimer interaction PES (E^{FRM}) of biphenyl, *p*-terphenyl, *p*-quaterphenyl, and *p*-quinquephenyl. The latter has been obtained[18,19] from ab initio calculations through the FRM approach.[24,25] The fitting was carried out by minimizing the integral **I**

$$\mathbf{I} = \sum_{k}^{N_g} w_k[E^{FRM}(\boldsymbol{R}_k, \Omega_k) - E^{inter}(\boldsymbol{R}_k, \Omega_k)]^2 \tag{16}$$

where $\boldsymbol{R}_k$ and Ω_k are, respectively, the intermolecular vector and the reciprocal orientation vector of the dimer in the kth geometry, and N_g is the number of employed geometries. Finally, w_k is a Boltzmann-like weighting factor which takes the form

$$w_k = e^{-\alpha E^{FRM}_k} \tag{17}$$

The molecular flexibility is taken into account by making each gGB disk able to rotate around the *p*-*n*-phenyl long axis. Such rotations are driven by an intramolecular torsional potential (E^{intra}) expressed by

$$E^{intra} = \sum_{i}^{n-1}[\sum_{j}^{M_i}C_{ij}\cos(m_{ij}\phi_i)] \tag{18}$$

where n is the number of rings, M_i is the number of cosine functions used for ring i, and C_{ij} and m_{ij} the potential parameters. The angle ϕ_i is the torsional dihedral between two neighboring rings (namely i and $i + 1$). For all the members of the series, this intramolecular potential has been parametrized in a previous work[28] by fitting the torsional energy profiles arising from accurate DFT calculations. The C_{ij} and m_{ij} parameters are reported in Tables 2, 4, and 6 of ref 28.

5CB. The model adopted for the 5CB molecule is shown in the lower panel of Figure 1. The 5CB's biphenyl core has been represented with two gGB oblate ellipsoids, whereas the N and C atoms (labeled N and C_n, see Figure 1) of the cyano group have each been modeled with a LJ potential. Following a standard UA description, all methyl and methylene groups of the lateral aliphatic chain have been also modeled with a single spherical LJ site. In this case, the term "hybrid" indicates that the model is composed by both isotropic (LJ) and anisotropic (gGB) interaction sites. Furthermore, a point charge was added to every site, so that the intermolecular model potential of two molecules i and j can be expressed by

$$E^{\text{inter}} = \sum_{A}^{N_i^{\text{gGB}}} \sum_{B}^{N_j^{\text{gGB}}} U_{\text{AB}}^{\text{gGB}} + \sum_{A}^{N_i^{\text{LJ}}} \sum_{B}^{N_j^{\text{gGB}}} U_{\text{AB}}^{\text{gGB/LJ}} + \sum_{A}^{N_i^{\text{gGB}}} \sum_{B}^{N_j^{\text{LJ}}} U_{\text{AB}}^{\text{gGB/LJ}} + \sum_{A}^{N_i^{\text{LJ}}} \sum_{B}^{N_j^{\text{LJ}}} U_{\text{AB}}^{\text{LJ}} + \sum_{A}^{N_i^{q}} \sum_{B}^{N_j^{q}} U_{\text{AB}}^{\text{Coul}} \quad (19)$$

where N_i^{LJ}, N_i^{GB}, and N_i^{q} are the number of LJ spheres, gGB ellipsoids, and point charges of molecule i.

The interaction energy between the gGB ellipsoids, $U_{\text{AB}}^{\text{gGB}}$, has the same expression given for the polyphenyls in eq 2, while the LJ potential energy, $U_{\text{AB}}^{\text{LJ}}$, has been slightly modified by the introduction of an additional parameter ξ:

$$U_{\text{AB}}^{\text{LJ}} = 4\epsilon^0{}_{\text{AB}}\left[\left(\frac{\xi_{\text{AB}}\sigma^0{}_{\text{AB}}}{r_{\text{AB}} + \sigma^0{}_{\text{AB}}(\xi_{\text{AB}} - 1)}\right)^{12} - \left(\frac{\xi_{\text{AB}}\sigma^0{}_{\text{AB}}}{r_{\text{AB}} + \sigma^0{}_{\text{AB}}(\xi_{\text{AB}} - 1)}\right)^{6}\right] \quad (20)$$

This has been done both to increase the flexibility of the function in the fitting procedure and to make it consistent with the expression of the gGB potential, $U_{\text{AB}}^{\text{GB}}$, given in eq 2. As done for the polyphenyl series, the three quantities $\sigma^0{}_{\text{AB}}$, $\epsilon^0{}_{\text{AB}}$, and ξ_{AB} can be computed from the parameters specifying each site (i.e. $\sigma^0{}_i$, $\epsilon^0{}_i$, and ξ_i, with i = A, B) through the following mixing rules

$$\epsilon^0{}_{\text{AB}} = \sqrt{\epsilon^0{}_{\text{A}}\epsilon^0{}_{\text{B}}};\ \sigma^0{}_{\text{AB}} = \frac{\sigma^0{}_{\text{A}} + \sigma^0{}_{\text{B}}}{2};\ \xi_{\text{AB}} = \sqrt{\xi_{\text{A}}\xi_{\text{B}}} \quad (21)$$

The hybrid interaction term $U_{\text{AB}}^{\text{gGB/LJ}}$, between an anisotropic gGB site A (B) and a spherical LJ one B (A), can also be described, as previously stated, by eq 2. Indeed, due to its spherical symmetry, the following relations stand for the ith (i = A, B) LJ site:

$$\sigma_i^0 = \sigma_i^{\text{ee}} = \sigma_i^{\text{ss}};\ \epsilon_i^0 = \epsilon_i^{\text{ee}} = \epsilon_i^{\text{ss}};\ \mu_i = \nu_i = 1 \quad (22)$$

In eqs 5−7, defining σ_{AB}, $\epsilon_{\text{AB}}^{(1)}$ and $\epsilon_{\text{AB}}^{(2)}$, this implies the annealing of the coefficients of the terms formally depending on the orientational vector $\hat{u}_i$ of the isotropic LJ site. For instance, in the interaction between a LJ site A and a gGB one B, eq 5 defining σ_{AB} simply reduces to

$$\sigma_{\text{AB}}(\hat{u}_{\text{B}}, \hat{r}_{\text{AB}}) = \sigma^0{}_{\text{AB}}[1 - \chi_{\text{AB}}\alpha_{\text{AB}}^{-2}(\hat{r}_{\text{AB}}\hat{u}_{\text{B}})^2]^{-1/2}$$

As one can see, this term correctly depends only on the distance vector between the isotropic (A) and the anisotropic (B) interaction site, $\hat{r}_{\text{AB}}$, and on the anisotropic site orientation, $\hat{u}_{\text{B}}$. Finally, the point charges interaction term $U_{\text{AB}}^{\text{Coul}}$ has been computed by a standard Coulomb potential.

The parametrization of the 5CB hybrid intermolecular potential has been carried out with the same procedure described, for the oligophenyl series, by eq 16. The reference ab initio PES for the dimer was already computed through the FRM approach and was recently[25] employed in an atomistic modeling of 5CB. The molecular flexibility of the 5CB molecule is certainly more complex than the polyphenyls one. Its description has been carried out by the intramolecular term E^{intra}, which results in a sum of stretching, bending, torsional, and intramolecular LJ contributions:

$$E^{\text{intra}} = E^{\text{stretch}} + E^{\text{bend}} + E^{\text{tors}} + E^{\text{LJintra}} \quad (23)$$

For all the terms in the above sum, the standard AMBER[36] expressions have been adopted. The parametrization of such potential terms for 5CB has been recently made by our group on the base of accurate quantum mechanical calculations,[26] and successfully validated through lengthy molecular dynamics simulations,[26,34] performed with the aforementioned atomistic modeling. All parameters of stretching, bending, and torsional intramolecular potentials can be found in Tables 1−5 of ref 26.

2.2. Computer Simulations. All simulation runs were carried out with the Monte Carlo method[37] (MC). To study phase transitions, the isothermal isobaric ensemble (NPT) has been preferred to the canonical one (NVT), although a bit computationally more expensive. In fact, MC NPT techniques, allowing fluctuations of the shape and the volume of the simulation box, favor the achievement of the natural structure of the system. Furthermore, the NPT ensemble is the closest to experimental conditions, since real experiments are usually performed at constant pressure.

The short-range intermolecular interactions have been truncated at R_c = 10 Å, employing the energy standard correction.[38] In the 5CB system, charge−charge long-range interactions have been treated with the Ewald method,[39] using a convergence parameter α of $5.36/2R_c$. All MC runs have been performed according to the usual rules of the NPT scheme, using systems of 600 molecules for the oligophenyl series and 192 for the 5CB simulations. The latter number has been chosen both for computational convenience and for a better comparison with the results of the previous atomistic model.[25] During the runs a molecule was selected at random, and trial displacements of its center of mass and inertia axes were performed. To sample the intramolecular

conformational space, a randomly selected gGB disk could also be rotated around the molecular long axis of both oligophenyls and 5CB. Moreover, for 5CB, attempts to move the chain's interaction sites were also performed, so to alter their stretching, bending, and torsional coordinates. Finally, the shape and the volume of the computational box have been changed during simulations, by attempting to vary a randomly selected edge of the box. With the aim of preserving the detailed balance condition, all the aforementioned trial moves have been selected randomly and not sequentially.

The equilibration of the systems studied has been assessed by monitoring the evolution of a number of observables such as enthalpy, density, and orientational order parameter, P_2. The latter was obtained by diagonalizing the Saupe ordering matrix **Q**, whose elements are defined as

$$Q_{ab} = \left\langle \frac{1}{2}(3u_a u_b - \delta_{ab}) \right\rangle$$

where the mean value $\langle \ldots \rangle$ is obtained averaging on all molecules composing the system, and $\hat{u}$ ($a = x, y, z$) is the eigenvector corresponding to the minimum eigenvalue of the molecular inertia tensor, i.e., the long principal axis. The maximum eigenvalue of Q is then taken as the principal order parameter P_2, and the corresponding eigenvector represents the phase director **n**.

Every equilibration has been followed by a production run for the evaluation of the thermodynamic properties. Positional order was monitored by calculating the correlation functions $g(r)$, $g(r_{\parallel})$, and $g(r_{\perp})$. $g(r)$ is the standard isotropic correlation function, while $r_{\parallel}$ and $r_{\perp}$ are the projections of the center of mass vectors along and normal to **n**, respectively. The orientational order was studied calculating the major order parameter P_2 and the orientational correlation function $G_2(r)$.[40] This function, computed as

$$G_2(r_{ij}) = \langle P_2(\hat{u}_i \cdot \hat{u}_j) \rangle (r_{ij})$$

describes the orientational correlation of two molecules *i* and *j* as a function of the distance between their center of mass. As shown by Bates et al.,[41] $G_2(r)$ reaches the asymptotic value of $\langle P_2 \rangle^2$ at large r_{ij}.

3. Results and Discussion

***p*-Polyphenyls.** The intermolecular PES of the *p*-*n*-phenyls have been computed through the FRM method,[24] which is based on the construction of a dimer interaction potential as a sum of fragment−fragment contributions. In particular, the polyphenyls PES can all be obtained as a sum of benzene−benzene interaction energies. The details of these FRM calculations have been already reported in refs 18 and 19. Furthermore, the accuracy of the ab initio calculations on the benzene dimers has been also extensively discussed in previous works.[24,42,43]

The computed *p*-*n*-phenyls PES have been fitted with the gGB model potential, reported in eq 1. In view of the poor significance of very repulsive values, all the energies $E_k^{\mathrm{FRM}} > 80$ kJ/mol were discarded from the integral **I** of eq 16, and the weighting coefficient α, defined by eq 17, was set to 0.84 (kJ/mol)$^{-1}$. Preliminary fittings performed over a wide α interval (0.42 (kJ/mol)$^{-1}$ < α > 8.37 (kJ/mol)$^{-1}$) showed a negligible dependence of the shape of the fitted curves over the weighting factor, at least in the considered conformations. Finally, the parametrization conforms to the molecular symmetry, i.e. equivalent rings are represented by identical gGB and quadrupolar potential terms. The obtained parameters are reported for all the series in Tables 1 and 2.

Table 1. Parameters of the GGB Model for the Oligophenyl Series[a]

n	*m*	ϵ_0	ϵ_{ss}	ϵ_{ee}	σ_0	σ_{ss}	σ_{ee}	μ	ν	ξ
2	1,2	1.26	4.18	28.38	6.83	6.83	3.18	1.0	1.20	0.86
3	1,3	1.59	5.02	13.60	6.69	6.69	3.41	1.0	2.07	0.89
3	2	0.71	1.80	48.03	6.86	6.86	2.83	1.0	−0.87	0.77
4	1,4	1.34	5.82	47.15	6.90	6.90	3.20	1.0	−0.86	0.83
4	2,3	1.67	2.34	22.13	6.56	6.56	3.16	1.0	−0.61	0.84
5	1,5	1.34	2.89	35.90	6.85	6.85	3.21	1.0	−1.24	0.79
5	2,4	2.26	2.34	25.06	6.51	6.51	3.19	1.0	−3.42	0.83
5	3	1.21	5.86	34.02	6.43	6.43	3.31	1.0	1.46	0.81

[a] *n* indicates number of phenyl rings in the molecule, and *m* indicates the position of the ring inside the polyphenyl; ϵ's are in kJ/mol and σ's in Å.

Table 2. Optimized Quadrupoles of the gGB Model for the Oligophenyl Series[a]

n	*Q*	SD	*n*	*Q*	SD
2	15.6	2.68	4	13.3	2.51
3	18.1	1.97	5	10.4	2.68

[a] All quadrupoles were imposed to be equal on each phenyl ring and are reported in C · m^2 · 10^{40}; in the last column, the standard deviations of the fitting procedures are reported in kJ/mol for each polyphenyl.

As it can be seen from the standard deviations reported in Table 2, although based on a reduced number of interaction sites with respect to an atomistic model, the present gGB model can describe to a good level of accuracy the main features of the reconstructed ab initio PES.

In Figure 2, cross sections of the computed FRM-PES are compared with the fitted potential for some selected conformations of *p*-quinquephenyl. In the first panel from the left, a sandwich parallel conformation is obtained by displacing the center of mass of one molecule along the translation vector ***R*** perpendicular to the central ring plane. All other conformations (i.e. cross, parallel displaced and T-shaped) are shown in the other panels of Figure 2 together with the selected ***R***. Similar results have been obtained for the other members of the series. It might be noted that aligned conformations (first and third panels from the left in Figure 2) result in deeper wells, thus favoring orientationally more ordered geometries. It may be worth also noticing as in the parallel displaced curve (third panel from the left), the model quadrupolar interaction allows for reproduction of the shift of the minimum toward slipped configurations ($R > 0$), found by the quantum mechanical results.[18]

Preliminary MC runs on the *p*-*n*-phenyls series had already been performed by our group,[19] by coupling the fitted intermolecular PES to a torsional ring−ring potential reported in the literature.[44] Unfortunately, it turned out that the height of the barrier for the planar conformation of biphenyl in ref 44 was overestimated compared to more recent experimental

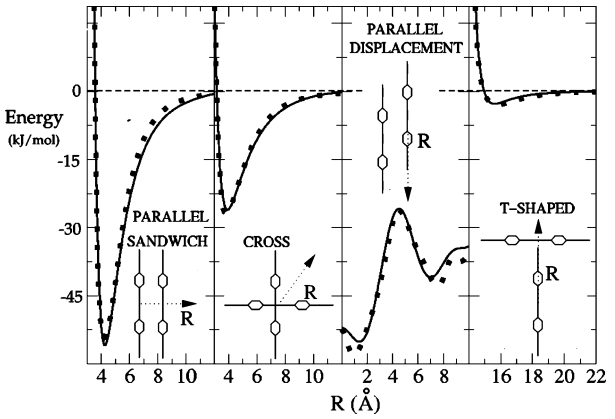


Figure 2. Selected cross sections of the FRM intermolecular PES (dotted black line) and the fitted gGB model (solid black line) for the polyphenyls series. The translation vector **R** is reported in the inner panels, for each of the considered geometries, with a dotted arrow. In the middle panels, the latter is perpendicular and parallel to both molecular long axis in cross and parallel displaced geometries, respectively. The intermolecular distance *R* is reported in Å and all energies in kJ/mol.

and theoretical data.[45,46] This led to inaccurate predictions of the thermodynamical properties of all the series.[19] In this work, MC NPT runs have been performed using the intermolecular hybrid model coupled with the intramolecular torsional potential reported in ref 28.

All simulations were started from the same configuration, created for all the omologues according to the following procedure. The polyphenyl centers of mass were disposed in a fcc crystal structure, elongated in the [111] direction, and chosen parallel to the *z* axis. Their molecular long axes were all aligned along the *z* direction, thus yielding a starting order parameter $P_2 = 1$; the dihedral angles between two contiguous disks were set to ∼45°, which corresponds to the minimum of the torsional potential as calculated in ref 28. Next, all crystals were expanded to a density $\rho \simeq 0.8$ g/cm^3, to favor the disordering processes. Starting from this expanded crystal, several MC runs were performed on each omologue at atmospheric pressure and varying the temperature by 50 K steps.

In Figure 3 the enthalpy, density, and orientational order parameter P_2 are shown as a function of temperature for the four systems. Some of the thermodynamical results reported in Figure 3 are compared with experimental data in Table 3.

The melting points (T_m) were determined in simulation by considering the average value between the highest temperature in the solid and the lowest in the liquid phase for each oligophenyl; clearing (T_{NI}) and boiling (T_b) points were determined analogously. In this way the computed values carry an uncertainty of ±25 K. At low temperatures, the four systems show a translationally and orientationally ordered phase. Raising the temperature, a liquid and a gas phase appear, with an almost vanishing orientational order parameter P_2, but with definitely different enthalpies and densities. With regards to the 5 ring member, a third fluid phase was found, with a partial orientational order ($P_2 \simeq$ 0.7). The lack of positional ordering, associated with a $P_2 > 0.4$, is typical of a nematic phase. The latter is experimentally found for *p*-quinquephenyl between 653 and 681 K,[49] thus very close to the computed range (625−675 K). In fact, Table 3 shows that the adopted model, although coarse grained, is capable of reproducing the range of stability of the phases with an error of ≃50 K. On the contrary, all transition enthalpies are significantly underestimated, with relative deviations increasing along the series.

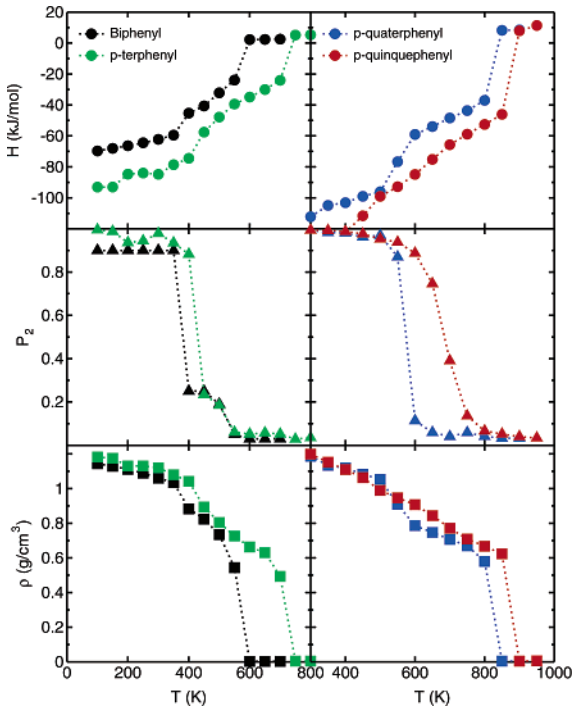


Figure 3. Thermodynamic properties of *p-n*-polyphenyls as a function of temperature for biphenyl (black symbols, left panels), *p*-terphenyl (green symbols, left panels), *p*-quaterphenyl (blue symbols, right panels), and *p*-quinquephenyl (red symbols, right panels). Enthalpies (circles) are in kJ/mol and densities (squares) in g/cm^3. All values have been averaged over production runs of 50 · 10^6 MC steps.

Table 3. Thermodynamic Results of the Employed Hybrid Model[d]

n	T_m	T_{NI}	T_b	ΔH_m	ΔH_{NI}	$\Delta H°_{vap}$
2	375		575	14.2		27.4
	344[a]		529[b]	18.8[a]		54.0[b]
3	425		725	17.1		34.6
	493[a]		658[b]	35.4[a]		79.0[b]
4	575		825	17.7		56.7
	587[a]		773[b]	37.8		117.0[b]
5	625	675	875	11.4	7.9	65.1
	660[a]	688[a]	823[b]	42.3[a]	0.9[a]	151.0[b]
	653[c]	681[c]				

[a] Reference 47. [b] Reference 48 and references therein. [c] Reference 49. [d] *n* indicates the number of phenyl rings in the *p-n*-phenyls. All temperatures are in K; all ΔH's are in kJ/mol.

The ordered phase shown at 650 K by *p*-quinquephenyl has been analyzed in more detail. First, in view of the known tendency of these systems to remain trapped in metastable states, the runs at $T = 600$, 650, and 700 K have been

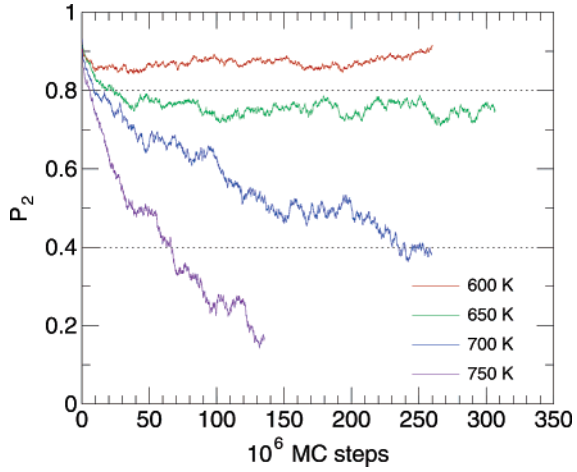


Figure 4. Equilibration of the order parameter of the *p*-quinquephenyl molecule at different temperatures. The nematic range [0.8−0.4] is indicated between dotted lines.

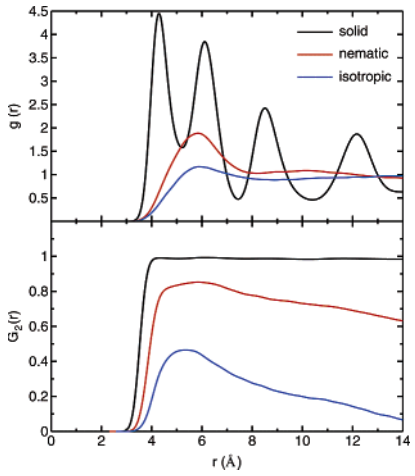


Figure 5. Correlation functions for *p*-quinquephenyl in the crystal (black line), in the nematic (red line) and liquid phase (blue line) at $T = 350$, 650, and 800 K respectively.

extended to over $250 \cdot 10^6$ MC steps. The P_2 was plotted over the step number and reported in Figure 4. The order parameter at 650 K (green line) remains stable for about 200 million configurations, resulting in an average value of $P_2 = 0.74$, whereas it grows up to 0.9 at 600 K (red line), and it falls under 0.4 (blue line) if the temperature is raised to 700 K. To assess the nematic nature of the phase at 650 K, the pair correlation functions $g(r)$ and $G_2(r)$ have been computed and reported in Figure 5. From the upper panel, it can be seen that, despite the high orientational order, no positionally ordered structure is present at 650 K, and the $g(r)$ correlation function shows a shape very similar to the one computed in the isotropic liquid phase. The $G_2(r)$ function at 650 K shows a long-range correlation, which decays to the expected $\langle P_2 \rangle^2$ value, consistent with the nematic nature of the orientationally ordered phase.

The proposed model accounts for flexibility through the intramolecular torsional potential computed in ref 28. The validation of the latter could be assessed if the proposed model results are able to reproduce the delicate interplay between intermolecular and intramolecular forces that drive

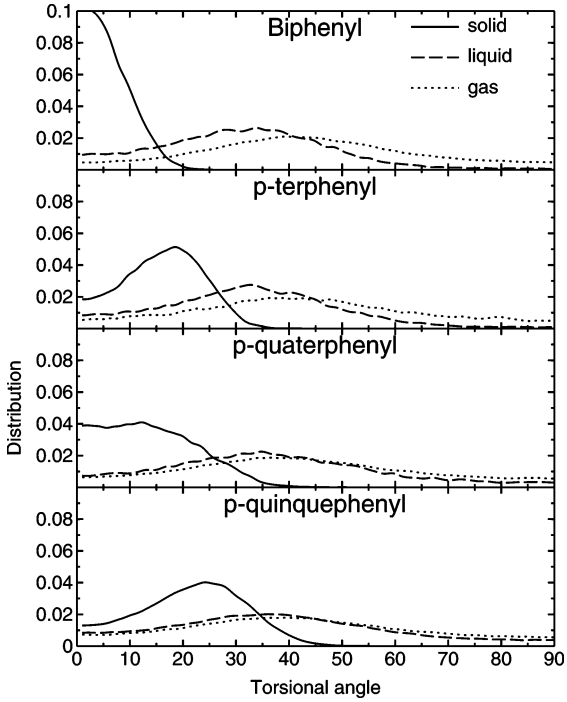


Figure 6. Population distribution of the first torsional angle along the *p*-oligophenyls series. For all molecules, distributions in solid, liquid, and gas phases are reported with solid, dashed, and dotted lines, respectively.

the phase transitions in the *p*-oligophenyls series. Indeed, the experimental inter-ring angle of biphenyl results to be around 40° in gas phase,[50,51] 32° in liquid phase,[52] and nearly 0° in crystalline phase.[53] This tendency is apparent also in the larger homologues,[54] where the molecular planarity is lost in going from the crystalline to the less ordered structures. In fact, this experimental trend is well reproduced by the simulation data, as one can see from Figure 6, where the distribution functions of the first dihedral angle are reported at different temperatures. For all the members of the series, a neat change in population can be observed at the transition between crystalline and liquid phases, in agreement with the experimental findings.[54] This can be easily explained by considering the gain in intermolecular energy due to the closer packing allowed for the planar conformations. Furthermore it is worth noticing the importance for the model to account for an accurate representation of internal flexibility. Indeed, the correction of the aforementioned overestimation in the torsional potential adopted in the preliminary runs[19] adjusts the balance between intermolecular and intramolecular contributions, thus leading to more accurate thermodynamic results, including transition temperatures. For instance, the melting temperatures of *p*-terphenyl and *p*-quaterphenyl for the present model are 425 and 575 K, respectively. These values clearly show a better agreement with the experimental measures (493 and 587 K[47]) with respect to those found in the preliminary runs,[19] namely 375 and 475 K.

5CB. The 5CB intermolecular potential energy surface, previously computed by the FRM method,[24,25] was fitted into the hybrid GB/LJ model. The fitting procedure, performed according to eq 16, has already been applied in ref 25 to an

Table 4. Fitted Intermolecular Parameters for the 5CB Hybrid Model[a]

site	ϵ_0	σ_0	ξ	ϵ_{ss}	ϵ_{ee}	σ_{ss}	σ_{ee}	μ	ν	q
Bz1	2.22	6.23	0.86	5.40	25.19	6.23	2.89	1.41	2.87	−0.070
Bz2	2.47	6.46	0.72	3.93	25.56	6.46	3.14	1.13	−1.02	0.127

	ϵ_0	σ_0	ξ	q
N	0.02	3.21	0.68	−0.369
Cn	0.07	2.68	0.15	0.318
Cp1	1.00	2.49	1.42	0.024
Cp2	0.33	4.16	0.70	0.000
Cp5	0.02	4.58	1.38	−0.030

[a] The charge on Cp2 site was imposed to be null. All ϵ's are reported in kJ/mol, σ's in Å, and all charges in fraction of e.

atomistic model potential. This 5CB atomistic model was used[26,34] in molecular dynamics (MD) simulations, and it was shown to reproduce with good approximation the thermodynamic, structural, and dynamic properties of 5CB crystalline, nematic, and liquid phases. In the present work, in fact, during both the parametrization and the validation procedures, the obtained results will be compared with those reported for the more "realistic" atomistic model.

Here, as done for the latter,[25] all repulsive energies larger than 40 kJ/mol were discarded from the sample of $N_g \simeq 2 \times 10^4$ dimer conformations, and the weighting coefficient α was set to to 1.67 $(kJ/mol)^{-1}$. Furthermore, the partial charges on the three Cp2 sites were imposed to be null, and their parameters were taken as equivalent. No other restriction but electroneutrality was imposed, obtaining in such a way a standard deviation of 2.8 kJ/mol. The latter well compares with the one obtained for the atomistic model,[25] i.e. 2.5 kJ/mol. The resulting intermolecular parameters are reported in Table 4.

Other coarse grained parametrizations, performed on biphenyl[18,19] or on cyanobiphenyl in 5CB,[20] report a strong dependence of the fitted parameters on the chosen weighting factors. In particular in ref 20, where a uniaxial single GB site is employed to model the cyanobiphenyl moiety in 5CB, the GB radii (σ_{ss} and σ_{ee}) are found to be very sensitive to the temperature employed in the fitting. In the present work, the value of the employed weighting factor α corresponds to a temperature of ≃200 K. As reported for the *n*-phenyls, other fittings performed with α set to 0.42 $(kJ/mol)^{-1}$ and 8.37 $(kJ/mol)^{-1}$ (corresponding to temperatures of 50 and 1000 K, respectively) did not show relevant differences in the shapes of the coarse grained PES. For instance, in the face-to-face conformation of the 5CB dimer (see Figure 7), the well depth (≃ −40 kJ/mol with $\alpha = 1.67$ (kJ/mol^{-1})) showed variations of less than 5%, while the distance where the fitted curves cross the $y = 0$ axis changes with α only by ≃2%. This different behavior with respect to other coarse grained parametrizations probably arises from the more accurate level of approximation introduced by the present model. Indeed, substituting each phenyl ring with a rigid gGB site implies only the remotion of "hard" internal potentials, which rule the ring hexagonal shape and planarity. Therefore, the degrees of freedom neglected by this model do not change much with temperature in the real molecule. Conversely, the substitution of both rings with a single (prolate) GB site[18,20]

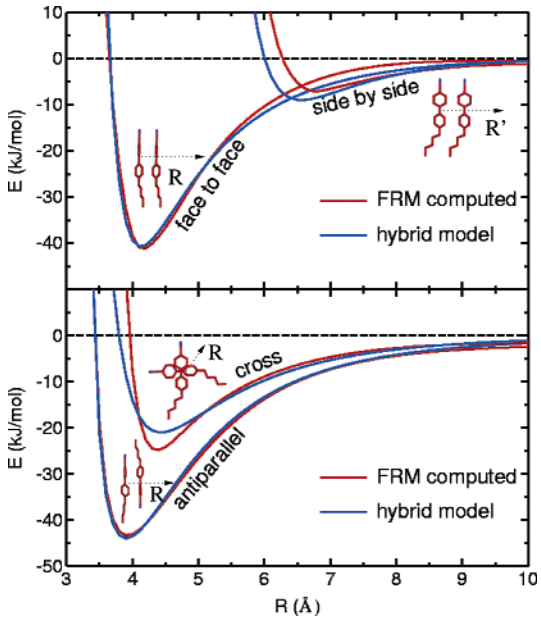


Figure 7. FRM reconstructed (red line) and hybrid model potential (blue line) energy curves. Translation vectors $\hat{\boldsymbol{R}}$ and $\hat{\boldsymbol{R}}'$, defined in the text, are reported in the inner panels with dotted black arrows. Energies are in kJ/mol and distances are in Å.

also removes the inter-ring "soft" potential, which drives molecular biaxiality and shows a rather flat shape in the minimum region.[28,25] In this case,[18,20] a rotational average on the reference PES is needed before performing the fitting, thus introducing the reported strong temperature dependence.

As done for the *p*-*n*-phenyls, some fitted potential curves are compared in Figure 7 with the corresponding FRM reconstructed energies, for selected dimer arrangements. The latter have been constructed as follows. In the parallel conformations, reported in the upper panel of Figure 7, the second molecule is moved along a vector **R** containing the midpoint of the ring−ring linkage, parallel to the $\hat{C}_6$ symmetry axis of the cyanophenyl group for the face-to-face geometries. The side-by-side arrangements are obtained by translating the second molecule along the vector **R′** containing the ring−ring linkage and perpendicular to both the molecular long axis and the $\hat{C}_6$ symmetry axis. In the antiparallel and cross geometries (lower panel of Figure 7), a rotation of 180° and 90°, respectively, is performed around $\hat{R}$, together with a translation along it. The agreement for all the four considered configurations is good; in particular the local parallel and antiparallel minima are well represented.

It is also worth noticing that the parallel dimer conformation is correctly less favorite than the antiparallel one. This is mainly due to the repulsive Coulomb contribution, which can be ascribed to an unfavorable dipole−dipole parallel arrangement. Indeed, the Coulomb contribution to the energy of the hybrid model in the parallel and antiparallel local energy minima result +7.1 kJ/mol and −0.1 kJ/mol, respectively. Moreover in the latter geometry, it may be of some interest to investigate the source of all energy contributions which concur to the local minimum and to compare the distribution arising in both the FRM calculation and the fitted hybrid model. The core−core, core−chain, and chain−chain

Table 5. FRM[24] and Fitted Energy Contributions in kJ/mol in the Antiparallel Local Energy Minimum

	FRM[24]	hybrid fitted model
r_0 (Å)	3.9	3.9
total energy	−43.4	−43.7
core−core	−33.6	−33.5
core−chain	−8.5	−9.6
chain−chain	−0.2	−0.5
Coulomb		−0.1

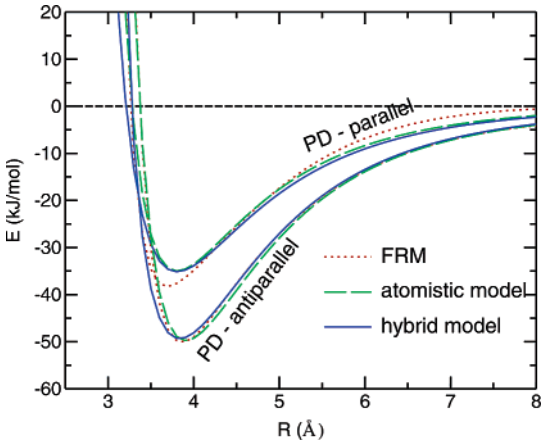


Figure 8. Comparison between FRM (red, dotted line) curve and atomistic[25] (green dashed line) and hybrid (blue solid line) models for some selected parallel displaced (PD) conformations. Energies are in kJ/mol and distances are in Å.

FRM terms, reported in Table 5, can be easily computed as sums of the appropriate fragment−fragment contributions. The core−chain term, for instance, arises from the sum of all benzonitrile−pentane and benzene−pentane interaction energies. Conversely, in the fitted model, the three contributions arise from different types of interactions: gGB-gGB, gGB-LJ, and LJ-LJ, respectively.

From the quantities reported in Table 5, it is evident that the hybrid model is capable of conserving the correct distribution of the energy contributions found by the FRM approach.[24] In fact it can be seen that the main source of attraction is the interaction between the aromatic cores, which accounts for almost 80% of the total interaction energy.

Finally, in Figure 8 a comparison between the 5CB atomistic[25] and hybrid models is presented. Two parallel displaced (PD) conformations were created with the same procedure, i.e. a displacement along the long molecular axis of 5.0 Å and −2.5 Å for the parallel and antiparallel dimers, respectively. The agreement between the two models is satisfactory, as the hybrid models do not lose much accuracy in representing the FRM curves, with respect to the more "realistic" atomistic one.

5CB simulations were carried out in the NPT ensemble, making use of the hybrid GB/LJ intermolecular model coupled to the intramolecular potential whose parametrization has been described in detail in ref 26.

Two configurations of 192 molecules, in the nematic and isotropic phase respectively, were extracted from the equilibrated trajectories of MD simulations previously performed by our group with the 5CB atomistic FRM model.[25,26,34]

Table 6. Calculated and Experimental Thermodynamic Properties[a]

	this work	MD-FA[26]	exp
ρ (300 K)	1.092 ± 0.04	1.085 ± 0.03	1.020[55]
ρ (320 K)	1.051 ± 0.05	1.063 ± 0.04	0.995[55]
P_2 (300 K)	0.546 ± 0.01	0.538 ± 0.03	0.54[56]
P_2 (320 K)	0.151 ± 0.01	0.176 ± 0.03	
U_{inter} (300 K)	−98.4 ± 0.1	−106.9 ± 0.5	
U_{inter} (320 K)	−86.7 ± 0.1	−103.7 ± 0.7	
CPU time/step	0.02	0.6	
$N_{steps} \cdot 10^6$	220	40	

[a] The data of this work, averaged over 50 millions MC steps, are compared with the corresponding MD values, reported in ref 26. Density ρ is expressed in g/cm^3, and intermolecular energies contributions U^{inter} are in kJ/mol. In the last two rows are reported the CPU time/step in seconds and the average number of steps needed to reach equilibrium, N_{steps}. Both MC and MD calculations were performed on Xeon 2.8 GHz processor.

These geometries were expanded to a density of 0.8 g/cm^3, and the interaction sites (six aromatic carbons and four hydrogens) of each phenyl ring were substituted with a rigid gGB disk, centered in the geometrical center of the phenyl plane and with its orientation axis $\hat{u}_{gGB}$ perpendicular to the ring (see panel b of Figure 1). Such hybrid configurations were eventually used as starting point for two MC-NPT runs at atmospheric pressure, at 300 and 320 K.

Equilibration, assessed by monitoring total energy, density, and order parameter of both systems, was reached after ≃280 $\cdot$ 10^6 and 160 $\cdot$ 10^6 MC steps for the lower and higher temperature, respectively. A comparison of the CPU usage of these simulations, with respect to the atomistic MD runs performed on the same system,[26,34] cannot be easily carried out, since different techniques as well as different modeling levels are concerned.

Nevertheless, in the last row but one of Table 6, the CPU time/step has been reported for the two methods. The cost of the MD technique is not surprising, considering that each step involves a collective move, the calculation of the forces, and the computation of the energies for a larger number of interaction sites. However, in the discussion of the CPU usage with the two different techniques, the number of steps performed to reach equilibrium should be also taken into account. The latter quantity has been reported in the last row of Table 6 and clearly shows the better efficiency of MD steps to reach equilibrium.

Once equilibrated, production runs of 50 $\cdot$ 10^6 MC steps have been carried out for both temperatures. The values of some average quantities are reported in the first rows of Table 6 and compared with those obtained with the atomistic MD runs. These results seem to indicate that the adopted hybrid model is still capable of reproducing stable isotropic and nematic phases in the experimental temperature range. In particular the orientational order parameter in the nematic phase at 300 K well agrees with both the MD and the experimental value. Conversely, by looking at the first rows of Table 6, one may note as the gain in energy from the isotropic phase at 320 K to the nematic at 300 K is higher for the hybrid model (11.7 kJ/mol) than for the atomistic one (3.2 kJ/mol), which better compares with the experimental clearing enthalpy ΔH^c of 0.54 kJ/mol.[57] Accordingly,

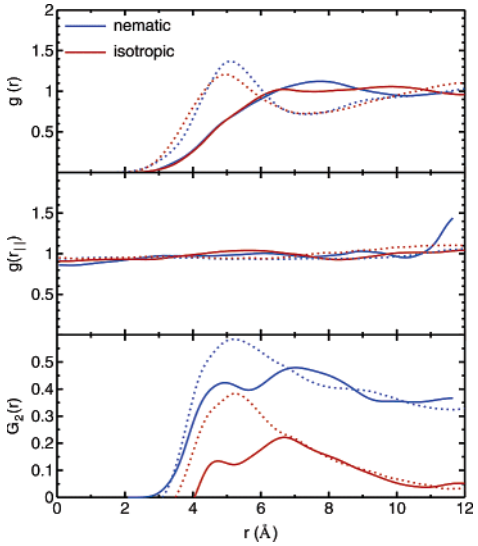


Figure 9. Correlation functions in the nematic (300 K, blue lines) and isotropic (300 K, red lines) phases. MC results of this work are reported with solid lines, while dashed lined are used for MD results of ref 26.

the MC ordered system at 300 K increases its density with respect to the MD value, while the isotropic phase behaves in an opposite manner.

A more detailed characterization of the structure of the two resulting condensed phases can be obtained from the computed correlation functions, $g(r)$, $g(r_{||})$, and $G_2(r)$, reported in Figure 9.

As can be seen from the behavior of $g(r)$ and $g(r_{||})$ functions (solid lines in Figure 9), no positional order is present in the two phases and both exhibit a liquidlike structure. More importantly, the $G_2(r)$ orientational correlation function at 300 K correctly decays to an asymptotic value of $\simeq$0.33, which is consistent with $\langle P_2\rangle^2 = 0.3$, thus confirming the nematic nature of the condensed phase at lower temperature. The $g(r)$ functions for the atomistic model[25,26] (reported with dashed lines) show a different short-range structure with respect to the hybrid model results. This is rather surprising since both models were fitted on the same FRM PES, and they show (see Figure 4) a very similar behavior, at least on the investigated configurations. Indeed, this extreme sensibility of the simulated properties of complex materials to the details of the molecular interactions reinforces the need of specifically computed intermolecular PES as well as models potential capable of accurately reproducing them.

Another important feature to verify is the capability of the hybrid coarse grained model to reproduce the molecular flexibility of the 5CB molecules. In Figure 10 are reported population distributions of the inter-ring dihedral (δ) and the first dihedral of the aliphatic chain (ϕ), i.e. defined by the sites Bz2-Cp1-Cp2-Cp2 (see Figure 1, panel b). With regards to the former, the average value of the dihedral between the two gGB disks 32.6° in the nematic phase well agrees with both experimental[56] (37°) and the atomistic model[26] results (31°). The distribution of conformers in the aliphatic chain, monitored by the dihedral ϕ, confirms the tendency of the 5CB molecules to assume a more elongated geometry in the

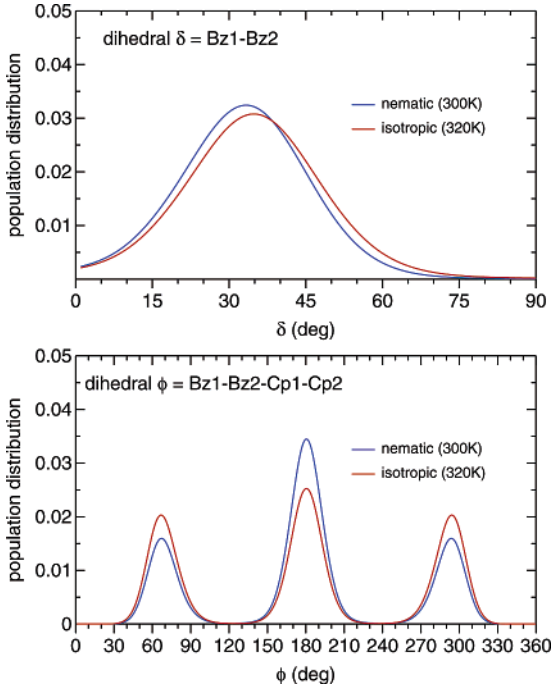


Figure 10. Population distribution functions of the inter-ring dihedral δ (upper panel) and the first chain torsional dihedrals ϕ (bottom) in the nematic (blue lines) and isotropic (red lines) phases.

nematic phase, by increasing the trans population with respect to the gauche one. This is in good agreement with both the experimental[56] findings and the atomistic model results.[26]

4. Summary and Conclusions

In this paper, coarse grained hybrid potentials, suitable for computer simulations of advanced materials, have been parametrized and validated through MC computer simulations. Thanks to the FRM approach,[24,25] pair potentials of large molecular dimers can be ab initio computed, constructing a reference PES on which the hybrid models can be sewn. In this way, the chemical detail entailed in a quantum mechanical description can be transferred, up to the chosen level of accuracy, to rather simple models that can be used in computer simulations. It may be worth stressing that the FRM/coarse grained approach can be employed for any large molecular system. Once obtained, the transferability of the FRM-derived parameters must be carefully evaluated in each case: in the polymer field, for instance, simulations of polyphenylene or similar molecules could be tempted with the adoption of the many gGB disks model, without the need of further parametrizations. Conversely, the specificity of the FRM derived potentials can become a drawback if one tries to extend the use of the same coarse grained parameters to similar fragments in very different chemical backgrounds.

In the cases studied, the proposed models have proven capable of reproducing the computed FRM PES with good approximation. The augmented tendency of the larger homologues of the *p*-*n*-phenyl series to assume parallel conformations is well reproduced by the gGB disks as well as the effect of the ring quadrupoles to shift the minimum energy toward displaced conformation. With regards to the

5CB molecule, comparison with the more realistic full atomic potential,[25,26] fitted on the same FRM PES, shows that most of the features of the molecular interaction are correctly accounted for. In a hierarchy of simulation models, the realism and computational complexity of such hybrid potentials lies between a single-site representation[3,7] and a fully atomistic modeling. Notwithstanding their simplicity, the proposed models have been proven capable of describing at the semiquantitative level the phase diagram of the *p*-oligophenyl series and reproducing to a good level of accuracy the structure and orientational properties of a typical mesogen.

Among the formers, three phases (solid, liquid, and gas) were reproduced at atmospheric pressure for the $n < 5$ members, with a maximum error of ≃50 K on the phase transition temperatures. Furthermore, the larger homologue, namely *p*-quinquephenyl, showed an orientationally ordered phase ($P_2 \simeq 0.7$) at 650 K and 1 atm, not far from the experimental range (653−688 K[47,49]). Conversely, the transition enthalpies are underestimated if compared with the relative experimental values, with an increasing error for the larger homologues. This might be due to the lack of intramolecular flexibility of the hybrid model, which neglects all internal degrees of freedom except the torsion between adjacent rings. Indeed, the importance of molecular flexibility on the resulting bulk properties is confirmed by the increased accuracy of the present results with respect to those obtained, for the oligophenyl series, by the use of single site models[18,19] or less accurate intramolecular potentials.[19] The inclusion (and the accurate description) of "flexible" internal coordinates appears to be a necessary feature for an accurate representation of the phase transitions. The hybrid model adopted for the 5CB molecule was shown to be able to reproduce a nematic and an isotropic liquid phase, in the correct temperature range. Both these phases, however, showed a short-range structure noticeably different from that shown by the correlation functions obtained[26] for the more realistic atomistic model. On the contrary, the internal distribution of the more flexible torsional angles (as the inter-ring or the chain dihedrals) is well reproduced and in agreement with both theoretical and experimental findings.

In conclusion, hybrid potentials turn out to be useful in modeling very large molecules, where the computational costs do not allow more complex functions. For example, the recently discover banana molecules,[58] which are formed from several phenyl groups, connected by benzoate linkages, could be a relevant test case for the capabilities of the proposed models.

Acknowledgment. The author would like to thank Prof. Ivo Cacelli for the many useful discussions and his careful reading of the manuscript. The University of Pisa is gratefully acknowledged for financial support.

Note Added after ASAP Publication. This article was released ASAP on March 28, 2006, with the incorrect Received Date. The correct version was posted on April 12, 2006.

References

(1) Advances in the Computer Simulations of Liquid Crystals. *NATO ASI series;* Pasini, P., Zannoni, C., Eds.; Kluwer: Dordrecht, 2000.

(2) Computer Simulations of Liquid Crystals and Polymers. *NATO ASI series;* Pasini, P., Zannoni, C., Zumer, S., Eds.; Kluwer: Dordrecht, 2005.

(3) Zannoni, C. *J. Mater. Chem.* **2001**, *11*, 2637.

(4) Crain, J.; Komolkin, A. *Adv. Chem. Phys.* **1999**, *109*, 39.

(5) Wilson, M. *Int. Rev. Phys. Chem.* **2005**, *24*, 421.

(6) Lebwohl, P.; Lasher, G. *Phys. Rev. A* **1972**, *6*, 426.

(7) M. P. Allen, G. T. Evans, D. F.; Mulder, B. *Adv. Chem. Phys.* **1993**, *86*, 1.

(8) Zewdie, H. *J. Chem. Phys.* **1998**, *108*, 2117.

(9) Gay, J. G.; Berne, B. *J. Chem. Phys.* **1981**, *74*, 3316.

(10) Cleaver, D. J.; Care, C. M.; Allen, M. P.; Neal, N. P. *Phys. Rev. E* **1996**, *54*, 559.

(11) Bates, M.; Luckhurst, G. *Struct. Bonding (Berlin)* **1999**, *94*, 65.

(12) Cross, C. W.; Fung, B. M. *J. Chem. Phys.* **1994**, *101*, 6839.

(13) LaPenna, G.; Catalano, D.; Veracini, C. A. *J. Chem. Phys.* **1996**, *105*, 7097.

(14) Wilson, M. *J. Chem. Phys.* **1997**, *107*, 8654.

(15) McBride, C.; Wilson, M. R. *Mol. Phys.* **1999**, *97*, 511.

(16) Cacelli, I.; Campanile, S.; Prampolini, G.; Tani, A. *J. Chem. Phys.* **2002**, *117*, 448.

(17) Fukunaga, H.; Takimoto, J.; Aoyagi, T.; Shoji, T.; Sawa, F.; Doi, M. *Mol. Cryst. Liq. Cryst.* **2001**, *365*, 739.

(18) Cacelli, I.; Cinacchi, G.; Geloni, C.; Prampolini, G.; Tani, A. *Mol. Cryst. Liq. Cryst.* **2003**, *395*, 171.

(19) Cacelli, I.; Cinacchi, G.; Prampolini, G.; Tani, A. *Computer Simulation of Mesogen with ab initio Interaction Potentials in Novel Approaches to the Structure and Dynamics of Liquids. Experiments, Theories and Simulation;* Samios, J., Durov, V., Eds.; Kluwer: Dordrecht, 2004.

(20) Fukunaga, H.; Takimoto, J.; Doi, M. *J. Chem. Phys.* **2004**, *120*, 7800.

(21) Lyunin, A.; Barwani, M.; Allen, M.; Wilson, M.; Neelov, I.; Allsopp, N. *Macromolecules* **1998**, *31*, 4626.

(22) Wilson, M.; Stimson, L.; Ilnytsky, J.; Hughes, Z. *Computer simulation of liquid crystalline polymers and dendrimers in Computer Simulations of Liquid Crystals and Polymers NATO ASI series;* Pasini, P., Zannoni, C., Zumer, S., Eds.; Kluwer: Dordrecht, 2005.

(23) Jorgensen, W. *BOSS 4.2;* Department of Chemistry: Yale, 2001.

(24) Amovilli, C.; Cacelli, I.; Campanile, S.; Prampolini, G. *J. Chem. Phys.* **2002**, *117*, 3003.

(25) Bizzarri, M.; Cacelli, I.; Prampolini, G.; Tani, A. *J. Phys. Chem. A* **2004**, *108*, 10336.

(26) Cacelli, I.; Prampolini, G.; Tani, A. *J. Phys. Chem. B* **2005**, *109*, 3531.

(27) Kovacic, P.; Jones, M. *Chem. Rev.* **1987**, *87*, 357.

(28) Cacelli, I.; Prampolini, G. *J. Phys. Chem. A* **2003**, *107*, 8665.

(29) Picken, S. J.; VanGunsteren, W. F.; VanDujen, P. T.; DeJeu, W. H. *Liq. Cryst.* **1989**, *6*, 357.

(30) Komolkin, A. V.; Laaksonen, A.; Maliniak, A. *J. Chem. Phys.* **1994**, *101*, 4103.

(31) Stevensson, B.; Komolkin, A.; Sandström, D.; Maliniak, A. *J. Chem. Phys.* **2001**, *114*, 2332.

(32) Zakharov, A.; Komolkin, A. V.; Maliniak, A. *Phys. Rev. E* **1999**, *59*, 6802.

(33) Zakharov, A.; Maliniak, A. *Eur. Phys. J. E* **2001**, *4*, 435.

(34) DeGaetani, L.; Prampolini, G.; Tani, A. *J. Phys. Chem. B* **2006**, *110*, 2847.

(35) Buckingam, A. *Adv. Chem. Phys.* **1978**, *12*, 107.

(36) Wiener, S. J.; Kollmann, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230.

(37) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. H.; Teller, A. H.; Teller, E. *J. Chem. Phys.* **1953**, *21*, 1087.

(38) Allen, M.; Tildesley, D. *Computer Simulation of Liquids;* Clarendon: Oxford, 1987.

(39) Ewald, P. P. *Ann. Phys.* **1921**, *64*, 253.

(40) *The Molecular Dynamics of Liquid Crystals NATO ASI series;* Luckhurst, G. R., Veracini, C. A., Eds.; Kluwer: Dordrecht, 1994.

(41) Bates, M.; Luckhurst, G. *J. Chem. Phys.* **1999**, *110*, 7087.

(42) Cacelli, I.; Cinacchi, G.; Prampolini, G.; Tani, A. *J. Am. Chem. Soc.* **2004**, *126*, 14278.

(43) Cacelli, I.; Cinacchi, G.; Prampolini, G.; Tani, A. *J. Chem. Phys.* **2004**, *120*, 3648.

(44) Tsuzuki, S.; Tanabe, K. *J. Phys. Chem.* **1991**, *95*, 139.

(45) Goller, A.; Grummt, U. *Chem. Phys. Lett.* **2000**, *321*, 399.

(46) Tsusuki, S.; Uchimaru, T.; Matsamura, K.; M., M. M.; Tanabe, K. *J. Chem. Phys.* **1999**, *110*, 2858.

(47) Smit, G. *Mol. Cryst. Liq. Cryst.* **1977**, *49*, 207.

(48) Irvine, P. A.; Wu, C.; Flory, P. *J. Chem. Soc., Faraday Trans.* **1984**, *80*, 1795.

(49) Dingrmans, T.; Murthy, N.; Samulski, E. *J. Phys. Chem. B* **2001**, *105*, 8845.

(50) Bastiansen, O.; Fernholt, L.; Cyvin, B. C. S.; Samdal, S.; Almenningen, A. *J. Mol. Struct.* **1985**, *128*, 59.

(51) Bastiansen, O.; Samdal, S. *J. Mol. Struct.* **1985**, *128*, 115.

(52) Steele, D.; Eaton, V. *J. Chem. Soc., Faraday Trans.* **1973**, *69*, 1601.

(53) Delugeard, Y.; Charbonneau, G. *Acta Crystallogr. B* **1976**, *33*, 1586.

(54) Baker, K.; Fratini, A.; Resch, T.; Knachel, H.; Adams, W.; Socci, E.; Farmer, B. *Polym. Pap.* **1993**, *69*, 1601.

(55) Sandmann, M.; Hamann, F.; Wurflinger, A. *Z. Naturforsh.* **1997**, *52*, 739.

(56) Adam, C.; Ferrarini, A.; Wilson, M.; Ackland, G.; Crain, J. *Mol. Phys.* **1999**, *97*, 541.

(57) Oweimreen, G.; Morsy, M. *Thermochim. Acta* **2000**, *346*, 37.

(58) Pelzl, G.; Diele, S.; Weissflog, W. *Adv. Mater.* **1999**, *11*, 707.

CT050328O

JCTC Journal of Chemical Theory and Computation

A Local Galilean Invariant Thermostat

Robert D. Groot*

Unilever Research Vlaardingen, P.O. Box 114, 3130 AC Vlaardingen, The Netherlands

Received November 3, 2005

Abstract: The thermostat introduced recently by Stoyanov and Groot (*J. Chem. Phys.* **2005**, *122*, 114112) is analyzed for inhomogeneous systems. This thermostat has one global feature, because the mean temperature used to drive the system toward equilibrium is a global average. The consequence is that the thermostat locally conserves energy rather than temperature. Thus, local temperature variations can be long-lived, although they do average out by thermal diffusion. To obtain a faster local temperature equilibration, a truly local thermostat must be introduced. To conserve momentum and, hence, to simulate hydrodynamic interactions, the thermostat must be Galilean invariant. Such a local Galilean invariant thermostat is studied here. It is shown that, by defining a local temperature on each particle, the ensemble is locally isothermal. The local temperature is obtained from a local square velocity average around each particle. Simulations on the ideal gas show that this local Nosé−Hoover algorithm has a similar artifact as dissipative particle dynamics: the ideal gas pair correlation function is slightly distorted. This is attributed to the fact that the thermostat compensates fluctuations that are natural within a small cluster of particles. When the cutoff range r_c for the square velocity average is increased, systematic errors decrease proportionally to $r_c^{-3/2}$; hence, the systematic error can be made arbitrary small.

1. Introduction

To simulate complex liquids on microsecond time scales and length scales intermediate between the atomistic scale and the macroscopic scale, several techniques are available. In some techniques, an underlying lattice is used like in time-dependent Ginzburg−Landau theory[1,2] and in Lattice Boltzman simulations.[3,4] Other techniques are particle-based, like Dissipative Particle Dynamics (DPD),[5−7] Voronoi Dissipative Particle Dynamics,[8] and Stochastic Rotation Dynamics.[9] In a particle-based simulation, the time evolution of a set of interacting particles is followed by integrating their equations of motion according to the laws of classical mechanics. When particles interact by a pairwise potential only, the natural thermodynamic ensemble is the microcanonical *NVE* ensemble, where energy E is conserved. However, in many cases, it is desirable that the simulation generates an *NVT* ensemble, where T is the absolute thermodynamic temperature. To this end, a "thermostat" is applied to the system.

The Nosé−Hoover thermostat[10,11] is a common choice for canonical molecular dynamic (MD) simulations. The basic idea behind this is to introduce a new internal degree of freedom into the Hamiltonian of the system H, representing the thermostat coupling. This in turn modifies the equations of motion and introduces one extra equation for the thermostat variable α, which has to be integrated together with the other equations:

$$m_i\frac{\mathrm{d}\mathbf{r}_i}{\mathrm{d}t} = \mathrm{d}\mathbf{p}_i;\quad \frac{\mathrm{d}\mathbf{p}_i}{\mathrm{d}t} = -\nabla_i U - \alpha\mathbf{p}_i;\quad \frac{\mathrm{d}\alpha}{\mathrm{d}t} = (T - T_0)/t_s \quad (1)$$

where $\mathbf{p}_i$ is the momentum of particle i with mass m_i at position $\mathbf{r}_i$, T is the momentary value of the system temperature as defined above, $U(\mathbf{r}_i)$ is the potential energy of the system, and t_s is the thermostat coupling parameter (the rise time), which controls energy transfer back and forth from the thermostat. In the last equation, α acts as an effective friction parameter. MD simulations, in this case, are actually performed in a microcanonical *NVE′* ensemble, with modified Hamiltonian H'. Nevertheless, the thermody-

* Author e-mail: rob.groot@unilever.com.

10.1021/ct050269e CCC: $33.50
Published on Web 04/14/2006

namic averages in this *NVE′* ensemble are equivalent to an average in the canonical *NVT* ensemble for the original Hamiltonian *H*, but with rescaled particle momentum. This thermostat is global because the momentary value of the temperature is based on a global definition. Obviously, the Nosé−Hoover thermostat is non-Galilean invariant, because all calculations are performed in a reference frame in which the center of mass of the system is at rest. This is the only frame where velocity rescaling and particle friction relative to the coordinate frame preserve total momentum. This restriction also means that the thermostat is not suitable when external forces (like external pressure or gravity) are acting on the system and accelerating the center of mass. The thermostat effectively brings an additional external friction force, which means that hydrodynamics is artificial. The same holds for total angular momentum; if it is nonzero, the thermostat introduces an external friction torque. Generally, this thermostat does not conserve local (angular) momentum. These drawbacks are due to the fact that effective thermostating forces are noncentral and nonpairwise additive. These disadvantages are removed in the pairwise noise and friction thermostat that is implemented in DPD.[5−8]

In the DPD method, pairwise additive and central forces are introduced: a dissipative force $\boldsymbol{F}_{ij}^{\mathrm{D}}$ and a random force $\boldsymbol{F}_{ij}^{\mathrm{R}}$, which act together with the conservative force $-\nabla_i U$. These forces are tuned in such a manner so that the system evolution in phase space is governed by the Liouville equation, so that it obeys the fluctuation−dissipation theorem.[6] By construction, the DPD thermostat is local and Galilean invariant and, therefore, preserves hydrodynamics. The main disadvantage of DPD is that the resulting stochastic equations of motion are difficult to integrate self-consistently.[12] Self-consistent integration requires several force calculations per time step, which decreases computational efficiency. An implementation that is not self-consistent[7] (single force calculation per time step) leads to some artifacts. One of these is that the pair correlation function of the ideal gas deviates from 1,[12] which means that nonphysical interactions are present between particles. The second disadvantage of DPD is that it simulates fluids with comparable diffusion coefficient D and kinematic viscosity $\nu = \eta/\rho$, resulting in Schmidt numbers $\mathrm{Sc} = \nu/D$ close to 1.[7] Depending on the application, this could be a disadvantage, because for most common liquids the Schmidt number is on the order of 10^3, but for diffusion-limited problems, this can be viewed as an advantage.

A completely different thermostat is the Andersen thermostat,[13] which implements a Monte Carlo scheme to sample the equilibrium velocity distribution. The velocity of a randomly chosen particle is replaced by a velocity drawn from a Maxwell distribution. This thermostat is local by nature, but it does not preserve hydrodynamics. Recently, Lowe[14] proposed a generalization of this as an alternative to DPD. The idea is to change the relative velocities of pairs of particles, rather than acting on single particles. To this end, the relative velocity is projected on the line connecting their centers, and this value is replaced by a value drawn from a Maxwell distribution. One might interpret this process as the exchange of a virtual particle between the two real particles. It is the momentum carried by these virtual particles that causes a viscosity increase. By construction, the Lowe−Andersen thermostat is Galilean invariant, local, and preserves hydrodynamics. The viscosity of the fluid sampled with this thermostat is linearly proportional to the exchange frequency, which in turn determines thermostating efficiency. To have good thermostating, a relatively high exchange frequency must be used, which in turn leads to very viscous fluids. In some cases, this can be a disadvantage, because a low Schmidt number cannot be accessed.

Recently, a thermostat that is able to simulate both at low and at high Schmidt numbers was proposed by Stoyanov and Groot.[15] This new thermostat also acts on pairs of particles, but at each time step, a random choice is made between the Lowe−Andersen thermostat and a local version of the Nosé−Hoover thermostat. Just as in DPD, a velocity-dependent pairwise force is introduced

$$\mathbf{F}_{ij} = \alpha\, \psi(r_{ij}/r_c)\,(1 - T/T_0)[(\mathbf{v}_i - \mathbf{v}_j)\cdot\mathbf{e}_{ij}]\mathbf{e}_{ij}/\delta t \qquad (2)$$

where $\psi(r)$ is smooth smearing function and α is a fixed coupling parameter. The idea behind this is that particles will experience a friction force if the temperature is too high, but they are accelerated when the temperature is too low. Generally, when the system is close to equilibrium, the force in eq 2 (nearly) vanishes, so that this thermostat avoids the problems encountered in DPD to integrate the equations of motion. There is, however, a subtle point in this formulation which leads to a nonlocal interaction. This is that the temperature T is defined as a *global* average over local squared velocity differences. The consequence of this is that the thermostat switches off when the global temperature equals the temperature T_0 that is aimed for, even if the local temperature is too high in one part of the system and too low in another part. We will show below that this situation may arise in an inhomogeneous system, where the initial state is not in pressure equilibrium. Because of adiabatic expansion, one part of the system then cools while the other part heats. Such an off-equilibrium temperature distribution may pertain for a relatively long time until heat diffusion has equilibrated the system spontaneously. To force the system to the desired temperature throughout the system in such situations, a truly local thermostat is needed, based on a local definition of temperature. This will be described below.

2. Pairwise Nosé−Hoover Thermostat

The basic idea behind the Stoyanov−Groot thermostat is to combine two thermostats coupled in parallel. The first is a thermostat *similar* to the Nosé−Hoover thermostat (NHT)[10,11] but which is Galilean invariant and acts on pairs of particles, rather than on single particles. The second is the Lowe−Andersen thermostat (LAT),[14] which is a pairwise analogue of the Andersen thermostat.[13] For each particle pair, a choice is made between NHT and LAT with probability $P = \Gamma\delta t$, where δt is the integration time step and Γ is the Lowe−Andersen exchange frequency. We will concentrate on the pairwise Nosé−Hoover thermostat here.

The pairwise analogue of the Nosé−Hoover thermostat is implemented by applying a thermostating force acting on

pairs of particles i and j within a cutoff distance r_c, as in eq 2. The coupling parameter α is constant during the simulation, unlike in the Nosé−Hoover thermostat. The advantage of this thermostat is that it conserves both total linear and angular momentum in the system, which is a necessary condition for restoring proper hydrodynamic behavior. A similar type of thermostating force was independently suggested recently by Phares and Srinivasa,[16] for implementing molecular internal degrees of freedom in energy-conserving molecular dynamic simulations. In that implementation, the local internal temperature of particles i and j follows from a differential equation describing heat flow. In the formulation by Stoyanov and Groot,[15] temperature is a local average of square velocity differences. To guarantee Galilean invariance, the temperature must be defined in a comoving frame for each particle. Within this frame, the local mean square velocity of the neighbors within a cutoff radius is determined. When the average is taken over all such determined local mean square velocities, the system temperature is defined as

$$kT = \frac{\sum_{i>j} \zeta(r_{ij}/r_c)\, M_{ij}(\mathbf{v}_i - \mathbf{v}_j)^2}{3\sum_{i>j} \zeta(r_{ij}/r_c)} \qquad (3)$$

where $\zeta(r)$ is a smearing function for the temperature, chosen such that $\zeta = 0$ for $r > 1$. The cutoff value r_c in the force (eq 2) and in the temperature (eq 3) could be different in principle. The parameter M_{ij} appearing in eq 3 is the reduced mass $m_i m_j/(m_i + m_j)$ of particles i and j. For particles of equal mass, $M_{ij} = {}^1/_2$. Because at equilibrium the velocities and particle positions are independent, it is straightforward to prove that the above temperature definition coincides with the equilibrium thermodynamic temperature.

The fact that the momentary value of the temperature (eq 3) is calculated on the basis of all particle velocities and coordinates of the system leads to a global thermostat. As we will see later on, this subtle global nature of this thermostat is important for systems far from equilibrium. Therefore, we also define a truly local momentary temperature as

$$kT_i = \frac{\sum_{j} \zeta(r_{ij}/r_c)\, M_{ij}(\mathbf{v}_i - \mathbf{v}_j)^2}{3\sum_{j} \zeta(r_{ij}/r_c)} \qquad (4)$$

The thermostating force then can be modified so that it contains only local particle temperatures; hence

$$\mathbf{F}_{ij} = \alpha\psi(r_{ij}/r_c)\,[1 - {}^1/_2(T_i + T_j)/T_0][(\mathbf{v}_i - \mathbf{v}_j)\cdot\mathbf{e}_{ij}]\mathbf{e}_{ij}/\delta t \quad (5)$$

However, in practice, MD simulations are usually performed with a relatively low particle density. This means that the local temperature defined through eq 4 is not very efficient in terms of temperature control, because of a large variance in the local temperature around its mean value. For such a local pairwise thermostat, the fluctuations around T_0 are typically an order of magnitude larger than that for the thermostat based on a global temperature, if in both cases the same smearing function is chosen.[15]

3. The Global Temperature Problem

Why a global temperature definition does lead to problems when the thermostatic force of eq 2 is used can be illustrated by the following example. Let us take a simulation box of dimensions $30 \times 10 \times 10\ r_c^3$, where r_c is the cutoff radius of the interaction that we shall take as our unit of length. The left half of this box is filled with 4500 A particles (leading to a mean density $\rho = 3$), and the right half of the box is filled with 1500 B_3 trimers. Each trimer is made by linking three B particles by harmonic springs with spring constant $C = 4$. The A and B particles interact with the soft potential

$$U_{ij} = \begin{cases} {}^1/_2 a_{ij}(1 - r_{ij}/r_c)^2 & \text{if } r_{ij} < r_c \\ 0 & \text{if } r_{ij} > r_c \end{cases} \qquad (6)$$

We take the repulsions $a_{AA} = a_{BB} = 25$ and $a_{AB} = 42$. Following Groot and Warren,[7] this corresponds to a phase-separating system with χ parameter $\chi \approx 5$. This is such a low miscibility that no trimers dissolve in the A phase and only a fraction $\phi = 0.2\%$ of the A beads dissolve into the B phase. This implies that the two phases do not exchange particles in practice; they only exchange heat and momentum.

Because the trimer liquid lacks entropic degrees of freedom as compared to the monomer liquid of the same density, the pressure of a trimer liquid is lower by ${}^2/_3\rho kT$, where ρ is the monomer density. Consequently, the equilibrium density of the B phase is higher than that of the A phase. The system is integrated using the standard velocity Verlet algorithm, using step size $\delta t = 0.05$, which is a safe step size for this potential. The pairwise Nosé−Hoover thermostat is used, based on a global temperature definition. When we start the simulation with equal densities of beads in both halves of the system, it starts off-equilibrium. In an adiabatic simulation (the natural ensemble when the thermostat is switched off), the A phase should cool under expansion, whereas the B phase should heat up. This is indeed observed to happen in the simulation. When the overall mean temperature, that is, the global temperature defined in eq 3, matches the desired temperature T_0, the thermostat force in eq 2 effectively vanishes for all particle pairs. Consequently, the thermostat does not remove any local deviations from the desired temperature, and such temperature variations can thus persist for a long time. This is illustrated in Figure 1, where the temperature is plotted as function of the x coordinate through the simulation box, for three different evolution times: $t = 14\tau_0$, $39\tau_0$, and $114\tau_0$. The temperature profiles are averaged over 450 time steps around the mean evolution times. Note that we use the units of length, mass, and energy as $r_c = m = k_B T_0 = 1$, and thus, we have the unit of time $\tau_0 = r_c(m/k_B T_0)^{1/2} = 1$.

The interpretation of this result is that the trimer phase on the right-hand side of the system is heated adiabatically early on in the simulation, and that heat is slowly spreading afterward. Thus, the temperature variation should follow Laplace's law, $\partial\theta/\partial t = D_h\nabla^2\theta$, where D_h is the heat diffusion

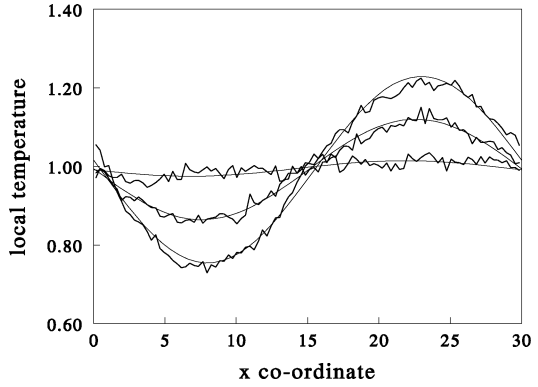


Figure 1. Temperature profile through the system at time t = 14, 39, and 114.

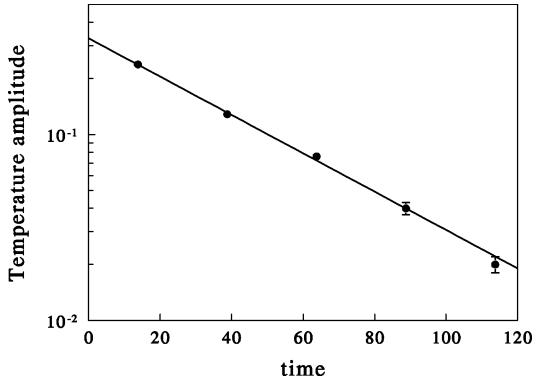


Figure 2. Amplitude of the temperature variation as a function of time. Error bars for the first points are within the size of the symbols.

constant and $\theta(x,t) = T(x,t) - T_0$. For a pure wave mode, this is solved by $\theta(x,t) = \theta_0 \exp(-t/\tau) \sin(kx)$, where $\tau^{-1} = D_h k^2$. To check this behavior, the local temperature variation is fitted to a sine function, and the amplitude is plotted as function of time in Figure 2. This shows that the amplitude indeed decays as a single exponential. From the decay time $\tau = 42 \pm 1$, we find a heat diffusion coefficient $D_h = L^2/4\pi^2\tau = 0.54 \pm 0.01$, where $L = 30$ is the system size in the x direction. This diffusion constant is comparable in size to the particle diffusion constant $D = 0.36 \pm 0.01$ and the kinematic viscosity $\nu = 0.26 \pm 0.01$ for single-particle systems with repulsion parameter $a = 25$. Although this behavior is physically quite reasonable, the problem now is that thermal equilibration in the system is essentially governed by heat transfer as far as temperature inhomogeneities are concerned. Thus, the response time depends on the system size as $\tau = L^2/4\pi^2 D_h \approx L^2/21$ for the present repulsion parameters. Hence, if a truly locally isothermal simulation is desired, the thermostat has to be adapted. This is described in the next section.

4. Local Nosé−Hoover Thermostat

As mentioned in section 2, a truly local pairwise Nosé−Hoover thermostat can be constructed on the basis of eqs 4 and 5, but this suffers from the large fluctuations in the local temperature T_i. To solve this problem, we need to define a slowly varying variable θ_i, which is a smoothed average over the local temperature T_i. Two options are open to obtain

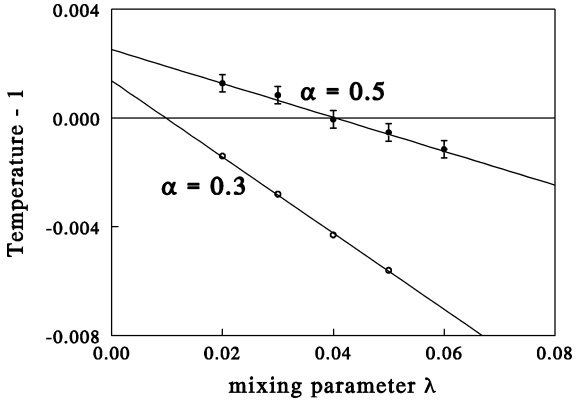


Figure 3. Temperature deviation from 1 for two values of the thermostat force, as a function of the mixing parameter.

such a slowly varying local temperature: by making a time average for each individual particle and by making a spatial average around each particle. The first option makes the thermostat slower. However, because the pairwise Nosé−Hoover thermostat based on a global temperature is an order of magnitude faster than DPD—it restores the correct temperature within a few MD steps—there is room to trade off thermostating speed for accuracy.

A time-averaged particle temperature can be defined following the update scheme:

$$\theta_i(t + \delta t) = (1 - \lambda)\, \theta_i(t) + \lambda T_i(t) \tag{7}$$

which is then substituted in the thermostat force

$$\mathbf{F}_{ij} = \alpha\psi(r_{ij}/r_c)\, [1 - {}^1/_2(\theta_i + \theta_j)/T_0][(\mathbf{v}_i - \mathbf{v}_j)\cdot\mathbf{e}_{ij}]\mathbf{e}_{ij}/\delta t \tag{8}$$

To test this thermostat, we used the weight functions $\psi(r) = \zeta(r) = (1 - r)$ for $r < 1$ and $\psi(r) = \zeta(r) = 0$ for $r > 1$; 3000 particles were simulated in a box of dimensions $10 \times 10 \times 10\ r_c^3$ for the potential given in eq 6, with $a = 25$ and time step $\delta t = 0.05$. Parameter λ in eq 7 is a mixing parameter that determines the rise time of the thermostat $t_s \sim 1/\lambda$. This is analogous to the ordinary Nosé−Hoover thermostat, eq 1. In the actual implementation, the velocity correction for each particle is stored in an array $d\mathbf{v}_i$. The sum over neighbors is done in the force loop, where for each pair of neighboring particles $\mathbf{F}_{ij}\delta t$ (from eq 8) is added to $d\mathbf{v}_i$ and subtracted from $d\mathbf{v}_j$. In the final velocity update step, $d\mathbf{v}_i$ is added to the velocity of each particle.

For $\alpha = 0.3$ and 0.5, the deviation of the temperature from its set value is shown in Figure 3. This shows that the deviations from the desired temperature are in the third digit. Moreover, the temperature decreases with λ, for both values of the thermostat force. This means that the thermostat is overdamped for too-small rise times and underdamped for too-large rise times. For a well-chosen value of the mixing parameter, the system is critically damped, which occurs at $\lambda \approx 0.01$ for $\alpha = 0.3$, $\lambda \approx 0.025$ for $\alpha = 0.4$, and $\lambda \approx 0.04$ for $\alpha = 0.5$

Obviously, we would like to have critical damping so that the temperature is exactly the desired value, and we would like to have a fast response of the thermostat. But these two requirements are incompatible within this scheme. Moreover,

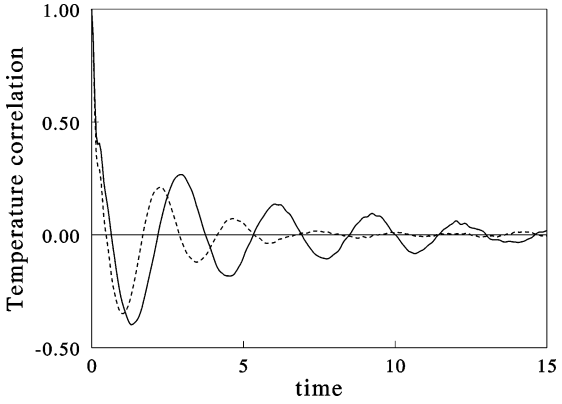


Figure 4. Temperature autocorrelation for $\alpha = 0.4$, averaged over 10^5 time steps, for $\lambda = 0.025$ and $\mu = 0.0$ (full curve) and for $\lambda = 0.05$ and $\mu = -0.01$ (dashed curve).

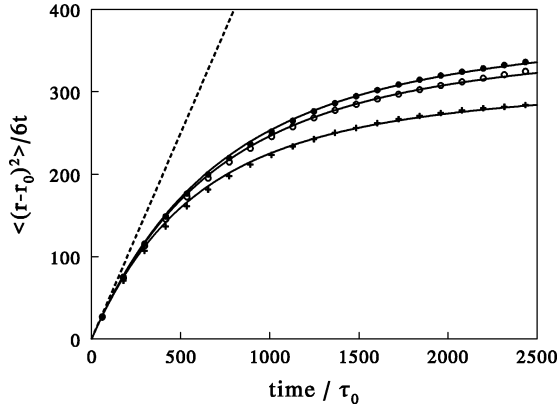


Figure 5. Mean square displacement for the ideal gas at density $\rho = 1$ (+), $\rho = 3$ (○), and $\rho = 10$ (●). The dashed curve is the ballistic result $\Delta r^2/6t = {}^1/_2 t$ for a divergent diffusion constant.

we find that the temperature autocorrelation function shows slowly decaying oscillations. An example is shown in Figure 4, for the parameters $\alpha = 0.4$ and $\lambda = 0.025$. To check if this problem is caused by oversteering, the force amplitude of the thermostat was lowered to $\alpha = 0.3$, and to maintain the same temperature, the mixing parameter was reduced to $\lambda = 0.01$. For these parameters, however, the oscillations in the correlation function do not disappear. To summarize, a straightforward generalization of the Nosé−Hoover thermostat to a purely local implementation does not lead to satisfactory results. The thermostat is typically slow and oscillatory.

To improve the performance of the thermostat, two scenarios have been studied. First, the local smoothed temperature on each particle is averaged over that of the neighboring particles to obtain a wider spatial average

$$\theta_i(t+\delta t) = (1-\lambda-\mu)\,\theta_i(t) + \lambda T_i(t) + \mu \sum_j \zeta(r_{ij})\,\theta_j(t-\delta t)/\sum_j \zeta(r_{ij}) \qquad (9)$$

The last term is an average over the local smoothed temperature of the neighbors of particle i. Mixing the local mean temperature with the average neighbor temperature has the advantage that a spatial average is generated iteratively. However, the influence of parameter μ on the mean temperature and on the temperature autocorrelation function is quite small. In practical terms, this scheme does not lead to an improvement over eq 7 for the range $0 < \mu < 1$, and the correlation time remains large. Therefore, a variation of eq 9 has been studied, where the neighbor temperature is the measured temperature of the previous time step, rather than the smoothed temperature variable:

$$\theta_i(t+\delta t) = (1-\lambda-\mu)\,\theta_i(t) + \lambda T_i(t) + \mu \sum_j \zeta(r_{ij})\,T_j(t-\delta t)/\sum_j \zeta(r_{ij}) \qquad (10)$$

This scheme works very well for $\alpha = 0.4$, $\lambda = 0.05$, and $\mu = -0.01$, that is, for a *negative* value of μ. The temperature thus obtained is $T = 0.999$, and the correlation time is $\tau = 2.1\tau_0$. This is 2.6 times faster than for $\mu = 0$, which is shown in Figure 4. For the same parameters, eq 9 leads to the same temperature, but to a longer correlation time, $\tau = 3.7\tau_0$. A negative value for μ implies that the change in the smoothed temperature variable θ_i depends not only on the deviation of the local temperature from the smoothed temperature but also on its time derivative. This can, in turn, be interpreted as the mechanical work acting on a particle, as $\mathrm{d}T_i/\mathrm{d}t \approx {}^2/_3\langle M_{ij}\mathbf{v}_{ij}\cdot\mathrm{d}\mathbf{v}_{ij}/\mathrm{d}t\rangle_i = {}^1/_3\langle \mathbf{v}_{ij}\cdot\mathbf{f}_{ij}\rangle_i$. Hence, this thermostat changes the local temperature by two physical effects: by heat flow and by mechanical work. A similar mechanical work term in the temperature update was put forward by Phares and Srinivasa.[16] In practical terms, by adding the second term in eq 10 with a negative value of μ, we gain the freedom to move the temperature up in Figure 3 and obtain the correct temperature for larger values of λ and, hence, for a shorter correlation time.

Does this mean that all problems with this thermostat are solved? A severe test is to study the ideal gas. For this system, the diffusion constant should diverge and the pair correlation should be 1. This is indeed the case with the Stoyanov−Groot thermostat, which is based on a global temperature. When a purely local temperature definition is used, however, this thermostat starts to compensate for the local temperature fluctuations that are natural within clusters of small numbers of particles. Just as with the Lowe−Anderson thermostat, this induces an interaction between the particles, which destroys their individual conservation of momentum. Consequently, the mean free path becomes finite and the diffusion constant no longer diverges. An example of this is shown in Figure 5, which gives the mean square displacement for an ideal gas at densities $\rho = 1$, 3, and 10 in a box of size $10 \times 10 \times 10$, with thermostat parameters $\alpha = 0.1$, $\lambda = 0.1$, and $\mu = 0$, over 50 000 time steps of size $\delta t = 0.05$. By fitting this to the solution of the Langevin equation,[15] $\langle [r(t) - r(0)]^2\rangle/6t = D[1 + D(e^{-t/D} - 1)/t]$, where we have put $m = kT = 1$, the diffusion constant is obtained. This gives diffusion constants $D \approx 327$, $D \approx 381$, and $D \approx 400$ for the three densities, respectively, that is, large but finite values.

As a further check, the pair correlation function was studied. This shows a slight deviation from $g(r) = 1$, as shown in Figure 6. The deviation is quite small, but this deviation indicates that the thermostat induces an artificial conservative force between the particles, just as DPD does.

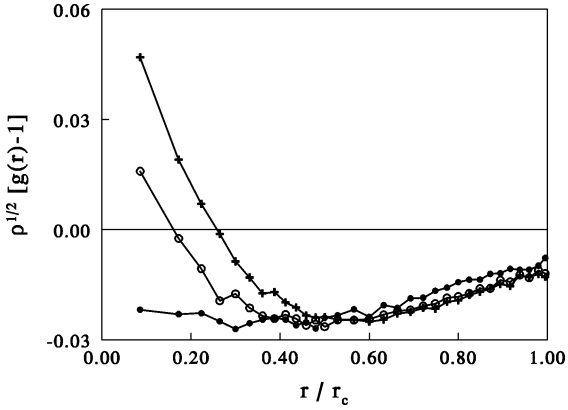


Figure 6. Total correlation function of the ideal gas at density $\rho = 1$ (+), $\rho = 3$ (○), and $\rho = 10$ (●).

For density $\rho = 10$, the maximum deviation is less than 0.009, but if for the same thermostat parameters the density is lowered, the deviations increase. This can be expected, because the mean number of neighbors of any particle is proportional to ρr_c^3. Because the relative noise in the measured temperature of each particle is inversely proportional to the square root of the number of neighbors N, and because the deviations in the pair correlation are driven by these fluctuations, one may expect that the ideal-gas total correlation function scales as

$$g(r) - 1 \propto N^{-1/2} \propto r_c^{-3/2} \tag{11}$$

Hence, for a fixed cutoff radius and variable density, we expect that $\rho^{1/2}[g(r) - 1]$ is independent of the density. A glance at Figure 6 illustrates that, for a density range varying over a factor of 10, this scaling roughly holds for large r values ($r > r_c/2$). This implies that the errors introduced by the thermostat can be made arbitrarily small by increasing the cutoff range r_c. In practical terms, a cutoff that contains some 40 neighbors leads to systematic errors of about 0.5%.

5. Discussion and Conclusions

The thermostat introduced recently by Stoyanov and Groot[15] is analyzed for inhomogeneous systems. In this thermostat, the temperature is obtained as an average over the whole system, and thus, it was anticipated that this could lead to a subtle nonlocal interaction.[15] This will not destroy hydrodynamic behavior, because momentum is locally conserved, but the consequence of such nonlocality should be investigated to ascertain the method. Here, it is found that, although the system as a whole is simulated in an *NVT* ensemble, the Stoyanov−Groot thermostat locally conserves energy rather than temperature. Consequently, local temperature variations can be long-lived, although they do average out by thermal diffusion. To obtain a faster local temperature equilibration, a truly local thermostat must be introduced. To conserve momentum and, hence, to simulate hydrodynamic interactions, the thermostat must be Galilean invariant.

Such a local Galilean invariant thermostat is introduced here. This scheme is very much comparable to Dissipative Particle Dynamics,[5−7] but the main difference is that no random numbers are used. This implies that the simulation method is faster than DPD, and because of the close similarity to DPD, it sheds light on the known artifacts in the results for the ideal gas. Generally, all thermostats produce some artifacts, and the choice of method is usually a tradeoff between the nature of the artifacts and the problem investigated.

In the present thermostat, a local temperature is defined on each particle. Thus, a locally isothermal ensemble can be constructed. The local temperature is obtained from the local mean-square velocity around each particle. To arrive at a smoothly varying local temperature, a convolution over the past square velocities is taken by mixing the actual measured value with the smoothed temperature by a few percent. This is very similar to the original Nosé−Hoover algorithm. For a fast and accurate thermostat, the local smoothed temperature must be coupled to the actual temperature and its time derivative. This can be done with an update scheme based on the actual local temperature and that of the previous time step.

Simulations show that this local Nosé−Hoover algorithm has a similar artifact as Dissipative Particle Dynamics: the pair correlation function of the ideal gas is slightly distorted. This is attributed to the fact that the thermostat compensates fluctuations that are natural within a small cluster of particles. Consequently, when the cutoff range r_c for the square velocity average is increased, systematic errors decrease proportionally to $r_c^{-3/2}$. In practice, however, the error is on the order of a few tenths of a percent.

As mentioned above, the choice for a particular thermostat depends on a tradeoff of requirements. If it suffices to simulate in the local *NVE* ensemble with a global temperature control, the Stoyanov−Groot thermostat can be used. One special feature of this thermostat is that the fluid viscosity can be given any desired value, by switching between a low-viscosity local Nose−Hoover thermostat and a high-viscosity exchange process with probability P.[15] For $P = 0$, the ideal gas has a divergent diffusion coefficient and zero viscosity, while for $P > 0$, it has a finite diffusion coefficient. Moreover, the temperature deviation from the required system temperature is at least an order of magnitude smaller than that for standard DPD, while the equilibrium properties of the system are very well reproduced. At the same time, it is computationally more efficient than self-consistent DPD,[12] by offering better temperature control and greater flexibility in terms of adjusting the diffusion coefficient and viscosity.

The same flexibility in simulated fluid viscosity is inherent in the present Galilean invariant thermostat, which simulates a truly locally isothermal system, but at the price of a slightly reduced accuracy in the temperature control. The presented thermostat can be applied in any conventional particle-based molecular dynamics simulation, including atomistic force fields. As a generalization, in a two-phase system, one can choose three different exchange frequencies P_{11}, P_{22}, and P_{12}. Thus, if phases 1 and 2 are mutually insoluble and if $P_{12} > P_{11} \geq P_{22}$, parameter P_{12} will introduce an excess surface viscosity of the resulting interface. With an extra surface-active component in the system, this would open up a new class of simulations where both bulk and surface viscosities and diffusion coefficients could be adjusted at will. Finally, simulation of dynamic intrusion and viscous fingering, as

well as spinodal decomposition in fluids of largely different viscosities, is straightforward.

References

(1) Koga, T.; Kawasaki, K. *Physica A* **1993**, *196*, 389−415.

(2) Qi, S.; Wang, Z. G. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* **1997**, *55*, 1682−1697.

(3) Shan, X. W.; Chen, H. D. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* **1993**, *47*, 1815−1819.

(4) Swift, M. R.; Orlandini, E.; Osborn, W. R.; Yeomans, J. M. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* **1996**, *54*, 5041−5052.

(5) Koelman, J. M. V. A.; Hoogerbrugge, P. J. *Europhys. Lett.* **1993**, *21*, 363−368.

(6) Español, P.; Warren, P. *Europhys. Lett.* **1995**, *30*, 191−196.

(7) Groot, R. D.; Warren, P. B. *J. Chem. Phys.* **1997**, *107*, 4423−4435.

(8) Serrano, M.; Español, P. *Phys Rev E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* **2001**, *64*, 046115.

(9) Malevanets, A.; Kapral, R. *J. Chem. Phys.* **1999**, *110*, 8605−8613.

(10) Nosé, S. *J. Chem. Phys.* **1984**, *81*, 511−519.

(11) Hoover, W. G. *Phys. Rev. A: At., Mol., Opt. Phys.* **1985**, *31*, 1695−1697.

(12) Pagonabarraga, I.; Hagen, M. H. J.; Frenkel, D. *Europhys. Lett.* **1998**, *42*, 337−382.

(13) Andersen, H. C. *J. Chem. Phys.* **1980**, *72*, 2384−2393.

(14) Lowe, C. P. *Europhys. Lett.* **1999**, *47*, 145−151.

(15) Stoyanov, S. D.; Groot, R. D. *J. Chem. Phys.* **2005**, *122*, 114112.

(16) Phares, D. J.; Srinivasa, A. R. *J. Phys. Chem. A* **2004**, *108*, 6100−6108.

CT050269E

JCTC Journal of Chemical Theory and Computation

Exploring Model Energy and Geometry Surfaces Using Sum of Squares Decompositions

Martin G. Burke and Sophia N. Yaliraki*

Department of Chemistry, Imperial College London, South Kensington Campus, London SW7 2AZ, United Kingdom

Received December 30, 2005

Abstract: The difficulty in exploring potential energy surfaces, which are nonconvex, stems from the presence of many local minima, typically separated by high barriers and often disconnected in configurational space. We obtain the global minimum on model potential energy surfaces without sampling any minima a priori. Instead, a different problem is derived, which is convex and hence easy to solve, but which is guaranteed to either have the same solution or to be a lower bound to the true solution. A systematic way for improving the latter solutions is also given. Because many nonconvex problems are projections of higher dimensional convex problems, Parrilo has recently shown that by obtaining a sum of squares decomposition of the original problem, which can be subsequently transformed to a semidefinite program, a large class of nonconvex problems can be solved efficiently. The semidefinite duality formulation also provides a proof that the global minimum of the energy surface has either been found exactly or has been bounded from below. It additionally provides physical insight into the problem through a geometric interpretation. The sum of squares polynomial representation of the potential energy surface may further reveal information about the nature of the potential energy surface. We demonstrate the applicability of this approach to low-dimensional potential energy landscapes and discuss its merits and shortcomings. We further show how to apply it to geometric problems by obtaining the exact distance of closest approach of anisotropic particles. Efficient molecular dynamics simulations of mixtures of ellipsoids are illustrated.

Introduction

The thermodynamic and kinetic properties of molecular assemblies in chemistry, physics, and biology are often linked to their underlying potential energy surface (PES).[3−5] Even in the simplest cases, exploring potential energy surfaces or even just important regions becomes computationally demanding very rapidly. This is primarily due to the nonconvexity of the underlying potential energy surface. A key step in understanding thermodynamic properties of systems as diverse as nanoscale clusters, proteins, and lipid membranes is the identification of the global minimum. Definite proof typically requires exhaustive sampling of every other minimum in configurational space, a prohibitive task given that the number of local minima on the surface typically grows exponentially with the number of particles in the system.[3,6]

A series of approaches has been proposed to tackle this problem with various degrees of success. Most stochastic methods are variations of the Monte Carlo (MC) method with additional steps introduced to aid barrier crossing between local minima,[7−9] although genetic-based algorithms[10,11] have also been used. Deterministic approaches usually derive from global optimization techniques[12−14] or molecular dynamics.[15] Methods that deform the potential in order to eliminate local minima include hypersurface deformation techniques,[16] such as the distance scaling method[17] and the very successful MC basin-hopping method.[18]

A commonality in most methods is that they require extensive sampling of the local minima of the potential energy surface itself prior to finding the global minimum.

* Corresponding author e-mail: s.yaliraki@imperial.ac.uk.

10.1021/ct050338p CCC: $33.50
Published on Web 04/26/2006

Not only is this process time-consuming, sensitive to initial conditions, and susceptible to trapping due to high barriers, it crucially provides no proof that a given point on the PES is indeed the global minimum other than that it is the lowest visited to date by the algorithm. It is usually not possible to evaluate how close such a minimum would be to the correct answer to further guide the search. Furthermore, all methods rely to a greater or lesser extent on the topology of the PES being "well behaved" in the sense that the global minimum is connected to other low lying local minima so that it can be reached from them.

A complementary strategy is to abandon the detailed atomic potentials and replace them instead with simpler models. Such coarse-grained models have a rich history and have had successes in recent years in the modeling of both proteins and lipid membranes (for recent reviews, see, for example, refs 19 and 20). These vary from simple to more detailed and have been successful in predicting structural properties as well as thermodynamic phase behavior. The successes of these models seem to share in common an accurate representation of the architecture and shape of the protein and/or lipid. Evidence is accumulating that shows that once this is well represented many of the properties of the units and their assemblies can be captured especially when their mechanical properties play a role.[21,22] At the same time, the shape of the units is often anisotropic. This may be more obvious in lipid molecules, but it turns out to be often the case in proteins as well. In describing such shapes, anisotropic coarse-grained potentials have an advantage over spherical ones in that they capture the interactions properly without introducing artificial long-range potentials or other schemes. Anisotropic potentials have a long history,[23−27] but their use has often been hampered by computing the distance of closest approach correctly and efficiently, a cumbersome task often solved by brute force. It would hence be desirable to capture the shape properties correctly and at the same time have a systematic way to check how close the approximation is from the true answer, that is to have an estimate of the error in comparison with the true problem.

Here we outline a different strategy to address such questions arising from exploring complex (nonconvex) landscapes. A key idea is that instead of sampling the underlying nonconvex function, we approach it from below. A different problem is sought, whose *solution* coincides with the solution of the problem we are interested in. For this approach to be successful, it must be coupled both with a rigorous and systematic way to find such problems that can be easily solved and with a way to evaluate how close the two solutions are. An important mathematical concept that provides such a link is that many nonconvex problems turn out to be projections of higher dimensional convex problems. The great advantage of the latter is that they are efficiently solved since the local minimum is the global minimum. Furthermore, they also allow for physical insight into the system because they possess duality properties. This duality provides geometric interpretations to the original problem and allows us to monitor how close the solution of the derived problem is to the true solution we are seeking. This is coupled with a systematic way to improve on our nonexact solutions. The combination of these features makes this approach appealing because physical insight is revealed about the system while seeking the solution itself.

The approach, originally developed independently by Parrilo[1,2] and Lasserre[28] by drawing together results from two seemingly unconnected fields, algebraic geometry and convex optimization, is based on obtaining a sum of squares (SOS) representation of the original problem. They observed that many nonconvex problems can be represented by finding a suitable globally non-negative polynomial function which is an accurate representation of the nonconvex problem in Euclidean space in which the problem is originally defined. Parrilo and Lasserre realized that although this function must by necessity be nonconvex in the Euclidean space, if well chosen, it may be convex in the vector space of monomials which form a basis for the polynomial. The obvious difficulty is how to choose such a function. Searching the entire space of globally non-negative polynomials is an impossible task; however, the space of SOS functions forms a significant subset of this space. A SOS polynomial is defined as any polynomial that has the following decomposition

$$f^{\mathrm{SOS}} = \sum_{i=1}^{r} (p_i(\mathbf{x}))^2 \geq 0 \qquad (1)$$

where the $p_i(\mathbf{x})$ are polynomials. Suitable subsets of the space of SOS functions can now be found because this search turns out to be convex in the coefficient space of the function and hence efficiently computed using modern convex optimization techniques such as semidefinite programming (SDP). An interesting question is what happens if a suitable SOS cannot be found. The answer, based upon an important theorem in algebraic geometry,[2] is that a lower bound will always be produced in a minimization problem of the type described here, and furthermore this bound can be improved asymptotically to the exact minimum by increasing the complexity of the SOS polynomial in a specific and systematic manner. A central idea is the systematic lifting of a polynomial function in its *coefficient* space (rather than its variables) to a SOS function.

Here we apply this approach in two different but important problems in molecular applications. First, in section II we give the main concepts of the theory and its relation to physical observables by showing how to locate the global minimum in a complex landscape without sampling other minima in configurational space. As a result, it is independent of initial conditions and of any requirement that the minimum be closely related to any other. Either the exact result is found together with a proof, or a lower bound is obtained. In no case is the original function deformed. In the Appendix, we show how to use the method, both in theory and in practice, to obtain a SOS decomposition. By going through each of the steps of an example, we show how the method is implemented and comment on possible choices available. In section III we apply it to a series of low-dimensional examples that exemplifies a range of difficulties encountered in complex landscape problems and which highlight the different capabilities and versatility of this method. In section II.B we show how the SOS approach can be further used to

elegantly and efficiently solve exactly the problem of distance of closest approach of anisotropic particles and interactions. We further show how this can potentially lead to efficient MD simulations by giving an illustration with ellipsoid particles. In section IV we discuss the advantages and shortcomings of the approach and outline future directions.

II. Theory

A. SOS Decompositions. A different approach to locating the global minimum U^* of a multivariable polynomial, $U(x_i,...x_n) \equiv U$ of degree d is to look for the largest real number λ such that $U - \lambda$ remains non-negative everywhere:

$$U(\mathbf{x}) - \lambda \geq 0, \quad \lambda \in \mathbf{R} \quad \forall \mathbf{x} \in \mathbf{R}^n \tag{2}$$

If we could find the largest λ for which this is true, then we would have found the global minimum of $U(\mathbf{x})$. In this formulation, this problem is at least as difficult to solve as the original problem of locating the global minimum. However, if instead of insisting that the above condition be satisfied, we ask instead that $U(\mathbf{x}) - \lambda$ can be decomposed as a SOS,[1,2,29] that is replace eq 2 with the following condition

$$U(\mathbf{x}) - \lambda \geq f^{\mathrm{SOS}}, \quad \lambda \in \mathbf{R}, \quad \forall \mathbf{x} \in \mathbf{R}^n \tag{3}$$

then the problem becomes considerably easier to solve. Using the definition of eq 1 for a SOS function leads to

$$U - \lambda = \sum_{i=1}^{r} p_i^2 \geq 0 \tag{4}$$

where p_i are polynomial functions of degree $d/2$ in the variables of U.

Being a SOS is a sufficient condition for a function to be non-negative, although not a necessary one. However, the set of SOS polynomials forms a large and significant subset of the set of non-negative polynomials and for reasonably low degree polynomials they are nearly identical. Therefore replacing the non-negativity condition of eq 2 with the SOS condition of eq 4 is a relatively minor relaxation to the problem. When the optimal λ is not the exact global minimum of the function, then it is a lower bound on the global minimum. In fact in a case where only a lower bound is obtained it is possible to improve that bound asymptotically toward the global minimum by increasing the dimensionality and/or degree of the sum of squares relaxation as we will show later. Indeed, both λ and f^{SOS} are guaranteed to bound the original problem from below.[29]

The next step then becomes obtaining the SOS decomposition together with the optimal λ that gives the closest answer to our original problem. The problem becomes tractable by additionally exploiting the theorem[1] that for globally non-negative polynomial functions, there exists a positive semidefinite matrix, $Q \geq 0$, such that

$$U(\mathbf{x}) - \lambda = f^{\mathrm{SOS}} = z(x)^T Q z(x), \quad \text{where } Q \geq 0 \tag{5}$$

so that the problem of eq 4 can now be formulated algebraically. $z(x)$ is the vector of monomials of U with degree less than or equal to half the degree of U and provides a basis for the SOS function decomposition. Monomials are the individual terms (without the coefficients) which when summed together make up a polynomial. Assuming the largest degree monomial in a polynomial is of even degree $2d$, then the monomials which are needed in the basis vector z are all monomials on x_1, ..., x_n of degree $\leq d$. It can be seen that the size of the monomial vector is therefore at most $\binom{n+d}{d}$. Since the variables in z are not algebraically independent, the matrix Q is not unique and may be positive semidefinite in some representations but not in others. The problem now becomes a search for a symmetric positive semidefinite matrix Q that satisfies the constraints imposed by eq 5. Note that the original nonconvex problem in the variables of U has been mapped to a different problem, namely obtaining the polynomial coefficients of the SOS representation, which is equivalent to obtaining the matrix elements of Q, while Q remains positive semidefinite and λ is optimal.

Formally the SOS optimization problem can then be written as

$$\begin{aligned} &\text{Max}: \ \lambda \\ &\text{Subject to}: \ Q \geq 0 \\ &\qquad \mathbf{Tr}(A_i Q) = b_i \ \forall i = 1, ..., m \end{aligned} \tag{6}$$

where b_i are the coefficients of the energy function $U - \lambda$, and the linear constraints on the components of Q are to ensure that the relevant entries in Q sum up to the coefficients of the polynomial. A concrete example of how the A_is appear is given explicitly in the Appendix.

B. SDP and Duality. It turns out that eq 6 is in the form of a primal SDP. SDPs[30,31] are a class of optimization problems over positive definite matrices solvable by polynomial time algorithms such as the very successful primal-dual interior point methods.[32,33] Such methods compute the optimal solution as well as provide certificates (proofs) of optimality. These certificates are based upon a duality theory in which two problems are solved simultaneously. The dual problem solution provides a lower bound on the primal problem. The difference between the primal and dual solutions, referred to as the duality gap, proves whether an exact solution has been found or provides the lower bound to the solution.

We have shown how the original polynomial minimization can be recast as the SDP given by eq 6. In the mathematical community the focus is typically on the primal, with the dual only used to provide the duality gap. However, as we show here there is relevant and interesting physics in the dual problem as well. In our case, the primal problem gives the lower bound to the global minimum as well as the SOS decomposition, while the dual provides the location of this bound in the configurational space.

What has been described so far can be regarded as the primal form of a SOS optimization, in which we have sought to bound the global minimum from below and hopefully exactly. In this primal form the basis has been the monomials that define the polynomial and the variables in looking for the sum of squares decomposition have been the coefficients.

There is also a corresponding dual form of SOS optimization. One way to consider the dual problem is that the basis is now a vector of coefficients and the variables are the values of the monomials. The new variables form the matrix $Y := zz^T$. The effect of this is that if sufficient monomials have been included in z the convex representation of the function is obtained and the coordinates of the global minimum can be found by minimization. Once a solution is obtained, it is possible to confirm if it is correct by verifying whether it is of rank 1.

Both the primal and the dual SOS optimization problems are exactly primal and dual semidefinite programs, respectively, and may be solved using polynomial time primal-dual interior point methods. This latter problem, although in higher dimensions, is now convex and amenable to efficient solutions by SDP. Additionally, the minimization of the potential is carried out concurrently with the search for the SOS coefficients. The theoretical elegance of this methodology coupled with efficient computational techniques allows for SOS decompositions to be found in polynomial time (as long as either the degree of the polynomial or the number of variables are fixed) and additionally provides the proof that the result is indeed a lower bound to the global minimum. If the answer is not exact, then a lower and upper bound on the true solution is known.

C. Convexity and Duality. Duality appears in many places in chemical physics, albeit without this name. It is intimately connected with convexity.[34] Multidimensional convex functions obey a range of useful properties, one of which is that at each point there is at least one hyperplane tangent to the graph, whose slope is the derivative of that function. In fact, this plane belongs to a family of planes that separate the entire configurational space in two parts, below and above the function. Convexity ensures that they never cross the function in more than one point. To demonstrate how duality is related to this, we restrict ourselves to a one-dimensional parabola $f(x) = x^2$, but the argument is completely general. The separating planes now become straight lines (Figure 1). There is a set of lines with slope y which lie below the parabola as long as $yx - d \leq x^2$ where d is the depth. The line will touch tangentially the parabola when $d = 1/4y^2$. Each point on the parabola can be represented as a line which touches the parabola with different slopes y. The envelope of these lines reproduces the parabola itself. In fact, by looking for the maximum slope y, we obtain simultaneously $f(x)$. This is in fact what the Legendre transforms do in thermodynamics where, for example, the Gibbs and Helmholtz free energy are coupled with pressure and volume, respectively. This interpretation was already recognized by Tisza.[35] The dual emphasizes the geometric interpretation of the problem, and in that sense its variables can be interpreted as generalized Lagrange multipliers. It further clarifies how we can obtain information about a function by looking *outside* and bounding it externally from below.

D. Different SOS Representations − Beyond Simple Functions. In many cases what makes finding the global minimum energy configuration difficult is as much related to the constraints imposed on the system as to the energy surface itself. In these cases it is necessary to include such constraints in the initial statement of the optimization problem. We shall call these constraints necessary constraints. There is also another type of constraint that is often used in optimization, called redundant constraints. These are constraints which repeat what is already stated in another constraint in the problem statement. Although appearing useless at first glance these play a very important role in duality theory in that they bring the lower bound produced by a dual problem closer to the exact solution. By adding these constraints which appear useless in the primal form of the problem, additional variables are added in the dual (lifting) which allows a better bound to be found on the primal objective function.

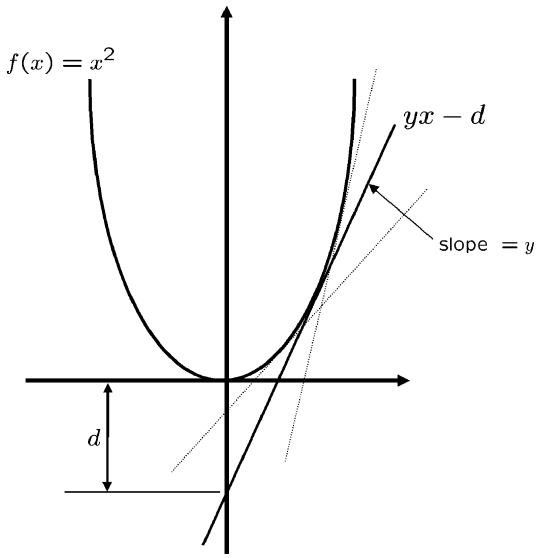


Figure 1. How convexity and duality are related. A set of lines with slope y lie below the parabola $f(x) = x^2$, as long as $yx - d \leq x^2$ where d is the depth. The line will touch tangentially the parabola when $d = 1/4y^2$. Each point on the parabola can be represented as a line which touches the parabola with different slopes y. The envelope of these lines reproduces the parabola itself. In fact, by looking for the maximum slope y, we obtain simultaneously $f(x)$. The relationship between convexity and duality appears in many places in chemical physics. An example of such a relationship are the Legendre transforms in thermodynamics.

Once again the analogy carries to SOS optimization. Additional constraints which initially may have appeared to add complexity to the problem allow the formation of progressively better SOS representations which although being higher in dimension, form better representations of the original function when projected back to the original space. Given a problem

$$\begin{aligned} &\min: U(\mathrm{x}) \\ &\text{subject to}: g_i(\mathrm{x}) \geq 0 \; i = 1, ..., m \end{aligned} \tag{7}$$

we can recast this problem as the following SOS problem

$$\begin{aligned} &\max: \lambda \\ &\text{subject to}: U(\mathrm{x}) - \lambda + \sum_{i=1}^{m} p_i(x) g_i(\mathrm{x}) \geq 0 \end{aligned} \tag{8}$$

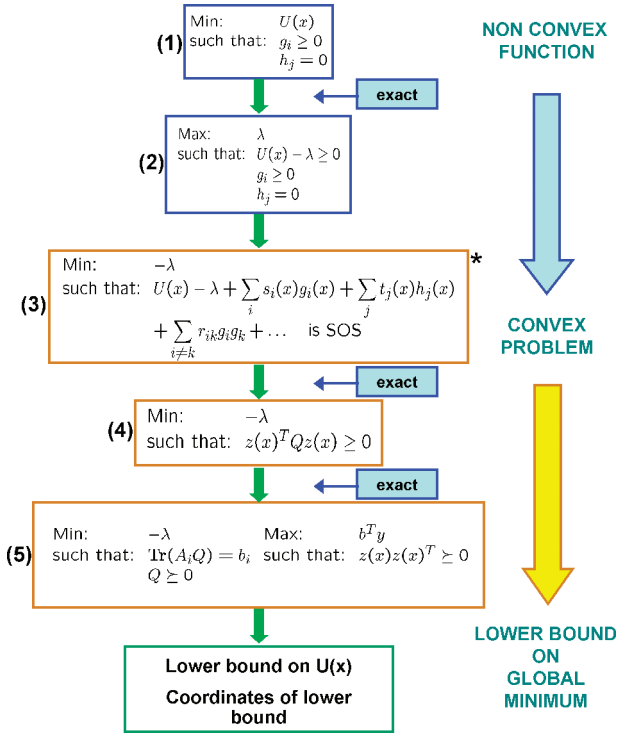


Figure 2. Schematic flowchart of a SOS optimization. The original global energy minimization problem is reformulated in a series of steps: (1) a polynomial optimization including any constraints is transformed to (2) a maximization of a real number λ which will keep the function non-negative. A sufficient but not necessary condition for this to be true is that the function is a SOS, so the problem is rewritten as (3). The constraints can be incorporated as well, and algebraic geometry guarantees the validity of this equation. The additional terms provide a series of hierarchical approximations to the true solution. (4) A function being an SOS is equivalent to finding a positive semidefinite matrix Q where z is a vector of monomials of U. Note that Q relates the coefficients of U and contains λ. Finding Q is equivalent to the semidefinite program given in (5), with the primal (left) and dual (right) problems. Each step is an exact reformulation of the previous apart from (3) which is guaranteed to give an optimal bound to the original problem. When expressed in the appropriate monomial basis the problem of locating the coefficients of Q becomes convex and hence efficiently solvable. However, the SOS function (3) is still a nonconvex function in configuration space and can be a good representation of the original function.

where $p_i(\mathbf{x})$ are undetermined SOS polynomials. These SOS polynomials can be thought of as generalized Lagrange multipliers in the space of non-negative polynomials. By fixing an upper bound on their degree we are again presented with having to solve a semidefinite program. Such representations allow for significantly more elaborate sum of squares relaxations to be found in a systematic manner. Indeed such liftings can be carried out until an exact solution has been obtained or the SDP becomes too large for the computer to solve.

The steps in the methodology are summarized in the flowchart in Figure 2. In the Appendix we show how the method works in practice through a simple example.

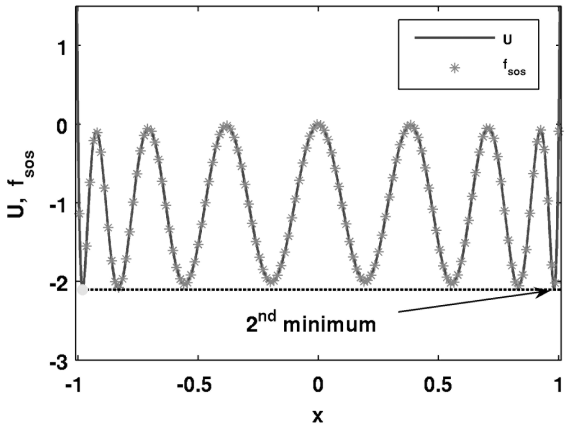


Figure 3. SOS optimization on a multiwell potential $U(-)$ with 8 minima, the leftmost being the global minimum. The second minimum, farthest away in domain space, is very close in value to the global minimum. The global minimum is correctly identified. The SOS function $f^{SOS}(*)$ is in excellent agreement with U. Here, $U = 0.01x - 128.1x^2 + 2688x^4 + -21504x^6 + 84480x^8 - 180224x^{10} + 212992x^{12} - 131072x^{14} + 32768x^{16}$.

Results

A. Global Minimum. To demonstrate the SOS approach in physical applications, we apply it now to model potentials that highlight the versatility of the method in dealing with typical aspects of landscapes that make them difficult to study. Once the problem is formulated theoretically as described above, we use third party Matlab toolboxes to carry out the SOS decomposition: SOSTOOLS, a tool for producing the SOS decompositions of polynomial optimization problems[36] and SeDuMi[37] or SDPT3,[38] SDP solvers which use primal-dual interior point methods. Other minimizations for comparisons were implemented in the Matlab optimization toolbox.

1. Multiwell Potential. To demonstrate the method we apply it first to a potential with many minima whose values are very close to the global minimum but are separated by high barriers (see Figure 3). The typical ratio of barrier height/local minima energy difference is 80. In addition, the minimum closest in energy to the global minimum is furthest apart in domain space. This type of poor connectivity behavior is not uncommon in potential energy surface problems where a number of nodes on the disconnectivity graph[39] have energy levels that are almost identical.

The SOS approach provides the global minimum which in this case is proven to be exact from the duality gap and correctly identifies its location in the coordinate space (Table 1). Additionally, the SOS decomposition we obtained is a faithful representation of the original function (Figure 3). A general feature of the SOS optimization, once formulated theoretically, is that it is carried out once, unlike other methods which may depend on a good initial guess or on the location of the initial conditions, and it provides all the information simultaneously.

We have compared to two other common minimization approaches: the downhill simplex method[40] and the BFGS quasi-Newton scheme,[41] which is used in the basin-hopping algorithm. Unlike SOS, both methods were applied iteratively from a grid of starting points spanning the domain space of

Table 1. Summary of SOS Optimization Results for the Multiwell, Müller, and Golflike Potentials[a]

	1D multiwell	10D multiwell	Müller	golflike
CPU time (s)	0.65	6840.00	10.06	0.80
lower bound	−2.106 [−2.106]	−193.685	−146.674 [−146.674]	−3.539 [−3.537]
configuration	−0.980	see below (*)	(−0.557,1.449)	(4.994,3.995)
duality gap	3.857E-06	7.56E-04	1.366E-4	0.002
primal	2.105998	193.6848	146.6743	3.537
dual	2.106002	193.6855	146.6745	3.539

[a] The CPU time in a standard PC in seconds, the lower bound value with the exact global minimum in brackets, the configurational coordinates of the minimum, the duality gap, and the final values of the primal and dual problem. The 10-dimensional multiwell example was defined as follows: $U = \sum_{i=1}^{10} a_1 x_i + a_2 x_i^2 + a_3 x_i^3 + a_4 x_i^4 + a_5 x_i^5 + a_6 x_i^6$ where $a_1 = 2$, $a_2 = 230$, $a_3 = -28$, $a_4 = -1000$, $a_5 = 1$, $a_6 = 1000$ and the global minimum of −193.685 was located among approximately 59 000 minima at (*)[0.7369, 0.7369, 0.7369, 0.7369, 0.7369, 0.7369, 0.7369, 0.7369, 0.7369, 0.7369].

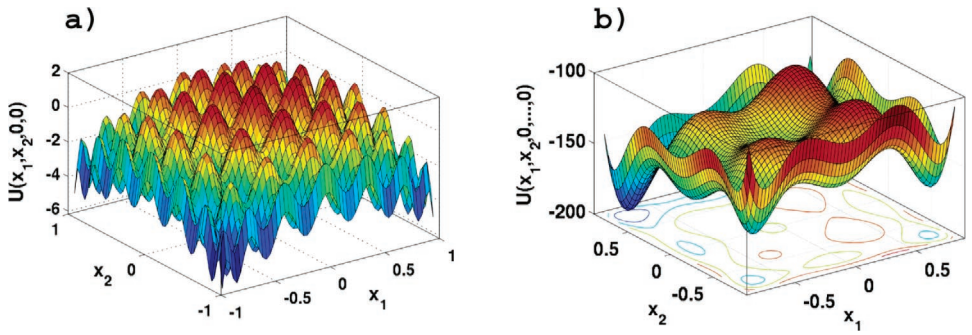


Figure 4. (a) A 2D projection of a four-dimensional 16th degree polynomial multiwell potential with 4096 minima given by $U(x_1, x_2, x_3, x_4) = \sum_{i=1}^{4} a_0 + a_1 x_i + a_2 x_i^2 + a_3 x_i^4 + a_4 x_i^6 + a_5 x_i^8 + a_6 x_i^{10} + a_7 x_i^{12} + a_8 x_i^{14} + a_9 x_i^{16}$ where $a_0 = 1$, $a_1 = 0.5$, $a_2 = -130.1$, $a_3 = 2688$, $a_4 = -21504$, $a_5 = 84480$, $a_6 = -180224$, $a_7 = 212992$, $a_8 = -131072$, $a_9 = 32768$. SOS optimization sucessfully identified the global minimum of $U = -13.977$ to within 0.026 at the point [−0.9792, −0.9792, −0.9792, −0.9792]. (b) A 2D projection of a 10-dimensional 6th degree polynomial multiwell potential given by $U(x_1, x_2, ..., x_{10}) = \sum_{i=1}^{5} a_1 x_i + a_2 x_i^2 + a_3 x_i^3 + a_4 x_i^4 + a_5 x_5 + a_6 x^6 + \sum_{i=6}^{10} b_1 x_i + a_2 x_i^2 + a_3 x_i^3 + a_4 x_i^4 + a_5 x_5 + a_6 x^6 + \sum_{j=i+1,...,10}^{i=1,...,9} r1_{ij} x_i x_j - r2 x_i^2 x_j^4$ where $r1_{ij}$ and $r2_{ij}$ are different numbers chose randomly between 0 and 10. The global minimum (−196.36) is located at [−0.7237, 0.7468, 0.7431, 0.7487, −0.7226, 0.7477, 0.7463, 0.7491, 0.7531, −0.7319]. There are 45 different projections onto different pairs of coordinates. The duality gap is $-2.316E-3$.

the problem with the grid density being increased until the global minimum was obtained. To succeed in finding the global minimum, at least 10 separate runs from starting points on an evenly spaced grid of points spanning the domain were required. This is in fact more attempts than the number of minima in the domain. The point however remains that these algorithms have no way of establishing that the minimum that they locate is in fact the global minimum.

We note that there is nothing special about the form of the potential presented here. We have tried successfully several multiwell examples with different polynomial representations, more asymmetric wells, and separations with identical success. Interestingly, the method can identify the global minimum even if the wells are nearly degenerate with arbitrarily close energy. In fact, a true degeneracy is reflected in the eigenvalues of the dual solution matrix and can hence be identified and located in the dual solution matrix. This can be very useful in locating often missed degenerate minima by looking at the dual.

The formalism holds for multidimensional problems, and so a generalization of this problem to higher dimensions works in the same fashion. The global minimum ($U = -14.036$, location = [−0.9798, −0.9798, −0.9798, −0.9798]) is identified among on the order of 4100 densely spaced minima (Figure 4 a) of the four-dimensional version of the same potential. At the same time, a 10-dimensional version with now disconnected minima dispersed in a higher 10-dimensional space (of the order of 59 000 minima) worked equally well for uncoupled (Table 1) or coupled variables (Figure 4b).

2. Golflike Potential. A different set of energy landscapes exhibit flat regions with deep and narrow wells which become especially challenging and rapidly intractable for searching algorithms. The reason is that the flat regions provide no energy differences that can guide any search. Exhaustive searching is the only option, which becomes intractable very rapidly as the domain increases. We were able to successfully locate the global minimum of a delta function embedded in a two-dimensional infinite region. Such an example is shown in Figure 5 where the function is almost flat everywhere except in two locations where almost identical delta function wells are present. Note that we have included the entire $\mathbf{R}^2$ domain. The SOS approach correctly identifies the minimum and its location (Table 1). This is also true when the higher energy well becomes wider. The power of this approach comes from the following general simultaneous reasons: (i) the solution is always approached from below so the flatness of the function does not enter, (ii) the minimum is very narrow in configurational space, however, in the lifted space of the coefficients this is no longer necessarily true, and (iii) the set of equations linking the coefficients of the function intersect the cone of positive

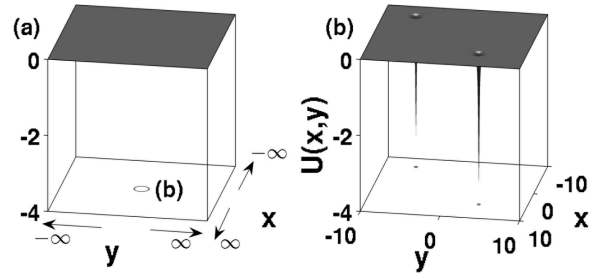


Figure 5. Plot of a golflike potential. U is given by $(-1/\pi)((e_1/((x-5)^2+(y-4)^2+e_1^2))+(e_2/((x+7)^2+(y+6)^2+e_2^2)))$, $e_1 = 0.09$, $e_2 = 0.1$ defined over the entire $\mathbf{R}^2$ domain (left). The area close to the origin (blue circle) is reproduced at greater magnification on the right. The two minima are $U(5, 4) = -3.5369$ (global) and $U(-7, 6) = -3.1832$. In this case the SOS optimization produces a lower bound of $U = -3.539$. The configurational coordinates are obtained correctly to within numerical accuracy.

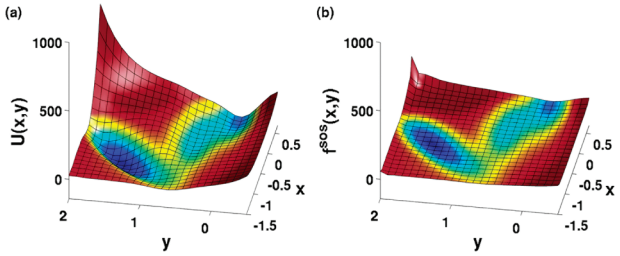


Figure 6. Contour plot of the original fitted Müller potential (a) and its SOS representation (b). The global minimum was identified correctly. f^{SOS} (b) is the projection to the original domain space from the lifted space of the SOS function including the inequality constraints. Much of the original function is preserved within the first order of eq 3, Figure 2.

definite matrices. The possible solutions are now on the boundary of this space and no longer on the interior and hence easy to distinguish near degeneracies even if disconnected in configuration space.

3. Müller Potential. The Müller potential[42] is commonly used[43,44] as a nontrivial test for reaction path methods due to the complexity of its minima structure that leads to a highly contorted reaction coordinate which requires a sharp change of direction at the saddle. The potential surface (Figure 6) contains three minima, with the global minimum of -146.7 at $(-0.558, 1.442)$

$$U(x, y) = \sum_{i=1,\ldots,4} A_i \exp[a_i(x - x_i)^2 + b_i(x - x_i)(y - y_i) + c_i(y - y_i)^2] \tag{9}$$

where $A = (-200, -100, -170, 15)$, $a = (-1, -1, -6.5, 0.7)$, $b = (0, 0, 11, 0.6)$, $c = (-10, -10, -6.5, 0.7)$, $x_i = (1, 0, -0.5, -1)$, $y_i = (0, 0.5, 1.5, 1)$. Its domain is constrained in order to exclude areas where the function is not well behaved. The SOS approach allows for constraints, both equality and inequality, to be incorporated on an equal footing as can be seen in eq 3 of Figure 2. We can take advantage of this powerful tool here by excluding parts of space that are not of interest in the formulation of the problem. The definition of this function includes an exponential so in order to apply the method here we fitted the function in a least squares sense to a polynomial of degree 16, U_f, which contains 153 monomials, 67 of which are odd. (The RMS error normalized by the data range was 0.0015 over the points sampled.) This leads to the minimization of U_f subject to a series of inequality constraints, g_i

$$\begin{aligned} &\max: \lambda \\ &\text{subject to}: U_f - \sum_{i=1}^{4} s_i g_i - \lambda \geq 0 \\ &g1: x \geq -1.5 \; g2: x \leq 0.7 \\ &g3: y \geq -0.5 \; g4: y \leq 2.0 \end{aligned} \tag{10}$$

where s_i are additional SOS functions with unknown coefficients.

The SOS approach succeeded in locating the minimum which is proven to be exact by the duality gap. It is interesting to point out that the final SOS decomposition of the potential is a good representation of the original function (see Figure 6) which may allow for additional properties of the function to be explored in different representations. For example, we were able to correctly locate the elusive saddle through eigenvector following on the SOS function. If more faithful representations were sought, a better bound to the function can be produced by systematically augmenting the dimension of the basis (eq 3, Figure 2) at higher computational cost.

B. Geometry: The Distance of Closest Approach of Anisotropic Coarse-Grained Molecular Units. 1. Formulation. One of the primary challenges in coarse-grained approaches to modeling of complex systems is to produce realistic interparticle potentials and, in particular, to incorporate the geometry or shape of the coarse-grained particles into such a representation of the physical system. In the spherical case this is straightforward as the potential remains isotropic after coarse graining and the interactions are simply functions of the interparticle radial distance. In the general nonspherical case, the interaction can be computed as a force acting at the points of minimum distance between the particles and in the direction of the minimum distance vector. This force can then be resolved into a radial part acting on the center of mass of the particles and a torque acting about the center of mass. The principal difficulty with this approach is how to compute the minimum distance between particles of an arbitrary shape.[25−27]

To formalize this we represent two arbitrary geometrical shapes as $\mathscr{A}$ and $\mathscr{B}$ in the real space $\mathbf{R}^n$, and we define the minimum distance problem as finding $a \in \mathscr{A}$ and $b \in \mathscr{B}$ such that the distance $||a - b||_2$ is a minimum:

$$\begin{aligned} &\min: ||a - b||_2 \\ &\text{such that}: a \in \mathscr{A}, b \in \mathscr{B}. \end{aligned} \tag{11}$$

Finding the minimum distance of closest approach between $\mathscr{A}$ and $\mathscr{B}$ can be formulated as a SOS problem, that is we

search for a function such that (see Figure 2, eq 3)

$$||a - b||_2 - s_1(a, b)g_1(a) - s_2(a, b)g_2(b) \geq 0 \quad (12)$$

where $s_i(a, b)$ are SOS polynomials and $g_1(a) \geq 0$, $g_2(b) \geq 0$ define $\mathscr{A}$ and $\mathscr{B}$, respectively. By "search for a function" we mean "find the coefficients of" such a function. Note that other SOS representations of eq 8 may be required depending on the problem, but we have chosen the most basic representation to make this introduction as clear as possible.

Finally, in analogy to obtaining the global minimum described in the previous section, we can rewrite eq 8 as

$$\begin{aligned} &\max : \lambda \\ &\text{such that} : ||a - b||_2 - s_1(a, b)g_1(a) - s_2(a, b)g_2(b) - \\ &\qquad \lambda \geq 0 \quad (13) \end{aligned}$$

where λ is a real number. The two steps of finding the SOS polynomial and finding its minimum can now be performed simultaneously by solving a semidefinite program. By using the duality, we can also obtain the proof if the minimum distance between the two geometrical shapes has been obtained.

This SOS approach provides a useful general mathematical framework for studying interactions between arbitrarily shaped coarse-grained particles, convex or nonconvex. A popular shape for anisotropic representations is the ellipsoid.[45−48] In this case, it turns out that the equations presented above simplify greatly since ellipsoids have a useful matrix representation, namely as symmetric positive definite matrices whose eigenvalues determine the length of its semiaxes and eigenvectors give their directions.[49]

Equation 7 now becomes

$$\begin{aligned} &\min : ||(c_1 + a) - (c_2 + b)||_2 \\ &\text{such that} : ||A^{-1}a||_2 \leq 1 \ (\mathscr{A}) \quad (14) \\ &\qquad ||B^{-1}b||_2 \leq 1(\mathscr{B}) \end{aligned}$$

where c_1 and c_2 are the centers of the ellipsoids, and A and B are the symmetric positive definite matrices corresponding to particles $\mathscr{A}$ and $\mathscr{B}$, respectively. The search for a globally non-negative representation is now trivial since the norm functions will always be quadratic forms and hence sum of squares.

2. Implementation to MD. We have obtained the minimum distance of closest approach exactly within numerical accuracy as confirmed by the duality gaps and infeasibility information. For an example see Figure 7. We found this to hold for different orientations and particles of varying and disparate size and eccentricity. Typical values reported for the primal and dual objectives were $5.3024e + 04$ with a difference in further than the 6 significant figure. Furthermore, the formulation of eq 10 allowed us to use a more efficient solver[50] since it met the definition of a "second-order cone program" (SOCP) which is a simpler subset of semidefinite programs. Another advantage of this approach is that we can obtain the minimum distances for all pairs of ellipsoids in any current configuration at once in a single step (see Figure 8). Here we show only 32 particles in order to illustrate all the distances at once. Typically, we calculate the minimum distance between approximately 12 000 pairs which takes of the order of 6 s.

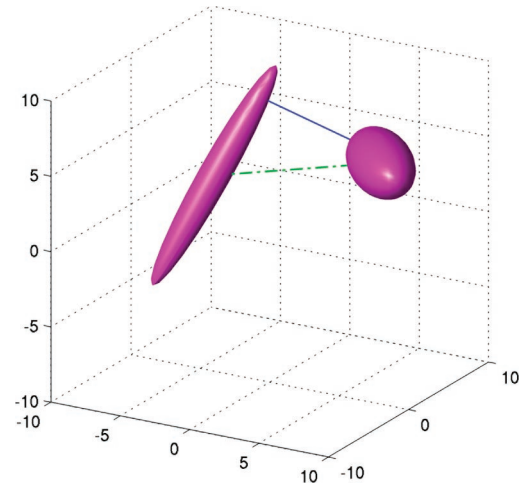


Figure 7. The exact distance of closest approach (6.355763) between two ellipsoid particles given by the SOS approach. Notice that this vector (blue continuous line) differs markedly both in direction and magnitude from the intercenter vector (green dashed line). The semiaxes lengths for the ellipsoids are {8, 1, 1} and {4, 2, 2}, respectively. Their center vectors are [−4, 0, 3] and [8, 0, 6], respectively. The minimum distance vectors are [−1.299203, 2.018481, 7.809342] and [4.874321, 1.910192, 6.302155]. The duality gap ($-1.18E - 7$) verifies the result. The approach works for arbitrary ellipsoids in any orientation.

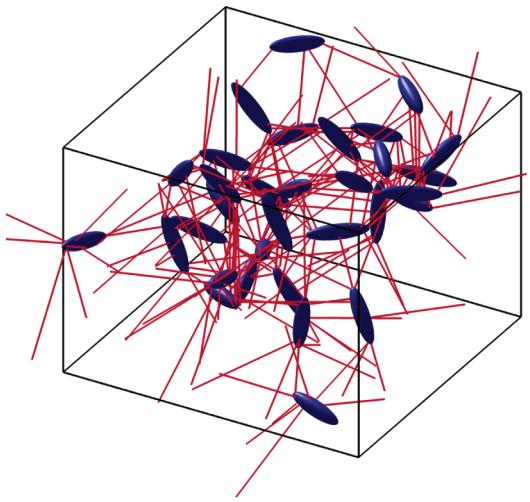


Figure 8. The exact distances of closest approach for all-to-all pair interactions between the particles in a periodic box. Here, only 32 particles are considered for illustrative purposes. The result is obtained at one step.

To show that this can be implemented in an MD simulation, we simulate an NVE ensemble of a binary A−B system of 240 ellipsoids interacting through a Lennard-Jones potential in a box of dimensions 21σ, 20σ, 16σ. The unit of length is σ, unit of energy is ϵ_{11}, while the mass of particle is 1, and the initial temperature is 0.5. Periodic boundary conditions are applied at the sides of the simulation box. The ellipsoids have semiaxis lengths of 3.0σ, 0.5σ, 0.5σ.

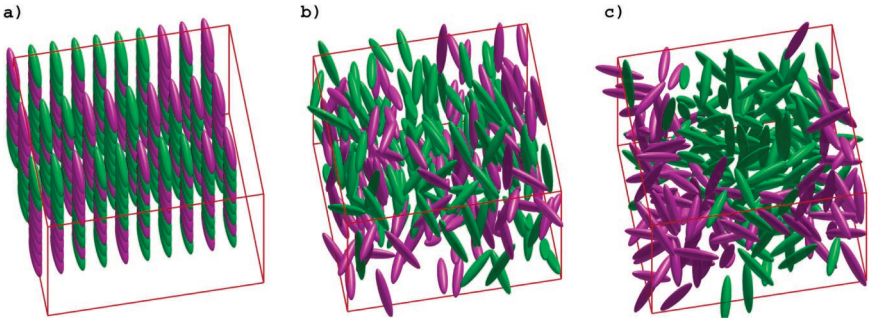


Figure 9. Snapshots from an NVE ensemble MD simulation of a mixture of 240 ellipsoids with semiaxes lengths $\{3.0\sigma, 0.5\sigma, 0.5\sigma\}$ in a box of dimensions {21, 20, 16}. The exact distances of closest approach of all-to-all particles are obtained in one step through the approach described in the text. There are typically 12 000 interactions per configuration. The simulation is initialized in a regular cuboidal array of randomly mixed A and B particles (a) (step 1). By step 500 much of the orientation correlation has disappeared (b), while by step 10 000 (c), a significant region of demixing has formed.

Lennard-Jones interactions are calculated based on distance of closest approach between pairs of ellipsoids using the following parameters: $\sigma_{ij} = 1.0$, $\epsilon_{11} = \epsilon_{22} = 1.0$, while $\epsilon_{12} = 2.0$. Two cutoff parameters are employed. The first ($\mathrm{rcut_{com}} = 9.0$) is a center of mass based spherical cutoff above which distances between ellipsoids are not calculated. A second ($\mathrm{rcut_{md}} = 4.0$) based upon the distance of closest approach between the two ellipsoids is then applied to this remaining set. The simulations were run with time step $\Delta t = 0.005$. Initially the randomized *A*/*B* particles were arranged in a cuboidal array with the principal semiaxis uniformly oriented parallel to the *x* axis. Maxwell−Boltzmann distributed linear velocities were assigned, while the initial angular velocities were zero. The forces and torques were then calculated following the standard way for rigid body dynamics. Time integration of the linear velocities was carried out using a standard leapfrog algorithm, while the angular velocities were integrated using Omelyan's algorithm.[51] The orientation of the ellipsoids was described using quaternions to avoid the singularity issue associated with Euler angles. By step 500 nearly all of the orientation correlation has disappeared (Figure 9b), while by step 15 900 (Figure 9c) significant regions of demixing have formed.

IV. Discussion

The SOS approach[1] is a rigorous and systematic way to approach complex problems by exploiting both their algebraic and geometric properties. We have applied it here to the identification of the global minimum of model potential energy surfaces which exemplify features that make this problem difficult in physical systems of interest such as narrow and disconnected wells embedded in flat or contorted regions. We have demonstrated the method on low-dimensional problems in which it is possible to gain an understanding of how the method works and how it is different to other approaches. We have also shown that it can be used to solve geometric problems such as obtaining the distance of closest approach of anisotropic coarse-grained particles. It is clear that no method will outperform all other methods for all problems, and it may well be that heuristic approaches may outperform the more formal SOS approach on a given subset of problems. However, the appeal of the SOS approach stems from its generality due to its deep theoretical foundations and efficient computational algorithms.

Another advantage comes from the added physical insight that can be gained by formulating the problem in this way where duality holds. An immediate consequence of duality is the ability to produce a proof when the exact answer has been reached without sampling or comparing to other minima, in contrast to almost all other methods. Second, the geometrical interpretation attached to the dual can lead to insight into the problem, including the identification of degeneracies and symmetries. Furthermore, there is freedom in how to pose a problem, so that constraints (equalities and inequalities) can be incorporated as variables or new variables can be introduced which can lead to efficient answers. This is particularly useful for dealing with more complex potentials that include trigonometric functions and so fitting could be avoided by introducing new variables. This may extend the applicability of the method to certain nonpolynomial functions.

Unlike other methodologies that approximate the functions either by subtracting or adding extra terms, which is known to often alter the behavior of the system especially in multidimensional problems, in this approach the original function is never deformed nor linearized. Rather, an entirely different and higher dimensional problem is sought, which is both convex solvable and whose solution in the space of the original variables coincides with that of the original question. Unlike gradient-based and Monte Carlo methods that use the original function, this approach bounds the function *externally* from below. This is one of the reasons that poorly connected and/or nearly degenerate minima separated by high barriers do not pose the difficulties encountered in the usual sense.

As the complexity of problems attempted increases, it is possible that the exact solution may not be found. In many cases, a "close enough" answer may be adequate. Here, we can quantify the gap between our answer and the exact solution. If the lower bound solution is not adequately close, then further refinement can be sought by seeking a higher dimensional monomial basis by introducing SOS functions, which can be thought here as a generalization of Lagrange multipliers. Accuracy can be systematically increased at the

expense of more computational cost, since there is a proof of the existence of the exact answer in such liftings to higher dimensions. In that respect, this method is closer in spirit to a variational approach but one which is guaranteed to be optimal and may lead eventually, at least in principle, to the right answer of the original problem. Unlike other global approaches, the method allows for insight into the nature of the PES rather than just providing a number. For example, having an explicit SOS polynomial representation of the problem can potentially be useful for studying other properties of the system besides its global minimum. The method also allows for exploiting the geometrical properties of the system which can lead to more efficient ways to simulate dynamics of coarse-grained systems which are often anisotropic.

A number of challenges remain due to the significant scaling of the number of the new variables that occurs as a result of the lifting process which means that, for physical problems, the current limitations on size of SDPs solvable using primal-dual interior point methods need to be overcome. Excluding symmetries and symmetry operations is a promising way in that direction. This is the focus of ongoing research.

Acknowledgment. We thank Pablo Parrilo and Mauricio Barahona for helpful discussions. This work has been funded by the ONR, U.S.A. and the BBSRC, U.K.

Appendix: How To Obtain the Global Minimum and a SOS Decomposition in Practice

We describe here in detail the process for obtaining a SOS decomposition as well as the global minimum of the function from the SOS decomposition in order to demonstrate how the method works in practice. Note that the decomposition and optimization are carried out simultaneously. We choose a very simple quartic polynomial function to exemplify the method by showing explicitly all the steps involved:

$$U = 2x^4 + 2x^3y - x^2y^2 + 5y^4, \quad x, y \in (-\infty, \infty)$$

For the purposes of this example we omit any constraints. The global minimum (0,0) is straightforward to obtain, but the decomposition is not immediate by standard techniques. The emphasis here is in the implementation of the method in the available software.

1. The first step is to rewrite the problem as a polynomial optimization problem in the form of eq 1 in the flowchart of Figure 2, where U is the function to be minimized subject to all the constraints:

$$\min: \ U(\mathbf{x})$$

$$\text{subject to}: \ g(\mathbf{x}) \geq 0$$

$$h(\mathbf{x}) = 0$$

Although in this example this is straightforward (U is given above and there are no constraints), in certain cases it may be the most difficult step in the calculation.

2. The next step is to introduce an additional real variable, $\lambda \in \mathbf{R}$, such that the minimization problem becomes now a maximization (eq 2, Figure 2):

$$\max: \ \lambda$$

$$\text{subject to}: \ U(\mathbf{x}) - \lambda \geq 0$$

$$g(\mathbf{x}) \geq 0$$

$$h(\mathbf{x}) = 0$$

Although this appears to be a trivial step, it has important theoretical consequences as described in the section Theory and also practical implications. For example, note that the function to be maximized now is a linear function, while all the nonlinear parts are in the constraints. There is no effect on the underlying problem itself, but the benefit is that solutions to the problem (optimal points) that may have been on the interior of the domain space are now on the surface. This makes it possible to find them more efficiently. Our example now becomes

$$\max: \ \lambda$$

$$\text{subject to}: \ 2x^4 + 2x^3y - x^2y^2 + 5y^4 - \lambda \geq 0$$

3. Solving this problem still remains very difficult and often impossible. However, requiring that $2x^4 + 2x^3y - x^2y^2 + 5y^4 - \lambda$ is a SOS instead turns out to be easier to solve. Although not necessarily exact, it often gives a good solution to the original problem we are interested in. In this example, this is equivalent to (eq 3, Figure 2)

$$\min: \ -\lambda$$

$$\text{subject to}: \ 2x^4 + 2x^3y - x^2y^2 + 5y^4 - \lambda \text{ is SOS}$$

4. We now show how to obtain the SOS decomposition. Once we have written our problem in the form above, we can submit it to the software SOSTOOLS which will do precisely that. To proceed, we must define the degree of the SOS basis we want to use. The maximum basis is d where $2d$ is the maximum degree of the function of interest including the constraints. (In this example, $d = 2$.) We now write out the entire vector of monomials whose total degree is less than d

$$z = [x_1, x_2, ..., x_n, x_1^2, ..., x_n^2, x_1x_2, ..., x_{n-1}x_n, ..., x_1^d ..., x_n^d]$$

which in our example is

$$z = [1, x, y, x^2, xy, y^2]$$

The ordering is not particularly important, but we have chosen the ordering that allows for the easiest interpretation of results from SOSTOOLS. If there are symmetry or physical reasons that exclude certain solutions, then we can take advantage of that and not include them in the monomial basis. Additionally, if the problem is of very high dimension, we can start by looking for an approximate solution with a reduced basis and subsequently refine the calculation by increasing the monomials in the basis.

5. The basis of monomials allows us to construct a matrix Q. Finding the SOS decomposition is equivalent to finding a suitable positive semidefinite matrix, $Q \geq 0$, such that

$$U - \lambda = z^T Q z$$

This Q matrix is an $r \times r$ symmetric matrix in unknown variables q_{ij}. The problem now becomes (eq 4, Figure 2)

$$\min : \ -\lambda$$

$$\text{subject to} : \ z^T Q z \geq 0$$

So in this example we want to find the entries of Q such that

$$U - \lambda = \begin{bmatrix} 1 \\ x \\ y \\ x^2 \\ xy \\ y^2 \end{bmatrix}^T \begin{bmatrix} q_{11} - \lambda & q_{12} & q_{13} & q_{14} & q_{15} & q_{16} \\ q_{12} & q_{22} & q_{23} & q_{24} & q_{25} & q_{26} \\ q_{13} & q_{23} & q_{33} & q_{34} & q_{35} & q_{36} \\ q_{14} & q_{24} & q_{34} & q_{44} & q_{45} & q_{46} \\ q_{15} & q_{25} & q_{35} & q_{45} & q_{55} & q_{56} \\ q_{16} & q_{26} & q_{36} & q_{46} & q_{56} & q_{66} \end{bmatrix} \begin{bmatrix} 1 \\ x \\ y \\ x^2 \\ xy \\ y^2 \end{bmatrix}$$

and Q remains positive semidefinite.

6. Note that the entries of Q and the coefficients of the monomials of the function of interest define a set of linear constraints. In our example, this leads to the following set of linear equations:

$$\begin{array}{lllll} q_{11} - \lambda = 0, & q_{13} = 0, & 2q_{16} + q_{33} = 0, & q_{36} = 0, & q_{66} = 5, \\ q_{12} = 0, & q_{15} = q_{23} = 0, & q_{35} + q_{26} = 0, & q_{56} = 0, & 2q_{14} + q_{22} = 0, \\ q_{43} + q_{45} = 0, & 2q_{46} + q_{55} = -1, & q_{24} = 0, & 2q_{45} = 2 & q_{44} = 2 \end{array}$$

In fact, our original problem can be rewritten as

$$\min : \ -\lambda$$

$$\text{subject to} : \ \mathbf{Tr}(A_i, Q) = b_i, \ \forall i = 1, ..., 15$$

$$Q \geq 0$$

where the A_i are matrices representing the coefficients in the linear constraints. For example, A_{10} is

$$\begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{11}$$

and $b_{10} = 0$. It turns out that the linear objective function to be minimized, the linear constraints, and the constraint $Q \geq 0$ form a statement of a general class of problems, semidefinite programs, that are solved efficiently and possess interesting duality properties. If we define C to be a 6×6 matrix with zeros everywhere except a -1 in the (1, 1) entry and b with

$$b = [0, 0, 0, 0, 5, 0, 0, 0, 0, 0, 0, -1, 0, 2, 2]$$

then we have the primal form of a semidefinite program[1]

$$\min : \ \mathbf{Tr}(C, Q)$$

$$\text{subject to} : \ \mathbf{Tr}(A_i, Q) = b_i, \ \forall i = 1, ..., m,$$

$$Q \geq 0$$

where m is the number of constraints. The dual form is easily formulated as

$$\max : \ b^T y$$

$$\text{subject to} : \ C - \sum_i A_i y_i \geq 0$$

SOSTools now invokes an SDP solver to obtain the solution we are interested in. Semidefinite programs are commonly solved using path following interior point methods in which the primal and dual formulations are solved simultaneously. Commonly used solvers include SeDuMi and SDPT3. Both solvers are indeed variations of the primal-dual interior point algorithm and will solve this problem very easily. For small to medium sized problems SeDuMi is extremely robust indeed, but in larger sparse problems we found SDPT3 to be more effective at finding solutions. For this example, we report solutions from SeDuMi. Once the SDP solver has completed its solution a number of checks must be made to ensure proper convergence. The checks are more or less standard for most SDP solvers such as SeDuMi or SDPT3.

(a) Check that the primal solution is feasible.
(b) Check that the dual solution is feasible.
(c) Check the duality gap.

The duality gap is defined as the difference between the primal and dual solutions. It must be checked manually, and in our example the primal solution is $-8.1518E - 10$, the dual solution is $-6.0227E - 10$, and the gap is $-2.1291E - 10$. This gives us the proof that we have found the exact minimum.

7. Upon solving the semidefinite program, a primal and dual solution matrix are produced. These will allow us to locate the global minimum and find the SOS decomposition. In our example, the solution matrices are

Primal:

$$\begin{bmatrix} 0.0000 & 0.0000 & 0.0000 & -0.0000 & -0.0000 & -0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & -0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & -0.0000 & -0.0000 \\ -0.0000 & 0.0000 & 0.0000 & 2.0000 & 1.0000 & -1.3214 \\ -0.0000 & -0.0000 & -0.0000 & 1.0000 & 1.6427 & -0.0000 \\ -0.0000 & 0.0000 & -0.0000 & -1.3214 & -0.0000 & 5.0000 \end{bmatrix}$$

Dual:

$$\begin{bmatrix} 1.0000 & 0.0000 & -0.0000 & 0.0000 & -0.0000 & 0.0000 \\ 0.0000 & 0.0000 & -0.0000 & -0.0000 & -0.0000 & 0.0000 \\ -0.0000 & -0.0000 & 0.0000 & -0.0000 & 0.0000 & -0.0000 \\ 0.0000 & -0.0000 & -0.0000 & 0.0000 & -0.0000 & 0.0000 \\ -0.0000 & -0.0000 & 0.0000 & -0.0000 & 0.0000 & -0.0000 \\ 0.0000 & 0.0000 & -0.0000 & 0.0000 & -0.0000 & 0.0000 \end{bmatrix}$$

From the primal solution matrix we can extract the global minimum of the function to be $\lambda = q_{11} = 0.000$. Interpretation of the dual solution matrix is obtained by noting that

the dual solution matrix in a sum of squares SDP solution should correspond to the matrix zz^T. Therefore we can conclude that the first row and the first column are equal to z provided the first entry in z is 1. Therefore we can read off the location of the global minimum to be the vector $(x, y) = (0, 0)$. It is important to check the rank of the dual solution matrix which should be 1 if the relaxation is exact.

8. Finding the SOS decomposition: Before we do this it is worth noting that the primal solution matrix can be divided into 4 (3×3) submatrices and that only one of these is nonzero. From this we can see that the SOS decomposition in fact only requires a smaller basis of monomials namely $z = [x^2, xy, y^2]$. There are in fact ways of identifying such reductions in basis size, and these will be discussed in a later publication. To see the SOS decomposition, we must factorize the matrix, and here we take the Cholesky factorization of the (3×3) matrix:

$$L = \begin{bmatrix} 1.4142 & 0.0000 & 0.0000 \\ 0.7071 & 1.0690 & 0.0000 \\ -0.9344 & 0.6180 & 1.9352 \end{bmatrix}$$

Each row contains a single square within the SOS with each entry in each row being the coefficient of the corresponding entry in z. So we can reconstruct the SOS as

$$f^{\mathrm{SOS}} = (1.4142x^2)^2 + (0.7071x^2 + 1.0690xy)^2 + (-0.9344x^2 + 0.6180xy + 1.9352y^2)^2$$

The SOS decomposition is not necessarily unique, and other decompositions may be also be found; however, the global minimum and location will be unique provided that this is the case for the energy function being minimized. If the global minimum is not unique, then this will be evident from the rank of the dual matrix which will be greater than 1. Often the nonuniqueness is as a result of a symmetry in the function which may not be apparent upon initial examination. Examining the eigenvalues of the dual solution matrix gives information as to in which degrees of freedom the degeneracies lie.

References

(1) Parrilo, P. Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimisation, Thesis, California Institute of Technology, 2000.

(2) Parrilo, P.; Sturmfels, B. In *Algorithmic and quantitative real algebraic geometry*, Basu, S., Gonzalez-Vega, L., Eds.; AMS: 2003; Chapter Minimising polynomial functions, Vol. 60, pp 83−99.

(3) Berry, R. S. *Chem. Rev.* **1993**, *93*, 2379−2394.

(4) Wales, D.; Scheraga, H. A. *Science* **1999**, *285*, 1368−1372.

(5) Munro, L. J.; Tharrington, A.; Jordan, K. D. *Comput. Phys. Commun.* **2002**, *145*, 1−23.

(6) Tsai, C. J.; Jordan, K. D. *J. Phys. Chem.* **1993**, *97*, 11227−11237.

(7) Kirkpatrick, S.; Gelatt, C.; Vecchi, M. P. *Science* **1983**, *220*, 671−680.

(8) Frantz, D. D.; Freeman, D. L.; Doll, J. D. *J. Chem. Phys.* **1990**, *93*, 2769−2784.

(9) Grubmueller, H. *Phys. Rev. E* **1995**, *52*, 2893−2906.

(10) Rabow, A. A.; Scheraga, H. A. *Protein Sci.* **1996**, *5*, 1800−1815.

(11) Deaven, D.; Tit, N.; Morris, J. R.; Ho, K. *Chem. Phys. Lett.* **1996**, *256*, 195−200.

(12) Maranas, C. D.; Floudas, C. A. *J. Chem. Phys.* **1992**, *97*, 7667−7678.

(13) Damsbo, M.; Kinnear, B. S.; Hartings, M.; Ruhoff, P. T.; Jarrold, M. F.; Ratner, M. A. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 7215−7222.

(14) Andricioaei, I.; Straub, J. *J. Comput. Chem.* **1998**, *19*, 1445−1455.

(15) Doye, J. P. K.; Wales, D. J.; Berry, S. R. *J. Chem. Phys.* **1995**, *103*, 4234−4249.

(16) Piela, L.; Kostrowicki, J.; Scheraga, H. A. *J. Phys. Chem.* **1989**, *93*, 3339−3346.

(17) Pillardy, J.; Piela, L. *J. Phys. Chem.* **1995**, *99*, 11805−11812.

(18) Wales, D.; Doye, J. *J. Phys. Chem. A* **1997**, *101*, 5111−5116.

(19) Tozzini, V. *Curr. Opin. Struct. Biol.* **2005**, *15*, 144−150.

(20) Nielsen, S.; Lopez, C.; Srinivas, G.; Klein, M. *J. Phys.: Condens. Matter* **2004**, *16*, R481−R512.

(21) Bahar, I.; Rader, A. J. *Curr. Opin. Struct. Biol.* **2005**, *15*, 586−592.

(22) Ming, D.; Kong, Y.; Wakil, S. J.; Brink, J.; Ma, J. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 7895−7899.

(23) Berne, B.; Pechukas, P. *J. Chem. Phys.* **1972**, *56*, 4213−4216.

(24) Gay, J.; Berne, B. *J. Chem. Phys.* **1981**, *74*, 3316−3319.

(25) Perram, J.; Rasmussen, J.; Præstgaard, E.; Lebowitz, J. *Phys. Rev. E* **1996**, *54*, 6565−6572.

(26) Allen, M.; Evans, G.; Frenkel, D.; Mulder, B. M. *Adv. Chem. Phys.* **1993**, *86*, 1.

(27) Paramonov, L.; Yaliraki, S. N. *J. Chem. Phys.* **2005**, *123*, 194111.

(28) Lasserre, J. B. *SIAM J. Optim.* **2001**, *11*, 796−817.

(29) Shor, N. Z. *Cybernetics (Engl. Transl.)* **1987**, *23*, 731−734.

(30) Vandenberghe, L.; Boyd, S. *SIAM Rev.* **1996**, *38*, 49−95.

(31) Boyd, S.; Vandenberghe, L. *Convex Optimization*; Cambridge University Press: 2004.

(32) Nesterov, Y. E.; Nemirovsky, A. S. *Ekon. Matem. Metody* **1988**, *24*, 1084−1091.

(33) Ben-Tal, A.; Nemirovski, A. *Lectures on modern convex optimization, analysis, algorithms and engineering applications;* MPS-SIAM: 2001.

(34) Strang, G. *Introduction to Applied Mathematics*; Wellesley-Cambridge Press: Wellesley, MA, 1986.

(35) Tisza, L. *Generalized Thermodynamics*; MIT Press: Cambridge, MA, 1966.

(36) Prajna, S.; Papachristodoulou, A.; Parrilo, P. *SOSTOOLS, Sum of Squares Optimization Toolbox for MATLAB, User's Guide, Version 1.00*; Department of Control and Dynamical Systems, California Institute of Technology: Pasadena, CA 91125, U.S.A., 2002.

(37) Sturm, J. *Using SEDUMI 1.02, A Matlab Toolbox for Optimisation Over Symmetric Cones*; Department of Econometrics, Tilburg University: Tilburg, The Netherlands, 2001.

(38) Tutuncu, R.; Toh, K.; Todd, M. *SDPT3 − a Matlab software package for semidefinite-quadratic-linear-programming*; 2001.

(39) Becker, O.; Karplus, M. *J. Chem. Phys.* **1997**, *106*, 1495−1517.

(40) Nelder, J. A.; Mead, R. *Comput. J.* **1965**, *7*, 308−313.

(41) Dennis, J. E.; Schnabel, R. B. *Numerical Methods for unconstrained optimization and nonlinear equations*; SIAM: Philadelphia, PA, 1983.

(42) Müller, K. *Angew. Chem., Int. Ed. Engl.* **1980**, *19*, 1−78.

(43) Olender, R.; Elber, R. *J. Chem. Phys.* **1996**, *105*, 9299−9315.

(44) Passerone, D.; Parrinello, M. *Phys. Rev. Lett.* **2001**, *87*, 108302.

(45) Liwo, A.; Oldziej, S.; Pincus, M. R.; Wawak, R.; Rackovsky, S.; Scheraga, H. *J. Comput. Chem.* **1997**, *18*, 849−873.

(46) Ayton, G.; Bardenhagen, S.; McMurty, P.; Sulsky, D.; Voth, G. A. *J. Chem. Phys.* **2001**, *114*, 6913−6924.

(47) Brannigan, G.; Tamboli, A.; Brown, F. L. H. *J. Chem. Phys.* **2004**, *121*, 3259−3271.

(48) Tsonchev, S.; Schatz, G. C.; Ratner, M. A. *Nano Lett.* **2003**, *3*, 623−626.

(49) Strang, G. *Linear Algebra and its applications*; Harcourt Brace Jovanovich: San Diego, CA, 1988.

(50) Andersen, E. D.; Roos, C.; Terlaky, T. *On implementing a primal-dual interior-point method for conic quadratic optimization*; Technical Report W-274; Helsinki School of Economics and Business Administration, Department of Economics and Management Science, PL 1210, FIN-00101 Helsinki, Finland, 2000.

(51) Omelyan, I. P. *Phys. Rev. E* **1998**, *58*, 1169−1172.

CT050338P

JCTC Journal of Chemical Theory and Computation

How Well Can Coarse-Grained Models of Real Polymers Describe Their Structure? The Case of Polybutadiene

Leonid Yelash,*,† Marcus Müller,‡ Wolfgang Paul,† and Kurt Binder†

Institut für Physik, WA 331, Johannes-Gutenberg Universität Mainz, Staudingerweg 7, D-55099 Mainz, Germany, and Institut für Theoretische Physik, Georg-August Universität Göttingen, Friedrich-Hund-Platz 1, D-37077 Göttingen, Germany

Received August 21, 2005

Abstract: Coarse-graining of chemical structure of macromolecules in the melt is investigated using extensive molecular dynamics simulation data which are based on a united atom force-field model of polybutadiene. Systematically increasing the number, *n*, of the united atoms approximated by an effective coarse-grained monomer, we study the influence of degree of coarse-graining on the structure functions such as the segment−segment intermolecular and intramolecular correlation functions. These results are compared to Monte Carlo simulations of the corresponding coarse-grained bead-spring model and Chen−Kreglewski potential for chain molecules. In contrast to the atomistic chemically realistic model of polybutadiene, the bending and torsional potentials are not included into the coarse-grained models. Nevertheless, for a range of intermediate values of *n* a good qualitative agreement between intra- and intermolecular coarse-grained correlations of the atomistic model and the coarse-grained bead-spring model is found on large and intermediate length scales, but deviations occur on length scales well below one nanometer. The structure functions obtained for the Chen−Kreglewski chains exhibit many artificial features.

1. Introduction

Understanding complex processes in real soft matter systems such as, e.g., relaxation processes in polymeric systems or the self-assembly and function of biological membranes is only possible with physically adequate and computationally effective models. Concerning different modeling approaches one can distinguish between microscopic, mesoscopic, and macroscopic models depending on the length and time scales of described objects and processes. At microscopic scales, the elementary degrees of freedom are due to atoms, and one deals with the intermolecular interactions and the atomistic structure of matter. Using molecular dynamics (MD) and Monte Carlo (MC) simulation methods,[1−4] one can obtain many properties, including the structure functions, which can be directly compared with corresponding experimental data. This microscopic structure in real systems can be experimentally probed by neutron or X-ray scattering.

The microscopic approaches require force-fields, which capture the nature of the different interactions. Some important types of interactions are the van der Waals (dispersion), polar, and ionic interactions. Accurate interatomic potential functions can be obtained from quantum chemical ab initio calculations of small molecules, e.g., using the Gaussian method. The force-field parameters can be also optimized in an automatic parametrization method[5] using experimental density and heat vaporization data of liquids. Such interaction potentials obtained for small molecules can also be used in atomistic simulations of oligomers and moderately long chains. In all-atom (AA) approaches,[4,6,7] for example, one explicitly describes all atoms in the simulation, e.g., the carbon atoms with the chemically bonded hydrogen atoms. However, this approach becomes inefficient for macromolecules, although it can be successfully applied to small and moderately large molecules. By ignoring hydrogen

* Corresponding author phone: +49-6131-3924104; e-mail: yelash@uni-mainz.de.
† Johannes-Gutenberg Universität Mainz.
‡ Georg-August Universität Göttingen.

10.1021/ct0502099 CCC: $33.50
Published on Web 02/04/2006

atoms in the united atom (UA) approach or by replacing the quantum mechanical potentials by computationally simpler functions such as the Lennard-Jones potential one can achieve an important improvement of the efficiency in computer simulations, i.e., a speedup of about 2 orders of magnitude in MD simulations. However, natural biological macromolecules as well as many polymers of industrial interest are usually consisting of several thousands up to some millions of atoms. A simulation of a microscopic UA model then often would require many more orders of magnitude of computer time to equilibrate the system than is available even on the fastest supercomputer. It is therefore necessary to explore more efficient computational techniques.

A promising approach, which can also increase the efficiency of simulations substantially and connect the microscopic and mesoscopic scales, is called "coarse-graining".[8−21] In this method one starts with a microscopic description of real molecules and reduces the degree of freedom, for example, by approximating several primary chemical units (e.g., carbon groups) by an effective monomer. Formally, a coarse-graining method can be related to the renormalization group approach for critical phenomena, in which an original Hamiltonian $H(\boldsymbol{x})$ of the system with $\boldsymbol{x}$ degrees of freedom has to be replaced by an effective Hamiltonian $F(\boldsymbol{m})$ with $\boldsymbol{m}$ coarse-grained degrees of freedom.[9,10] One tries to carry out a mapping[11−21] between the interactions on the atomistic scale and suitable effective interactions between the effective segments on the mesoscopic scale, attempting to create as much similarity between the physical properties of the atomistic model and the coarse-grained model as possible. Although this approach has had promising successes, it is clear that fine-scale structural details are lost, and as a consequence the desired physical properties of the system can only approximately be described. The question arises which level of coarse-graining is optimal, i.e., how many atoms (or united atoms, respectively) along the backbone of a chain should be combined into an effective segment. Some answers to this question can be found in the quoted literature, but they are still incomplete.

Similar approximations as in the coarse-graining method are used in deriving analytical models, where the details of atomistic structure of real molecules are usually neglected. In the first-order thermodynamic perturbation theory (TPT1),[22−25] for example, the theoretical monomer approximately represents several carbon groups of a real chain, e.g., alkanes.[26] In TPT1, a dimer consisting of two tangent spheres can model alkanes such as butane, which has four carbon groups. In a coarse-grained description one can tune the "resolution" of the model by merging different number of primary units. In this way it is possible to cover different physical scales bridging the gap between the micro- and mesoscopic scales. Hence, the coarse-graining approach is an extensively developed area of modern science dealing with molecular modeling.

Recently, the dissipative particle dynamics (DPD) method, which is a mesoscopic particle-based simulation technique intensively developed during the past decade,[27−30] has been applied to *cis*-polybutadiene.[31] In contrast to the standard molecular simulation techniques such as molecular dynamics and Monte Carlo, the particles in DPD represent a small amount of fluid which can include a large number of atoms. The DPD-particles interact via a soft-repulsive potential (a conservative force in DPD). In this method, however, it is still a difficult problem to relate the interactions of the DPD-particles to a microscopic interaction potential (which can be obtained from the first principles). A further step beyond DPD modeling of polymers is the soft-ellipsoid approach,[32] in which the volume occupied by a chain molecule is approximated by an ellipsoid.

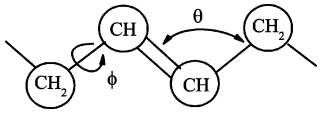


Figure 1. Schematic united atom (UA) model of polybutadiene molecules. The carbon groups of polybutadiene, which include one carbon atom and up to three hydrogen atoms (*CH*, CH_2, and endgroup CH_3), are approximated by spheres − the united atoms. The bond lengths between the united atoms are set to fixed values depending on the type of the bond: $r_0 = 1.34$ Å for $CH{=}CH$, $r_0 = 1.5$ Å for $CH{-}CH_2$ (α-bond), $r_0 = 1.53$ Å for $CH_2{-}CH_2$ (β-bond). The bending potential (eq 4) describes the bond angle θ which can harmonically change near the values 112° and 126° for $CH_2{-}CH_2{-}CH$ and $CH_2{-}CH{=}CH$, respectively. The torsional potentials (eq 5) describe the cis, gauche, and trans conformations, which can be distinguished by the torsion angle ϕ.

In this paper, we study the influence of the scale over which the coarse-graining is performed on the segment−segment correlation functions and the bond angle distribution of an atomistic UA model of polybutadiene.[33] We compare these results with Monte Carlo simulations of two coarse-grained models for chain molecules as an input. Polybutadiene has been chosen due to the availability of extensive molecular dynamic simulations[34−37] and experimental data.[35,36] Furthermore, a simple chemical structure and its importance for the chemical industry makes polybutadiene attractive to study.

2. Modeling Polybutadiene

2.1. United-Atom Model of Polybutadiene. The high-frequency motions, e.g., the hydrogen atoms vibrations, occur on very short time and length scales if compared to the typical scales for the molecular motions. An exact treatment of such motions can be very costly in computational simulations on long time scales. In a first step, one can approximate the entire functional group by one effective atom located at the position of the carbon atom. This effective group is then called united atom (UA). For example, each carbon group of polybutadiene, i.e., a carbon atom with bonded hydrogen atoms, can be replaced by one effective atom, which is shown schematically in Figure 1. Using this united atom model one can extend the investigation to longer time scales. Furthermore, there is a possibility to refine the results of the united-atom modeling later by reintroducing the neglected hydrogen atoms.

We use the united atom force field model of polybutadiene developed by Smith and Paul[33] which is based on the results of quantum chemical ab initio calculations. This model was

Table 1. Bond Lengths of Polybutadiene

bond type	bond length r_0 [Å]
CH═CH	1.34
CH−CH$_2$ (α)	1.5
CH$_2$*−CH*$_2$ (β)	1.53

extensively tested in many molecular dynamic simulations and has been shown to reproduce the experimental data for dielectric spectroscopy,[34] dynamic neutron scattering,[35] and relaxation processes[36] in polybutadiene melts very well.

The potential energy of polybutadiene can be written as a sum of the different contributions

$$U = U_{\text{bond}} + U_{\text{angle}} + U_{\text{tors}} + U_{\text{LJ}} \tag{1}$$

whereas U_{bond} is the bond length potential, U_{angle} is the bending energy for the angle between two neighboring bonds, U_{tors} is the energy of the torsion angles, and U_{LJ} is the Lennard-Jones interaction.

In the polybutadiene model, the bond lengths are rigid, and the distances between the carbon groups have been set to $r_0 = 1.34$, 1.5, and 1.53 Å for *CH═CH*, *CH*$_2$*−CH*, and *CH*$_2$*−CH*$_2$ bonds, respectively (see Table 1).

The potential energy of such bonds can be formally given by the Dirac's delta function

$$U_{\text{bond}}(r_{ij}) = C\delta(r_{ij} - r_0) \tag{2}$$

where the amplitude prefactor C is used here for dimensional reasons: $\delta(\mathbf{r})$ has dimension of inverse volume.

The bending contribution U_{angle} can be given by a harmonic deviation of the angle θ from an equilibrium angle θ_0

$$U_{\text{angle}}(\theta) = \frac{k}{2}(\theta - \theta_0)^2 \tag{3}$$

where θ is the angle between three bonded carbon groups. For computational reasons it is convenient to express the bending potential eq 3 in terms of cosθ:

$$U_{\text{angle}}(\theta) = \frac{k}{2\sin^2\theta_0}(\cos\theta - \cos\theta_0)^2 \tag{4}$$

For modeling polybutadiene, two bond angles were used: $\theta = 2.1973$ rad (125.9°) for the *CH*$_2$*−CH═CH*-type bonds and $\theta = 1.9487$ rad (111.6°) for the *CH*$_2$*−CH*$_2$*−CH*-type bonds.

The rotations around the bonds involve the torsional potential

$$U_{\text{tors}}(\phi) = \frac{1}{2}\sum_{n=1}^{6} k_n(1 - \cos n\phi) \tag{5}$$

where ϕ is the torsional angle (shown in Figure 5). The torsional potential is a four-particle potential. One can distinguish between five types of the chain conformations in polybutadiene: the cis- and trans-types double bonds *CH*$_2$*−CH═CH−CH*$_2$, the cis- and trans-types α-bonds *CH═CH−CH*$_2$*−CH*$_2$, and the β-bond *CH−CH*$_2$*−CH*$_2$*−CH*.

The interactions between the atoms which are arranged along the backbone of a chain at least four atoms away from each other or between the atoms of different chains are modeled via the Lennard-Jones potential:

$$U_{\text{LJ}}(r_{ij}) = \epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - 2\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \tag{6}$$

The pair interaction parameters, ϵ_{ij} and σ_{ij}, which depend on the type of the interaction between the different carbon groups, can be found elsewhere[33,37] (see also Tables 1−4).

For our purpose we utilize the molecular dynamics simulation configuration trajectories of polybutadiene at atmospheric pressure obtained by Krushev[37] at temperatures 240 and 353 K, which are well above the glass transition temperature of polybutadiene $T_g \approx 180$ K. The simulation box consisting of 40 chains with 116 united atoms per molecule has the edge size 47.665 and 49.33 Å, which at these temperatures yields the polybutadiene density 0.9652 g/cm^3 and 0.8676 g/cm^3, respectively.

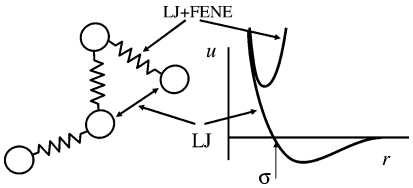


Figure 2. Bead-spring model for chain molecules based on the Lennard-Jones and FENE potentials. Circles represent the coarse-grained monomers. Springs represent the bonded interactions between monomers via the LJ+FENE potential. Nonbonded monomers interact via the Lennard-Jones potential. The corresponding potential curves are shown schematically on the right-hand side.

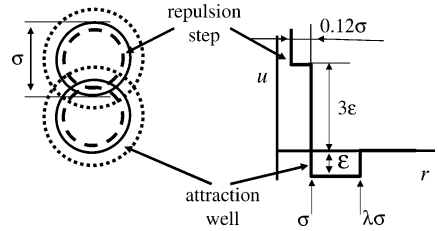


Figure 3. The Chen and Kreglewski interaction potential.[61] The solid circles represent the "softened" segment diameter σ, at which the molecules repel each other with interaction energy $+3\epsilon$. The interaction energy diverges at the hard-core diameter 0.88σ (dashed circles). The corresponding potential curve is shown on the right-hand side. For comparison, the interaction energy in the hard-sphere and square-well models diverges at diameter σ.

2.2. Coarse-Grained Bead-Spring Description of Polybutadiene. Coarse-grained models have been applied to different problems such as surfactant and lipid systems,[38−45] to study the self-organization of rod-coil molecules,[46] surfactant oligomers,[47] diblock copolymers,[48,49] chemically reacting systems,[50] nanoparticles,[51] liquid bisphenol A-polycarbonate,[15,16,21] and phospholipids.[52,53]

A widely utilized coarse-grained bead-spring model has been developed by Kremer and Grest.[54] The Lennard-Jones potential describes the repulsion and dispersion forces

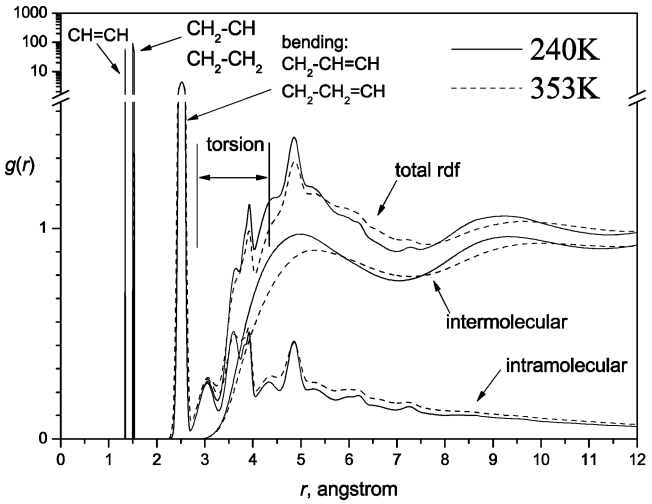


Figure 4. Segment−segment correlation functions from atomistic simulations of polybutadiene at 240 K (solid) and 353 K (dashed) and atmospheric pressure.[37] The intramolecular correlations are the segment−segment correlations along a polymer molecule. The intermolecular correlations are the correlations of segments which belong to different molecules. The total radial distribution function is a sum of the intra- and the intermolecular functions. The sharp (δ-function-like) peaks at $r = 1.34$, 1.5, and 1.53 Å correspond to the double (*CH═CH*) and single (*CH*$_2$−*CH* and *CH*$_2$−*CH*$_2$) bonds of polybutadiene. The correlation peak at $r \approx 2.53$ Å represents the distribution of the united atoms interacting via the bending potential (eq 4). Several correlation peaks at 3 Å $\leq r \leq$ 4 Å are due to the torsional cis and trans interactions (bonding four carbon groups along a backbone).

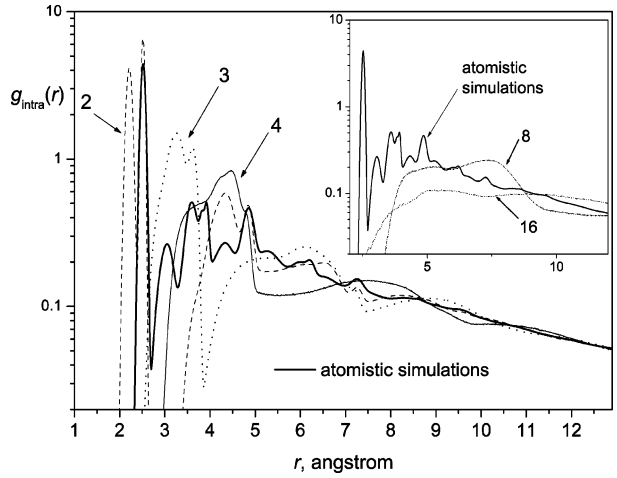


Figure 5. Coarse-graining of the intramolecular segment−segment correlation function of polybutadiene (240 K). The numbers are for the different degrees of the coarse-graining. A bold curve is the correlation function calculated using the UA MD data[37] (also shown in Figure 4, however, the δ-type peaks at $r = 1.34$ Å, 1.5 Å, 1.53 Å are omitted here).

between the segments for both the intermolecular and nonbonded intramolecular interactions:

$$U_{\mathrm{LJ}}(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] + \frac{127}{4096}\epsilon \qquad (7)$$

The interaction range of the Lennard-Jones potential is infinite. In the simulations, however, it is computationally effective to reduce the interaction range of the potential, which does not change the qualitative phase behavior.[55] In the literature one can find different ways how to truncate the interaction range of the potential. In this work we use a cutoff at $r_{\mathrm{cutoff}} = 2\sqrt[6]{2}\sigma$,[56,57] which is close to the minimum of the pair correlation function beyond the second correlation shell in the dense phase, thus minimizing the error of cutoff. The potential is also shifted (by the last term in eq 7) in order to avoid an energy discontinuity at the cutoff distance.

Table 2. Bending-Potential Parameters of Polybutadiene

type	energy k [K]	angle θ_0 [rad]
CH$_2$−*CH*$_2$−*CH*	57899	1.9487
CH$_2$−*CH═CH*	45010	2.1973

The interactions of the segments bonded along a backbone are modeled via the FENE potential in addition to the Lennard-Jones potential. FENE stands for finitely extensible nonlinear elastic. This model is shown schematically in Figure 2. In the coarse-grained bead-spring model, the distance between the bonded segments can change to a certain extent, e.g., upon variation of the pressure, yielding an equilibrium distribution of bond lengths for fixed thermodynamic parameters. The stretching of the bonds is restricted by a bond-energy penalty, which diverges at the bond length $r = 1.5\sigma$:

$$U_{\mathrm{FENE}}(r) = -33.75\epsilon \ln\left[1 - \left(\frac{r}{1.5\sigma}\right)^2\right] \qquad (8)$$

With this choice of the constants in the FENE potential the maxima of the distance distribution for the bonded and nonbonded monomers in a dense melt occur at $r \approx 0.96\sigma$ and $r \approx 1.1\sigma$, respectively. The parameters of the Lennard-Jones, bond length and bending potentials can be further optimized using results of the atomistic simulations and experimental data to reproduce the observables.[58] An optimization of the bond angle parameter can be important to reproduce, for example, the peaks of the pair correlation function.[59] For modeling flexible macromolecules of polybutadiene on the coarse-graining level no bending potential (which is the coarse-grained analogue of the angular potential eq 4) is included here, unlike for the modeling of lipids and surfactants.

This bead-spring model is very well studied and has already been applied to many different problems including, recently, the coarse-grained modeling of polymer solutions.[56,57,60] For example, it has been shown that phase behavior predicted by this model for mixtures of carbon dioxide/hexadecane is in good agreement with experimental data.

2.3. Chen−Kreglewski Potential for Chain Molecules. Another coarse-grained model of chain molecules which we study here is based on the Chen−Kreglewski (CK) potential.[61] This interaction potential is the theoretical basis of some analytical equations of state such as BACK[61] and PC-SAFT[62,63] equations. The "softened" repulsion introduced into the Chen−Kreglevski potential is also used in the SAFT equation.[26] The potential is shown schematically in Figure

Table 3. Torsion-Potential Parameters of Polybutadiene[37]

type	k_1 [K]	k_2 [K]	k_3 [K]	k_4 [K]	k_5 [K]	k_6 [K]
$CH_2-CH=CH-CH_2$ (cis or trans)	0	12083.3	0	0	0	0
$CH=CH-CH_2-CH_2$ (α- cis)	432.984	−20.1388	584.025	80.5552	191.319	−60.4164
$CH=CH-CH_2-CH_2$ (α- trans)	−120.833	−367.533	1198.26	40.2776	45.3123	−30.2082
$CH-CH_2-CH_2-CH$ (β)	−498.435	−312.151	−2034.02	−35.2429	−125.868	−95.6593

Table 4. Lennard-Jones Parameters of Polybutadiene

interaction type	energy ϵ [K]	diameter σ [Å]
$CH \leftrightarrow CH$	50.347	3.8
$CH \leftrightarrow CH_2$	51.102	4.257
$CH_2 \leftrightarrow CH_2$	47.125	4.5

3. It is similar to the square-well potential; however, it has an additional repulsive step (inverse well):

$$U_{CK}(r) = \begin{cases} +\infty & : \quad r < \sigma - 0.12 \\ +3\epsilon & : \quad \sigma - 0.12 < r < \sigma \\ -\epsilon & : \quad \sigma < r < \lambda\sigma \\ 0 & : \quad \lambda\sigma < r \end{cases} \tag{9}$$

The height $+3\epsilon$ and the width 0.12σ of the repulsion step were empirically determined by Chen and Kreglewski in order to improve the description of the equation of state for some small molecules, e.g., noble gases and short alkanes.[61] Here we study the structure which can be obtained for a dense melt in computer simulations using the Chen−Kreglewski potential and compare these results to those from the coarse-grained bead-spring model and from the atomistic simulations. For modeling chain molecules using the Chen−Kreglewski potential, the monomers are connected by rigid bonds as it is assumed in many analytical theories for the equations of state.

2.4. Simulation Details. The Monte Carlo simulations of the coarse-grained bead-spring model and the Chen−Kreglewski potential are performed using the computer code which has been developed in the Condensed Matter Theory Group at the University of Mainz for the Lennard-Jones+FENE model[64] and generalized to different interaction potentials.[65,66] For these potentials we generate chains with $m = 29$ beads per molecule. This number of monomers represents a 4:1 mapping of an atomistic polymer with 116 united atoms per chain onto the coarse-grained chain molecules. This mapping ratio has been found to be an optimal degree of the coarse-graining for polybutadiene, which we discuss later on in this paper. In the simulations of the LJ+FENE chains, the simulation cell contains 160 chain molecules. For the CK model we use 40 chains per cell. Initial configurations of the chain systems are generated using the configurational bias method.[1] After equilibrating the chains, the MC simulations are performed in an *NpT* ensemble, in which in addition to volume change (once per MC step) we include the local displacement of randomly chosen monomers (every monomer once per MC step) as well as "reptation moves"[67] where a monomer at one end of the chain, which is chosen randomly, is cut off and added at the other end of the same molecule (typically 100 attempts per MC step).

The choice of the reduced temperature and pressure can have a significant influence on the local packing of the monomers in the coarse-grained simulations.[68] Unfortunately, it is not possible to simply utilize the values of the molecular interaction parameter, ϵ, from a united atom model for a coarse-grained one since due to a different representation of the structure on the length scale of few monomers the energy scales of these modeling approaches can be different by 1 order of magnitude: In the atomistic model, the monomers are bonded at distance $r_0 \approx 1.5$ Å overlapping strongly and forming a kind of flexible tube of the diameter ≈4.5 Å. In the coarse-grained model, a bead accounts for all possible interactions of approximated united atoms with other ones; therefore, the interaction energy of a bead is an average interaction energy, which is very different from that in the united atom model.

The reduced temperature and pressure for the coarse-grained simulations are obtained using the results of our recent investigation of phase diagrams of chain molecules[65,66] which is based on Monte Carlo simulations, analytical equations of state, and experimental data of polybutadiene. The correlation of the experimental data of polybutadiene at temperatures between 300 and 460 K and pressures from the atmospheric one to 200 MPa has yielded an estimate of $\epsilon/k_B \approx 269$ K and $\sigma = 3.4$ Å, from which one can calculate reduced temperatures $T^* = 240\ \text{K}/(\epsilon/k_B) \approx 0.892$ and $T^* = 353\ \text{K}/(\epsilon/k_B) \approx 1.312$ and reduced pressure $p^* = 10^5\ \text{Pa}\cdot(\sigma^3/\epsilon) \approx 1.06\cdot10^{-3}$. Hence, the coarse-grained systems are simulated here at the reduced temperatures $T^* \equiv k_BT/\epsilon = 0.9$ and 1.3 and reduced pressure $p^* \equiv p\sigma^3/\epsilon = 0.001$ yielding very similar thermodynamic states (i.e., a melt at ambient pressure and temperature) for both the atomistic and the coarse-grained models. The reduced monomer density, $\rho^* \equiv \rho\sigma^3$, obtained at these temperatures is 0.893 and 0.781 in LJ+FENE and 0.863 and 0.789 in CK simulations, respectively.

3. Results and Discussions

3.1. Pair Correlation Functions of Polybutadiene. The segment−segment correlation functions calculated from the united atom molecular dynamics simulation data of polybutadiene[37] are shown in Figure 4. The detailed structure revealed in this figure represents different kinds of interactions between functional groups of polybutadiene and different energy scales for these interactions, if compared to the thermal fluctuations at the simulated conditions. At 240 K the structure is more pronounced (compare the height of the minima and maxima of $g(r)$) and more compact (compare r at the extremes) than at 353 K. It is convenient to separately investigate the segment−segment correlations for monomers which belong to the same backbone (intramolecular) and to

different molecules (intermolecular). The total segment−segment correlation function is the sum of the intra- and intermolecular functions.

At first, we analyze the intramolecular correlations. The three peaks at distances $r = 1.34$, 1.5, and 1.53 Å correspond to the covalent bonds between the carbon groups: *CH═CH* (double bonds), CH_2-CH (α-bonds), and CH_2-CH_2 (β-bonds), respectively. Of course, these sharp peaks are a somewhat artificial feature of the fixed bond length constraint used in the simulation. A more realistic description allowing for a more accurate potential for the length of the covalent bonds would result in peaks of finite (albeit small) width. Therefore it is not relevant to keep such features for a coarse-grained modeling. The fourth peak at distance $r = 2.53$ Å is somewhat broader due to the bending potential eq 4 which restricts the relative positions of the next-nearest-neighbor carbon groups along the backbone. Several peaks between $r = 3$ Å and $r = 4$ Å are due to the four-body interactions which are modeled by the torsional potential eq 5 in the simulation.[37] At large length scales, the intramolecular correlation function becomes smoother as the monomer positions become uncorrelated and decays on the length scale of the molecular extension.

The intermolecular correlation functions of polybutadiene, which are also shown in Figure 4, exhibit the main peak at distance $r \approx 5$ Å, a wide correlation hole at $r \leq 3$ Å caused by the excluded volume interaction of the Lennard-Jones potential and typical oscillations for long length scales. The intermolecular pair correlation function is much smoother than for the intramolecular one because the nonbonded interactions described by the Lennard-Jones interactions are much softer.

3.2. Coarse-Graining of the Atomistic Model. Here we investigate how the correlation functions develop if several united atoms of the atomistic model are replaced by one effective segment. Our aim is to find out the number, *n*, of united atoms taken to correspond to one effective segment which yield an optimal coarse-grained representation within the bead-spring model, in the sense that similar conditions are obtained. We denote *n* as a "degree of coarse-graining" in the following. Since there is no unique way of mapping a polymer molecule onto a coarse-grained molecule we calculate the average positions of centers of mass without accounting for the molecular masses of different carbon groups, i.e., *CH*, CH_2, and CH_3 are equivalently treated in our approach.

Figure 5 shows the intramolecular correlation functions obtained in the atomistic simulations of polybutadiene at 240 K and atmospheric pressure as well as for different coarse-graining degrees, *n*, given by the numbers in the figure. One can see that different choices for *n* have different impact on the structure. For low degree of coarse-graining (e.g., $n = 2$ and 3), additional peaks appear which are not present in the atomistic structure. At high degree of coarse-graining ($n \geq 8$ shown in the inset), the structure is nearly uniform, and correlations can also be found on length scales shorter than the size of the coarse-grained segments due to the overlapping of the effective segments.

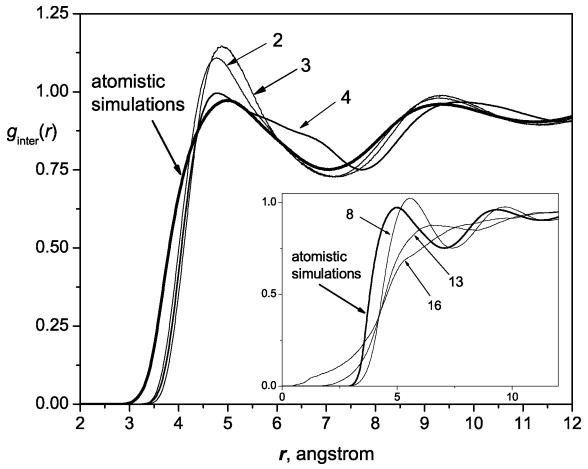


Figure 6. Coarse-graining of the intermolecular segment−segment correlation function of polybutadiene at 240 K. For $n \leq 8$, the correlation function changes quantitatively. For larger degrees of the coarse-graining (e.g., $n > 13$), the correlation peaks and minima disappear and the correlation hole at $r < 3$ shrinks. A bold curve is the correlation function calculated using the UA MD data[37] (also shown in Figure 4).

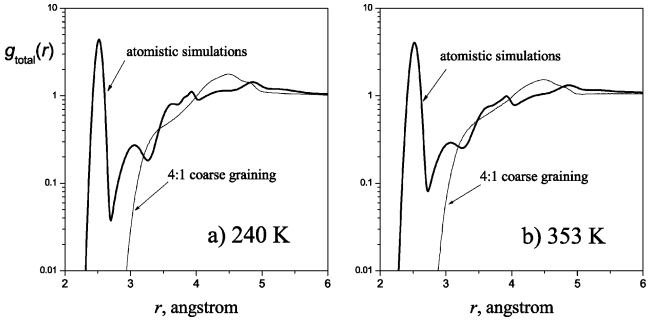


Figure 7. Total segment−segment correlation functions of polybutadiene for the optimal value $n = 4$ of the coarse-graining degree (thin curve). Bold curves are the correlation functions calculated using the UA MD data[37] shown in Figure 4.

The intermolecular correlations are less sensitive to the coarse-graining for small values of *n*, in contrast to the intramolecular correlations. In Figure 6 one can see that the intermolecular correlation function only slightly changes for $n < 4$. At first, the intermolecular correlations in the first correlation shell (the peak at $r \approx 4.5$ Å) increase due to the loss of local internal structure. For the coarse-graining degree $n = 4$, the intermolecular structure becomes nearly the same as in the atomistic model. In this approximation, butadiene, which is the monomer of polybutadiene, is replaced by one effective monomer. For a larger degree of coarse-graining, the structure decreases, and we can observe the fusion of the effective segments (e.g., for $n = 16$, $g(r)$ becomes nonvanishing down to $r = 0$).

The total correlations function for the optimal choice, $n = 4$, of the coarse-graining degree is shown together with the atomistic total correlations function in Figure 7. Although, the correlation functions change with temperature, this effect is not very strong for polybutadiene at 240 K to 353 K in contrast to the variation of *n*, which has a much stronger impact on the structure functions. One can conclude that for

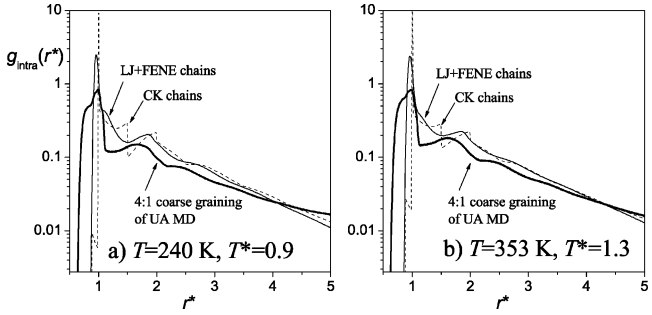


Figure 8. Intramolecular segment−segment correlation functions obtained from the Monte Carlo simulations of the LJ+FENE bead-spring chains (thin solid curve) and the Chen−Kreglewski chains (dashed curve) for chains with 29 beads/molecule at reduced pressure $p^* \equiv p\sigma^3/\epsilon = 0.001$ and reduced temperatures $T^* = 0.9$ and 1.3. Bold curves are from the UA MD simulations of polybutadiene[37] at 240 and 353 K. (The bold curve in (a) is the curve 4 in Figure 5; however, here the distance $r^* = r/\sigma$ has been scaled by the effective Lennard-Jones diameter $\sigma = 4.5$ Å.)

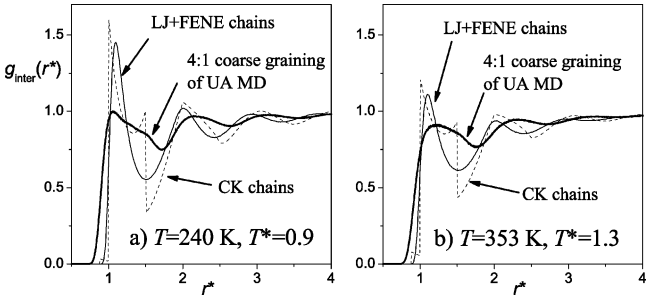


Figure 9. Intermolecular pair correlation functions obtained from the Monte Carlo simulations of the bead-spring chains (thin solid curves) and Chen−Kreglewski chains (dashed curves) for the same systems as in Fugure 8. Legends as in Figure 8.

the pair correlation functions there is an optimal value of the coarse-graining degree beyond which the intermolecular structure simplifies drastically and the intermolecular correlations nearly vanish. In this limit the polymer melt is described as if it would result from an ensemble of noninteracting ideal random walks, i.e., the well-known Flory picture of dense polymer melts. On this extremely simplified level of the description, all effects of chemical detail would be lost. In the case of polybutadiene studied here this value of n when correlations are essentially lost is between 8 and 16, which corresponds to approximation of three to four monomers (butadiene) by one effective segment. Too small values of coarse-graining degree, however, can produce unphysical correlations such as curve 3 in Figure 5.

3.3. Structure of LJ+FENE and CK Chains. Discontinuous potentials, e.g., the hard sphere and square-well potentials, are widely used in statistical theories of molecules. The reason for using such simplified potentials is, of course, not a better representation of the physics of molecular interactions in real substances by such models but is rather pragmatic and dictated by the complexity of molecular theories which have to deal, for example, with multiple integration of the potential functions. Even for simple potentials these theories frequently can only be solved in analytical form using several approximations. However, such a discontinuity of the steplike potential influences the structure of fluids.

In Figure 8 the intramolecular segment−segment distribution functions are compared for the Chen−Kreglewski and the LJ+FENE potentials for chain molecules with $m = 29$ monomers/chain at two temperatures. This number of monomers corresponds to the 4:1 coarse-graining of the atomistic simulation data of polybutadiene (bold curve in the figure). The molecular parameters, ϵ and σ, and the reduced state parameters, $p^* = 0.001$ and $T^* = 0.9$ and 1.3, are chosen the same for the CK and LJ+FENE potentials.

At two temperatures the overall agreement for the intramolecular correlation functions from the coarse-grained and atomistic models does not change significantly: Some change of the structure occurs on length scales between $r^* = 1$ and $r^* = 2$ (i.e., in the range of the potential attractive part). However, there are significant qualitative differences in the correlation functions of the LJ+FENE (thin solid) and CK (dashed) models. The distribution functions from the LJ+FENE potential and the UA MD simulations are smooth and continuous in the whole range of molecular distances, as one would expect. However, the Chen−Kreglewski model exhibits steplike jumps in the intramolecular correlations at several distances: at $r^* = 0.88$ due to the excluded volume interactions with a repulsive step, at $r^* = 1.5$ and $r^* = 2$ due to the square-well attraction and a strong peak at $r^* = 1$ (the bond length). Such peaks have been already observed in the atomistic simulations (Figure 4), which correspond to the nearly unchanged bond lengths between the carbon groups in real polymers. For the CK chains, however, this peak is due to the implementation of rigid bonds between the coarse-grained segments. The LJ+FENE model exhibit a continuous, smooth maximum at bond distances ($r^* \approx 0.96$). This peak can rather well be compared with the peak in the coarse-grained structure of polybutadiene (bold curve), although the last one is lower and broader.

The intermolecular correlation functions for the LJ+FENE and CK potentials are compared with the 4:1 coarse-graining of the atomistic data in Figure 9. The CK correlation function (dashed) shows similar discontinuities as discussed above for the intramolecular correlation function. The LJ+FENE correlation function (thin solid) yields a correct qualitative description of the 4:1 coarse-grained atomistic correlation function (bold curve in the figure), although the LJ+FENE structure is more pronounced. At higher temperature, the agreement improves because the height of the LJ+FENE main peaks decreases, and it approaches the atomistic one.

3.4. Bond Angle Distributions from Atomistic Modeling. The atomistic simulations yield the distribution of the bond angle with two very pronounced peaks at 112° for the CH_2-CH_2-CH groups and 126° for the $CH_2-CH=CH$ groups. This distribution decreases rapidly for other values of the bond angle. Figure 10 shows this bond angle distribution calculated from the atomistic simulation data for polybutadiene at 240 and 353 K and how this distribution changes if several carbon groups are approximated by an effective segment (for the 240 K curve). For increasing degree of coarse-graining, the peaks shift to higher values

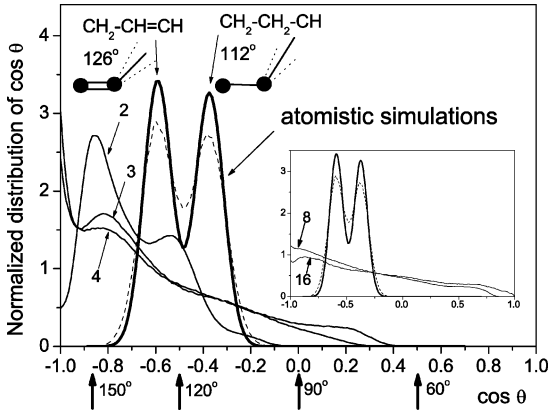


Figure 10. Bond-angle distribution of polybutadiene at 240 K (bold) and 353 K (dashed) and atmospheric pressure obtained using the UA MD simulation data.[37] The atomistic simulations at two temperatures yield qualitatively similar distributions with two peaks at $\theta \approx 112°$ and $\theta \approx 126°$ for CH_2-CH_2-CH and $CH_2-CH=CH$ groups, respectively. These distributions change significantly for the coarse-grained models (numbered curves). For clarity, only the coarse-grained results for 240 K are shown. The 4:1 curve for 353 K is shown in Figure 12.

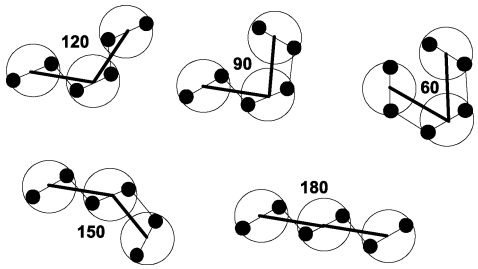


Figure 11. Schematic examples illustrating the bond angles which can be found for effective chains in a coarse-grained model if in the atomistic model only one bond angle is allowed (for instance, the bond angle 120° between the connected filled circles). Bold lines connecting the centers of the big, open circles are bonds in a coarse-grained model.

of the bond angle and the structure becomes smeared-out. For a coarse-graining degree $n > 4$ (i.e., if several butadiene groups of polybutadiene are approximated by one monomer), the peaked structure cannot be recognized any more in the coarse-grained model. The bond angle distribution becomes nearly uniform in the whole range of bond angle values. Figure 11 schematically illustrates that different bond angles can arise in a coarse-grained model even if in an atomistic model only one bond angle exists. For $n = 4$, which we so far identified as the best choice for the degree of coarse-graining, a residual structure at $\theta \approx 150°$ is, however, clearly visible in the angle distribution function in Figure 10. This means that on this coarse-graining level an effective bending potential should be used in simulations trying to describe polybutadiene, as in case of polyethylene.[69]

3.5. Bond-Angle Distributions from the LJ+FENE and CK Models. For simplicity, the coarse-grained models employed for the investigation here do not include explicitly a bending interaction potential, which however can be very important, for example, in the modeling of polymer crystallization.[69] As further motivation for this omission we note that in the first-order thermodynamic perturbation theory of Wertheim the chain molecules are modeled by flexible chains of tangent spheres. The stiffness of the molecules is considered as a second-order contribution to the free energy and is ignored in that analytical theory of chain molecules. To estimate the impact of simplifications due to ignoring the stiffness, the bond angle distribution is investigated here for polymer molecules.

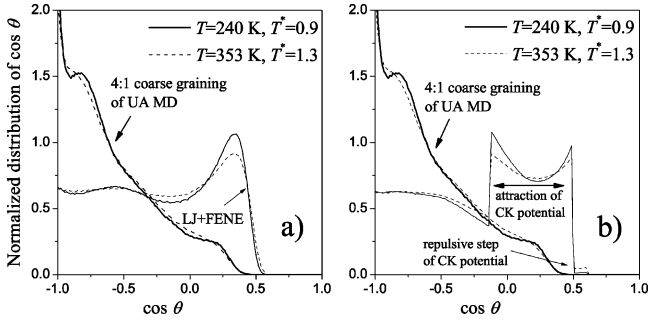


Figure 12. (a) Bond-angle distributions obtained in the Monte Carlo simulations of the bead-spring chains with 29 beads/molecule at reduced temperatures $T^* = 0.9$ (solid) and $T^* = 1.3$ (dashed) and reduced pressure $p^* = 0.001$. The 4:1 coarse-grained UA MD simulations of polybutadiene are at 353 K (dashed) and 240 K (solid, this is curve 4 in Figure 10). (b) Same as (a) but for the Chen−Kreglewski potential instead of the bead-spring potential. Although, no bending interaction is explicitly included in the coarse-grained LJ+FENE and CK simulations, the distributions of the bond angles become very inhomogeneous due to the excluded volume interactions.

The bond angle distributions for the coarse-grained LJ+FENE and Chen−Kreglewski potentials are compared in Figure 12 with the distribution functions obtained for the 4:1 coarse-graining of the UA MD data of polybutadiene at two temperatures. The bending interaction in polybutadiene is explicitly accounted for in the underlying atomistic model. The LJ+FENE bond angle distribution is nearly homogeneous over a wide range of large values of the bond angle. At $\theta \approx 70°$ ($\cos\theta_0 \approx 0.33$), it exhibits a maximum and vanishes rapidly for smaller bond angles. Hence, one can speak about stiffness of the LJ+FENE molecules with the equilibrium angle $\theta_0 \approx 70°$, although no bending potential has been included in the simulations. From the examination of the correlation hole and the main peak in Figures 8 and 9 one concludes that these are effects of the excluded volume interaction, which is explicitly present in the LJ+FENE potential.

The bond angle distribution for the Chen−Kreglewski chains exhibits steplike discontinuities similar to those discussed in detail for the pair correlation functions. It is worth mentioning here that in contrast to the bead-spring model the CK angle distribution jumps to zero for small values of the bond angle ($\theta \approx 52°$), which is due to the hard sphere repulsion in the CK potential. For the LJ+FENE model, this property decreases smoothly. Furthermore, there is a minimum of the CK distribution function at $\theta \approx 76°$ ($\cos\theta \approx 0.24$), which yields two equilibrium bond angles at $\theta \approx 96°$ ($\cos\theta \approx -0.1$) and $\theta \approx 62°$ ($\cos\theta \approx 0.47$). This

artifact is caused by the square-well attraction. The LJ+FENE potential yields a distribution function with one maximum and, hence, with only one equilibrium bond angle at $\theta_0 \approx 70°$.

4. Conclusions

Mapping of an atomistic model of polymers onto a coarse-grained one is a popular step to tackle the problem of bridging in structure from the subangstrom scale to mesoscopic scales, involving at the same time a considerable speed-up of the simulation codes. However, such a mapping necessarily is somewhat approximate, some information is necessarily lost, and the question arises at which level of coarse-graining one obtains the most reasonable compromise between atomistic accuracy and simulational efficiency. This is addressed in the present paper by a systematic comparison between different levels of coarse-graining.

We conclude that there is a reasonable correspondence, at least for the static structure functions, between correlations of LJ+FENE chains and suitable coarse-grained atomistic models, e.g., the coarse-graining degree $n = 4$ for polybutadiene, i.e., when 4 united atoms yield one effective subunit. Further investigations, for example, including dynamics, would be of great interest in order to complete the understanding of the coarse-graining effect on different properties.

In contrast, the Chen−Kreglewski model for chain molecules yields many artificial features in their correlations which have no counterpart in reality. As is well-known, the CK model is widely used mostly as a starting point for the approximate equation-of-state theories such as PC-SAFT. However, our recent analysis of this theory has revealed a variety of artifacts such as artificial gas−gas and liquid−liquid equilibria causing also systematic deviations between the equation of state and real data in regime of experimental interest,[65,66] while other theories based on the LJ+FENE potential exhibit a much better description of chain molecules. Therefore, we suggest abandoning the discontinuous potentials such as the CK potential as a starting point for the description of polymers.

Acknowledgment. We benefited from stimulating discussions with H. Weiss, U. Nieken, W. Hültenschmidt, Ch. Ch. Liew, A. Moreira, and O. Evers (all BASF Aktiengesellschaft, Ludwigshafen). We thank P. Virnau (MIT, Cambridge) for providing us with the Monte Carlo LJ+FENE code and for interesting discussions. CPU time was provided by the computing center of the Johannes-Gutenberg University of Mainz.

References

(1) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications;* Academic Press: 2002.

(2) Allen, M. P.; Tildesley, D. J. *Computer simulation of liquids;* Clarendon Press: Oxford, 1991.

(3) Landau, D. P.; Binder, K. *A Guide to Monte Carlo Simulations in Statistical Physics;* Cambridge University Press: Cambridge, 2000.

(4) *Simulation Methods for Polymers*; Kotelyanskii, M., Theodorou, D. N., Eds.; Marcel Dekker: New York, 2004.

(5) Faller, R.; Schmitz, H.; Biermann, O.; Müller-Plathe, F. *J. Comput. Chem.* **1999**, *20*, 1009−1017.

(6) Polson, J. M.; Frenkel, D. *J. Chem. Phys.* **1999**, *111*, 1501−1510.

(7) Das, C.; Frenkel, D. *J. Chem. Phys.* **2003**, *118*, 9433−9440.

(8) Nielsen, S. O.; Lopez, C. F.; Srinivas, G.; Klein, M. L. *J. Phys.: Condens. Matter* **2004**, *16*, R481−R512.

(9) Paul, W. A Mapping from Atomistic Polymer Models to Coarse-Grained Models. In *Multiscale Computational Methods in Chemistry and Physics*; Brandt, A., Bernholc, J., Binder, K., Eds.; IOS Press: 2001.

(10) Binder, K.; Paul, W.; Santos, S.; Suter, U. W. Coarse-Graining Techniques. In *Simulation Methods of Polymers*; Kotelyanskii, M., Theodorou, D. N., Eds.; Marcel Dekker: New York, 2004.

(11) Paul, W.; Binder, K.; Kremer, K.; Heermann, D. W. *Macromolecules* **1991**, *24*, 6332−6334.

(12) Tries, V.; Paul, W.; Baschnagel, J.; Binder, K. *J. Chem. Phys.* **1997**, *106*, 738−748.

(13) Rapold, R. F.; Mattice, W. L. *Macromolecules* **1996**, *29*, 2457−2466.

(14) Doruker, P.; Mattice, W. L. *Macromolecules* **1997**, *30*, 5520−5526.

(15) Tschöp, W.; Kremer, K.; Batoulis, J.; Bürger, T.; Hahn, O. *Acta Polym.* **1998**, *49*, 61−74.

(16) Tschöp, W.; Kremer, K.; Batoulis, J.; Hahn, O.; Bürger, T. *Acta Polym.* **1998**, *49*, 75−79.

(17) Baschnagel, J.; Binder, K.; Doruker, P.; Gusev, A. A.; Hahn, O.; Kremer, K.; Mattice, W. L.; Müller-Plathe, F.; Murat, M.; Paul, W.; Santos, S.; Suter, U. W.; Tries, V. *Adv. Polym. Sci.* **2000**, *152*, 41−156.

(18) Kremer, K.; Müller-Plathe, F. *MRS Bull.* **2001**, *26*, 205−10.

(19) Hahn, O.; delle Site, L.; Kremer, K. *Macromol. Theory Simul.* **2001**, *10*, 288−303.

(20) Müller-Plathe, F. *ChemPhysChem* **2002**, *3*, 754−769.

(21) Abrams, C. F.; Kremer, K. *Macromolecules* **2003**, *36*, 260−267.

(22) Wertheim, M. S. *J. Stat. Phys.* **1984**, *35*, 19−34.

(23) Wertheim, M. S. *J. Stat. Phys.* **1984**, *35*, 35−47.

(24) Wertheim, M. S. *J. Stat. Phys.* **1986**, *42*, 459−476.

(25) Wertheim, M. S. *J. Stat. Phys.* **1986**, *42*, 477−492.

(26) Huang, S. H.; Radosz, M. *Ind. Eng. Chem. Res.* **1990**, *29*, 2284−2294.

(27) Hoogerbrugge, P. J.; Koelman, J. M. V. A. *Europhys. Lett.* **1992**, *92*, 155−160.

(28) Español, P.; Warren, P. *Europhys. Lett.* **1995**, *30*, 191−196.

(29) Groot, R. D.; Warren, P. B. *J. Chem. Phys.* **1997**, *107*, 4423−4435.

(30) Kranenburg, M.; Venturoli, M.; Smit, B. *J. Phys. Chem. B* **2003**, *107*, 11491−11501.

(31) Guerrault, X.; Rousseau, B.; Farago, J. *J. Chem. Phys.* **2004**, *121*, 6538−6546.

(32) Murat, M.; Kremer, K. *J. Chem. Phys.* **1998**, *108*, 4340−4348.

(33) Smith, G. D.; Paul, W. *J. Phys. Chem. A* **1998**, *102*, 1200−1208.

(34) Smith, G. D.; Borodin, O.; Paul, W. *J. Chem. Phys.* **2002**, *117*, 10350−10359.

(35) Smith, G. D.; Paul, W.; Monkenbusch, M.; Richter, D. *J. Chem. Phys.* **2001**, *114*, 4285−4288.

(36) Smith, G. D.; Borodin, O.; Bedrov, D.; Paul, W.; Qiu, X.; Ediger, M. D. *Macromolecules* **2001**, *34*, 5192−5199.

(37) Krushev, S. Computersimulationen zur Dynamik und Statik von Polybutadienschmelzen, Dissertation Thesis, Universität Mainz, 2002.

(38) Haas, F.; Hilfer, R.; Binder, K. *J. Phys. Chem.* **1996**, *100*, 15290−15300.

(39) Stadler, C.; Lange, H.; Schmid, F. *Phys. Rev. E* **1999**, *59*, 4248−4257.

(40) Stadler, C.; Schmid, F. *J. Chem. Phys.* **1999**, *110*, 9697−9705.

(41) Düchs, D.; Schmid, F. *J. Phys.: Condens. Matter* **2001**, *13*, 4853−4862.

(42) Marrink, S. J.; de Vries, A. H.; Mark, A. E. *J. Phys. Chem. B* **2004**, *108*, 750−760.

(43) Nielsen, S. O.; Lopez, C. F.; Moore, P. B.; Shelley, J. C.; Klein, M. L. *J. Phys. Chem. B* **2003**, *107*, 13911−13917.

(44) Müller, M.; Katsov, K.; Schick, M. *J. Polym. Sci.: Polym. Phys. B* **2003**, *41*, 1441−1450.

(45) Stevens, M. J. *J. Chem. Phys.* **2004**, *121*, 11942−11948.

(46) Sayar, M.; Stupp, S. I. *Macromolecules* **2001**, *34*, 7135−7139.

(47) Maiti, P. K.; Lansac, Y.; Glaser, M. A.; Clark, N. A.; Rouault, Y. *Langmuir* **2002**, *18*, 1908−1918.

(48) Srinivas, G.; Shelley, J. C.; Nielsen, S. O.; Discher, D. E.; Klein, M. L. *J. Phys. Chem. B* **2004**, *108*, 8153−8160.

(49) Binder, K.; Müller, M. *Curr. Opin. Colloid Interface Sci.* **2000**, *5*, 315−323.

(50) Yingling, Y. G.; Garrison, B. J. *J. Phys. Chem. B* **2004**, *108*, 1815−1821.

(51) Izvekov, S.; Violi, A.; Voth, G. A. *J. Phys. Chem. Lett. B* **2005**, *109*, 17019−17024.

(52) Shelley, J. C.; Shelley, M. Y.; Reeder, R. C.; Bandyopadhyay, S.; Klein, M. L. *J. Phys. Chem. B* **2001**, *105*, 4464−4470.

(53) Shelley, J. C.; Shelley, M. Y.; Reeder, R. C.; Bandyopadhyay, S.; Moore, P. B.; Klein, M. L. *J. Phys. Chem. B* **2001**, *105*, 9785−9792.

(54) Kremer, K.; Grest, G. S. *J. Chem. Phys.* **1990**, *92*, 5057−5086.

(55) Smit, B. *J. Chem. Phys.* **1992**, *96*, 8639−8640.

(56) Binder, K.; Müller, M.; Virnau, P.; MacDowell, L. G. *Adv. Polym. Sci.* **2005**, *173*, 1−104.

(57) Virnau, P.; Müller, M.; MacDowell, L. G.; Binder, K. *Comput. Phys. Comm.* **2002**, *147*, 378−381.

(58) Nielsen, S. O.; Lopez, C. F.; Srinivas, G.; Klein, M. L. *J. Chem. Phys.* **2003**, *119*, 7043−7049.

(59) Reith, D.; Meyer, H.; Müller-Plathe, F. *Macromolecules* **2001**, *34*, 2335−2345.

(60) Virnau, P.; Müller, M.; MacDowell, L. G.; Binder, K. *J. Chem. Phys.* **2004**, *121*, 2169−2179.

(61) Chen, S. S.; Kreglewski, A. *Ber. Bunsen-Ges.* **1977**, *81*, 1048−1049.

(62) Gross, J.; Sadowski, G. *Ind. Eng. Chem. Res.* **2001**, *40*, 1244−1260.

(63) Gross, J.; Sadowski, G. *Ind. Eng. Chem. Res.* **2002**, *41*, 1084−1093.

(64) Virnau, P. Monte Carlo-Simulationen zum Phasen- und Keimbildungsverhalten von Polymerlösungen, Dissertation Thesis, Universität Mainz, 2003.

(65) Yelash, L.; Müller, M.; Paul, W.; Binder, K. *J. Chem. Phys.* **2005**, *123*, 14908−15.

(66) Yelash, L.; Müller, M.; Paul, W.; Binder, K. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3728−3732.

(67) Binder, K. In *Monte Carlo and Molecular Dynamics Simulations in Polymer Science*; Binder, K., Ed.; Oxford University Press: New York, 1995.

(68) Heine, D. R.; Grest, G. S.; Curro, J. C. *Adv. Polym. Sci.* **2005**, *173*, 209−249.

(69) Vettorel, T.; Meyer, H. *J. Chem. Theory Comput.* **2006**, *2*, 616−629.

CT0502099

JCTC Journal of Chemical Theory and Computation

Multiscale Modeling of Poly(ethylene oxide)−Poly(propylene oxide)−Poly(ethylene oxide) Triblock Copolymer Micelles in Aqueous Solution

Dmitry Bedrov,* Chakravarthy Ayyagari, and Grant D. Smith

Departments of Materials Science & Engineering and Chemical Engineering, University of Utah, 122 South Central Campus Drive, Room 304, Salt Lake City, Utah 84112

Received December 28, 2005

Abstract: We present a multiscale modeling approach for simulation of poly(ethylene oxide)−poly(propylene oxide)−poly(ethylene oxide) triblock copolymer micelles in aqueous solution. We rely on systematic elimination of computationally expensive degrees of freedom yet retain implicitly their influence on the remaining degrees freedom in a coarser-grained model. Quantum chemistry (QC) calculations, atomistic explicit solvent (AES) molecular dynamics (MD) simulations, and coarse-grained implicit solvent (CGIS) simulations have been applied to investigate physical properties of these important self-assembling triblock copolymers. High-level QC calculations have been used to parametrize classical atomistic force fields that implicitly take into account and reproduce the important energetic and structural features due to correlations of electronic degrees of freedom. AES MD simulations utilizing the QC-based force fields have been used to provide structural and conformational properties of polymers in aqueous solution which were subsequently used for parametrization of the CGIS model using the Inverted Boltzmann method. The CGIS simulations were then employed to investigate structural properties of two PEO−PPO−PEO micelles (EO_{13}−PO_{30}−EO_{13} and EO_{99}−PO_{65}−EO_{99} also known as Pluronic L64 and F127, respectively) in aqueous solution.

I. Introduction

Poly(ethylene oxide) (PEO) and PEO-based copolymers, including PEO−poly(propylene oxide)−PEO or PEO−PPO−PPO, triblock copolymers (also known as Pluronics), are utilized in a wide variety of applications in aqueous environments including protein crystallization,[1,2] modification of surfaces for biocompatibility,[3,4] control of particle aggregation in solutions,[5,6] and drug delivery.[7] At room and physiological temperatures PEO of any molecular weight is soluble in water at any concentration, while PPO has very limited solubility in water for short oligomers only. In dilute solutions at room-temperature PEO−PPO−PEO copolymers can usually dissolve in water. As the temperature or concentration of polymer increases PEO−PPO−PEO triblocks form isolated micelles (usually spherical in shape but with other shapes possible depending on the triblock architecture) with PPO units in the core and PEO chains forming the corona. Upon further increase in polymer concentration or temperature, the supramicellar structures such as disordered gels and crystals as well as cylindrical and lamellar morphologies can form. The formation of micelles and supramicellar structures is quite sensitive to solvent quality. Slight changes in temperature,[8] small addition of cosolvent,[9] or change in pH[10] can significantly alter the phase behavior of PEO−PPO−PEO aqueous solutions by shifting the delicate balance between hydrophilic and hydrophobic interactions in the system.

A typical Pluronic chain contains about 50−300 ether monomers, while a Pluronic micelle formed in aqueous solution consists of up to 100 chains and has a diameter of several nanometers.[11] Dynamic light scattering measurements of micellization kinetics for Pluronic chains in water upon

* Corresponding author e-mail: bedrov@cluster2.mse.utah.edu.

10.1021/ct050334k CCC: $33.50
Published on Web 03/28/2006

increasing solution temperature have shown that formation of micelles is a complex multistage process.[12,13] For example, for L64 Pluronic ($EO_{13}-PO_{30}-EO_{13}$) the micellization was found to consist of three processes: (1) The fast process associated with unimers joining/leaving incipient micelles with characteristic relaxation time on the order of microseconds. (2) The relaxation of micelle size distribution, a process during which the initially formed micelles adjust their size in accordance with equilibrium micelle size distribution corresponding to imposed thermodynamic conditions. During this process, the initially formed micelles can collide and merge into one larger micelle, or a large micelle can split into smaller micelles, or a micelle can dissolve into unimers which can then join other micelles. While the nature of operative mechanism for this process is a topic of ongoing debate, the characteristic relaxation time for this process for L64 Pluronic is on the order of milliseconds. (3) A process observed only for some Pluronics at temperatures much higher than the critical micellization temperature and has been associated with clustering of micelles. This process has operative relaxation times on the order of seconds. It was also found that all these kinetic processes slow dramatically with increasing molecular weight of the Pluronic chain.[12]

Taking into account the slow kinetics of formation and relatively large dimensions of Pluronic micelles it is impractical to expect conventional, brute force atomistic molecular simulations (molecular dynamics (MD) or Monte Carlo (MC)) to be able to model these systems. For example, a simulation of a single L64 micelle consisting of 40 triblock chains and hydrated by water would involve about 150 000 atoms. For such a large system accessing the operative time scales for the first (fast) micellization process is unrealistic even with utilization of massively parallel architectures. To date the modeling of Pluronics solutions has been limitted either to self-consistent field lattice models[14−16] or mean field density functional theory approaches (MesoDyn).[17−19] In both cases, an ideal Gaussian chain representation of Pluronic chains and mean-field approximations were employed, while thermodynamic parameters (interaction between species) were either determined semiempirically[17,18] or based on incomplete physical models.[14−16] While these mean field approches allow computationally expedient prediction of solution morhology and phase behavior, their ability to incorporate important atomistic scale phenomena (i.e., hydrogen bonding, changes in hydration structure, hydrophobic interactions, local conformations, solvent clustering, etc.) operative in these systems is very limitted.

In this paper we discuss an alternative approach to study Pluronic micelles in aqueous solution using a multiscale modeling hierachy in which computationally expensive degrees of freedom are systematically eliminated, while their contribution/influence on thermodynamics, structure, and conformations is implicitly retained in the coarser-grained model. We demonstrate for L64 and F127 Pluronics that our multiscale modeling approach can be used to efficiently study Pluronic micelles both at atomistic and coarse-grained levels.

II. Development of the Multiscale Model for Pluronic Chains in Aqueous Solution

The strategy of our multiscale modeling relies on systematic elimination of computationally expensive degrees of freedom while retaining implicitly their influence on the remaining degrees freedom in the coarser-grained model. For each level of coarsening we employ the corresponding most accurate and effective method/simulation technique available to investigate physical properties of the system at that level. Then, using the obtained information we parametrize a coarser model that incorporates all essential physics/phenomena observed at the finer level. Here, we briefly outline the general strategy of our multiscale modeling approach:

(1) We perform high-level quantum chemistry (QC) calculations on model compounds and their clusters to determine relative conformational and binding energies. Due to computational expense of high-level QC calculations this investigation is limited to short (one-three repeat units) oligomers and clusters with relatively few molecules (e.g., an ether oligomer with two water molecules). Optimal geometries, relative conformational energies, and cluster binding energies as a function of molecules separation and orientation obtained from QC calculations are subsequently utilized in parametrization of fully atomistic force fields for classical MD simulations where all subatomic scale degrees of freedom and correlations are averaged out and represented implicitly through effective nonbonded van der Waals interactions and through simplified intramolecular potentials that describe bond and bend vibrations and dihedral rotations.

(2) Using parametrized QC-based force fields we conduct extensive atomistic, explicit solvent (AES) MD simulations of low molecular weight polymeric systems (e.g., PEO and PPO oligomers in aqueous solutions). These simulations serve two purposes: they (a) allow us to validate the developed force fields by direct comparison of thermophysical, structural, and dynamical properties obtained from MD simulations with data available from experiments and (b) provide us with a vast amount of molecular scale structural, conformational, and dynamic properties/correlations that help to identify important interactions/correlations which can be used to parametrize coarse-grained models.

(3) Using conformational and structural properties obtained from AES MD simulations we parametrize the coarse-grained, implicit solvent (CGIS) model in which polymer chain segments or monomers are represented as single force centers (beads) and solvent is treated implicitly. Langevin dynamics or Lattice-Boltzmann simulations are then conducted using the CGIS representation of polymer chains. These simulations are about 3 orders of magnitude computationally less expensive than simulations with the AES model, therefore allowing us (a) to simulate realistic size polymers (e.g. Pluronic chains) and (b) to significantly extend the accessible time scales (e.g., extend the simulation time window up to the range where the first process of micellization can be resolved for some Pluronics).

(4) Finally, equilibrium configurations of polymeric systems obtained from CGIS simulations can be reverse mapped back to the AES model, and AES MD simulations

can be conducted in order to investigate atomistic scale structural and dynamic properties/correlations in the system.

Below we discuss in detail each step of our multiscale modeling strategy as applied to Pluronics in aqueous solution.

A. Quantum Chemistry Calculations and Force Field Fitting. To parametrize the atomistic force field for PEO and PPO in aqueous solution previously we have conducted quantum chemistry calculations for 1,2-dimethoxyethane (DME) and 1,2-dimethoxypropane (DMP) compounds in the gas phase and studied the conformational energetics and rotational energy barriers in these oligomers. We also have calculated complexation energy of these ethers with water as a function of their separation for two different paths.[20] Details of these calculations can be found in refs 20−23. Molecular geometries, conformational and complexation energies, and electrostatic fields surrounding the model molecules determined from QC were then utilized to parametrize intra- and intermolecular interactions for the classical MD simulations in which each atom is represented as a single point force center. Local intramolecular interactions were described by bonds, bends, and dihedrals, while nonbonded intra- and intermolecular interactions were described by Lennard-Jones or Buckingham potentials and Coulomb interactions. The parameters for bonds, and bends, were fitted to reproduce optimal geometries of compounds obtained from QC calculations. Partial atomic charges were fitted to reproduce the dipole moment of and an electrostatic field around the molecule obtained from QC calculations. Nonbonded van der Waals interactions were then fitted to reproduce QC binding energies for the molecular clusters (e.g. paths between DME or DMP and water). Finally, the dihedral potentials for ethers were parametrized to reproduce relative conformational energies and rotational barriers obtained from QC calculations for model compounds such as DME and DMP. The force field parameters and details regarding force field parametrization in general and specifically for ether/water systems can be found in our previous work.[20−23] In this study, the parameters for the Lennard-Jones interactions between PPO atoms and water were the same as used for PEO and water[20] for identical atom types (e.g. ether oxygen, ethylene carbons, ethylene hydrogen, etc.), while the distribution of atomic partial charges for each ether varied due to difference in their chemical structure.

B. Atomistic, Explicit Solvent MD Simulations. The force fields described above have been utilized extensively in MD simulations of PEO and PPO melts[24] as well as PEO in aqueous solution.[25,26] Specifically for the purpose of parametrization of the CGIS model for Pluronic chains in aqueous solution additional AES MD simulations of PEO and PPO oligomers in water were performed at atmospheric pressure and 298 K:

- PEO/water solutions with PEO $M_w = 530$ (12 repeat units) at 0.17 and 0.52 weight fraction of PEO
- PPO/water solutions with PPO $M_w = 349$ (6 repeat units) at 0.17 and 0.52 weight fraction of PPO
- PPO/PEO/water solutions with one PPO $M_w = 349$ (6 repeat units) chain and several PEO $M_w = 215$ (6 repeat units) chains at 0.52 weight fraction of ether molecules.

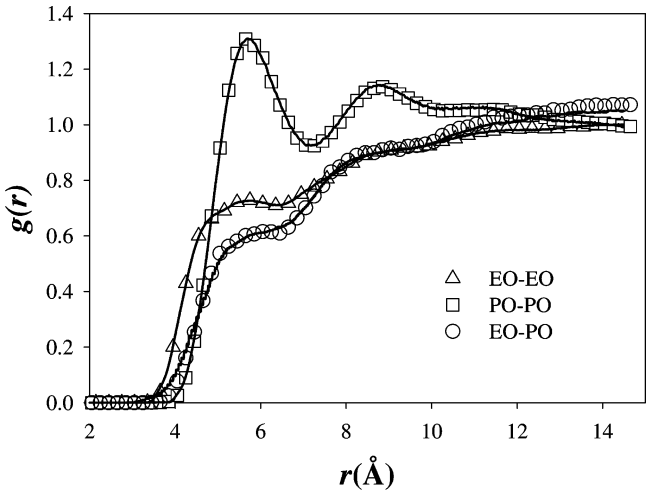


Figure 1. Intermolecular monomer−monomer pair distribution functions for ethers in aqueous solution at 298 K and 0.52 weight fraction of ethers as obtained from AES MD (symbols) and CGIS (lines) simulations.

All AES MD simulations were performed using the simulation package *Lucretius*.[27] Each system contained several ether chains and up to 1000 water molecules (depending on concentration) represented using the TIP4P water model.[28] Bond lengths have been constrained using the SHAKE algorithm.[29] A reversible multiple time step integrator[30] was employed with a 0.75 fs time step for bends and torsions, a 1.5 fs time step for nonbonded and the real part of electrostatic interactions within 6.0 Å, and a 3 fs time step for nonbonded interactions for separations between 6.0 Å and 10.0 Å as well as for the reciprocal part of the Ewald summation.[31] Systems were equilibrated over 3 ns using an *NPT* ensemble, while production runs of 10 ns were performed in an *NVT* ensemble with *V* equal to the average value obtained from the *NPT* simulations.

These AES MD simulations provided us with structural and conformational properties of ethers in aqueous solution. In particular we focused on intermolecular pair correlation functions between monomer (defined as $-CH_2-O-CH_2-$ for EO, and $-O-CH(CH_3)-CH_2-$ for PO) centers-of-mass in solution. In Figure 1, we show EO−EO, PO−PO, and EO−PO intermolecular pair correlation functions, $g(r)$, as obtained from AES MD simulations for solutions with 0.52 weight fraction of ethers. The difference between EO−EO and PO−PO $g(r)$ at short separations reflects the larger size of the PO monomer. Particularly striking is the EO−PO $g(r)$ which clearly shows unfavorable interaction of EO monomers with PO monomers. Our initial expectation was that the EO−PO $g(r)$ would lie somewhere between EO−EO and PO−PO $g(r)$, which is clearly not the case. This indicates that despite the amphiphilic nature of PEO, PEO prefers to minimize its interaction with the hydrophobic PPO chain in the presence of water. This result, while unexpected, appears to be consistent with density functional theory modeling of Pluronic solutions where it was found that a relatively large positive χ (effective Flory−Huggins parameter) is necessary to describe PEO−PPO interaction in water.[17,18] The PEO−PPO χ used in these studies was an order of magnitude larger than the one estimated from SANS measurements for melts

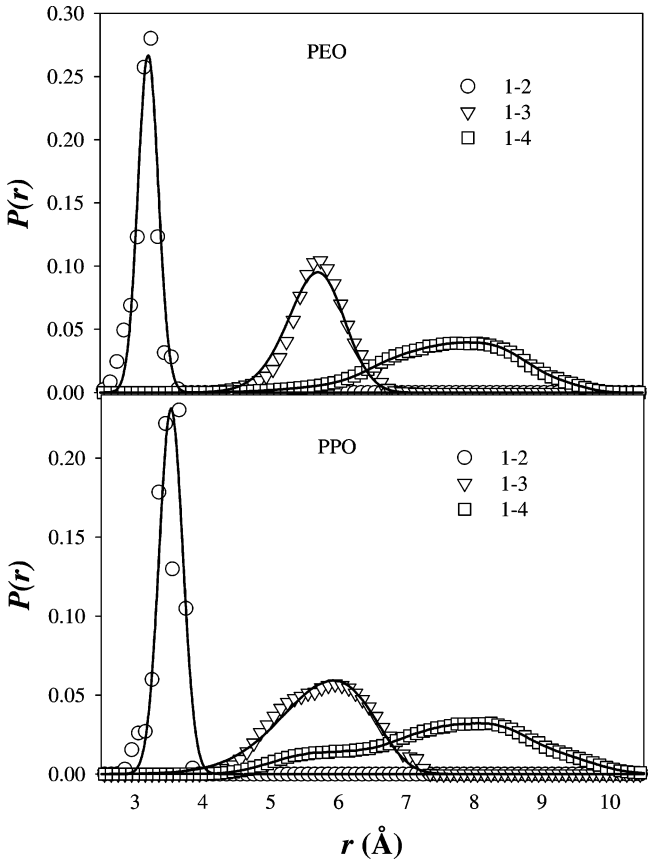


Figure 2. Intramolecular distributions of distances between neighboring monomers (1−2), monomers separated by one monomer (1−3), and monomers separated by two monomers (1−4) for PEO and PPO chains in water as obtained from AES (symbols) and CGIS (lines) simulations.

of PEO−PPO block copolymers[32] and about two times larger than the PPO−PPO χ in water.

In addition to intermolecular correlations we have analyzed intramolecular correlations for PEO and PPO chains in water. In Figure 2, we show the distribution of intramolecular distances between centers-of-mass of two adjacent (1−2 interaction), separated by one (1−3 interaction), and separated by two (1−4 interaction) EO or PO monomers. These distributions characterize local conformations of PEO and PPO chains in water and need to be accurately reproduced in the CGIS model. More detailed discussion of local structure and conformations in ether/water solutions as a function of temperature and concentration can be found in our previous work.[25]

C. Parametrization of the Coarse-Grained, Implicit Solvent Model. In the CGIS model water is treated as an implicit viscous solvent, while polymer chains are represented as bead-spring polymers where each bead represents a EO or PO monomer. Our initial guess was to represent EO ($-CH_2-O-CH_2-$) and PO ($-O-CH(CH_3)-CH_2-$) monomer as a single spherical force center with the center located at the center-of-mass of the corresponding monomer in the atomistic model. While this was the simplest and intuitively obvious geometric mapping between two models, we were prepared to use a more complicated mapping (e.g. representing a monomer with two or three force centers)

Table 1. Intramolecular Parameters for Ethers Bonds and Bends in the CGIS Model

$U_{bond}(R) = 0.5*K_{bond}*(R_0-R)^2$		
bond type	K_{bond} (kcal/mol/Å^2)	R_0 (Å)
EO−EO	30.0	3.20
PO−PO	20.0	3.52
EO−PO	25.0	3.36
$U_{bend}(R) = 0.5*K_{bend}*(\theta_0-\theta)^2$		
bend type	K_{bend} (kcal/mol)	θ_0 (deg)
EO−EO−EO	10.0	128.0
PO−PO−PO	5.0	110.5
PO−EO−EO	7.5	119.25
PO−PO−EO	7.5	119.25

between two models. However, as we illustrate below, the obtained agreement between CGIS and AES models did not require any additional and more complex structural mapping. The harmonic bonds and bends that connect ether monomer centers-of-mass have been introduced, and the strength of stretching and bending constants has been adjusted to reproduce the 1−2 and 1−3 intramolecular distributions obtained from AES MD simulations as illustrated in Figure 2. The force constants, equilibrium bond lengths, and bend angles in the CGIS model for PEO and PPO are given in Table 1. To reproduce 1−4 interactions one can either introduce a dihedral potential or a specific nonbonded potential that can be different from other nonbonded interactions. In this work, we chose the latter approach, and the resulting nonbonded potentials for intramolecular EO−EO and PO−PO 1−4 interactions are shown in Figure 3. We note that, in principle, the local intramolecular interactions (1−2, 1−3, and 1−4) can depend on nonbonded intra- and intermolecular interactions and vice versa, necessitating iterative fitting of intramolecular and nonbonded interactions. However, in this work, we found that the description of the nonlocal (1−5 and greater) nonbonded intramolecular and nonbonded intermolecular interactions had essentially no influence on the description of the local (1−2, 1−3, and 1−4) intramolecular interactions shown in Figure 2.

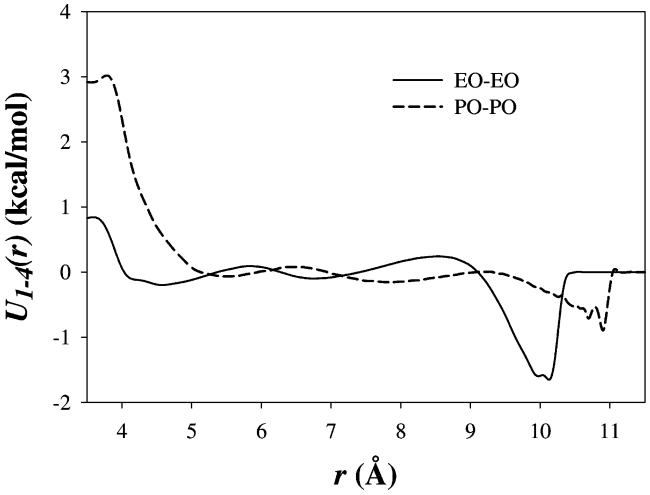


Figure 3. Converged 1−4 nonbonded interactions in the CGIS model.

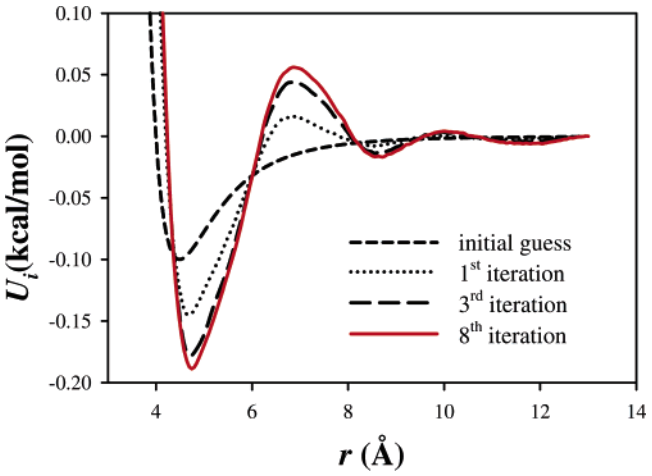


Figure 4. The intermolecular EO−EO nonbonded interaction in the CGIS model obtained after several iterations using the Inverted Boltzmann method.

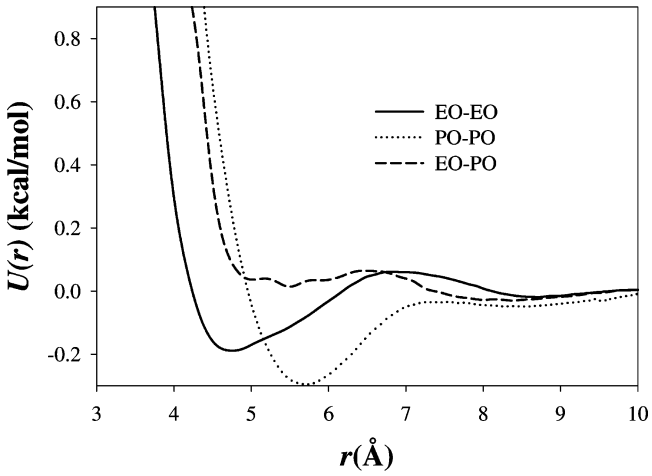


Figure 5. Converged nonbonded interactions in the CGIS model obtained using the Inverted Boltzmann method.

To parametrize nonlocal (1−5 and greater) nonbonded intramolecular and nonbonded intermolecular interactions we have employed the Inverted Boltzmann approach[33] in which a numerical potential is determined iteratively based on deviation of properties obtained from short CGIS simulations using the current iteration of the potential, $U_i(r)$, and the desired target properties obtained from AES MD simulations. The intermolecular $g(r)$ shown in Figure 1 for EO−EO, PO−PO, and EO−PO monomers obtained from AES MD simulations were chosen as the target functions, $g_{target}(r)$, that we wish to reproduce from simulations using the CGIS model. These $g_{target}(r)$ reflect all important correlations between ethers in aqueous solution, and therefore the coarse-grained potentials that reproduce these target functions would implicitly incorporate important water-induced interactions between the ethers. The initial guess for all interactions was a Lennard-Jones type potential as illustrated in Figure 4 for the EO−EO interaction. Short CGIS simulations using Langevin Dynamics (LD) were performed for the system exactly equivalent to the one used in AES MD simulations (the same number of ether chains and volume of the simulation cell), and the corresponding $g_i(r)$ were determined. These simulations must be sufficiently long to allow accurate determination of $g_i(r)$. In our CGIS model all nonbonded interactions were represented by numerical functions which are not restricted to any particular analytical form. Based on deviations between $g_i(r)$ and $g_{target}(r)$ we predict the new numerical potential $U_{i+1}(r)$ for the $i+1$ iteration of the CGIS model using the following expression

$$U_{i+1}(r) = U_i(r) + w(r)k_BT\ln\left(\frac{g_i(r)}{g_{target}(r)}\right) \qquad (1)$$

where $w(r) = \exp(-r)$ is a weighting function and k_B is the Boltzmann constant. Then, we repeat the CGIS simulations to obtain the new $g_{i+1}(r)$ pair correlations and continue this procedure until $g_i(r)$ reproduces $g_{target}(r)$ as accurately as possible. In Figure 4, we illustrate this process by showing $U_i(r)$ for EO−EO at several iterations, while in Figure 1 we show the agreement between the final (8th) iteration $g_i(r)$ obtained from CGIS simulations and $g_{target}(r)$ from AES MD simulations. In the fitting nonbonded interactions we first fitted the EO−EO and PO−PO interactions to match $g_{target}(r)$ from AES MD simulations of solutions containing only PEO or PPO chains, respectively. Using these potentials for interactions between beads of the same type the cross-term EO−PO nonbonded interactions were fitted to reproduce EO−PO $g_{target}(r)$ obtained from the PEO/PPO/water system.

In Figure 5 the final ether−ether nonbonded interactions obtained for the CGIS model are shown. We point out that all potentials have a nontrivial shape with oscillatory modulations due to structure of the hydrating water around the ethers that results in water-induced interactions which are manifested in $g_{target}(r)$ and therefore reflected in the CGIS potentials. We also note that the EO−EO monomer−monomer interaction has a reasonably deep minimum at separations around 4.8 Å indicating an apparent short-range attraction between EO monomers in water. However, this minimum is followed by a positive maximum and then by a very shallow second minimum at larger separations which results in an overall positive second virial coefficient $B = 71$ Å^3 for EO monomers consistent with the solubility of PEO in water. The PO−PO potential shows much stronger attraction resulting in a negative second virial coefficient $B = -56$ Å^3 consistent with the poor solubility of PPO in water. Finally, the EO−PO monomer−monomer potential is repulsive for almost all separations yielding $B = 193$ Å^3.

Finally, the monomer friction coefficient for CGIS LD simulations of Pluronic micelles was set to yield a viscosity of the effective solvent 100 times smaller than that of water thereby facilitating equilibration and sampling of the micelle structure. To establish an exact correspondence between time scales in the CGIS simulations and real time units a head-to-head comparison of motion and relaxation in the micelles should be conducted for CGIS and AES simulations which are currently underway. In this paper, we report all dynamical properties as a function of the number of integration steps. All systems were simulated using integration time step $\Delta t = (\epsilon/m\sigma^2)^{1/2} = 0.003$ where $\epsilon = 0.189$ kcal/mol is the lowest energy in the EO−EO nonbonded interaction shown in Figure 5, m is the mass of the EO monomer, and $\sigma = 4.213$ Å is the size of the EO monomer defined as the smallest

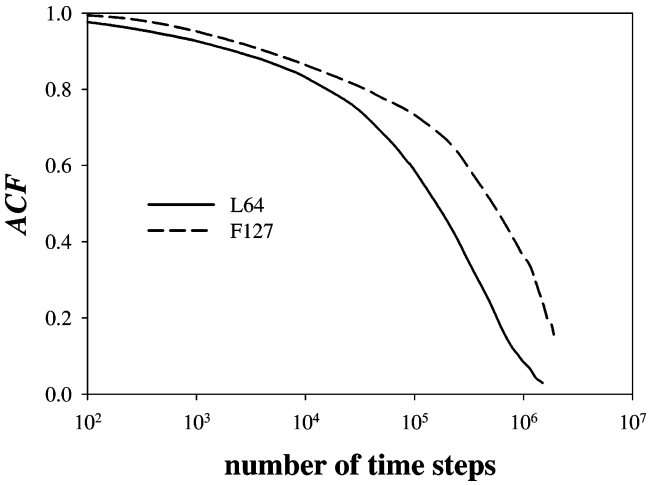


Figure 6. The average time autocorrelation function for relaxation of angles formed by two intermolecular vectors connecting the micelle center-of-mass and PEO−PPO junction point for each chain as obtained from CGIS LD simulations of L64 and F127 micelles.

separation between two EO−EO monomers at which their nonbonded interaction is zero (see Figure 5).

III. Multiscale Modeling of Pluronic Micelles

A. Coarse-Grained, Implicit Solvent Simulations of Pluronic Micelles. Following parametrization of the CGIS model the single micelle simulation for L64 ($EO_{13}-PO_{30}-EO_{13}$) and F127 ($EO_{99}-PO_{65}-EO_{99}$) Pluronics has been set up. For each Pluronic micelle we used a number of chains within the range of aggregation numbers estimated from SANS measurements.[34−36] Specifically, we used $N_{agg} = 40$ and 30 for L64 and F127, respectively. Initially, randomly generated configurations of Pluronic chains were simulated with an external field that imposed an attractive interaction between the center of the simulation cell and all PO monomers. This allowed us to quickly assemble a micelle with all PO monomers within a spherical micelle core and EO segments extending into the solution making the micelle corona. Following this forced "micellization" we turned off the external field and allowed the micelles to equilibrate their structure during the CGIS simulation. Our primary objective was to obtain equilibrium configurations of a micelle with a fixed N_{agg}, and, therefore, it was necessary to confirm that the complete decorrelation/relaxation of the micelle structure occurs on time scales accessible to the CGIS LD simulations. To characterize relaxation of the micellar structure we have defined vectors that connect the micelle core center-of-mass and PEO−PPO junction points for each chain. We then calculated the time autocorrelation function (ACF)

$$\mathrm{ACF}(t) = \frac{\langle \cos\theta(t)\cos\theta(0)\rangle - \langle\cos\theta\rangle^2}{\langle\cos^2\theta\rangle - \langle\cos\theta\rangle^2} \qquad (2)$$

where θ is the angle between two intermolecular vectors, and $\langle\ \rangle$ represents averaging over all possible intermolecular vector pairs in the micelle. When the ACF, shown in Figure 6, decays to zero, the micelle configuration has completely decorrelated relative to the time zero orientation of Pluronic chains in the micelle. Figure 6 illustrates that even for the relatively long chains of the F127 the micelle relaxation time is on the order of 10^6 integration time steps, while simulations using the CGIS model can be routinely performed over 10^8-10^9 integration steps.

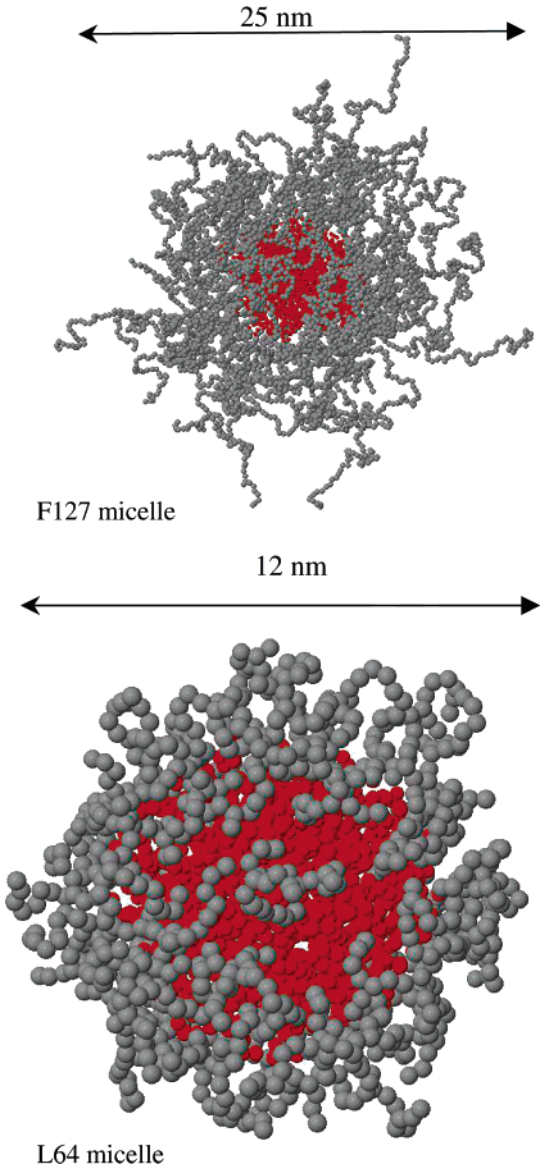


Figure 7. Representative snapshots of L64 and F127 micelles in the CGIS model.

The CGIS simulations carried out over multiple characteristic micellar relaxation times (defined above) provide us with equilibrium configurations of Pluronic micelles independent of how the initial micelle configuration has been created for a given N_{agg}. Representative configurations of L64 and F127 Pluronic micelles obtained from CGIS simulations are shown in Figure 7. We note that over the entire simulation for both micelles (10^8 time steps) none of the Pluronic chains was able to leave the micelles despite the fact that the structure of the micelles has relaxed multiple times. In Figure 8 we show volume fractions of EO and PO monomers (ϕ_{EO} and ϕ_{PO}) as a function of distance from the micelle core center-of-mass averaged over the entire CGIS simulation for both micelles. To calculate these profiles we used the AES simulations of pure water and ether/water

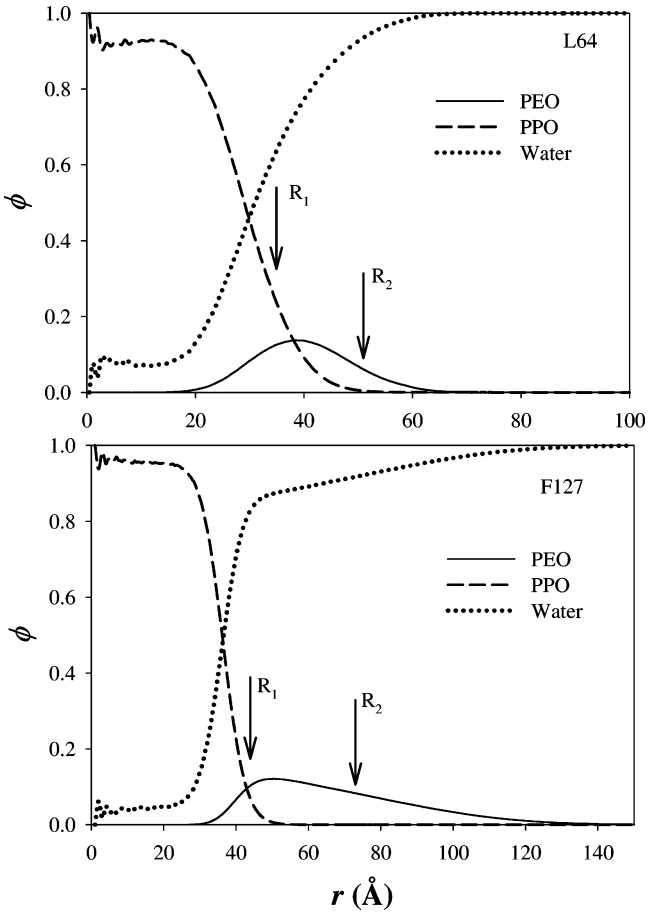


Figure 8. Volume fraction profiles of EO and PO monomers and water for L64 and F127 micelles obtained from CGIS LD simulations.

solutions to determine specific volumes occupied by a bulk water molecule (υ_W=29.8 Å^3) and by EO and PO monomers in aqueous solutions (υ_{EO}=62.3 Å^3 and υ_{PO}=105.7 Å^3). Then, using configurations obtained from CGIS simulations the local number density of EO and PO monomers (and hence their volume fractions) have been determined, while the local volume fraction of water, ϕ_w, has been calculated as $\phi_w = 1 - (\phi_{EO}+\phi_{PO})$ and is also shown in Figure 8. For both micelles we observe a noticeable fraction of water in the micellar core. Also shown in Figure 8 are values for the core (R_1) and corona (R_2) radii reported for L64 and F127 micelles of similar sizes by fitting the static structure factors from SANS measurements using a core-corona form factor model and hard-sphere approximation for the micelle interactions.[35,36] The best way to make a direct comparison between experiments and our CGIS simulations would be to calculate the static structure factor for micellar solutions at finite concentrations and use the same models and fitting protocol as in interpretation of SANS measurements. These efforts are currently underway. In addition to volume fraction profiles the CGIS simulations can provide detailed information on configuration of polymer chains, fluctuations of micellar shape, and intermicellar interactions/correlations.

B. Reverse Mapping onto Atomistic, Explicit Solvent Model. The representative configurations of Pluronic micelle obtained from CGIS simulations can be reverse mapped back to the AES model for further, more detailed investigation of

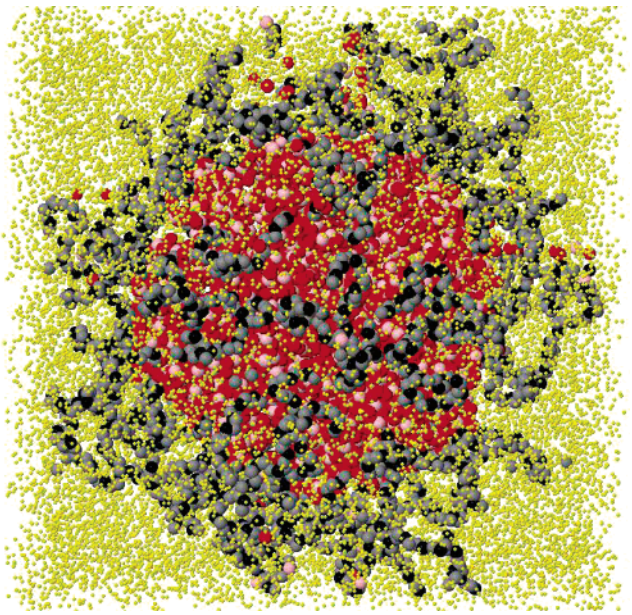


Figure 9. A representative snapshot of the L64 micelle in water using the AES representation reversed mapped from the CGIS configuration. Hydrogen atoms are not shown.

structural, conformational, and short-time dynamical properties. In the reverse mapping procedure we begin with placing the atomistic Pluronic chains along the contour of the coarse-grained chains. To do this, we first generate atomistic Pluronic chains in random configurations randomly placed in a simulation cell. Then, for the center-of-mass of each ether monomer in the atomistic model a corresponding monomer in the CGIS model with a defined position in space is assigned. An external field is then applied that forces the center-of-mass of the atomistic monomer to occupy the same position as the corresponding monomer in the CGIS model. A short atomistic, implicit solvent LD simulation in the presence of these external fields in which all intermolecular interactions are turned off (phantom chains) allows all atomistic chains to be placed along the contours of corresponding CGIS chains for a given micelle configuration. Since chains arrangement in the configuration from the CGIS simulation already takes into account the excluded volume interactions between ether monomers, the mapped atomistic configuration has only minor overlap between atoms. In the next step, the water molecules are inserted around and inside the micelle proportionally to the water density profile (shown in Figure 8). The unphysical overlap between water and ether atoms is then removed by a subsequent short (several picoseconds) AES MD simulation with upper bounds on nonbonded forces. Figure 9 shows a snapshot of the AES system of an L64 micelle reversed mapped from the configuration obtained from CGIS simulations. The AES system contains 30 000 water molecules.

Following the reverse mapping procedure, we are currently conducting extensive AES MD simulations of single micelles hydrated by water. Because the AES system is extremely large the observation times will be limited to a few nanoseconds which is sufficient for redistribution of water molecules inside the micelles and accurate sampling of chain local conformations, characteristics of hydrogen bonding, and

other atomistic scale properties in the core and corona of the Pluronic micelle. If the micellar structure obtained from AES MD simulations using the reverse-mapped configuration significantly changes/differs from predictions using the CGIS model, the latter will be further revised based on the new structural and conformational data obtained from AES MD simulations and the cycle will be repeated.

C. Further Extension of Multiscale Modeling of Pluronic Micelles. The CGIS simulations of preformed individual micelles or ensemble of micelles can provide equilibrium sampling of structural and conformational properties of those micelles; however, they can hardly access the time scales for the second process of micellization which establishes the equilibrium micelle size distribution in the system. While at some thermodynamic conditions it is possible to observe the individual events of Pluronic chains leaving/joining the micelle or micelle breaking up into two smaller ones, the CGIS simulations are not able to effectively access the milliseconds time scales which would allow the sufficient number of those events to adequately sample the equilibrium micelle size distribution. Therefore, the multiscale modeling protocol described above has to be further enhanced to effectively access larger length and longer time scales and to allow the prediction of the equilibrium distribution of micelle sizes as a function of Pluronic concentration and temperature. One of the approaches which we are currently investigating is the Parallel Tempering method,[37] modification of which we have recently developed and successfully applied for equilibration of model self-assembling polymer networks.[38] Other potential extensions of the multiscale modeling of Pluronic micelles include coupling with single-chain mean field approach,[39,40] density functional theory,[17,18] and self-consistent theory approaches.[15] However, for all these methods it is important to establish the reliable method to incorporate implicitly all important atomistic and molecular scale physics and phenomena.

Acknowledgment. The authors would like to acknowledge the support through NSF ITR, CHE0312226, and MRSEC (University of Colorado), DMR-9809555, grants.

Note Added after ASAP Publication. This article was released ASAP on March 28, 2006, with the incorrect Received Date. The correct version was posted on April 12, 2006.

References

(1) McPherson, A. *Methods Enzymol.* **1985**, *114*, 112.

(2) Cudney, B.; Patel, S. *Acta Crystallogr.* **1994**, *D50*, 414.

(3) Harris, J. M. In *Poly(Ethylene Glycol) Chemistry: Biotechnical and Biomedical Applications*; Harris, J. M., Ed.; Plenum Press: New York, 1992, p 1.

(4) Andrade, J. D.; Hlady, V.; Jeon, S. I. *Polym. Mater.: Sci. Eng.* **1993**, *69*, 60.

(5) Sakai, T.; Alexandridis P. *Nanotechnology* **2005**, *16*, S344.

(6) Alexandridis, P. *Curr. Opin. Colloid Interface Sci.* **1997**, *2*, 478.

(7) Adams, M. L.; Lavasanifar, A.; Kwon, G. S. *J. Pharm. Sci.* **2003**, *92*, 1343.

(8) Almgren, M.; Brown, W.; Hvidt, S. *Colloid Polym. Sci.* **1995**, *273*, 2.

(9) Pandit, N. K.; McIntyre, H. J. *Pharm. Dev. Technol.* **1997**, *2*, 181.

(10) See, A. S.; Gast, A. P.; Butun, V.; Armes, S. P. *Macromolecules* **1999**, *32*, 4302.

(11) Wanka, G.; Hoffman, H.; Ulbricht, W. *Macromolecules* **1994**, *27*, 4145.

(12) Goldmints, I.; Holzwarth, J. F.; Smith, K. A.; Hatton, T. A. *Langmuir* **1997**, *13*, 6130.

(13) Hositza, M. J.; Bohne, C.; Alexandridis, P.; Hatton, T. A.; Holzwarth, J. F. *Langmuir* **1999**, *15*, 322.

(14) Linse, P. *Macromolecules* **1994**, *27*, 6404.

(15) Hurter, P. M.; Scheutjens, J. M. H. M.; Hatton, T. A. *Macromolecules* **1993**, *26*, 5592.

(16) Svensson, M.; Linse, P. *Macromolecules* **1998**, *31*, 1427.

(17) Lam, Y. M.; Goldbeck-Wood, G. *Polymer* **2003**, *44*, 3593.

(18) Fraaije, J. G. E. M.; Zvelindovsky, A. V.; Sevink, G. J. A. *Mol. Sim.* **2004**, *30*, 225.

(19) Guo, S. L.; Hou, T. J.; Xu, X. J. *J. Phys. Chem. B* **2002**, *106*, 11397.

(20) Smith, G. D.; Borodin, O.; Bedrov, D. *J. Comput. Chem.* **2002**, *23*, 1480.

(21) Smith, G. D.; Borodin, O.; Bedrov, D. *J. Phys. Chem. A* **1998**, *102*, 10318.

(22) Bedrov, D.; Pekny, M.; Smith, G. D. *J. Phys. Chem. B* **1998**, *10*, 996.

(23) Smith, G. D.; Borodin, O. In *Molecular Simulation Methods for Predicting Polymer Properties*; Galiatsatos, V., Ed.; John Wiley & Sons: New York, 2005.

(24) Londono J. D.; Annis, B. K.; Habenschuss, A.; Smith, G. D.; Borodin, O.; Tso, C.; Hsieh, E. T.; Soper, A. K. *J. Chem. Phys.* **1999**, *110*, 8786. Annis, B. K.; Borodin, O.; Smith, G. D.; Benmore, C. G.; Soper, A. K.; Londono J. D. *J. Chem. Phys.* **2001**, *115*, 10998. Borodin, O.; Douglas, R.; Smith, G. D.; Trouw, F.; Petrucci, S. *J. Phys. Chem. B* **2003**, *107*, 6813.

(25) Smith, G. D.; Bedrov, D.; Borodin, O. *J. Am. Chem. Soc.* **2000**, *122*, 9548. Smith, G. D.; Bedrov, D.; Borodin, O. *Phys. Rev. Lett.* **2000**, *85*, 5583. Smith, G. D.; Bedrov, D. *Macromolecules* **2002**, *35*, 5712. Smith, G. D.; Bedrov D. *J. Phys. Chem. B* **2003**, *107*, 3095.

(26) Borodin, O.; Bedrov, D.; Smith, G. D. *Macromolecules* **2002**, *35*, 2410. Borodin, O.; Bedrov, D.; Smith, G. D. *J. Phys. Chem. B* **2002**, *106*, 5194. Borodin, O.; Trow, F.; Bedrov, D.; Smith, G. D. *J. Phys. Chem. B* **2002**, *106*, 5184.

(27) http://lucretius.mse.utah.edu.

(28) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. *J. Chem. Phys.* **1983**, *79*, 926.

(29) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.

(30) Martyna, G. J.; Tuckerman, M.; Berne, B. J.; Martyna, G. J. *J. Chem. Phys.* **1992**, *97*, 1990.

(31) Allen, M. P.; Tildesley, T. J. *Computer Simulations of Liquids*; Oxford University Press: Oxford, 1987.

(32) Hamley, I. W.; Castelletto, V.; Yang, Z.; Price, C.; Booth, C. *Macromolecules* **2001**, *34*, 4079.

(33) Reith, D.; Putz, M.; Muller-Plathe, F. *J. Comput. Chem.* **2003**, *24*, 1624.

(34) Wu, C.; Chu, B.; Schneider, D. K. *J. Phys. Chem.* **1995**, *99*, 5094.

(35) Yang, L.; Alexandridis P.; Steytler, D. C.; Kositza, M. J.; Holzwarth, J. F. *Langmuir* **2000**, *16*, 8555.

(36) Thurn, T.; Coulderc, S.; Sidhu, J.; Bloor, D. M.; Penfold, J.; Holzwarth, J. F.; Wyn-Jones, E. *Langmuir* **2002**, *18*, 9267.

(37) Yan, Q.; de Pablo, J. J. *J. Chem. Phys.* **1999**, *111*, 9509.

(38) Ayyagari, C.; Bedrov, D.; Smith, G. D. *J. Chem. Phys.* **2005**, *123*, 124912.

(39) Muller, M.; Smith, G. D. *J. Polym Sci., Part B: Polym. Phys.* **2005**, *43*, 934.

(40) Szleifer, I. *J. Chem. Phys.* **1990**, *92*, 6940.

CT050334K

JCTC Journal of Chemical Theory and Computation

Systematic Coarse-Graining of a Polymer Blend: Polyisoprene and Polystyrene

Qi Sun and Roland Faller*

Department of Chemical Engineering and Materials Science, University of California—Davis, Davis, California 95616

Received February 16, 2006

Abstract: The Iterative Boltzmann Inversion technique (also known as the Inverse Boltzmann Method) is generalized to polymer blends. We systematically optimize a mesoscale model against the structure of the blend. A polyisoprene−polystyrene blend is used as an example. Atomistic simulations of a blend of short chains in the miscible regime under melt conditions are taken as a starting point. We optimize the mesoscale model and study the onset of phase separation with increasing chain length. The mesoscale model phase separates at a chain length of 15 monomers where it was optimized, whereas the atomistic model shows only a preference of chains to aggregate to neighborhoods of like chains. We discuss the differences of the optimization between a blend and a homopolymer system in detail.

I. Introduction

Multiscale modeling of homopolymers has been successfully applied using a variety of techniques.[1−6] One particularly successful technique is known as the Inverse Boltzmann Method or Iterative Boltzmann Inversion (IBI).[7−11] Reliable models for a number of polymers have been developed using this technique, and the corresponding static and dynamic behaviors have been investigated. For example, recently, we could show that we can semiquantitatively estimate the entanglement length of polystyrene using this technique.[11]

The state of the art in the systematic modeling of polymer blends is much less satisfactory. The main obstacle is the lack of so-called "mixing rules". Unlike in the case of atomistic simulations, the equivalence of, say, Lorentz−Berthelot rules cannot be applied in structurally coarse-grained blends. It becomes immediately clear that the unlike interactions between different constituents have to be treated completely independently from the like interactions because the nature of these interaction stems from the interplay of the energetics and the local chain packing. Especially, the latter cannot be predicted for a mixture even if the behaviors of the pure components are known. Moreover, the resulting potentials may be concentration- and temperature-dependent because any systematically optimized model depends on the thermodynamic state against which it was optimized. We recently showed for a polystyrene model that the reliable temperature range of an optimized model may indeed be very limited.[12]

In this light, we elucidate the possibility to optimize a coarse-grained model for a blend of polyisoprene (PI) and polystyrene (PS) as a test case. Experimentally, this system is known to be miscible at short chain lengths,[13,14] and demixing is observed at longer chain lengths.[14] We are embarking on this specific system because there are systematically coarse-grained models of the pure components available.[8,10] This system has, for short chain lengths, recently attained some attention due to its interplay of short and large lengths and time-scale dynamics.[13,15,16] Because atomistic simulations are not able to equilibrate the terminal relaxation even of miscible short-chain systems, a mesoscale approach will be helpful for understanding such dynamic questions as well.

This paper is organized as follows: We first briefly describe the atomistic simulations on which the mesoscale modeling is based. Then, we explain in detail how the mesoscale modeling is performed with a focus on optimizing interactions between different polymers. The large-scale simulations are analyzed, and we finish with conclusions.

II. Simulation Investigations

A. Atomistic Simulations of a *cis*-PI and PS Melt Mixture. Polymer blends of *cis*-PI and atactic PS are investigated

* Corresponding author fax: 530-752-1031; e-mail: rfaller@ucdavis.edu.

10.1021/ct600065v CCC: $33.50
Published on Web 04/08/2006

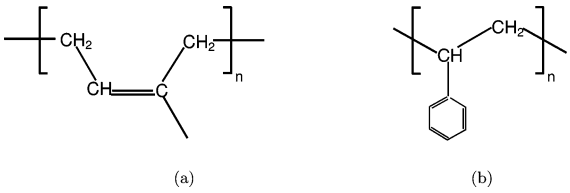


Figure 1. (a) PI monomer and (b) PS monomer.

in atomistic detail to obtain the input parameters for the mesoscale model. Figure 1 shows the structures of the atomistic monomers of *cis*-PI and PS. We choose to simulate chains of a length of 15 monomers in atomistic detail. These are clearly oligomers; however, it has been shown that for such lengths the chains are not completely dominated by end effects[17] such that a reliable structure optimization should be possible. Longer chains would definitely not equilibrate in reasonable computer time for an atomistic model. Moreover, experiments suggest that chains of this length should be miscible, but an effect of the demixing at longer chain lengths is expected.[13] Our atomistic simulations are performed at a variety of concentrations where the number of PS or PI chains varies between 24 and 72 for the two constituents. The subsequent optimization is based on a system with 24 PS and 36 PI chains, corresponding to a blend of 50 wt %. For the force field of PS, we use the Lennard-Jones parameters of Jorgensen and Severance.[18] The bending potential from all of the angles, the torsion potentials for all of the backbone and phenyl ring carbons, and the improper potentials which keep the phenyl groups planar are taken from refs 19 and 20. All 1−4 nonbonded interactions of PS atoms are excluded. The bond lengths and angles for the atomistic PI simulation and the potentials associated with the bonds, angles, and improper dihedrals are adapted from previous simulations.[21,22] Atoms connected by any bonding potentials do not interact by the Lennard-Jones potential. Further nonbonded interactions are excluded following the rules addressed in ref 23. All of the bond lengths are constrained using the LINCS algorithm.[24] The simulations are conducted using molecular dynamics (MD) at 450 K and a pressure of 101.3 kPa for 20 ns. Constant temperature and pressure are ensured using the weak coupling method[25] with time constants of 0.2 ps for temperature and 1.0 ps for pressure. A compressibility of 1.12×10^{-6} kPa^{-1} is applied to the three Cartesian directions independently. Simulations are performed using the GROMACS software[26] with a time step of 2 fs and a cutoff for nonbonded interactions at 1.2 nm. Configurations are saved every 2000 steps. A detailed analysis of the atomistic simulations in view of the local dynamics will appear elsewhere.[27]

B. Mapping from Atomistic Scale to Mesoscale. Mapping simulation models derives the input parameters of a mesoscale model from an atomistic simulation. The following tasks need to be solved during the mapping. First, we need to determine the location of superatom centers along the chains. A superatom is the unit containing a number of atomistic atoms which are grouped together and are represented by a single interaction center on the larger scale. The principle of selecting the superatom center is that the superbonds between superatoms can be represented with a single harmonic potential.[5,6] Moreover, we want the potentials for superangles and -torsions to be simple or, ideally, to vanish completely. For PS, we tried several possible positions on the backbone or the side-group carbons. We calculated the bond distributions along the polymer chains for all of them. Only the superatom locations which result in narrow Gaussian distributions for the bond lengths fulfill the above requirement. The height-to-width ratio of the bond distributions defines the harmonic bond strength. The distance corresponding to the maximum of the distribution defines the superbond length. The superatom centers of PS were finally chosen on the backbone carbons connecting to the phenyl rings.[10] The superatom centers of PI are placed on the centers of the single bonds connecting two neighboring atomistic monomers.[8] The angle distribution and angle potentials were determined from the atomistic simulation as well. Three consecutive superatom centers form a superangle. All angle distributions were normalized and averaged, excluding the three superatoms closest to either chain end. The resulting distributions are divided by the sine of the corresponding angles as Jacobian and normalized.[5,6] A running average of length 3 is additionally applied. The angle distributions are resolved to an accuracy of 1°. The angle forces are calculated as the logarithmic difference between two consecutive angle distributions. Again, a running average of 3 is applied. A torsion potential is not needed for the polymers under study.

In MD, generally, interactions between particles are chosen to be analytic functions with a few adjustable parameters. Our work derives a numerical potential for the nonbonded interaction which is dependent on the radial distribution functions (RDFs). These distribution functions are a good choice to describe the structure of polymers, keeping the identity of the polymer structure. We use IBI[9] to reproduce the structure by means of RDFs. Because this technique has been discussed in detail in a number of publications,[8−10] we focus only on the modifications due to the blend. The generalization to binary blends requires three potentials from the corresponding distribution functions: PI−PI, PS−PS, and PI−PS. We obviously first need two sets of superbond length and superangle distributions, together with bending potentials and angle potentials, which are directly transferred from our earlier work.[8,10]

The IBI method obtains a numerical force field; that is, the potential values are specified by a tabulation on a grid of equidistant distances or angles. For the nonbonded interaction, we use a grid resolution of 0.01 times the repulsive core of the interaction as measured at its zero passage (i.e., 0.01σ in the case of a Lennard-Jones potential). The whole procedure relies on serially executed iterations.

In dense systems, individual distributions depend on the full set of potentials through higher-order correlation. For technical reasons, one can keep the majority of potentials constant while iterating a particular one; the rest has to be readjusted afterward, one by one. For practical purposes, it is effective to start with those potentials that are least affected by changes in the others. This is not a trivial a priori knowledge but typically found during the initial steps of the

optimization. In our case, we found that the PI−PI and PS−PS pairs are relatively independent of each other compared to the degree to which they interdepend on the PI−PS potential. This is not too surprising because we expect a tendency to demix, and therefore, the PI−PI potential should not change the PS−PS structure and vice versa. Basically, the like pairs can be optimized in parallel. Especially, the PI−PI structure is quasi independent of PS−PS or, to a lesser extent, PI−PS potentials. The PS−PS structure is influenced more strongly by the PI−PS potential. Examples of RDFs during the optimization are shown in Figure 2. All optimization focuses on the local structure, that is, up to 1 nm for PI−PI and 1.8 nm for both PS−PS and PI−PS. It has been shown in our earlier work that the PI and PS structures have different sensitivities to the neighborhood between 1 and 1.8 nm. After each iteration, the radial distribution function $g(r)$ was checked carefully to make sure all of the target structures were present, as well as if there were any nonphysical structures involved. These may stem from, for example, cutoffs. The potential at the cutoffs should be set to 0 to ensure energy conservation. We adjust the potential to 0 by shifting all values, keeping the shape fixed. Our optimization obtains the correct structure by removing unwanted features and optimizing to the desirable structure as close as possible. First, we focus on removing undesired features. Small humps or spurious corners, as, for example, seen in 15th iteration of the PS−PS RDF (cf. 4), are likely associated with humps or spurious corners in the potential. Smoothing gets rid of the above problems. When the general features are reproduced and only smaller changes are needed, adjusting the initial slope of the potential is a good starting point. This means we do not optimize all length scales of the potential at one time but focus on smaller windows, starting from shorter length scales.[10] During the optimization of the PS−PS pair, we ran into a problem that the first peak in the $g(r)$ around $r \approx 1.2$ nm was clearly wider than the target peak. We found in this case that it was more efficient to change the slope of the potential manually instead of continually optimizing consistently according to the IBI

$$\Delta V(r) = -k_B T \ln\left[\frac{g(r)}{g_{\text{target}}(r)}\right] \quad (1)$$

where ΔV is the correction potential from a direct Boltzmann inversion of the difference between the RDF of the corresponding iteration $g(r)$ and the atomistic target g_{target}.

The speed of convergence is influenced by the order in which one optimizes the various potentials. We started the optimization with the PI−PI potential, followed by the optimization of PS−PS, and the interaction of PI−PS is dealt with during the late stages of the PS−PS iteration.

C. Adapting the Pressure. We started with a purely structural optimization of the PI−PS system. Our best potential, however, yields a positive pressure of $P^* = 1.92$ in dimensionless units defined equivalently to the standard Lennard-Jones units. $P^* = 1.92$ corresponds to $P = 1.192 \times 10^7$ Pa as calculated by $P = P^*\epsilon/\sigma^3$.[28] This does not reflect the ambient conditions of the parent atomistic system. This

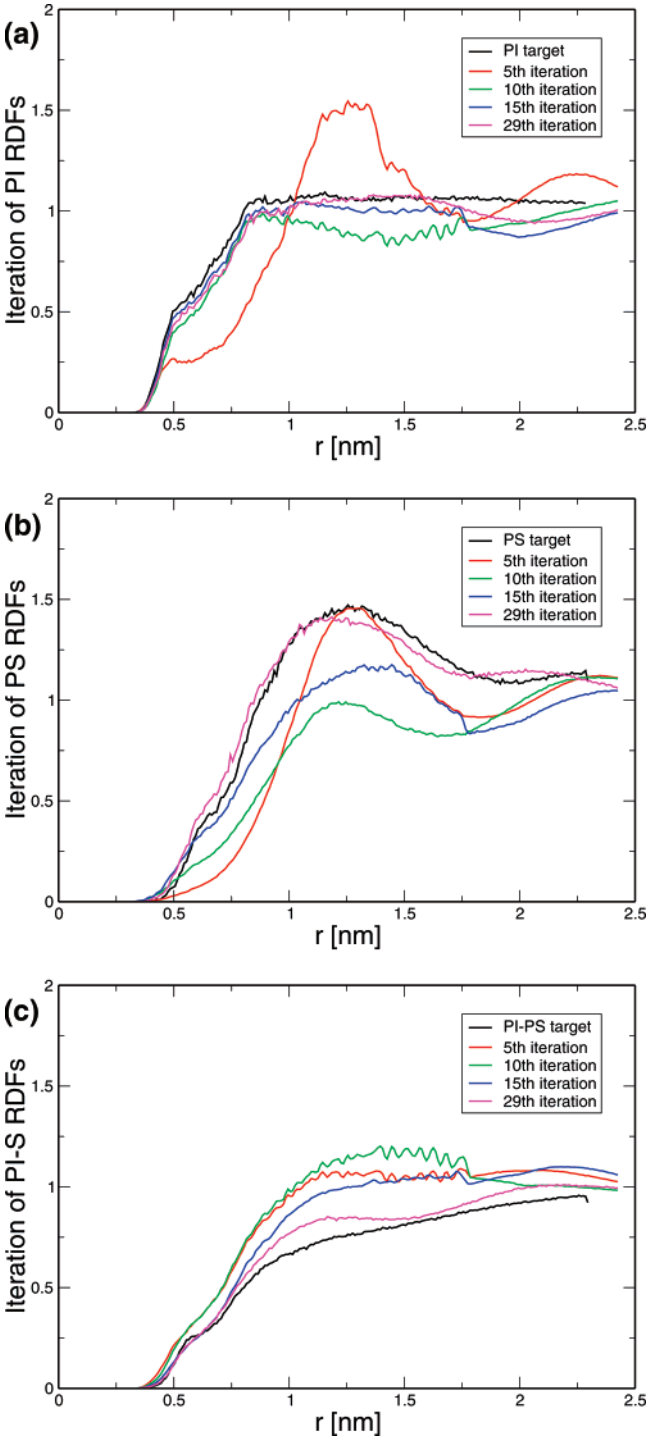


Figure 2. (a) Iteration of PI RDFs, (b) iteration of PS RDFs, and (c) iteration of PI−PS RDFs.

is a consequence of the simulation being run at a constant volume and thermodynamic properties not being used in the optimization. As a proof of concept that pressure correction is possible in a blend, we tried to postoptimize the mesoscopic system without lowering the quality of RDFs. The potential of a system without charges is always attractive at long ranges because of the dispersion interactions, and a corresponding pressure correction technique has been suggested.[7] Consequently, we choose an attractive linear tail function as a weak perturbation to the three potentials previously optimized without pressure correction:

$$\Delta V_{\mathrm{lim}}(r) = A_i\left(1 - \frac{r}{r_{\mathrm{cutoff}}}\right) \qquad i = 1, ..., 3 \qquad (2)$$

with A_i taking small values ranging from 0 to $-0.1\ k_BT$. The A_i values are different for the three interactions. The corrected potentials were taken as initial guesses in a reoptimization of the potentials against the structure using IBI. After that, the pressure was re-evaluated. Combined with optimum results of the RDFs, an improvement to $P^* \approx 1.3$ is obtained. This condition was reached after 10 iteration cycles. The pressure correction focuses in the potential region beyond $r > 1$ nm, which is not very important for the structural fit. It shifts the whole potentials downward, thus providing for the previously missing long-range attraction. However, we did not find significant differences in the overall structure of the system, and the fundamental results below are the same with and without pressure correction.

The optimization took about 35 runs to obtain the resulting RDFs. Each iteration runs for 3000 τ, after which the system with a new set of potentials could attain an equilibrium. We confirmed that, between 1000 and 3000 τ, the RDFs did not change.

D. Mesoscale Simulation. The PI−PS blends are investigated in the *NVT* ensemble for the mesoscale simulations. The reduced temperature ($T^* = k_BT/\epsilon$) is set to 1, where ϵ is the interaction energy between pairs of atoms. σ^* is set to 1 nm. The density is 6.4175 monomers/σ^3 according to the atomistic parent simulation. The temperature is maintained by a Langevin thermostat with a friction constant $\Gamma = 1.0\ \tau^{-1}$, where $\tau = \sqrt{\sigma^2 m/\epsilon}$.[28] We use an orthorhombic box under periodic boundary conditions. The initial condition is directly transferred from an atomistic simulation. The equation of motion is integrated with a time step $\Delta t = 0.005\ \tau$, which is the smaller of the time steps used in the simulations of pure PI and PS.[8,10,11] Trajectories are stored every 2000 steps. The superbond length of PI is 0.2345 nm with a strength of 0.0015 k_BT, while that of PS is 0.2545 nm with a strength of 0.03 k_BT. The cutoff for the nonbonded interactions is 1.910σ, which is the bigger of the PI and PS cutoffs of the pure systems. The numerically optimized potential from the system of the 36 PI and 24 PS chains of length 15 is applied to all systems. The details of the investigated systems are listed in Table 1.

III. Results and Discussion

A. Structure. To test the transferability of applying bond or angle structures to polymer blends, the bond length and bond angle distributions and potentials of PS from both 48 PS chains systems and the blend of 36 chains of PI and 24 chains of PS were compared and plotted in Figure 3. The bond distributions clearly show the same pattern in the pure PS and in a polymer blend with 36 chains of PI; however, the PS−PS−PS bond angle force shows some discrepancies. These stem from the effect of local packing because a number of the PS−PS intermolecular interactions are replaced by PS−PI contacts. Here, we kept the intramolecular potentials of the pure systems for transferability purposes.

During all of our simulations for optimization and analysis, we ensured that the chains diffuse at least $2\langle R_g^2\rangle$ such that the configurations are fully relaxed. We then checked each system by monitoring the radius of gyration (R_g) and $\langle R_g^2\rangle$ to make sure that the systems attained equilibration. Our results for the persistence lengths l_p given by $l_p \approx 0.65$ nm are independent of the chain length for $10 \leq N \leq 80$.

The RDFs of the atomistic and optimized mesoscale simulations are plotted in Figure 4. These are both shown for 15 monomers. This is the best potential after 35 iterations. The optimization result is judged by the sum of the squared difference between the target and the optimized $g(r)$. Compared to a pure polymer melt, the iteration of blends is technically more demanding in that mapping takes into account the effects of three sets of potentials. Therefore, the quality of the optimized RDFs is somewhat weaker than that for pure systems. Because the PI−PI or PS−PS pair interacts with the inter-PI−PS pair, iterating each pair induces deviations in other pairs. Examples of radial distribution functions during the iteration are plotted in Figure 2. Our work provides the first application in the system of a binary blend with reasonable consistency.

B. Phase Separation. Using our mesoscale model, we can embark now on a study of the phase behavior of the blend focusing on the following questions: What morphologies do we obtain, and how do they depend on concentration? How does an initially randomly distributed configuration evolve into a phase-separated system? By increasing chain length, where is the crossover between miscibility and phase separation? The visualization of PI−PS blends with a fixed mass concentration ratio of 1:1 shows the evolution of phase separation in Figure 5. In the initial configuration of the system with 36 PI and 24 PS chains, both of length 60, the components are well-mixed, as shown in Figure 5a. Increasing times of 300, 600, 1000, 2000, and 20 000 τ show the evolution of phase separation. Starting from the beginning, the PS chains (dark red spheres) show a tendency toward aggregation. At 600 τ, the aggregated PS chains take on a cylindrical shape with a few PS chains in the PI majority phase. The PS region is only separated by several PI chains in the middle (Figure 5d). The final snapshot at 20 000 τ shows the equilibrium. We now have a clearly lamellar pattern as expected for an equiconcentrated mixture.

Besides the lamellar morphology shown for the 36 PI 60-mer−24 PS 60-mer system, we investigate the phase separation dependent on the PI chain length with a fixed PI/PS weight ratio (50:50). Figure 6 shows lamellar shapes for the 72 PI 10-mer−48 PS 10-mer and 36 PI 30-mer−24 PS 30-mer systems. Cylindrical shapes are observed in Figure 7 for 36 PI chains of length 45 with the 24 PS 30-mer and for the 36 PI 60-mer−4 PS 30-mer system. In the systems of 72 PI 10-mer−48 PS 10-mer, 36 PI 30-mer−24 PS 30-mer, and the above 36 PI 60-mer−48 PS 60-mer, the weight ratio of PI to PS monomer weight concentration is kept fixed at 50:50. We conclude that at balanced concentrations the morphology of PI−PS blends is lamellar. As the concentration of PI to PS monomers changes to 60:40, as in the case of 36 PI 45-mer−24 PS 30-mer, or even higher to 66:34, as in the case of 36 PI 60-mer−24 PS 30-mer, the systems prefer a cylindrical morphology, as shown in Figure

Table 1. Characterizations of PS and PI Chains at Various Concentrations

		PI				PS			
N_c	box size	R_g	R_g^2	l_p (bond)	l_p (nm)	R_g	R_g^2	l_p (bond)	l_p (nm)
72PI7/48PS7	5.0774	0.3448	0.1206	1.5921	0.3733	0.3906	0.1539	1.9341	0.4922
72PI10/48PS10	5.7184	0.4795	0.2330	2.2002	0.5159	0.5272	0.2804	2.3639	0.6016
72PI14/48PS15	6.4574	0.6229	0.3940	2.5806	0.6052	0.7153	0.5188	2.6351	0.6706
72PI15/48PS15	6.5459	0.6234	0.3936	2.6317	0.6171	0.7139	0.5152	2.8987	0.7377
36PI30/24PS30	6.5459	1.0789	1.1930	2.7503	0.6449	1.1519	1.3598	2.6861	0.6836
36PI30/24PS45	6.9560	1.0801	1.1964	2.6889	0.6305	1.3821	1.9759	2.4283	0.6180
36PI30/24PS60	7.3228	1.0808	1.1944	2.6590	0.6235	1.7505	3.1953	2.6070	0.6635
36PI30/24PS80	7.7610	1.0777	1.1884	2.7108	0.6357	2.0348	4.2990	2.5715	0.6544
36PI45/24PS30	7.1441	1.4239	2.0921	3.0102	0.7059	1.1473	1.3491	2.6334	0.6702
36PI60/24PS30	7.6561	1.6746	2.9071	2.9246	0.6858	1.1192	1.2888	2.6019	0.6622
36PI80/24PS30	8.2473	2.0446	4.3507	2.8551	0.6695	1.0716	1.1775	2.7668	0.7041
36PI60/24PS60	8.2473	1.6873	2.8422	2.7252	0.6391	1.7793	3.2170	2.4545	0.6247
36PI80/24PS80	9.0773	1.9113	3.9538	2.9011	0.6803	1.9321	4.0789	2.6153	0.6656

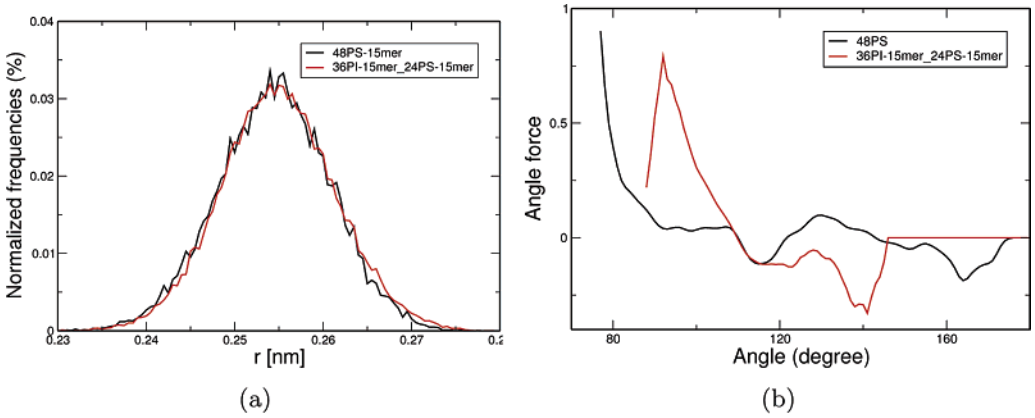


Figure 3. Comparison of the intramolecular potentials in the homopolymer melt and the 50 wt % blend. (a) PS bond distribution and (b) PS angle force.

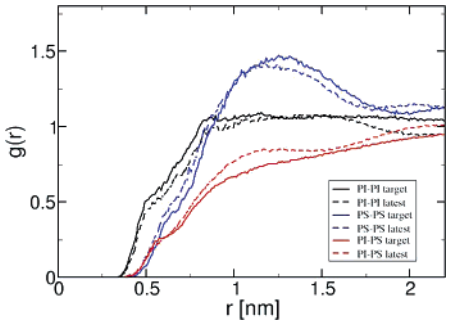


Figure 4. RDFs of the PI−PS mixture melt.

7. The radius of the PS cylinder decreases as the PI chain length increases from 45 to 60.

Figure 8 shows a well-mixed configuration of the 72 PI 7-mer−48 PS 7-mer system at 10 000 τ. Compared to chains of lengths 10, 30 (in Figure 6), and 60 (in Figure 5), the phase separation starts above 7 monomers. Note that, both in atomistic simulations and in experiments, the system is still mixed at 15 monomers, whereas we observe a phase separation tendency. However, the atomistic simulations are only 20 ns long, and the radial distribution functions for the 15-mers are by construction consistent between the atomistic and mesoscale systems. This study shows again that one structural characteristic alone leads typically to good potentials but some discrepancies between the atomistic and mesoscale models remain and the finite-size of the system may play a role.

C. RDFs of Superatoms. Intermolecular pair distributions of PI−PI and PS−PS monomers are calculated and shown in Figure 9. Figure 9a illustrates that, at constant conditions of 36 PI chains of chain length 30, the inter-PI−PI $g(r)$ increases with the PS chain length, which means the PI monomers increasingly prefer a PI neighborhood. Phase separation can explain this observation, as suggested by previous results. Figure 9b shows that the correlation hole dominates as the number of inter-PI−PI pairs decreases with the increase of the PI chain length at a constant PS chain length. A correlation hole means that with increasing chain length the chain itself provides a non-negligible number of neighboring monomers and, by this, suppresses the interchain RDF.[30,31] The interchain radial distribution functions of like monomers here probe for long chains, essentially a homopolymer system, because of the phase behavior. Figure 9d illustrates that at the same condition of 36 PI chains of the 30-mer, as in Figure 9a, the PS−PS pairs decrease with the increase of the PS chain length, opposite of the increasing trend of the PI−PI pairs. Figure 9c demonstrates that the number of PS−PS pairs increases generally with the increase of the PI chain length in the mixture with 24 PS 30-mers,

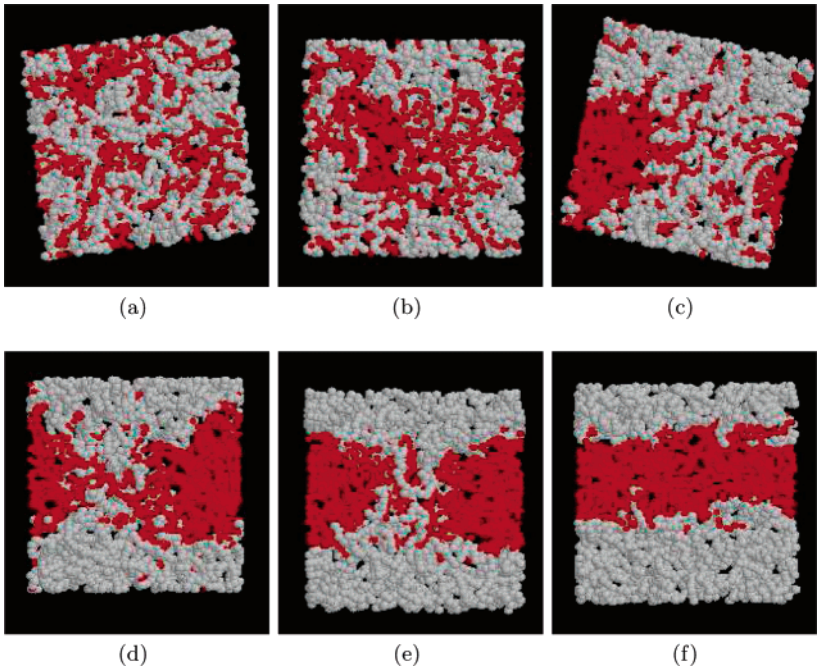


Figure 5. Snapshots of 36 PI 60-mer−24 PS 60-mer systems at various times: (a) starting configuration (time 0), (b) $t^* = 300\ \tau$, (c) $t^* = 600\ \tau$, (d) $t^* = 1000\ \tau$, (e) $t^* = 2000\ \tau$, and (f) $t^* = 20\ 000\ \tau$. The light spheres represent the polyisoprene superatoms, and dark red ones are polystyrene. The program RASMOL[29] was used for the visualization.

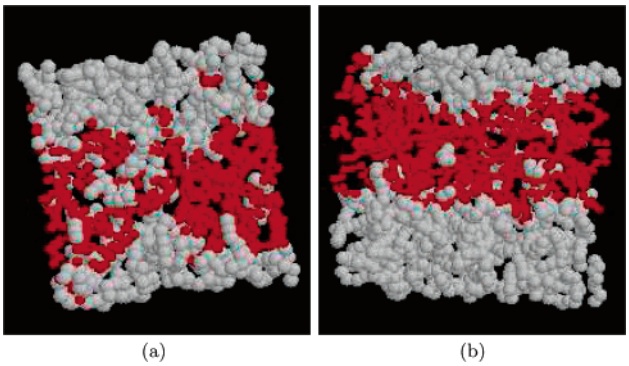


Figure 6. Snapshots of polymer mixture systems at various chain lengths: (a) 72 PI 10-mer−48 PS 10-mer at a unit time of 10 000 τ and (b) 36 PI 30-mer−24 PS 30-mer at a unit time of 10 000 τ.

except for the 80-mer case, which we again attribute to a correlation hole effect. Overall, the trends show that the number of pairs of the fixed chain length and the number of chain polymers increase with the chain length of the opposite component, while those of the varying polymers themselves decrease.

To test the dependence of phase separation on the chain length, we increase the chain length of both polymers from 10 to 30, 60, and 80 monomers. The concentration of PI and PS is kept fixed at 50:50 by weight. Pair distributions of inter-PI−PI, inter-PS−PS, and PI−PS are calculated to obtain a general picture of distribution. For comparison, the atomistic parent simulations at length 15 are shown as well. All intrapairs are excluded. Figure 10 shows that the interchain radial distribution function of PI−PI does not change as much as that of PS−PS or PI−PS. At the region of distance $r < 1.0$ nm, the $g(r)$ values of PI−PI pairs decrease as the chain length increases from 10 to 30, 60,

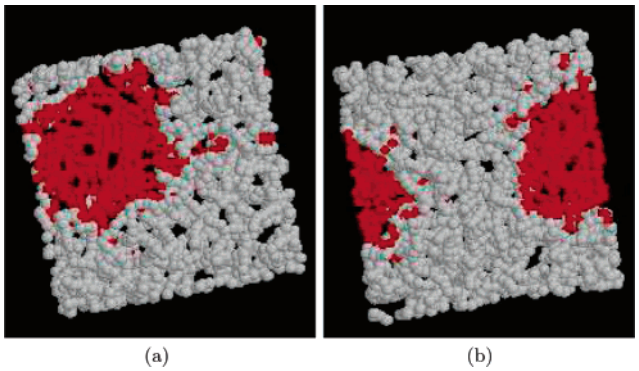


Figure 7. Snapshots of polymer systems (a) 36 PI 45−24 PS 30 at a unit time of 10 000 τ and (b) 36 PI 60−24 PS 30 at a unit time of 10 000 τ.

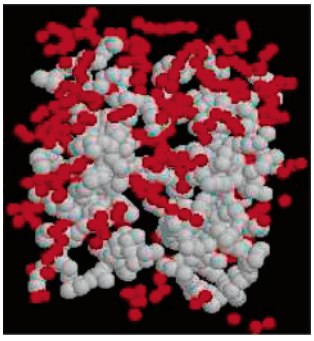


Figure 8. 72 PI 7-mer−48 PS 7-mer at a unit time of 10 000 τ.

and 80 monomers. The reason is the increasing radius of gyration. The PI−PI interactions show that PI monomers prefer their own neighborhood from $r > 1.0$ nm up to 3.0 nm. The PS−PS RDF (Figure 10b) shows a consistent trend of moving closer together as the chain length of PS increases.

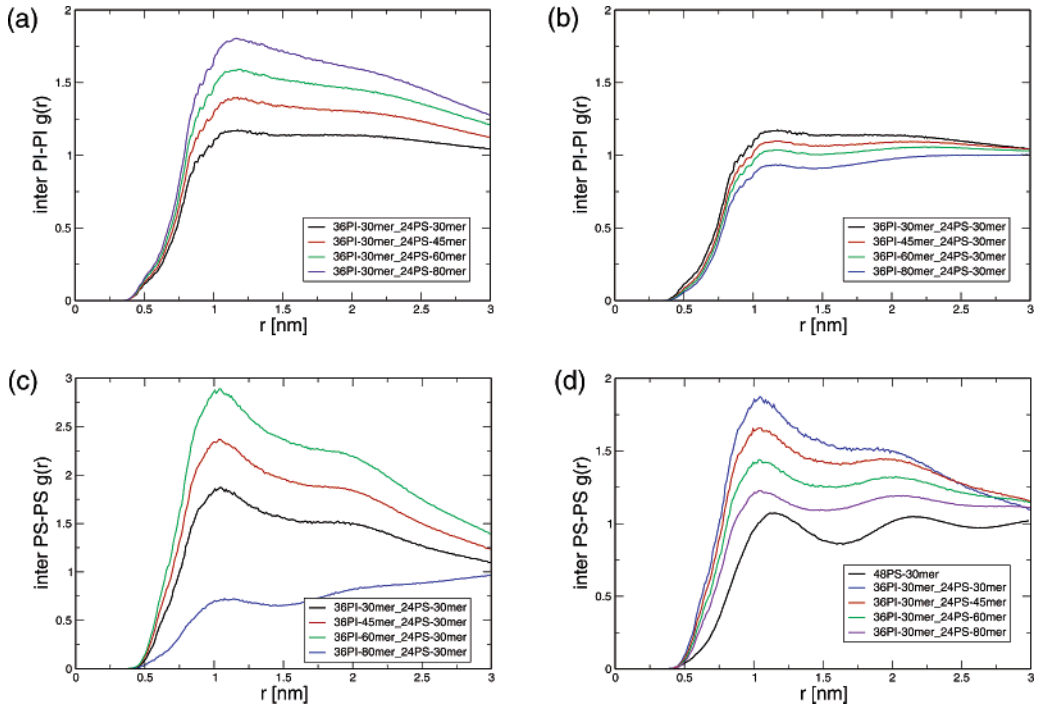


Figure 9. (a) Inter-PI−PI $g(r)$ at fixed 36 PI 30-mer, (b) inter-PI−PI $g(r)$ at fixed 24 PS 30-mer, (c) inter-PS−PS $g(r)$ at fixed 24 PS 30-mer, and (d) inter-PS−PS $g(r)$ at fixed 36 PI 30-mer.

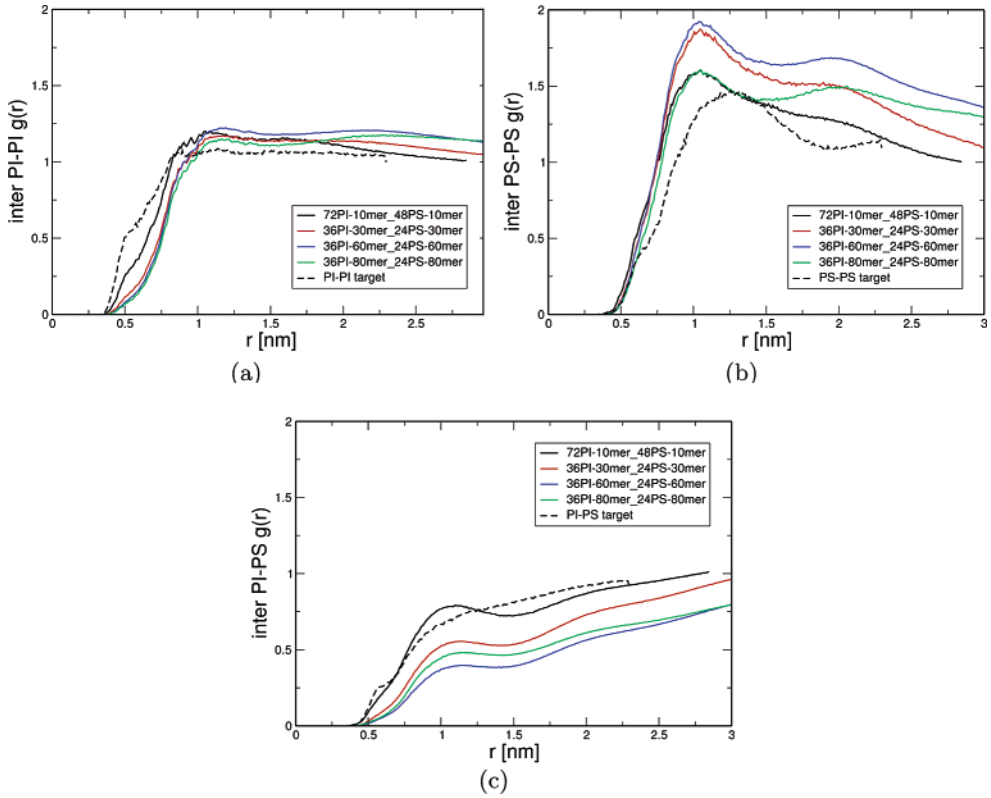


Figure 10. (a) Inter-PI−PI $g(r)$ at fixed PI/PS ratio of 1:1, (b) inter-PS−PS $g(r)$, and (c) inter-PI−PS $g(r)$.

Consider that the phase morphology does not change except for a chain length of 80; increasing the chain length leads to fewer contacts of PI−PS and correspondingly more PS−PS interactions. Moreover, we expect the phase separation to become stronger, that is, the PS region to have an even higher PS concentration at long chain lengths. PI−PS (in Figure 10c) pairs have a peak and a valley at short distances $r <$ 1.5 nm and increase gradually to a value of 1.0 near 3.0 nm. However, the number of PI−PS pairs decreases with the increase of the chain length, indicating that the PI and PS chains are more separated from each other. At the chain length of 80 monomers, all $g(r)$ functions show a different trend; PI−PS pairs increase instead of decrease. This can be explained by a cylindrical instead of a lamellar morphol-

ogy. Take the three RDFs as a whole; the PS−PS structures have two peaks throughout the entire chain length. The peak positions are at 1.5∼2.0 nm, which are at greater distances than those of PI−PI. It indicates that PS−PS has a stronger tendency to its own neighborhood. The PI−PS RDFs show consistently unfavorable interactions.

IV. Conclusion

We have shown here for the first time that the Iterative Boltzmann Inversion can be applied to a binary blend of polymers. The resulting potentials lead to results in qualitative agreement with expected experimental behavior. We are able to equilibrate systems which are phase-separating and observe the dynamics of the phase separation for the first time by a systematically coarse-grained blend.

It is worth stressing that we expect that the resulting potentials are dependent on concentration and temperature, and further studies to elucidate the phase diagram are under way. They are, however, stable under a change of chain length.

It is furthermore worth mentioning that there is no clear and obvious order in which of the optimizations should be conducted. The various potentials are strongly interdependent, and we observed that changing the order of optimization can have a serious influence on the efficiency of the convergence.

The phase separation sets in at around seven monomers, which is shorter than in experiments and also shorter than in the parent atomistic simulations. This again shows that we cannot expect all observables to be in agreement when we did not specifically optimize against them. It additionally shows that we have to develop such polymer blend models at short chain lengths in order to increase the number of unlike contacts for optimization. For a close to equimolar concentration, we find predominantly lamellar morphologies as expected. Other morphologies can be reproduced as well.

Acknowledgment. Financial support from the U.S. Department of Energy, Office of Advanced Scientific Computing, under grant DE-FG02-03ER25568 is gratefully acknowledged.

References

(1) Tschöp, W.; Kremer, K.; Batoulis, J.; Bürger, T.; Hahn, O. Simulation of Polymer Melts. I. Coarsegraining Procedure for Polycarbonates. *Acta Polym.* **1998**, *49*, 61−74.

(2) Baschnagel, J.; Binder, K.; Doruker, P.; Gusev, A. A.; Hahn, O.; Kremer, K.; Mattice, W. L.; Müller-Plathe, F.; Murat, M.; Paul, W.; Santos, S.; Suter, U. W.; Tries, V. Bridging the Gap between Atomistic and Coarse-Grained Models of Polymers: Status and Perspectives. *Adv. Polym. Sci.* **2000**, *152*, 141−156.

(3) Müller-Plathe, F. Coarse-Graining in Polymer Simulation: From the Atomistic to the Mesoscopic Scale and Back. *ChemPhysChem* **2002**, *3*, 754−769.

(4) Müller-Plathe, F. Scale-Hopping in Computer Simulations of Polymers. *Soft Mater.* **2003**, *1*, 1−31.

(5) Faller, R. Automatic Coarse Graining of Polymers. *Polymer* **2004**, *45*, 3869−3876.

(6) Faller, R. Coarse-Grain Modeling of Polymers. *Rev. Comput. Chem.* **2006**, *23*, in press.

(7) Reith, D.; Meyer, H.; Müller-Plathe, F. CG-OPT: A Software Package for Automatic Force Field Design. *Comput. Phys. Commun.* **2002**, *148*, 299−313.

(8) Faller, R.; Reith, D. Properties of Polyisoprene Model Building in the Melt and in Solution. *Macromolecules* **2003**, *36*, 5406−5414.

(9) Reith, D.; Pütz, M.; Müller-Plathe, F. Deriving Effective Mesoscale Potentials from Atomistic Simulations. *J. Comput. Chem.* **2003**, *24*, 1624−1636.

(10) Sun, Q.; Faller, R. Systematic Coarse-Graining of Atomistic Models for Simulation of Polymeric Systems. *Comput. Chem. Eng.* **2005**, *29*, 2380−2385.

(11) Sun, Q.; Faller, R. Crossover from Unentangled to Entangled Dynamics in a Systematically Coarse-Grained Polystyrene Melt. *Macromolecules* **2006**, *39*, 812−820.

(12) Ghosh, J.; Wong, B. Y.; Sun, Q.; Pon, F. R.; Faller, R. Simulations of Glasses: Multiscale Modeling and Density of States Monte Carlo Simulations. *Mol. Simul.* **2006**, in press.

(13) He, Y.; Lutz, T. R.; Ediger, M. D.; Pitsikalis, M.; Hadjichristidis, N.; von Meerwall, E. D. Miscible Polyisoprene/Polystyrene Blends: Distinct Segmental Dynamics but Homogeneous Terminal Dynamics. *Macromolecules* **2005**, *38*, 6216−6226.

(14) Koningsveld, R.; MacKnight, W. J. Liquid−Liquid-Phase Separation in Multicomponent Polymer Systems XXVII. Determination of the Pair Interaction Function for Polymer Blends. *Polym. Int.* **1997**, *44*, 356−264.

(15) He, Y.; Lutz, T. R.; Ediger, M. D. Segmental and Terminal Dynamics in Miscible Polymer Blends. *J. Chem. Phys.* **2003**, *119*, 9956−9965.

(16) Lutz, T. R.; He, Y.; Ediger, M. D.; Pitsikalis, M.; Hadjichristidis, N. Dilute Polymer Blends: Are the Segmental Dynamics of Isolated Polyisoprene Chains Slaved to the Dynamics of the Host Polymer? *Macromolecules* **2004**, *37*, 6440−6448.

(17) Faller, R. Correlation of Static and Dynamic Inhomogeneities in Polymer Mixtures: A Computer Simulation of Polyisoprene and Polystyrene. *Macromolecules* **2004**, *37*, 1095−1101.

(18) Jorgensen, W. L.; Severance, D. L. Aromatic−Aromatic Interactions: Free Energy Profiles for the Benzene Dimer in Water, Chloroform. *J. Am. Chem. Soc.* **1990**, *112*, 4768−4774.

(19) Müller-Plathe, F. Local Structure and Dynamics in Solvent-Swollen Polymers. *Macromolecules* **1996**, *29*, 4782−4791.

(20) Sun, Q.; Faller, R. Molecular Dynamics of a Polymer in Mixed Solvent: Atactic-Polystyrene in a Mixture of Cyclohexane and N,N-dimethylformamide. *J. Phys. Chem. B* **2005**, *109*, 15714−15723.

(21) Faller, R.; Schmitz, H.; Biermann, O.; Müller-Plathe, F. Automatic Parametrization of Force Field for Liquids by Simplex Optimization. *J. Comput. Chem.* **1999**, *20*, 1009−1017.

(22) Schmitz, H.; Faller, R.; Müller-Plathe, F. Molecular Mobility in Cyclic Hydrocarbons—A Simulation Study. *J. Phys. Chem. B* **1999**, *103*, 9731−9737.

(23) Faller, R.; Müller-Plathe, F.; Doxastakis, M.; Theodorou, D. Local Structure and Dynamics of *trans*-Polyisoprene Oligomers. *Macromolecules* **2001**, *34*, 1436−1448.

(24) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18*, 1463−1472.

(25) Berendsen, H. J. C.; Postma, J. P. M.; Gunsteren, V.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.

(26) Lindahl, E.; Hess, B.; van der Spoel, D. Gromacs 3.0: A Package for Molecular Simulation and Trajectory Analysis. *J. Mol. Model.* **2001**, *7*, 306−317.

(27) Pon, F. R.; Sun, Q.; Faller, R. Detailed Molecular Modelling Study on the Local Dynamics in Polyisoprene−Polystyrene Blends. **2006**, manuscript in preparation.

(28) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press Inc.: New York, 1987.

(29) Sayle, R. A.; Milnerwhite, E. J. RASMOL: Biomolecular Graphics for All. *Comput. Corner* **1995**, *20*, 374−376.

(30) Kremer, K.; Grest, G. S. Dynamics of Entangled Linear Polymer Melts: A Molecular-Dynamics Simulation. *J. Chem. Phys.* **1990**, *92*, 5057−5086.

(31) Faller, R.; Pütz, M.; Müller-Plathe, F. Orientation Correlations in Simplified Models of Polymer Melts. *Int. J. Mod. Phys. C* **1999**, *10*, 355−360.

CT600065V

JCTC Journal of Chemical Theory and Computation

Coarse Graining of Short Polythylene Chains for Studying Polymer Crystallization

Thomas Vettorel[†,‡] and Hendrik Meyer*[,†]

Institut Charles Sadron, 6 rue Boussingault, 67083 Strasbourg, France, and Institut für Physik, WA 331, Johannes-Gutenberg Universität, Staudinger Weg 7, D-55099 Mainz, Germany

Received December 21, 2005

Abstract: We derive coarse-grained models of polyethylene in the melt state with the aim to study polymer crystallization. This requires a low level of coarse-graining: We use a mapping of two CH_2 groups onto one bead. The coarse-grained beads are connected with harmonic springs, an optimized angular potential, and an optional torsional potential. Coarse-grained potentials are derived from detailed all-atom simulations, and an optimized form of the force field is then derived which achieves a good accuracy in reproducing the static properties of the chains. We address the question over which temperature range such models can be used, and in particular if the model is capable of reproducing the phase transition to an ordered state; it is found that the qualitative behavior of short polyethylene chains is well described, and the experimental melting temperature of $C_{44}H_{90}$ is approached when using the most accurate optimized model.

1. Introduction

Despite the progress of computer power, the simulation of complex materials with all atomic and electronic details remains unfeasible for many relevant applications. So, much effort is spent to construct simplified models retaining properties of the original chemistry.[1−3] Depending on the properties one is interested in, more or less details can be averaged out. Classical all-atom simulations can already be considered as a coarse-grained model with respect to quantum chemistry calculations by abandoning the electrons. The next steps are united-atom models which eliminate the pending hydrogens. A natural extension to coarser scales is to lump groups of atoms into one sphere. One now needs effective interaction parameters between these spheres which can be obtained via distribution functions from simulations on a more detailed level. The Boltzmann relation $P(x) = \exp(-F(x)/k_BT)$ relates the distribution with the associated free energy, the so-called potential of mean force $F(x) \sim -k_BT \ln P(x)$. One applies this to different distributions which may occur, as bond lengths, angles, and torsions along chains and radial distribution functions. One wants to avoid cross-correlations between these degrees of freedom for reasons of efficiency. So one should choose mapping centers such that cross-correlations are as small as possible.[4]

There are several further difficulties by applying this procedure: When $P(x)$ is the radial distribution function, $F(x)$ is used as a pair potential. However, $P(x)$ includes the effect of many-body interactions. Take for example a dense Lennard-Jones or even a hard sphere fluid: The potential has a single minimum or even no minimum; however, the radial distribution function exhibits long-range oscillations. This means that complex behavior is generated by a simple-looking potential. It is thus justified to make some simplifying assumptions about the interaction potential, e.g. about the cutoff of the interaction, and an optimization starting from the potential of mean force may be necessary. Different procedures for such an optimization process have been proposed to work e.g. with functional forms[4,5] or complete distribution functions.[6] Self-consistent optimization has been shown to work with rather long cutoff distances.[6−8] As we are working in dense melts, and to optimize the computational performance, we here choose rather short cutoff distances.

A related problem is that the distributions have been determined at a certain state point and that it is a priori not

* Corresponding author e-mail: hmeyer@ics.u-strasbg.fr.
† Institut Charles Sadron.
‡ Johannes-Gutenberg Universität.

10.1021/ct0503264 CCC: $33.50
Published on Web 04/08/2006

clear how far the temperature or density can be modified to still yield reasonable results. This is one question we address in the present paper.

Another difficulty is the question of transferability when different chemical species are mixed. This is already well-known from the construction of atomistic force fields. Recent attempts have been made with simultaneous optimization of hydrocarbon substances[9] and an example to handle tacticity in polystyrene on a united monomer level.[10] When going to higher degrees of coarse graining, chemical specificity is lost very rapidly.

Finally, the question arises of which level of coarse-graining one should choose. Actually, when many particles are lumped together, their effective interaction becomes very soft.[11,12] This is very popular in the context of so-called dissipative-particle dynamics[13] and may be very efficient to handle phase separation of copolymer systems. For polymers, the softness may result in the problem that bonds of neighboring chains can cross each other, thus violating the fundamental property of topological constraints. Some effort has been spent by Padding and Briels[14] to circumvent this problem for the simulation of polymer rheology.

The present work is motivated by research on the polymer crystallization process. In this context, the role of possible precursor or transient phases is debated.[15−18] Polymer crystals exhibit not only positional order but also orientational and conformational order. This means that a simulation model must be able to represent the melt as well as a crystalline state. This restricts the possibility of coarse graining to a rather low level. We choose in this work the "united monomer" representation, i.e., two backbone carbons per coarse-grained bead.

The paper is organized as follows: Section 2 describes the atomistic simulations we performed to get reference data of distribution functions and describes then a coarse-grained model of poly-vinyl alcohol "CG-PVA" which has already been used to model polymer crystallization.[19,20] Section 3 describes the derivation of coarse-grained models of polyethylene on the "united-monomer" basis where still a lot of choices are possible. Finally, we discuss in section 4 how the optimized models perform when extending the temperature range to get crystallization.

2. Simulation Methods on Different Scales

2.1. All-Atom Model: *n*-Alkanes. We performed all-atom simulations of short polyethylene chains with the YASP simulation program[24] using a model derived from the OPLS force-field. The OPLS-AA force field[21] was developed on the basis of preexisting force fields in order to simulate organic compounds accurately. Parameters describing bond stretching and angle bending interactions were taken from the AMBER and CHARMM/22 force fields,[22,23] ab initio molecular orbital calculations were used for torsions, and Monte Carlo simulations were used for nonbonded interactions in adjusting the thermodynamic and structural properties.

The interactions described by this force field as given in ref 21 are the following: connectivity interaction (harmonic potential $U_{\mathrm{bond}}(b)$ constraining the distance b between covalently bound atoms); angular interaction (harmonic potential $U_{\mathrm{ang}}(\theta)$ that determines the angle θ between two consecutive bonds); torsional interaction (3-term Fourier series)

$$U_{\mathrm{tors}}(\phi) = \sum_{n=1}^{3}\left(\frac{1}{2}k_n^{\mathrm{tors}}(1-\cos(n\phi-\phi_n^0))\right) \qquad (1)$$

for the angle ϕ formed by two consecutive dihedrals (similar potential terms exist for torsion angle involving both carbon and hydrogen atoms); and nonbonded interactions (Lennard-Jones potential $U_{\mathrm{mol}}(r)$ for atoms of different chains or of the same chain provided they are far enough from each other along the chain). The electrostatic interaction due to the partial charges q_a at every atom is also taken into account via the Coulomb interaction, which is cut off for distances larger than 0.9 nm.

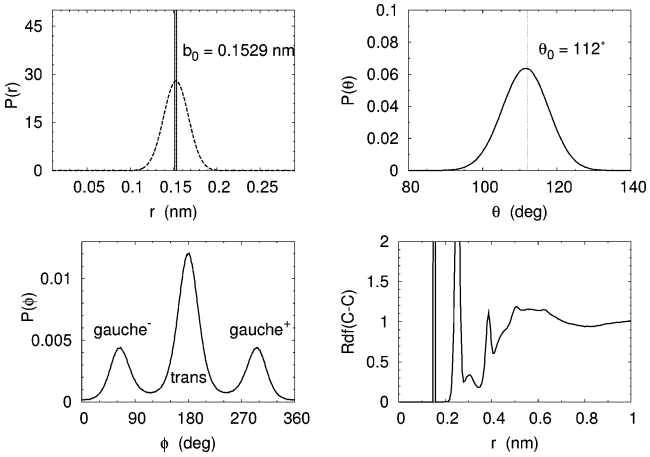


Figure 1. Probability distributions for the following quantities: bond length, bond angle, torsion angle, and radial distribution function (RDF). Only carbon atoms are considered here; these quantities are the result of the interactions of the force field. For the bond length distribution, the OPLS-AA prescription (dashed-line Gaussian curve) is shown in comparison with the delta function enforced by the constraint algorithm. The structure of the torsion potential exhibits the three favorable states trans, gauche$^+$, and gauche$^-$ which determine the flexibility of the chains. The first four peaks one can observe in the RDF are related to the three bonded interactions.

Figure 1 shows the distributions of the quantities that are directly or mostly influenced by the above mentioned interactions, considering the carbon atoms of the chains. The bond lengths, bond angles, and torsion angles are mainly determined by their corresponding potentials; however, some influence by the other degrees of freedom exists as we shall see later on. The radial distribution function (RDF) is determined by all the different interactions, more or less importantly depending on the distance. The quantities of Figure 1 are the ones we shall consider while discussing the coarse-grained description of the *n*-alkane chains.

The YASP simulation package implements a constraint algorithm (SHAKE) that has been used instead of the harmonic potential parametrized in the OPLS-AA force field, to access larger time steps by neglecting the rapid (and

unimportant) vibrations around the average value b_0 of the bond length. The YASP program makes use of the Berendsen thermostat (velocity rescaling; the coupling time used is 0.2 ps) and barostat (coupling factor 2 ps $\times$ 10^{-6} kPa) to keep temperature and pressure constant. Runs have been performed at constant atmospheric pressure and various temperatures. We also implemented a Langevin thermostat which essentially yielded the same results.

The parameters for the torsional potential of eq 1 have been adjusted (optimization of the k_n^{tors} coefficients) in order to reproduce the probabilities of trans and gauche states as reported in other simulation studies where the parameters of the force field had been optimized to recover accurate static properties of short polyethylene chains.[25] The optimized values of the parameters are given as follows (no changes were made to the terms involving both carbon and hydrogen sites):

$$k_1^{tors} = 5; \quad k_2^{tors} = -1; \quad k_3^{tors} = 0.6 \tag{2}$$

This optimization appeared necessary in view of the very strong influence of the torsional potential on the chains' conformation. Since torsions are the most flexible degree of freedom of the chains, they are mainly responsible for significant movements of the atoms on the local scale, and they also statistically determine the characteristics on the scale of the chain's size.

Atomistic simulations are calibrated in absolute units; the appropriate units here are nanometers (nm), picoseconds (ps), and kiloJoule (kJ) per mole. The time step used is 0.002 ps. We studied *n*-alkane molecules in the melt at atmospheric pressure and $T = 500$ K for chains containing 8, 10, 13, 14, 15, 16, 30, and 44 carbon atoms, in simulations runs lasting up to 20 ns. Table 1 summarizes a few quantities characterizing the simulation of systems with chains of different molecular weight at $T = 500$ K.

The model described above has been simulated at various temperatures, both above and below the experimental melting temperatures of the different *n*-alkanes. A continuous cooling protocol has been applied in order to trigger the transition toward the crystalline state, that comprises several different phases well characterized experimentally.[26] During cooling (at rates between -0.05 and -0.5 K ps^{-1}), no ordering of the chains was observed. The dynamics of the systems rather slowed as a glassy state was approached. Isothermal relaxation below the experimental melting point also could not reproduce the crystalline structure. One possible reason for this failure is a lack of precision in the model considered here, but it can also be related to the fact that nucleation barriers are rather high and thus makes nucleation in the small simulation volume a rare event.

There are many force fields one could use in order to simulate alkanes; the optimization of the parameters of such data sets is a difficult task. It is therefore not straightforward to find the most appropriate force field for a particular study. Moreover the optimization used here to *tune* the torsional interaction is questionable. However, the results obtained for static quantities show that the model still provides a reasonable description of *n*-alkane melts in the liquid state. Furthermore, using this model to simulate the melting of perfectly crystalline arrangements of the chains proved that such structures are stable at low temperatures, and a "rotator" phase (transitory phase showing a hexagonal symmetry; see ref 26 for experimental details and ref 27 for early simulation studies) was observed while the system transformed from the solid to the liquid state. This fact demonstrates the ability of the present model to account for the behavior of the low-temperature phases of short polyethylene chains.

On the other hand, it is possible to observe a phase transition using continuous cooling from the melt for artificially stiff chains. Modifying the torsional interaction to make the trans state much more favorable turned out to enforce the crystallization of a $C_{16}H_{34}$ system at a temperature $T_{cryst} \approx 350$ K, which is large compared to the experimental value (around 300 K), whereas one would rather expect to find a crystallization temperature lower than the experimental value when applying such high cooling rates. At $T = 500$ K, this unrealistically stiff model has a probability of trans states more than 20% higher than our regular *n*-alkane model, which results in a persistence length and a squared radius of gyration, respectively, 35% and 15% larger.

There is another reason our all-atom simulations cannot reproduce a crystalline configuration, at least for a reasonably stiff model: One is restricted to studying very small systems over short time periods, otherwise the computational time needed becomes prohibitive. Therefore, the probability to form a crystal is drastically lowered: As mentioned above, the smaller the box, the lower the probability of a nucleation event, and the very fast cooling rate hinders the rearrangement of the chains in ordered conformations. This suggests that there might be other possible simulation schemes that would be more appropriate to approach the liquid−crystal transition; more precisely, as crystallization implies a modification of the chain's conformation it is more important to consider what happens on the length scale of the monomers rather than on that of the individual atoms. This is the idea the coarse-grained models rely on. We shall see how the data obtained from all-atom simulations can be used to develop a coarse-grained model that retains prominent features of the molecules while making their simulation more efficient and thus complex processes such as crystallization accessible.

2.2. Coarse-Grained Model: CG-PVA. The CG-PVA model is a coarse-grained model that allows a qualitatively accurate description of a crystallizable polymer, in that the typical semicrystalline structure is obtained after cooling simulations.[19,20] Figure 2 shows a schematic representation of both the coarse-grained model and the underlying PVA molecule: It can be seen that the coarser model lumps one monomer (i.e. 7 atoms here) into one bead, thus making the simulation much faster. The figure also represents the different interactions between such coarse-grained beads.

Besides the gain in computational efficiency this scheme is particularly interesting because the coarse-grained chain still has a close link to the underlying atomistic description via the definition of the effective potentials governing the behavior of the coarse-grained monomers. These interactions are determined using the all-atom simulation data, and thus

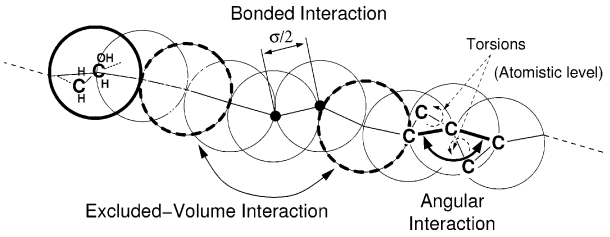


Figure 2. Schematic representation of the CG-PVA model. Atoms of the underlying molecule are shown to emphasize the mapping of one monomer onto one coarse-grained particle. The effective interactions between the latter are also depicted.

the coarse-grained model is really designed so as to reproduce the properties of the more detailed all-atom model.[4]

The interactions are the following: connectivity (harmonic potential $U_{\text{bond}}(b)$), angular interaction (tabulated potential $U_{\text{ang}}(\theta)$, see below), and nonbonded interaction (soft Lennard-Jones 6−9 potential $U_{\text{mol}}(r)$, without attractive part). There is no electrostatic interaction. Electrostatic effects accounted for by an explicit potential in the all-atom model are here absorbed in the effective nonbonded interaction.

The coarse-grained simulations were carried out using a simulation program similar to the YASP package but optimized for linear polymer chains and working with LJ units (see below). It is based on a velocity-Verlet integrator (time step 0.005 τ) and uses the Langevin thermostat instead of the Berendsen algorithm, with friction 0.5 τ^{-1}. The pressure is kept constant via position rescaling (Berendsen, coupling constant $10^{-5}\ k_{\text{B}}T\sigma^{-3}\tau$).

Figure 3 presents the three effective potentials. The main difference when compared to the all-atom situation is the absence of a torsional potential and the form of the angular potential. The angular degree of freedom in the coarse-grained model actually plays the role of torsions on the atomistic scale: More precisely, the coarse-grained angle θ is exactly determined by two consecutive torsional states of the atomistic backbone. This is reflected in the structure of the angular potential which displays three minima corresponding to trans−trans, trans−gauche, and gauche−gauche states of the underlying chain. Since in the all-atom case the *torsional* degree of freedom is most important for changes in the chains' conformations, the *angular* degree of freedom becomes the most important ingredient of the coarse-grained model. The torsion angle *on the coarse-grained level* is far less relevant and shows a weak structure only. The narrow minimum at 180 ° in $U_{\text{ang}}(\theta)$ is responsible for the possibility that parts of the chains stretch (forming stems that constitute the crystalline lamellae), and the other minima allow that the folds between two stems are locally stable. In the melt state at high temperature, the barriers can be crossed easily. The procedure used to derive the coarse-grained effective potentials was similar to what is presented in the next section 3 for the case of polyethylene.

The units used in the coarse-grained simulation are the following (the same units will be used for the description of the coarse-grained models of PE): length scale: $\sigma = 0.52$ nm, average bond length between the coarse-grained monomers $b_0 = \sigma/2$ (please note that there is no relation between

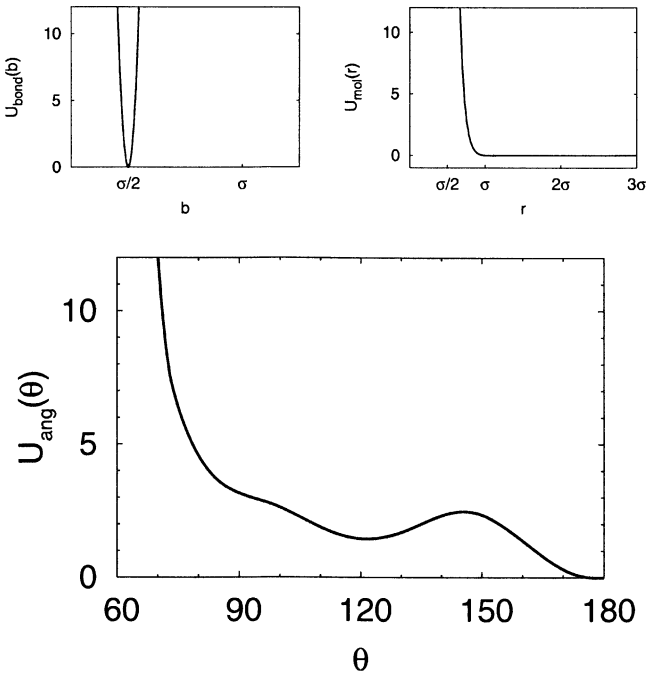


Figure 3. Effective interaction potentials between monomers in the CG-PVA model. The harmonic binding potential and soft Lennard-Jones nonbonded potential have simple analytical expressions, while the angular potential, also determined from all-atom simulations of PVA, has a more complex structure and has to be tabulated. The minima correspond to trans−trans, trans−gauche, and gauche−gauche states of the underlying atomistic backbone and lead to the formation of both stretched and folded parts of the chains at low temperature.

σ and the characteristic length σ_0 entering the Lennard-Jones potential definition, see eqs 4 and 6; mass scale: mass of one monomer taken to be 1 (all particles identical); and energy scale: $k_{\text{B}}T = 1$ for a reference temperature, here $T = 550$ K (temperature at which the data from atomistic simulations of PVA were taken to perform the mapping). The time scale cannot be determined in a straightforward manner as for the length scale, which is directly related to the way the mapping is performed. One possible way of setting this scale consists of a mapping involving the diffusion coefficient.

The properties of the CG-PVA model which allow for describing crystallizable polymer melts and to recover the correct qualitative structure are discussed in refs 19 and 28.

3. Coarse-Graining of the All-Atom Model

3.1. Derivation of Different Coarse-Grained Models. The coarse-graining procedure consists of defining effective interactions governing the behavior of particles on a larger length scale than the original model (thus also enforcing a longer time scale). The coarse-grained interactions are to be adjusted such that the static properties of the two models coincide. Obviously, since the coarse-graining process is devoted to "losing" some (irrelevant) information contained in the original model in order to make the simulation more efficient, the definition of the coarse-grained model cannot be unique, and choices have to be made so as to create an

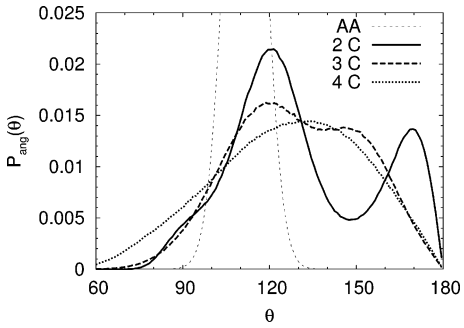


Figure 4. Comparison of the angular distributions obtained for different choices of coarse-graining level at $T = 550$ K. The original (harmonic) bond angle distribution is also indicated (AA). The distribution's structure weakens for mapping centers located 3 carbons away and completely disappears for centers separated by 4 bonds.

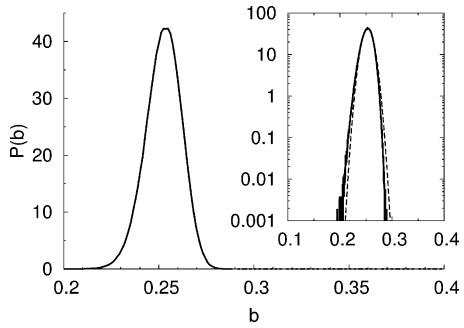


Figure 5. Distribution of bond lengths between two consecutive mapping centers measured from the all-atom simulation data. The inset shows the good agreement with the Gaussian distribution obtained using the harmonic potential of eq 4. The target distribution is slightly shifted toward smaller bond lengths.

appropriate model. In the case we are interested in, polyethylene (PE) melts and their behavior at low temperature experience acquired using the CG-PVA model tells us that a suitable coarse-grained model to access the semicrystalline structure through cooling of a melt has to retain enough features from the original chain. Moreover, the similarity of the PVA and PE chains prompts us to define a coarse-grained PE model in a way similar to the case of PVA. Since our studies of CG-PVA have shown that the key parameter in the model is the angular potential, it is natural to make the same choice here.[29] Thus we place the mapping centers on every other carbon atom of the atomistic backbone, which will yield a similar angular potential for the CG-PE model.

Figure 4 shows the influence of the degree of coarse graining on the structure of the angular potential: We computed the bond angle distributions for several choices of mapping centers. It can be seen that, when defining coarser models in which the interaction sites lie 2 or 3 carbons away from each other, the structure of the distribution (and thus of the effective angular potential to be derived) vanishes, leading to a bead-spring model with a generic, average rigidity as considered in the wormlike chain model.

Once the position of the mapping centers is chosen, one still has to define the type and characteristics of the effective interactions between such centers. In the case of simple linear polymer chains as for PVA and PE, the most natural choice is to define the very same type of interactions as for the standard all-atom models. Then the parameters of the bond, angular, torsional, and nonbonded interactions remain to be determined. To this end, using the assumption that the quantities constrained by the different effective interactions are weakly correlated, one can invert the probability distribution measured from the detailed simulation data via the Boltzmann formula in order to get the corresponding effective potential:

$$U(x) = -k_B T \ln P(x) \tag{3}$$

Of course, this method can only be approximate, since the quantity x is necessarily correlated to the other degrees of freedom of the system, and therefore $U(x)$ is not, strictly speaking, a potential energy. Furthermore, $U(x)$ as defined in eq 3 is only determined up to an additive constant, and a particular reference can be chosen for the energy. The inversion of the probability distribution is straightforward and accurate for the stiff degrees of freedom like the bond length. It becomes less obvious as the energy level of the effective potential decreases and the target quantity becomes more influenced by the medium. For the definition of the nonbonded interaction from the measured RDF for instance, more choices will have to be made for the form of the effective interaction.

With this methodology, we can now create several different coarse-grained models and adjust their parameters so as to reproduce first the target distributions $P(x)$ and then the other properties. Besides this, an obvious and very simple choice for a coarse-grained model of PE consists of taking the parameters from the CG-PVA model and modifying only the most important interaction from the all-atom simulation data, i.e., the angular potential. One can justify such an approach by assuming that, the structures of PVA and PE being overall very similar, the other parameters of CG-PVA should constitute an acceptable approximation for alkanes. We call this model "CG-PE_0"; the modified angular potential is derived in section 3.1.2, as part of the "regular" coarse-graining procedure that we shall present now.

In the following, we focus on simulations of 100 chains of $N = 22$ monomers which match our results for $C_{44}H_{90}$ obtained with the all-atom model.

3.1.1. Connectivity. The stiffest interaction in the system corresponds to the potential that ensures the connectivity of the consecutive particles inside a chain. As for covalent bonds between atoms in all-atom simulations (which in the case of our alkane model are treated as rigid constraints, see section 2.1), the fluctuations of coarse-grained beads around the average bond length is very rapid and relatively unaffected by the other interactions. Therefore the simplest description of the bonded interaction consists of assuming a harmonic form of the interaction

$$U_{bond}(b) = \frac{1}{2} k_{bond} (b - b_0)^2 \tag{4}$$

and determining the appropriate coupling constant k_{bond} via a fit of the data obtained from the all-atom simulations. Figure 5 presents the target distribution $P_{bond}(b)$ and the

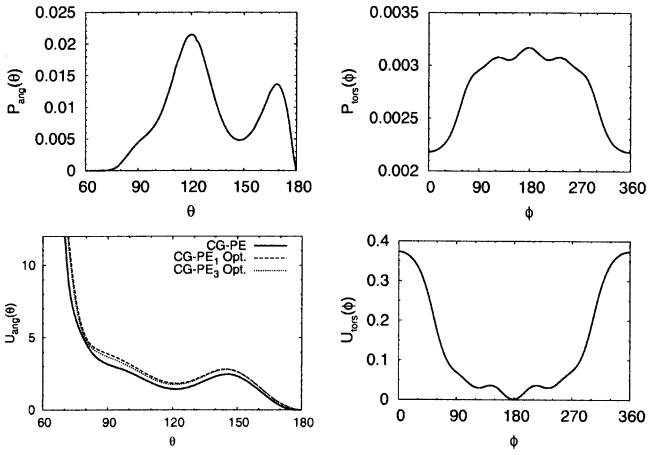


Figure 6. Bond angle and torsional angle distributions measured from the all-atom simulation and the corresponding effective potential derived from them using eq 3.

Gaussian fit reproducing well the all-atom data, except for a slight shift of the distribution toward the lower bond lengths. This effect is unimportant for our coarse-grained models. The fitting parameter takes the value $k_{bond} = 11476\ k_B T\sigma^{-2}$; the average bond length $b_0 = 0.2535$ nm $\approx \sigma/2$ is the reference length in the coarse-grained simulations, see section 2.2.

3.1.2. Bond Angles. The bond angle distribution measured between every other carbon atom in the atomistic simulations exhibits a more complex structure than the bond length distribution or angular distribution on the atomistic level, as a result of the influence of the underlying torsional states as already mentioned (cf. Figure 6). Here it is thus not possible to find a simple analytic expression for the effective potential $U_{ang}(\theta)$, and one has to use the Boltzmann inversion

$$U_{ang}(\theta) = -k_B T \ln\frac{P_{ang}(\theta)}{\sin\theta} + U^0_{ang} \qquad (5)$$

The resulting potential is also shown in Figure 6; the arbitrary constant U^0_{ang} has been set so that the reference angular energy is 0 for trans−trans states. The structure of the coarse-grained angular potential obtained for *n*-alkanes is much less pronounced than the corresponding CG-PVA potential. Only two of the three expected favorable states are observable, the gauche−gauche state being almost totally absorbed into the neighboring trans−gauche state. This comes from the fact that the angular distribution is sharper in the case of PVA, with very weak probabilities of transition between two states. This is due to intrachain hydrogen bonds that stabilize the conformations of the PVA chains. The absence of hydrogen bonds in the PE melt causes the chains to become more flexible. Nevertheless, we shall see that the specific shape of the angular potential will strongly favor stretched states and thus drive the crystallization at low temperature.

3.1.3. Torsions. The distribution of torsion angles measured from the conformation of four consecutive mapping centers is shown in Figure 6 and can be inverted in the same fashion as has been done for the angular degree of freedom (except there is no sine from the jacobian in the Boltzmann inversion) to yield an effective torsional potential $U_{tors}(\phi)$.

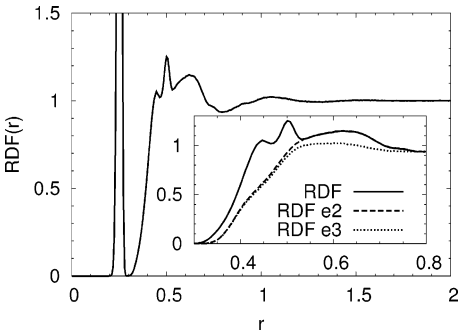


Figure 7. RDF measured from the position of every other carbon atom along the atomistic backbone. In the inset, the same distribution is compared to variants calculated with the 2 closest neighbors excluded (RDF e2) and the 3 closest neighbors excluded (RDF e3). The latter functions are used in the Boltzmann inversion formula when deriving the nonbonded effective potential in the case of a completely repulsive potential and a potential with an attractive part.

Although much less stiff than the angular potential (smaller energy scale), it still clearly favors the trans configuration of the coarse-grained chains. The trans state of the coarse-grained backbone should not be confused with the trans state of the original atomistic model, which determines the angular interaction. The definition of a torsional interaction on the coarse-grained level can be viewed as an attempt to fine-tune the description of the chains' flexibility on a larger length scale.

The torsional potential derived for CG-PE appears weak but still clearly favors the stretched conformations, thus it seemed worth including it in the coarse-grained model. To be sure about the influence, we constructed two models, one with and another without torsions.

3.1.4. Nonbonded Interactions. The effective interactions determined above are rather straightforward to parametrize, using either an analytic expression or direct Boltzmann inversion to map the data from all-atom simulations. Whether using these mappings in coarse-grained simulations will yield distributions that match the target functions is of course not obvious, but the method is clear. In the case of the nonbonded interaction $U_{mol}(r)$, it is not possible to neglect the influence of the other degrees of freedom anymore, and more choices have to be made. This is apparent when looking at the RDF presented in Figure 7: The first peaks are directly related to the positions of the nearest neighbors of one "monomer" (i.e. an ethylene group in the case of the PE chains considered here), and thus RDFs have to be calculated excluding these contributions the effect of which has already been taken into account when deriving the previous effective interactions. Figure 7 also shows RDFs where respectively the first two and three neighbors of one particle have not been considered in the calculation, in comparison with the full RDF.

One possible solution to determine $U_{mol}(r)$ consists of guessing a functional form as we did in the case of the binding interaction in section 3.1.1; CG-PVA uses a softer Lennard-Jones potential of the form

$$U_{mol}(r) = \epsilon_0\left(\left(\frac{\sigma_0}{r}\right)^9 - \left(\frac{\sigma_0}{r}\right)^6\right) \qquad (6)$$

which is truncated and shifted at the position of the minimum, yielding a fully repulsive interaction. It is possible to choose such a functional form and then try to adjust the values of the coefficients ϵ_0 and σ_0 in order to reproduce schematically the trend exhibited by the RDF. However, this method cannot be very precise, and it is possible to attempt to match the RDF more closely either with an optimization of (ϵ_0,σ_0) as presented in ref 4 or using the Boltzmann inversion on part of the target RDF.

In any case, the first choice to be made is related to the torsional interaction: Depending on whether an effective interaction is included in the coarse-grained model or not (in the case of CG-PVA, it was found unnecessary), the mapping for the nonbonded interaction is different. The reason for this is that we take the viewpoint that all interactions which are explicitly taken into account by bonded interaction potentials should not be considered in the nonbonded potential. Thus, if torsions are included, the 1−4 interaction (i.e. particles separated by three consecutive bonds) should be excluded from the nonbonded interaction and from the target distribution. If no torsions are considered, the 1−4 interaction should be also included in the RDF target distribution.

Since these RDFs take into account the correlations between any two particles in the system (except the close neighbors along the chain), they contain several peaks indicating the average positions of the neighbor shells. A direct Boltzmann inversion of such a function would yield a long-range potential which takes many-body interactions into account. To avoid this, a cutoff distance has to be chosen beyond which the effective potential is zero (see also ref 6). Here again, many possibilities arise. Since an essential aim of the coarse-graining procedure is to save computer time, we wish to have a potential with the shortest possible interaction range. Thus, the most natural possibilities consist in cutting off the potential either at the first *minimum* (yielding a fully repulsive interaction, as in the case of the CG-PVA model) or at the first *maximum* (so as to include an attractive part). The latter options appear equally sensible, since in the molten state the high density of the system screens the nonbonded interactions beyond the nearest-neighbor distance. Thus, keeping the density of the system constant either in constant-volume or constant-pressure simulations should yield equivalent results, at least for simulations at the given state point. However, note that the thermodynamic properties involving derivatives as the equation of state are strongly influenced by the attractive tail of the potential. Thus, to ensure a better transferability in the temperature or pressure range, a potential with an attractive part is probably the better choice.

The simulation procedure applied consisted in a first equilibration run at constant temperature and volume, using for the density the value extracted from the all-atom simulations. During this NVT simulation, the average pressure is measured. This pressure is applied in the following constant-pressure simulations for a second equilibration step and the cooling-heating cycles. The value of the pressure is arbitrary in our simulations and differs depending on the model. The values of the pressure (in $k_BT\sigma^{-3}$) for the different coarse-grained models are indicated in Table 2, together with a summary of the force-field characteristics. It can be seen that a lower pressure is sufficient to achieve the right density in the case of the models with an attractive nonbonded interaction.

Table 1: Static Properties of Different *n*-Alkanes Obtained by All-Atom Simulations[a]

system	chains	duration (ns)	l_p/b_0	l_p (nm)	R_g^2 (nm^2)	R_e^2/R_g^2	D (nm^2 ps^{-1})
C_8H_{18}	256	2	2.17	0.332	0.0707	7.78	0.01273
$C_{16}H_{34}$	84	20	2.33	0.357	0.205	7.93	0.00353
$C_{44}H_{90}$	100	4	2.35	0.359	0.831	6.98	0.0080

[a] For the three different chain lengths presented here, simulation details (number of chains and length of the simulation runs) are indicated, together with the following properties: Persistence length l_p (measured as the characteristic length of the vector orientation decorrelation along one chain, for vectors joining every other carbon atom), squared radius of gyration R_g^2 and end-to-end distance R_e^2, and diffusion coefficient.

Table 2: Summary of the Force-Field Characteristics for the Different CG-PE Models

model	nonbonded potential	torsional potential	applied pressure (at $T=1$)
CG-PE$_1$	repulsive	explicit	3.73
CG-PE$_2$	repulsive	implicit	4.36
CG-PE$_3$	attractive	explicit	2.74
CG-PE$_4$	attractive	implicit	1.80

In case surfaces are involved in the study, the model must be optimized such that the applied pressure is 0. References 6 and 10 solved this by adding a linear term to the nonbonded potential. However, one should probably include the pressure (as well as other thermodynamic quantities as mentioned above) in the optimization procedure. Here, we are only interested in simulations of the melt and can treat the pressure as a force-field parameter.

3.2. Optimization of the Coarse-Grained Models. The prescription given above for the different coarse-grained models has been tested in simulation; the first requirement of the coarse-graining approach is that the target distributions extracted from the all-atom simulation data are reasonably well reproduced with the simplified scheme. This should ensure that the static properties are very close in the two simulation models.

The bond length distribution, as expected, exactly matches (within the line width) the Gaussian function obtained by inversion of the prescribed potential, whatever model is used. This is not surprising since the latter potential is very stiff, and therefore the bond interaction is not influenced by the other degrees of freedom.

Contrary to that, the angular distribution shows significant deviations from the target, for all models (cf. Figure 8). This is very different from what has been observed for the CG-PVA model for which the measured angular distribution reproduces almost exactly the prescribed one (with only minor differences which do not significantly affect the population of the different, well-separated states). The reason for that is the much smoother form of the angular distribution for PE, as already mentioned above. In particular, it can be

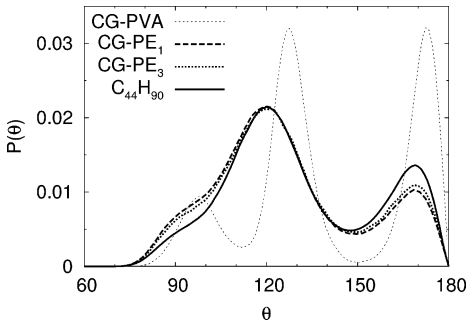


Figure 8. Comparison of the angular distributions obtained with the various models at $T = 1$. The probability of trans−trans states appears reduced compared to the target distribution from all-atom models.

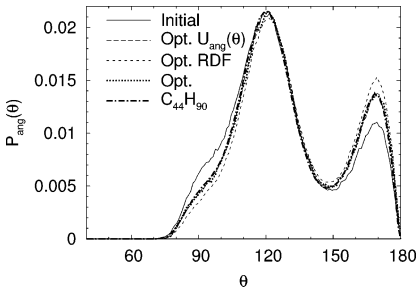


Figure 9. Evolution of the angular distribution $P_{ang}(\theta)$ during optimization of $U_{ang}(\theta)$ for the CG-PE$_3$ model. The *initial* distribution is obtained directly from the potential inverted from the atomistic $P_{ang}^{target}(\theta)$, i.e., without any optimization. After optimizing $U_{ang}(\theta)$ the distribution matches the target (Opt. $U_{ang}(\theta)$) but deviates again after the optimization of the RDF (Opt. RDF). The agreement is fine again after a second optimization of both $U_{ang}(\theta)$ and $U_{mol}(r)$ (Opt.).

noticed that the population of the trans−trans state is significantly lowered for all coarse-grained models, whereas the shoulder at 90° is higher than in the target distribution. This effect seems slightly more pronounced in the case of models having no explicit torsional potential, surprisingly (one would expect the torsional potential to make the chains stiffer, but this probably only occurs on larger length scales). This lowering of the probability of trans−trans states is probably due to the density of the melt which tends to compress the chains and makes folded angular states more favorable than the stretched states.

The torsional distribution that is enforced via an explicit torsional potential appears to be well reproduced for the models CG-PE$_1$ and CG-PE$_3$, whereas the other models have a less pronounced distribution of torsional angles, showing nevertheless the same trend (more favorable trans state), see also Figure 10.

The RDF measured on the data obtained with the various coarse-grained models proved to depart significantly from the original function. These discrepancies are directly related to the mismatch described above for the angular interaction: The peaks associated with the angular states are too low, and a shoulder appears between the first and second peak in the global RDF, corresponding to the too large proportion of folded states already observed. Apart from these differences concerning the short-range particle−particle

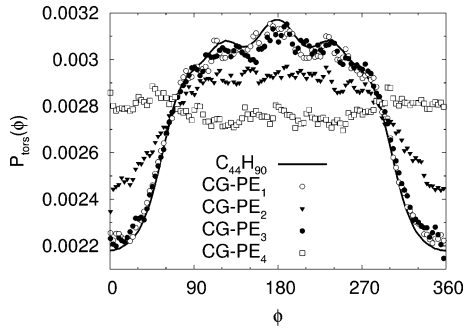


Figure 10. Torsional distribution obtained for the different coarse-grained models. The results for CG-PE$_1$ and CG-PE$_3$ which both include an explicit torsional potential are very close to the target distribution, whereas the other two models exhibit significant deviations; for CG-PE$_4$, the distribution is almost completely flat, indicating that no particular torsion angle is favored. CG-PE$_2$ has some structure identical in trend to the stiffer models. (Gliding averages have been applied to the data.)

correlations, one can notice that the subsequent peaks of the RDF, associated to the neighbor shells, are shifted. These long-range differences are more difficult to relate to the shape of the effective potentials, which all necessarily have short ranges. The mismatch between the coarse-grained and the all-atom RDF varies with the type of coarse-grained model but not qualitatively.

The previous observations show that the coarse-grained models "simply" derived, i.e., with effective potentials directly inverted from the target distributions, are not sufficient to reproduce the correct behavior of a dense polymer melt, even at high temperatures where one can hope that the different degrees of freedom are relatively independent from each other. This only works well in the case of very sharp distributions. Thus, an additional step is needed in the coarse-graining procedure: The effective potentials associated with distributions that are not well reproduced in the coarse-grained simulations have to be optimized.

The optimization of the effective potentials is an iterative procedure during which one coarse-grained potential is modified until the corresponding probability distribution matches the target distribution.[2,3] This has to be performed for all the degrees of freedom that are not well described compared to the reference all-atom simulation. In doing so, it is sensible to consider the interaction in decreasing order of intensity, i.e., the strongest potential should be optimized first, and the method should be applied in turn to the other interactions, ideally refining the model at every step (provided the different optimizations do not have contradictory effects).[4] The order in which the different potentials should be optimized is therefore, in our case, the following:

$$U_{bond}(b) \rightarrow U_{ang}(b) \rightarrow U_{mol}(b) \rightarrow U_{tors}(b) \qquad (7)$$

Optimization of the binding interaction was not needed for any of the models. The procedure would have been slightly different than for the other potentials, since for this degree of freedom we used an analytic form for the potential (whereas it is tabulated for the other interactions). In that

case, an optimization of the parameters (here, b_0 and k_{bond}) using e.g. a simplex method is needed.

The angular potential is the potential that has to be optimized first, and this is a particularly important step since we know its prominent influence on the crystallization process. The method consists of modifying $U_{ang}(\theta)$ iteratively[6] via

$$U_{ang}^{n+1}(\theta) = U_{ang}^{n}(\theta) + k_B T \ln \frac{P_{ang}^{n}(\theta)}{P_{ang}^{target}(\theta)} \qquad (8)$$

with $U_{ang}^{n}(\theta)$ being the potential at step n, $P_{ang}^{n}(\theta)$ the probability distribution of angles when applying this potential in the simulation, and $P_{ang}^{target}(\theta)$ the target distribution from the all-atom simulation. The optimization procedure was observed to converge extremely fast, with most of the changes already accomplished after the first iteration, subsequent iterations only slightly refining the agreement between $P_{ang}^{n}(\theta)$ and $P_{ang}^{target}(\theta)$. Ten iterations were performed, yielding a satisfactory agreement; the resulting distributions at various steps during the optimization procedure are shown in Figure 9.

Since the measured torsional distribution was very close to the prescription already, it was not necessary to apply the optimization algorithm to this interaction; this is shown in Figure 10.

The same procedure as described in the case of the angular potential was applied to optimize the effective nonbonded potential, and it also converged very fast—again one iteration was sufficient to yield a very small discrepancy between $RDF^{n}(r)$ and $RDF^{target}(r)$ (the total RDF were used as a means of comparison between the nth and n + first distributions). After 10 iterations, the RDF did not change appreciably, and the mismatch with the target RDF was small (but still nonzero, especially in the case of CG-PE_1 for which the nonbonded interaction has a very short range and therefore cannot strongly influence the large-r region behavior of the RDF).

After optimizing $U_{mol}(r)$, it appeared that the angular distribution has been modified as well and no longer matched the target $P_{ang}^{target}(\theta)$. This is not surprising, since the structure of the melt is necessarily strongly influenced by the angular distribution. Therefore convergence of both the angular distribution and the RDF at the same time cannot be guaranteed. However, repeating the iterative process for the angular distribution first, then for the RDF yielded satisfactory results for both distributions, in the case of CG-PE_1 and even better for CG-PE_3 (see Figure 11)—which could reproduce very accurately the RDF on a larger length scale due to the larger cutoff in $U_{mol}(r)$; the optimized nonbonded potential thus has a stronger influence on the large-scale structure and facilitates convergence to the target distribution. Again, 10 iterations were performed for each degree of freedom, while most changes occurred after the first step.

The pressures that were used to maintain the density to its value obtained from the all-atom simulation were, respectively. 2.39 and 1.21 $k_B T\sigma^{-3}$ for CG-PE_1 Opt. and CG-PE_3 Opt., which is lower than the values of the unoptimized models indicated in Table 2. This means that it is easier to constrain the target density in the case of optimized models. Figure 12 presents the final form of the angular and nonbonded potentials, compared to the original CG-PE model, i.e., before and after optimization. As a conclusion, optimization of the coarse-grained potentials of the CG-PE models appeared quite successful in improving the ability to reproduce the static structure of the original all-atom model, as can be seen from the values of the squared radius of gyration R_g^2 which are 2.59 for CG-PE_1 and 2.70 for CG-PE_3 and become 2.90 and 2.94, respectively, for CG-PE_1 Opt. and CG-PE_3 Opt.; this is comparable to the value 3.03 obtained for R_g^2 with the corresponding detailed simulations of $C_{44}H_{90}$.

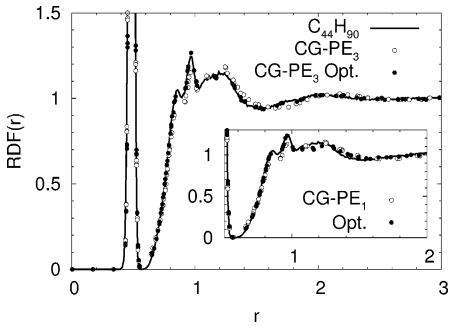


Figure 11. Comparison of the (total) RDF for the original $C_{44}H_{90}$ model, CG-PE_3, and its optimized form (after the second optimization procedure). The inset shows the same quantities in the case of CG-PE_1. It can be seen that the RDF is very well reproduced using optimized effective potential, even better in the case of CG-PE_3 which has an attractive part and therefore a longer range.

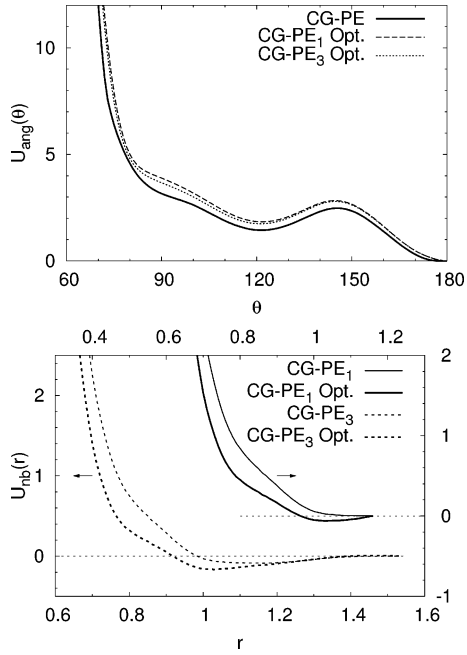


Figure 12. Optimized angular and nonbonded potentials for the CG-PE_1 and CG-PE_3 models. The original potential $U_{ang}(\theta)$ is unique, whereas $U_{mol}(r)$ depends on the CG-PE model (smaller or larger cutoff). One can notice that the optimization process for CG-PE_1 which is originally completely repulsive, enforces the apparition of a slightly attractive part.

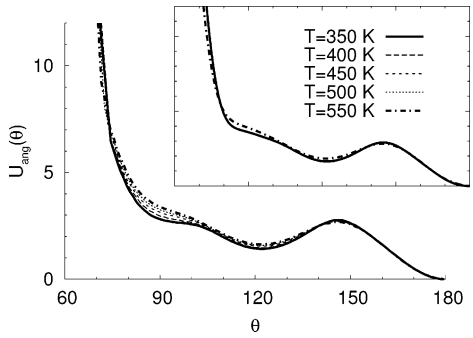


Figure 13. Comparison of the effective angular potentials $U_{ang}(\theta)$ (no optimization) inverted from the atomistic simulations data corresponding to equilibrium at several temperatures. The curves differ particularly for the gauche−gauche state at around 90° for the extreme temperatures, whereas for the optimized angular potentials (inset) they are found to be closer to each other (potentials for the models CG-PE_1 and CG-$PE_1^{T=350}$).

3.3. Effect of the Temperature on the Mapping. In this study of coarse-grained models for *n*-alkane melts, we make the assumption that the effective potentials describing the interactions between the beads are valid at *any* temperature, even though they are determined from the atomistic simulations *at one temperature only*. This is certainly justified in case we only consider the melt at this particular temperature or for small variations around it but is obviously wrong when we use the simple coarse-grained model over a large range of temperatures, as for a cooling toward the crystalline state of the polymer system, as presented in section 4. Therefore we will consider hereafter the effect of temperature changes on the parameters in the models and check whether it can be valid to make use of such a simplistic approach.

To check the validity of our simplified approach, we can have a look at the different effective potentials and monitor their behavior as the temperature is modified. The binding potential $U_{bond}(b)$ is a harmonic potential [see eq 4] whose parameters were determined by a fit: One can calculate the parameters b_0 and k_{bond} at other temperatures as well. Doing so between $T = 350$ K and $T = 550$ K, it was found that the parameters are temperature independent, i.e., that the change in the bond length distribution observed for two different temperatures T_1 and T_2 is completely determined by the ratio T_1/T_2. This comes from the independence of the bond length from the other parameters in the system and is not surprising when considering that, in this case, the harmonic interaction on the coarse-grained level is directly related to the harmonic interaction constraining the bond angles on the atomistic level. Thus the effective binding potential is a true potential energy for our models.

The angular potential is determined via the Boltzmann inversion of an angular distribution. Figure 13 shows the comparison of the effective potentials that are derived directly from the atomistic simulations, without optimization, for various temperatures. It is observed that in this case the potentials are different; there is a temperature dependence of the angular potential. To obtain another comparison of the angular potentials and the effect of temperature, we proceeded to the derivation of coarse-grained effective potentials using atomistic simulation data at a lower temperature, i.e., $T = 350$ K (corresponding to $T = 0.64$ for coarse-grained simulations). The method used was exactly the same as presented in sections 3.1 and 3.2: After the first derivation, the effective potentials were optimized to yield a model CG-$PE_1^{T=350}$. This model again significantly improves the reproduction of the structure of the atomistic model, with for instance a squared radius of gyration $R_g^2(T = 0.64) \simeq 3.13$ instead of 3.62 before optimization; this is to be compared to the value measured in the atomistic simulation, $R_g^2(T = 350\text{K}) \simeq 3.19$. The model CG-$PE_1^{T=350}$ is the equivalent of the CG-PE_1 model, i.e., it has an explicit torsional potential and no attractive part. The model CG-$PE_3^{T=350}$ with an attractive part in the nonbonded potential could not be optimized nor used, since at the temperature $T = 0.64$ it already started to crystallize during the equilibration phase of the simulation.

The inset in Figure 13 compares the *optimized* effective angular potentials of CG-PE_1 (mapping done at $T = 550$ K) and CG-$PE_1^{T=350}$: The two functions only slightly differ, compared to the differences observed for the nonoptimized potentials. The two potentials most notably differ (by approximately 10%) at around $\theta = 90°$ and $\theta = 120°$. This shows that the optimized potentials are much closer to being potential energies than the simply derived potential which are in fact free energies and take into account a certain amount of entropy corresponding to the influence of the other degrees of freedom in the system.

The torsional potential was also computed from the atomistic simulations at different temperatures; here the changes were more drastic, since the probability distribution of the angle ϕ is qualitatively different depending on the temperature: The local minima on each side of the maximum at $\phi = 180°$ (see for instance Figure 10) are more pronounced at temperatures lower than 400 K (not shown). The coarse-grained models do not reproduce this tendency, and this discrepancy could only be treated by introducing the appropriate temperature dependence in the definition of the effective potentials. However, one can argue that this is a weak effect and that the torsional degree of freedom does not play a key role for our models. Actually, cooling simulations of CG-PE_1 and CG-PE_2 show that the two models differing by the torsional potential have essentially the same behavior and similar crystallization temperatures.

The observations for the nonbonded interaction are the same as for the angular potential: The inverted potentials are indeed rather different depending on the temperature chosen for the mapping, but when optimized the mismatch tends to disappear, as observed when comparing CG-PE_1 to CG-$PE_1^{T=350}$. So once again the effective potentials determined for the optimized models are a good approximation to potential energies.

Thus, it is possible as a first approximation to use coarse-grained potentials mapped at one single temperature from a reference (here an all-atom simulation) but only when using *optimized* models. For the latter, the correlations that might exist between the different degrees of freedom are taken into account during the optimization procedure, and the effective potentials therefore correspond more closely to the ideal

potential energies. For our CG-PE models, it could be checked that the effective potentials are not too different at different temperatures (above T_{cryst} of course), which means that they are a good approximation to the "real" potentials. However, this is still not a sufficient condition for the coarse-grained models to reproduce the phase transition to the crystalline state at low temperature; this is what is considered in the next section.

4. Crystallization Simulations

Crystallization of the different models of PE could be observed in nonequilibrium simulations consisting in a continuous cooling, i.e., a steady decrease of temperature imposed by the thermostating algorithm. Continuous cooling has been performed at a constant rate of $-5 \times 10^{-6}\ \tau^{-1}$ and was followed by continuous heating (rate $2 \times 10^{-5}\ \tau^{-1}$).

4.1. Time Scales. The use of identical cooling and heating rates for the nonequilibrium simulations of the different models is justified by the similarity of the time scales that can be derived for CG-PE: The characteristic time $\tau_R = R_g^2/D$ takes the value 372 τ, $\pm 54\ \tau$ depending on the model. This allows us to estimate the time unit τ in the coarse-grained simulations: Comparing $\tau_R = 870$ ps for $C_{44}H_{90}$ to $\tau_R = 372\ \tau$, one obtains $\tau \approx 2.3$ ps. It is possible with this figure to compare more quantitatively the "efficiency" of the two simulation schemes: The all-atom simulations' integration time-step is 0.002 ps, whereas it is $0.005\ \tau \approx 0.01$ ps for the coarse-grained method.

This estimation implies that during cooling at the rate used here, the temperature changes by about 10^{-3} during the chain's relaxation time τ_R, meaning that the system is not so far from the equilibrium, at least at the beginning of the cooling run. At $T = 0.6$, the relaxation time is increased by almost a factor 20, i.e., the chain relaxes during cooling of 0.02. This is still very rapid and means that the slowing down should not be an important driving force of the crystallization.

Upon optimization of the potentials, D is observed to increase slightly. The behavior of the diffusion coefficient with temperature is examined in Figure 14; it is seen that data for the coarse-grained models exhibit the same trend typical of an Arrhenius law, $D \sim \exp(-U/k_BT)$. Using the length and time scales σ and τ, it is possible to compare the coarse-grained and all-atom data and this shows that both simulations yield a similar behavior, at least for high temperatures. This also corroborates the statement that the systems are not far from equilibrium during cooling. The decrease of D with $1/T$ is faster for the all-atom simulations than for the coarse-grained models; this points out the limitations of the simple coarse-graining procedure that does not allow for reproducing the dynamics of the system on a large temperature scale (note that the deviation has the same trend as the deviation of the thermal expansion coefficient shown in Figure 17).

4.2. Crystal Structure. Figure 15 shows the results for cooling of the models with nonoptimized and optimized potentials; they crystallize and melt at the temperatures indicated in Table 3 (melting and crystallization temperatures are determined as the inflection point in the behavior of the volume as a function of temperature). The experimental value

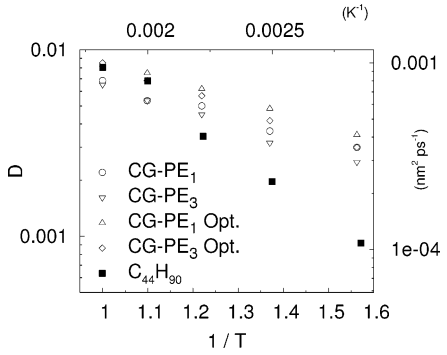


Figure 14. Evolution of the diffusion coefficient as a function of the inverse temperature. The "coarse-grained" units are used; the values for $C_{44}H_{90}$ have been rescaled by a factor σ^2/τ with $\tau = 2.3$ ps as derived from the mapping of the relaxation times τ_R, i.e., $\tau \approx 2.3$ ps. The points corresponding to the coarse-grained models show very similar behavior, whereas the data for $C_{44}H_{90}$ depart from it at temperatures lower than 500 K. This is to be connected to the mismatch in the expansion coefficient shown by Figure 17. (The top and right-hand axes show for reference the corresponding values in units appropriate to all-atom simulations, i.e., K^{-1} and $nm^2\ ps^{-1}$.)

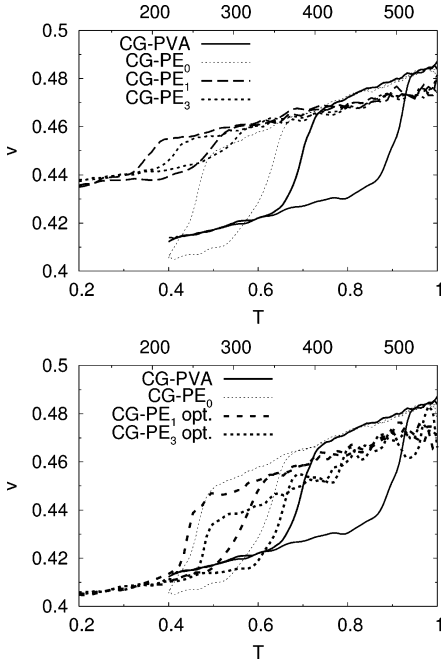


Figure 15. Phase diagrams for the different coarse-grained models and for CG-PVA: Evolution of the volume per monomer as a function of temperature during continuous cooling and subsequent reheating. The upper figure shows the results for the nonoptimized models, and the phase diagrams for the models with optimized effective potentials are presented below. It can be seen that the model which matches the high-temperature structure of the atomistic model the best (CG-PE_3 Opt.) also yields the melting point that is closest to the experimental data for $C_{44}H_{90}$ ($T_{melt} \approx 360$ K). (The simulation data have been smoothed.)

for the melting temperature of $C_{44}H_{90}$ is $T_{melt} \approx 360$ K, which corresponds to $T = 0.65$ in the units relevant to our models. It has to be pointed out that the crystalline structures formed

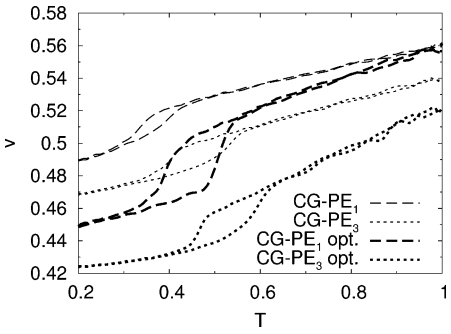


Figure 16. Phase diagrams obtained for chains of 50 monomers simulated with the different coarse-grained models for PE. The hierarchy of the different coarse-grained models is unchanged: The optimized CG-PE_3 has the higher melting point, and CG-PE_1 the lowest. It is also worth noticing that, for this longer chain length, the densities measured for the different models do not match anymore, as was the case for the system used for the mapping of the potentials ($N = 22$, density 2.1 at $T = 1$). The phase diagram obtained for $N = 100$ is extremely similar to what is observed here.

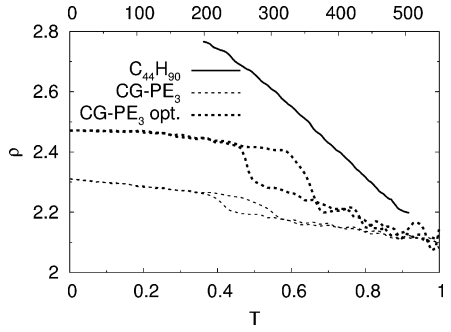


Figure 17. Density of the melt vs temperature measured during continuous cooling for the original atomistic model and CG-PE. The expansion coefficients for the latter are much smaller than for $C_{44}H_{90}$.

Table 3: Crystallization and Melting Temperatures Measured during Continuous Cooling and Heating of the CG-PE Models

	$N = 22$		$N = 50$		$N = 100$	
model	T_{cryst}	T_{melt}	T_{cryst}	T_{melt}	T_{cryst}	T_{melt}
CG-PE_1	0.35	0.49	0.33	0.39	0.33	0.42
CG-PE_2	0.40	0.52				
CG-PE_3	0.41	0.54	0.40	0.53	0.40	0.52
CG-PE_1 Opt.	0.44	0.56	0.39	0.51	0.40	0.49
CG-PE_3 Opt.	0.48	0.65	0.47	0.59	0.47	0.62

during the cooling simulations of coarse-grained models cannot reproduce all features of *n*-alkane systems: The crystals have a hexagonal symmetry due to the cylindrical shape of the chains' stretched parts. This is qualitatively different from the orthorhombic (or triclinic, depending on the chain length[26]) phase that is characteristic of *n*-alkanes and for the description of which the explicit treatment of the hydrogen atoms is mandatory. However, the larger scale features of the crystals, i.e., the lamellar-like structure, is well reproduced, and so is the apparition of folds along the chains for higher molecular weight molecules, as we shall see later on.

The temperatures of crystallization/melting are correlated to the probability of finding stretched states in the melt *at high temperature* ($T = 1$), showing the prominent influence of the bond angle degree of freedom on the properties of the model and particularly its low-temperature behavior.[29] The nonoptimized models which have a too low probability of trans−trans states compared to the original atomistic model yield lower crystallization and melting points, since chain rigidity favors formation of the crystalline phase and also prevents it from melting at lower temperature. The pronounced hysteresis found not only relates to the large rates for cooling and heating but also depends on the activation barrier to be overcome when melting the crystalline phase.

Since the results obtained from the simulations of the coarse-grained models with $N = 22$ (corresponding to $C_{44}H_{90}$) appeared reasonable, it was interesting to extend the simulation to longer chains, as *n*-alkanes are just short PE molecules. The phase diagrams obtained for $N = 50$ are shown in Figure 16, and Table 3 summarizes the crystallization and melting temperatures also for $N = 100$. Simulations were performed with 72 chains for $N = 50$ and 192 chains for $N = 100$; the cooling and heating rates applied are the same as in the case of $N = 22$.

Note that the typical relaxation times increase as N^2 in this length region. Thus the temperature difference corresponding to the relaxation time becomes significant for the longest chains, and they may be already out of equilibrium. The different models crystallize and melt at about the same temperatures as $N = 22$, whereas one might have expected that the crystallization temperature increases with chain length. The phase diagrams appear very similar for the longer chains $N \geq 50$, indicating a fast saturation of T_{cryst} and T_{melt}. It is to be noted that the densities of the melt at $T = 1$ are not identical anymore depending on the model; the density was a constraint in the derivation of the coarse-grained models, as the value of the pressure has been adjusted so as to reproduce the density of the all-atom simulation.

The density varies linearly with the inverse chain length $1/N$ due to chain-end effects (this has been checked for the CG-PVA model, with chains from $N = 10$ to $N = 1000$). However, the proportionality coefficient necessarily depends on the force field and therefore need not be the same for different models. Another constraint could have been introduced during the coarse-graining procedure to enforce the density variation as a function of the chain length, or temperature.

The morphology of the crystals depends on chain length and on the force field used. For short chains, the melt turns into a crystal of rods as the chains can still fully stretch upon cooling, whereas for longer chains the qualitative behavior of polymer systems is recovered: The chains are constrained, and parts aggregate during crystallization before the whole chain has had the possibility to stretch in order to reach the all-trans state. This gives rise to ordered lamellar regions and amorphous zones surrounding them. This is observed in our simulations of CG-PE as well as for CG-PVA;[20] the crossover between the rod and the polymer regime is located

around $N = 20$ at such cooling rates. For the nonoptimized (and therefore not stiff enough) models most chains are folded once in the crystal (for $N = 22$), and for the optimized models all-trans states are obtained, with chains slightly tilted as a result of the small size of the simulation box.

For $N = 50$, the CG-PE_1 Opt. model still has fully elongated chains, CG-PE_3 Opt. yields one well-ordered lamella surrounded by amorphous chain parts, and the nonoptimized models form more disordered structures. The models having an attractive nonbonded potential prove to be significantly denser at low temperature. The structure of the crystals obtained for the different models can be characterized by the proportion of trans−trans states at low temperature, which is above 90% for $N = 22$ (92% for the nonoptimized models and 98% for the stiffer optimized models) and slightly smaller for $N = 50$ as a result of more disorder in the final configurations.

5. Conclusions

We have constructed coarse-grained models of PE in the same spirit as the CG-PVA model, this means on a "united-monomer" level with two CH_2 groups per CG-bead, and including at least an angular potential beyond the bead-spring interactions. It has been shown that such models yield good results in the description of the melt in the sense that *optimized* coarse-grained models successfully reproduce the structure obtained through the more detailed (all-atom) simulations of $C_{44}H_{90}$. This simplification scheme yields a reasonable agreement with the structural properties of the original model in equilibrium at high temperature (it was optimized to this end). It also allowed us to reproduce qualitatively the features of the crystals that develop at low temperature, whereas this could not be achieved using the all-atom simulations. Furthermore, the refinement of the coarse-grained models via optimization of the angular and nonbonded potentials made it possible to approach the experimental value of the melting point. Even though the agreement with this value does not necessarily mean that the model describes completely the system under consideration, it is still very interesting that with a rather simple method one is able to obtain an effective model which achieves very good qualitative agreement with experiment.

The results of the present study show that the coarse-graining procedure used to address the crystallization of PVA can be applied to the even more generic PE and yield both the same qualitative results for the crystalline state (which is interesting because of the abundant literature about crystallization of PE) and a simplified model that reproduces accurately many static properties of the melt at high temperature.

One of the main reasons for the success of our coarse-graining approach lies in the chosen "scale" at which the mapping is performed, i.e., the position of the mapping centers. The choice to retain every other carbon atom from the original backbone is indeed crucial, since this leads to a shift of the different interactions (the binding potential is related to the atomistic bond angles, the angular potential to the torsions, and the torsional potential has very little influence): Thus the global structure of the model is preserved, while the unimportant degrees of freedom are averaged out without losing the connection to the underlying all-atom description. The angular potential in its structure is reminiscent of the torsional states, and this connection captures the most important features of the chains we are modeling. Simulations of CG-PVA-like models with a structureless angular distribution such as the ones obtained for coarser models (see Figure 4) only led to the glassy state upon continuous cooling since nothing drives the crystallization anymore in this case. To be consistent when deriving a coarse-grained model the way we described it here in the case of more remote mapping centers, one would need to take into account the bond length distribution which is not Gaussian anymore but presents several peaks. However, this is not likely to lead to the formation of an ordered state at low temperature since nothing would then force the chains to stretch.

The process of coarse-graining consists of discarding part of the information available from a system; by doing so one makes the simulation more efficient but also introduces possible inconsistencies in the new models. Even though the high-temperature structure has been shown to match rather precisely the properties of the original all-atom model, the continuous cooling simulations revealed that the expansion coefficient of the CG-PE models do not correspond to the values measured from the atomistic simulations: Figure 17 compares the particle density for both simulation schemes, indicating the mismatch. The value obtained for $C_{44}H_{90}$ (all-atom data yield a result close to 9×10^{-4} K^{-1} between 350 and 450 K) is rather close to the experimental data (6×10^{-4} K^{-1} is reported in the melt[30]), while the coefficient measured for the different CG-PE models is much smaller; this also explains why crystallization is easier to observe using the coarse-grained than the atomistic models. To improve the model, one certainly has to better reproduce the thermodynamics of the model, in particular the compressibility and the thermal expansion coefficient. This was also found to be a prerequisite to successfully reintroduce atomistic details into the simulation.[31]

This work shows that except for extreme cases, the optimization of the coarse-grained potentials is both needed and very efficient (optimization was found unnecessary e.g. for very flat potentials as in the case of torsions for CG-PE or for very sharp potentials such as the angular potential of CG-PVA). Concerning the other choices for the force-field parameters, a compromise between the speed and accuracy of the simulations has to be adopted. The CG-PVA model was derived with the idea of providing a very efficient model for polymer melt simulations, i.e., with short-range potentials, tabulated values with linear interpolation. The torsional potential was discarded because of its weak influence on the structure of the melt. In the case of CG-PE, it was found that taking into account a torsional potential (which has a somewhat more pronounced structure) slightly improves the accuracy of the static properties without influencing notably the performance of the simulation code. This is the reason we focused on the models *with* explicit torsional potential here (CG-PE_1 and CG-PE_3). The form of the nonbonded interaction is a more complicated issue. The choice of a

smaller cutoff distance (CG-PE_1) obviously makes the simulation faster, but since the model with a larger cutoff (CG-PE_3) achieves a higher accuracy in reproducing the structure of the melt it cannot be deemed "less efficient". And even though the (optimized) CG-PE_3 takes roughly twice as much CPU time as CG-PE_1, it remains much faster than the all-atom scheme.

Acknowledgment. The authors are grateful to Jörg Baschnagel for helpful remarks; T.V. acknowledges financial support from the International Research Training Group "Soft Matter", and from the European Science Foundation ('Experimental Theoretical Design of Stimuli-Responsive Polymeric Materials'). Computer time for simulation was granted by IDRIS, Orsay, France.

References

(1) Baschnagel, J.; Binder, K.; Doruker, P.; Gusev, A. A.; Hahn, O.; Kremer, K.; Mattice, W. L.; Müller-Plathe, F.; Murat, M.; Paul, W.; Santos, S.; Suter, U. W.; Tries, V. *Adv. Polymer Sci.* **2000**, *152*, 41−156 review.

(2) Müller-Plathe, F. *ChemPhysChem* **2002**, *3*, 754 review.

(3) Müller-Plathe, F. *Soft Mater.* **2003**, *1*, 1−31 review.

(4) Reith, D.; Meyer, H.; Müller-Plathe, F. *Macromolecules* **2001**, *34*, 2335.

(5) Meyer, H.; Biermann, O.; Faller, R.; Reith, D.; Müller-Plathe, F. *J. Chem. Phys.* **2000**, *113*, 6264−75.

(6) Reith, D.; Pütz, M.; Müller-Plathe, F. *J. Comput. Chem.* **2003**, *24*, 1624−36.

(7) Faller, R.; Reith, D. *Macromolecules* **2003**, *36*, 5406.

(8) Sun, Q.; Faller, R. *Comput. Chem. Eng.* **2005**, *29*, 2380.

(9) Bourasseau, E.; Haboudou, M.; Boutin, A.; Fuchs, A.; Ungerer, P. *J. Chem. Phys.* **2003**, *118*, 3020−34.

(10) Milano, G.; Müller-Plathe, F. *J. Phys. Chem. B* **2005**, *109*, 18609−19.

(11) Ashbaugh, H. S.; Patel, H. A.; Kumar, S. K.; Garde, S. *J. Chem. Phys.* **2005**, *122*, 104908.

(12) Yelash, L.; Müller, M.; Paul, W.; Binder, K. *J. Chem. Theory Comput.* **2006**, *2*, 588−597.

(13) Guerrault, X.; Rousseau, B.; Farago, J. *J. Chem. Phys.* **2004**, *121*, 6538−46.

(14) Padding, J. T.; Briels, W. J. *J. Chem. Phys.* **2002**, *117*, 925−943.

(15) Strobl, G. *Eur. Phys. J. E* **2000**, *3*, 165.

(16) Lotz, B. *Eur. Phys. J. E* **2000**, *3*, 185.

(17) Muthukumar, M. *Adv. Chem. Phys.* **2004**, *128*, 1−63.

(18) Ungar, G.; Zeng, X.-B. *Chem. Rev.* **2001**, 4157.

(19) Meyer, H.; Müller-Plathe, F. *J. Chem. Phys.* **2001**, *115*, 7807.

(20) Meyer, H.; Müller-Plathe, F. *Macromolecules* **2002**, *35*, 1241.

(21) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225.

(22) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. R.; Cheatham, T. E., III; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. *Comput. Phys. Commun.* **1995**, *91*, 1.

(23) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.

(24) Müller-Plathe, F. *Comput. Phys. Comm.* **1993**, *78*, 77.

(25) Yoon, D. Y.; Smith, G. D.; Matsuda, T. *J. Chem. Phys.* **1993**, *98*, 10037.

(26) Dorset, D. L. *Crystallography of the polymethylene chain. An inquiry into the structure of waxes;* Oxford University Press: Oxford, 2004.

(27) Ryckaert, J.-P.; Klein, M. L. *J. Chem. Phys.* **1986**, *85*, 1613.

(28) Meyer, H. Structure Formation and Chain-Folding in Supercooled Polymer Melts. Some Ideas from MD Simulations with a Coarse-Grained Model. In *Polymer Crystallization: Observations, Concepts and Interpretations*; Sommer, J.-U., Reiter, G., Eds.; Springer: Berlin, 2003; Vol. 606.

(29) Vettorel, T. Ph.D. Thesis, University Louis Pasteur, Strasbourg, France, 2005. http://eprints-scd-ulp.u-strasbg.fr: 8080/archive/00000334/.

(30) Van Krevelen, D. W. *Properties of Polymers*; Elsevier: Amsterdam, 1997.

(31) Praprotnik, M.; Delle Site, L.; Kremer, K. *J. Chem. Phys.* **2005**, *123*, 224106.

CT0503264

JCTC Journal of Chemical Theory and Computation

Adsorption and Dynamics of a Single Polyelectrolyte Chain near a Planar Charged Surface: Molecular Dynamics Simulations with Explicit Solvent

Govardhan Reddy,[†] Rakwoo Chang,[‡] and Arun Yethiraj*,[†]

Department of Chemistry, University of Wisconsin, Madison, Wisconsin 53706, and Department of Chemistry, Kwangwoon University, Seoul, 139-701, Republic of Korea

Received October 27, 2005

Abstract: The effect of solvent quality on the behavior of a polyelectrolyte chain near a charged surface is studied using molecular dynamics simulation with explicit solvent. The polyion adsorbs completely on the surface for a high enough surface charge density, and the surface charge required for complete adsorption becomes lower as the solvent quality is decreased. Several static and dynamic properties display a nonmonotonic dependence on surface charge density and solvent quality. For a given value of solvent quality the component of the radius of gyration (R_g) parallel to the surface is a nonmonotonic function of the surface charge density, and for a given surface charge density the component of R_g perpendicular to the surface is a nonmonotonic function of the solvent quality. The center-of-mass diffusion coefficient and rotational relaxation time are nonmonotonic functions of the surface charge density. Translational diffusion coefficient increases, and the rotational relaxation time decreases as solvent quality is decreased for a fixed surface charge density.

I. Introduction

The ease of processing makes polymers useful coating materials. An intriguing case is when the polymer molecules contain charged groups along a hyrophobic backbone[1−3] because the balance between hydrophobic and electrostatic interactions can, in principle, be exploited to tune the adsorption and conformational properties. Charged polymers have a number of applications, for example, in electrooptic devices, semiconducting films, and drug delivery. Our understanding of the behavior of adsorbed polyions is not yet complete, however, and this is a problem of current interest.[4] In this work we study the adsorption of a charged polymer to an oppositely charged surface using molecular dynamics simulations.

There is a large body of theoretical and computational work on the adsorption of charged polymers to surfaces. These include scaling and meanfield theories for dilute[5−17] and semidilute[10,12,13,16−25] solution, and computer simulations using Monte Carlo[26−40] and molecular dynamics[41−44] methods. There has been much less attention, in computer simulations, on the effect of solvent quality on polyion adsorption. Since in many cases, e.g. polystyrene sulfonate, the polyions are composed of hydrophobic backbones for which water is a bad solvent, the effect of solvent quality can be important. In fact, it is often suggested that the solvent plays a crucial role[45−47] in the polyion adsorption on surfaces.

In the theoretical study of polymeric materials some level of coarse-graining in the model is necessary, given the size of the molecules. This is particularly true of charged polymers because the molecules can be significantly stretched. Several levels of coarse-graining are possible. At the united atom level a small group of atoms, e.g., CH_2 in polyethylene, is treated as a single site. This model preserves the carbon−carbon bond angle and bond length characteristics but coarse grains over hydrogen atoms. A more computationally convenient model, which has played an important role in our understanding of polymer solutions and melts, is when several of these united atoms are treated as a single site. This results in a class of models that can be loosely referred

* Corresponding author e-mail: yethiraj@chem.wisc.edu.
[†] University of Wisconsin.
[‡] Kwangwoon University.

10.1021/ct050267u CCC: $33.50
Published on Web 04/21/2006

to as bead-spring models, an example of which is the freely jointed tangent hard sphere model. Information about local chemical details is lost in this class of models, but they provide a computationally convenient way of investigating the effect of parameters, for example the degree of polymerization, on the physical properties.

In the study of polymer solutions, the properties of the solvent molecules are generally not of interest, and it is attractive to be able to coarse-grain out the solvent coordinates. The result is an effective solvent-induced potential between polymer sites. A further approximation is to assume that this potential is pairwise additive. For polymers in poor solvents, i.e., when the polymer−solvent interactions are not favorable, the solvent-induced interaction is attractive in nature and tends to collapse the chain for sufficiently strong solvent-induced interactions. Note that in general this solvation potential should be many-body in nature, as is clear from the McMillan-Mayer theory,[48] and a pairwise additive form can never be exact.

The treatment of the solvent can play an important role in simulations of polymers in poor solvents. When the polymer chain is collapsed, the solvent is excluded from the globule to a large extent. In simulations with explicit solvent, the interactions between the interior polymer sites is just a weak van der Waals attraction between the hydrocarbon sites. With a pairwise additive implicit solvent, however, there is a very strong attraction between these interior sites which artificially increases the stability of the globule. The physical origin of this effect is that a potential, which is appropriate for two isolated monomers in the solvent, has been used under conditions where there is no solvent present. The consequence is that simulations of polymers in poor solvents often get trapped in meta-stable states unlike corresponding simulations with explicit solvent. In fact, the collapse dynamics of a homopolymer[49] and the phase behavior of polyelectrolyte solutions[50] are *qualitatively* different in pairwise additive implicit solvent simulations, when compared to explicit solvent simulations.

In this work we study the role of the solvent on the properties of a single polyion adsorbed on a planar surface using computer simulations. We include the solvent explicitly and perform molecular dynamics simulations on a system composed of a polyion, counterions, co-ions (counterions to the surface), and solvent particles, confined between two surfaces. The polyion is composed of freely jointed hard spheres with a negative charge on each sphere and the surface as atomically smooth with a uniform positive charge spread. The counterions to the polyion and the co-ions (ions which balance the charge on the surface) are monovalent hard spheres with positive and negative charges, respectively.

We find that the polyion adsorbs completely (flat) onto the surface for sufficiently high surface charge densities or poor enough solvents. Several static and dynamic properties have a nonmonotonic dependence of surface charge density and solvent quality. For example, the component of the radius of gyration parallel to the surface, the center-of-mass diffusion coefficient, and the rotational relaxation time show a nonmonotonic dependence on the surface charge density (for a given quality of solvent), while the component of the radius of gyration perpendicular to the surface shows a nonmonotonic dependence on solvent quality. In all cases the chain center-of-mass exhibits normal diffusive behavior, but in good solvents the end-to-end vector autocorrelation function does not decay over the course of the simulation. Interestingly the chain dynamics become faster as the solvent quality is decreased.

The rest of the paper is organized as follows. The molecular model and simulation details are presented in section II, results are presented and discussed in section III, and some conclusions are presented in section IV.

II. Molecular Dynamics Simulations

A. Molecular Model. The molecular model is similar to the model used by Chang and Yethiraj[50] in their study of polyelectrolytes in poor solvents. The polyion is modeled as a flexible bead spring chain composed of monomers with one negative charge each, and the counterions to the polyion are monovalent positively charged spheres. The charged surface is modeled as atomically smooth with a uniform positive surface charge density. The co-ions (which balance the surface charge) are monovalent negatively charged spheres. Solvent molecules are incorporated explicitly and are modeled as uncharged spheres.

The potential of interaction $V_{ij}(r)$ between any two particles i and j is given by the sum of an electrostatic and nonelectrostatic part, i.e.,

$$V_{ij}(r) = k_B T \frac{q_i q_j \lambda_B}{r} + \lambda_{ij} V_{LJ}(r) + (1 - \lambda_{ij}) V_{WCA}(r) \quad (1)$$

where k_B is the Boltzmann's constant, T is the temperature, q_i is the charge valence of site i, $\lambda_B \equiv e^2/\epsilon k_B T$ is the Bjerrum length, e is the electronic charge, ϵ is the dielectric constant of the solvent, and the parameters λ_{ij} control the quality of the solvent. The potentials $V_{LJ}(r)$ and $V_{WCA}(r)$ are the full and purely repulsive Lennard-Jones (LJ) potentials, respectively, and are given by

$$V_{LJ}(r) = \begin{cases} 4\epsilon_{LJ}\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} - c(2.5\sigma)\right] & r \le 2.5\sigma, \\ 0 & r > 2.5\sigma, \end{cases} \quad (2)$$

and

$$V_{WCA}(r) = \begin{cases} 4\epsilon_{LJ}\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} - c(2^{1/6}\sigma)\right] & r \le 2^{1/6}\sigma, \\ 0 & r > 2^{1/6}\sigma, \end{cases} \quad (3)$$

where ϵ_{LJ} is the attractive well depth, and the function $c(r)$ is chosen such that the value of the potential is zero at the cutoff, i.e., $c(r) = (\sigma/r)^{12} - (\sigma/r)^6$.

The λ_{ij} parameters determine the quality of the solvent. In this work, we set $\lambda_{mm} = \lambda_{ss} = \lambda$, where m denotes monomer and s denotes solvent, and $\lambda_{ij} = 0$ for all other pairs of i and j. The nonelectrostatic contribution to the potential to all the pairs is purely repulsive except for the monomer−monomer and solvent−solvent interactions. The parameter λ introduces an attractive well depth between monomer−monomer and solvent−solvent interactions as it is increased from 0 to 1. When $\lambda = 0$, all nonelectrostatic interactions are identical, and therefore the solvent quality

is good. On the other hand, when λ is large and positive the attractions between monomers and between the solvent molecules makes the solvent quality poor.

The nonelectrostatic interaction between the particles and the surface is purely repulsive and is given by[51]

$$V_{\text{wall}} = \begin{cases} 2\pi\epsilon_{\text{w}}\sigma_{\text{w}}^2\left[\frac{2}{5}\left(\frac{\sigma_{\text{w}}}{z}\right)^{10} - \left(\frac{\sigma_{\text{w}}}{z}\right)^4 + \frac{3}{5}\right] & z \le \sigma_{\text{w}}, \\ 0 & z > \sigma_{\text{w}}, \end{cases} \quad (4)$$

where z is the perpendicular distance from the particle to the solid surface and the subscript w denotes the surface. This potential is obtained from the integration of the 12−6 LJ potential of a planar solid continuum in the two directions parallel to the surface. We set $\epsilon_{\text{w}} = \epsilon_{\text{LJ}}$ and $\sigma_{\text{w}} = \sigma$ for all the particles present in the simulation box. Both the surfaces interact with the particles in the simulation cell using the soft core repulsive LJ potential given by eq 4, but only one of the surfaces is electrically charged. The electric field, E_j, acting on a charged particle j due to the uniform charge density, σ_{SCD}, on a nonconducting surface is given[52] by $E_j = (\sigma_{\text{SCD}}/2\epsilon)$, and the direction of E_j is perpendicular to the surface.

The bonding potential between neighboring beads in the polyion is given by the FENE (finitely extensible nonlinear elastic) potential[53]

$$V_{\text{FENE}}(r) = -\frac{1}{2}k_{\text{FENE}}R_{\text{o}}^2\ln\left(1 - \frac{r^2}{R_{\text{o}}^2}\right) \quad (5)$$

where k_{FENE} is the spring constant and R_{o} is the maximum extension of the bond. In this work, we set $k_{\text{FENE}} = 30.0k_{\text{B}}T/\sigma^2$ and $R_{\text{o}} = 1.5\sigma$; these parameters prevent[53] the crossing of the chains.

Some limitations of the model merit discussion. The effect of the solvent on the electrostatic interactions is taken into account implicitly through the dielectric constant ϵ. The change in local dielectric properties as well as polarization effects are therefore neglected. The interaction of the surface with the solvent and the polyion is purely electrostatic and there is no short-range attraction, which implies that the interfacial tensions between the surface and solvent and between the surface and polyion beads are similar. The surface in the simulations is assumed to be a nonconducting surface with a dielectric constant identical to the solution. In general, surfaces consist of low dielectric constant materials, and polar solvents near such surfaces give rise to like image charges. These effects are not taken into account in this work.

B. Simulation Method. The simulation cell is a cubic box of side length $L = 16\sigma$ with periodic boundary conditions in the x and y directions and surfaces in the z direction. The system consists of a single polymer chain with $N = 16$ sites, $N_{\text{ctr}} = 16$ counterions to balance the charge on the polyion, and N_{coi} co-ions to balance the surface charge, with $N_{\text{coi}} = 16 \times 16\sigma^2\sigma_{\text{SCD}}$. The monomer density $\rho_{\text{m}}\sigma^3 \equiv N\sigma^3/L^3$ is approximately 0.004, and the total site density in the simulation cell, $\rho_{\text{tot}}\sigma^3 \equiv (N + N_{\text{ctr}} + N_{\text{coi}} + N_{\text{s}})\sigma^3/L^3 = 0.864$. The polyion, counterions, and co-ions are added first, and then the required number of N_{s} solvent molecules are added so that $\rho_{\text{tot}}\sigma^3 = 0.864$. Lengths are measured in units of σ, time in units of $\tau_{\text{MD}} = (m\sigma^2/\epsilon_{\text{LJ}})^{1/2}$, and temperature in units of $\epsilon_{\text{LJ}}/k_{\text{B}}$. In this calculation we set all three parameters m, σ, and ϵ_{LJ} to unity, and a reduced temperature is defined as $T^* \equiv k_{\text{B}}T/\epsilon_{\text{LJ}}$ which is set to 1.

Initial configurations are generated with the atoms on the lattice points of a face centered cubic structure with N_{m} adjacent vertices chosen for the polyion. Initial velocities are generated using a Gaussian random number generator and scaled to the desired temperature. The system is propagated in the canonical ensemble (NVT constant) using an explicit reversible integrator,[54] and the temperature is maintained constant using a Nose-Hoover thermostat.[55,56] An integration time step of $0.005\tau_{\text{MD}}$ is used, and the Nose-Hoover coupling constant is set to 5. First the system is equilibrated without the charge on the surface, and then the surface charge is switched on and the required number of solvent molecules are converted to co-ions to balance the charge on the surface. The system is then re-equilibrated before properties are averaged. The data are obtained by averaging over 4−6 trajectories, and each trajectory is run for approximately 1 million time steps.

III. Results and Discussion

A. Static Properties. The adsorption behavior of the chain demonstrates an interesting interplay between solvent quality and surface charge density. As expected, increasing the surface charge density promotes adsorption of the polymer chain. Decreasing the solvent quality also promotes adsorption but, in addition, promotes chain collapse. These effects can be quantified via the components of the mean-square radius of gyration parallel and perpendicular to the surface, denoted $R_{\text{g,para}}^2$ and $R_{\text{g,perp}}^2$, respectively, and the molecular axis orientational correlation function G_2. These quantities are defined as

$$R_{\text{g,para}}^2 = \frac{1}{N}\left\langle\sum_{i=1}^{N}(x_i - x_{\text{cm}})^2 + (y_i - y_{\text{cm}})^2\right\rangle \quad (6)$$

$$R_{\text{g,perp}}^2 = \frac{1}{N}\left\langle\sum_{i=1}^{N}(z_i - z_{\text{cm}})^2\right\rangle \quad (7)$$

where x_i, y_i, and z_i are the Cartesian coordinates of the monomer i, and the subscript cm stands for the center of mass of the polyion chain, and

$$G_2 = (3\langle\cos^2\gamma\rangle - 1)/2 \quad (8)$$

where γ is the angle between the molecular axis and the surface. The molecular axis is defined[57] as the eigenvector corresponding to the smallest eigenvalue of the moment of inertia tensor of the molecule. If the molecular axis aligns parallel to the surface, then $G_2 = -0.5$; if it is perpendicular to the surface, then $G_2 = 1.0$; and if it is isotropic, then $G_2 = 0$.

For a given solvent quality, increasing the surface charge density causes the chain to adsorb flat on the charged surface, and this is manifested in a decrease in G_2 and $R_{\text{g,perp}}^2$. Figure 1 depicts G_2 as a function of $\sigma^*_{\text{SCD}}(= \sigma_{\text{SCD}}\sigma^2/e)$ for $\lambda = 0$

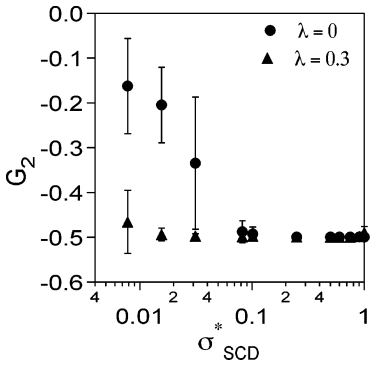


Figure 1. Orientation of the polyion plotted as a function of surface charge density, σ^*_{SCD} for $\lambda = 0.0$ and 0.3.

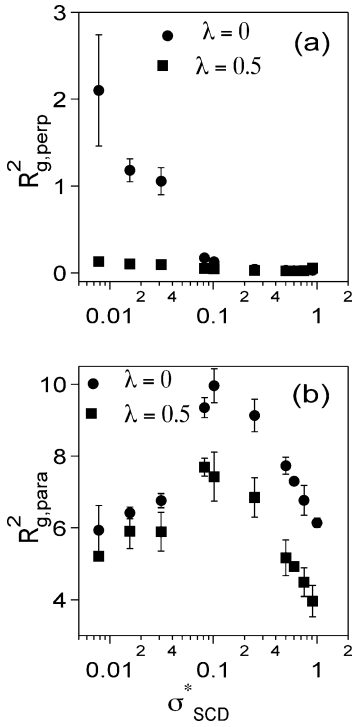


Figure 2. Averaged mean square radius of gyration of the polyion perpendicular to the surface, $R^2_{g,perp}$, and parallel to the surface, $R^2_{g,para}$, as a function of surface charge density, σ^*_{SCD}, for $\lambda = 0$ and 0.5.

and 0.3. In good solvents ($\lambda = 0$) and for low values of σ^*_{SCD}, the polyion adopts an almost isotropic orientation near the surface. although it is still slightly biased parallel to the surface, i.e., $G_2 < 0$. As σ^*_{SCD} is increased, the polyion adsorbs to the surface and for $\sigma^*_{SCD} \approx 0.1$ the orientation is completely parallel to the surface ($G_2 \approx -0.5$). When the solvent quality is slightly poor ($\lambda = 0.3$) the polyion is adsorbed flat on the surface even for low values of σ^*_{SCD} because of the unfavorable monomer−solvent interaction. This behavior is consistent with recent experiments that investigated the size of an adsorbed polyelectrolyte layer exposed to solvents of varying quality[46] where the size of the adsorbed polyelectrolyte layer decreased as the solvent quality was decreased.

Figure 2 depicts $R^2_{g,perp}$ and $R^2_{g,para}$ as a function of σ^*_{SCD} for $\lambda = 0$ and 0.5. The variation of $R^2_{g,perp}$ as a function of σ^*_{SCD} reflects that of G_2 discussed earlier, i.e., as the surface charge density is increased the chain adsorbs flat onto the surface resulting in a sharp decrease in $R^2_{g,perp}$. Note that for $\lambda = 0.5$, the chain is strongly adsorbed on the surface even for σ^*_{SCD} close to zero. On the other hand, for a given value of λ, $R^2_{g,para}$ displays a nonmonotonic dependence on σ^*_{SCD}, i.e., $R^2_{g,para}$ increases until the chain is completely flat on the surface, which occurs for $\sigma^*_{SCD} \approx 0.1$, and then begins to decrease. The increase in $R^2_{g,para}$ with σ^*_{SCD} can be explained by noting that the polyion is being confined from three dimensions to two dimensions. The decrease in $R^2_{g,para}$ for large values of σ^*_{SCD} can be explained in terms of the effect of co-ions (counterions to the surface). As σ^*_{SCD} increases the number of co-ions in the system also increases, and these are found preferentially near the charged surface. This can be seen in snapshots of configurations or density profiles (not shown). Since the co-ions and the polyion are similarly charged, the repulsion from the co-ions results in a decrease in the size of the polyion.

These results are different from previous simulations for polyelectrolytes in good solvents. Lattice simulations[26] found that $R^2_{g,para}$ was relatively insensitive to σ^*_{SCD} although the behavior of $R^2_{g,perp}$ was similar to what is seen in this work. In the off-lattice Monte Carlo simulations of Kong and Muthukumar[27] and in the brownian dynamics simulations by Panwar and Kumar[43] $R^2_{g,para}$ increased monotonically and $R^2_{g,perp}$ decreased monotonically as σ^*_{SCD} was increased. These differences can be attributed to the neglect of co-ions in previous simulations. If the effect of poor solvent and co-ions is not incorporated, then the adsorption behavior of a polyelectrolyte chain is identical[58,59] to the behavior of neutral polymer chains adsorbed on a surface (as a function of adsorption energy).

For some values of the surface charge density, $R^2_{g,perp}$ is a nonmonotonic function of solvent quality. Figure 3 depicts $R^2_{g,perp}$ and $R^2_{g,para}$ as a function of λ for $\sigma^*_{SCD} = 0.1$ and 0.6. $R^2_{g,para}$ is a monotonically decreasing function of λ in all cases which is expected because the chain becomes more compact as the solvent quality is decreased. For $\sigma^*_{SCD} = 0.1$, $R^2_{g,perp}$ is a nonmonotonic function of σ^*_{SCD}. The decrease in $R^2_{g,perp}$ occurs because the polyion collapses onto the surface to avoid the unfavorable monomer−solvent interactions. As λ is increased further, however, the attraction between the monomers overcomes the attraction between the surface and the monomers, and this causes the polyion to form a globule instead of a pancake. This does not happen when the monomer−surface interactions are very strong, which is the case for $\sigma^*_{SCD} = 0.6$. These simulations confirm the assumptions of the scaling theory of Borisov[15] et al. where they assumed that the polyion in poor solvent is adsorbed as a pancake structure when σ^*_{SCD} is high but adopts a hemispherical shape when σ^*_{SCD} is low. The results from the scaling theory of Borisov[8] et al. for the behavior of single chain near a charged surface cannot be confirmed from these simulations because of the small size of the polyion chain.

B. Dynamic Properties. The rotational and translational dynamics of the polyion become *faster* as the solvent quality is decreased, for a given value of the surface charge density. Figure 4(a),(b) depicts the mean-square displacement parallel

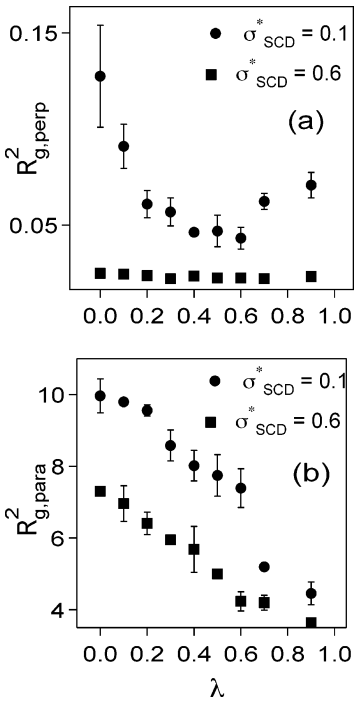


Figure 3. Averaged mean square radius of gyration of the polyion perpendicular to the surface, $R^2_{g,perp}$, and parallel to the surface, $R^2_{g,para}$, plotted as a function of solvent quality, λ for $\sigma^*_{SCD} = 0.1$ and 0.6.

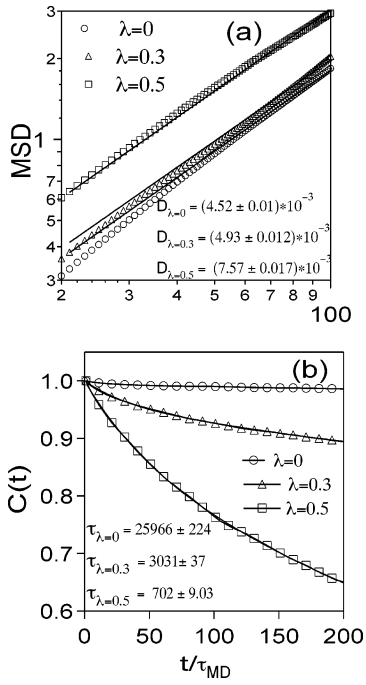


Figure 4. (a) Log−log plot of the lateral mean-square displacement of the center of mass of the polyion versus time for $\sigma^*_{SCD} = 1.5$. Solid lines are fit to the data. (b) End-to-end time autocorrelation function plotted versus time for $\sigma^*_{SCD} = 1.5$. Solid lines are fit to the data, and it overlaps with the line of the markers.

to the surface (on a log−log plot) and end-to-end vector correlation function as a function of time for $\sigma^*_{SCD} = 1.5$ and $\lambda = 0$, 0.3, and 0.5. MSD is the mean-square displacement of the center of mass of the adsorbed polyion chain parallel to the surface, i.e., MSD $= \langle|\boldsymbol{R}_{||}(t) - \boldsymbol{R}_{||}(0)|^2\rangle$, and the lateral diffusion coefficient, $D_{||}$, is defined via $\langle|\boldsymbol{R}_{||}(t) - \boldsymbol{R}_{||}(0)|^2\rangle = 4D_{||}t$, where $\boldsymbol{R}_{||}(t) = x_{cm}(t)\hat{i} + y_{cm}(t)\hat{j}$. The end-to-end vector autocorrelation function, $C(t)$ is defined as $C(t) = \langle\boldsymbol{R}(0).\boldsymbol{R}(t)\rangle/\langle\boldsymbol{R}^2(0)\rangle$ where $\boldsymbol{R}(t)$ is the end-to-end vector of the chain at time t. In all cases the center of mass of the polyion shows subdiffusive behavior at shorter time scales and normal lateral diffusion at a longer time scale, i.e., MSD is a linear function of time. The crossover time from subdiffusive to normal diffusion decreases as the solvent quality is decreased. And as the solvent quality is decreased, the translational diffusion is faster, i.e., $D_{||}$ *increases*, taking on values (in units of σ^2/τ_{MD}) of 4.52×10^{-3}, 4.93×10^{-3}, and 7.57×10^{-3} (statistical uncertainties are in the last significant digit) for $\lambda = 0$, 0.3, and 0.5. The faster translational diffusion is accompanied by faster rotation. For $\lambda = 0$, $C(t)$ does not decay over the course of the simulation, but the decay becomes faster as the solvent quality is decreased.

We characterize the rotational relaxation via the initial (short time) decay of $C(t)$. At short times, i.e., $t \leq 200\tau_{MD}$, the decay of $C(t)$ is well-fit by a single exponential, and we define a relaxation time τ as the time constant of this exponential. We are unable to reliably extract the long time relaxation behavior because $C(t)$ does not decay in some cases and in others the statistics are poor even with trajectories of length $5000\tau_{MD}$. The rotational relaxation time τ for $\sigma^*_{SCD} = 1.5$ shown in Figure 4(b) takes on values of (in units of τ_{MD}) 25 966 ± 224, 3031 ± 37, and 702 ± 9 for $\lambda = 0$, 0.3, and 0.5, respectively.

An interesting feature from Figure 4 is that although $C(t)$ does not decay for some values of λ, the chain exhibits normal center-of-mass diffusion. This decoupling of rotational and translational dynamics is often viewed as a signature of "glassy" behavior, since it is reminiscent of what is observed in the dynamics of supercooled liquids. It has been suggested[60−63] that the strong interactions between a polymer and the surface can result in the polyion becoming trapped in local minima. In the present work the behavior exhibited by the polyion is not due to the interaction of the polyion with the surface but rather because of the interaction between the polymer and the solvent.

There are several possible explanations for the faster dynamics as the solvent quality is decreased. First of all, as the solvent is made poorer, the hydrodynamic radius decreases, and this can result in an increase in the diffusion coefficient. Another possibility is that this is caused by the confinement of the chain from three dimensions to two dimensions, and the dynamics of two-dimensional chains are expected to be faster than three-dimensional chains. Finally, as the solvent quality is decreased the density of the solvent molecules near the surface decreases, and this results in a lower friction on the adsorbed chain.

Figure 5(a),(b) depicts the total density of sites and the solvent density as a function of distance from the charged surface. As the solvent quality is decreased the solvent density near the surface decreases. This is not very prominent in Figure 5(a) but is more clearly seen in Figure 5(b) where the value of the peak near $z = 2$ decreases by about 75%

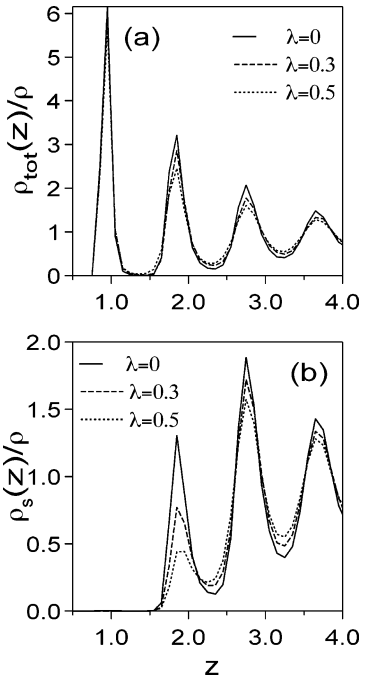


Figure 5. (a) Total density profile of the species as a function of the distance from the surface for $\sigma^*_{SCD} = 1.5$. (b) Density profile of the solvent as a function of the distance from the surface for $\sigma^*_{SCD} = 1.5$.

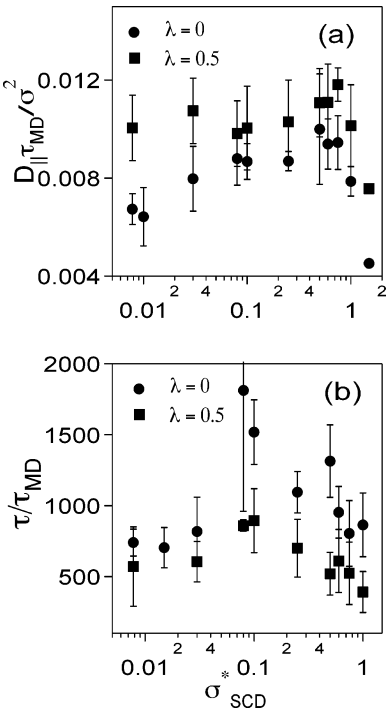


Figure 6. (a) $D_{||}$ plotted as a function of σ^*_{SCD} for $\lambda = 0.0$ and 0.5. (b) τ plotted as a function of σ^*_{SCD} for $\lambda = 0.0$ and 0.5.

for λ going from 0 to 0.5. It is possible that the decrease in site density in the layer adjacent to the polyion results in a weaker friction and therefore faster dynamics of the adsorbed polyion.

The diffusion coefficient and end-to-end vector relaxation time are nonmonotonic functions of the solvent quality in some cases. Figure 6(a),(b) depicts $D_{||}$ and τ, respectively, as a function of σ^*_{SCD} for $\lambda = 0$ and 0.5. For $\lambda = 0$, $D_{||}$ is a nonmonotonic function of σ^*_{SCD} with a peak that occurs at a value of σ^*_{SCD} where the chain is adsorbed flat to the surface. One can attribute the increase in $D_{||}$ with increasing σ^*_{SCD} for lower surface charge densities to a confinement of the chain from three to two dimensions and the decrease in $D_{||}$ with increasing σ^*_{SCD} for higher surface charge densities to the strong wall-polymer interaction which increases the friction with the surface. For a given σ^*_{SCD}, $D_{||}$ is always higher in poorer solvents. This is because the chain size is smaller and the solvent density at the surface is lower as the solvent quality is decreased, both of which tend to increase $D_{||}$. In poorer solvents, however, $D_{||}$ is relatively insensitive to the surface charge density (within statistical uncertainties) because the polyion is always adsorbed completely on the surface for all surface charge densities. For both $\lambda = 0$ and 0.5 τ is a nonmonotonic function of σ^*_{SCD} with a peak at approximately $\sigma^*_{SCD} \approx 0.1$ which is when the polyion is completely adsorbed to the surface. Again the effect of changing the surface charge is more significant in good solvents than in poor solvents. One can attribute the increase in τ for low values of σ^*_{SCD} to a confinement of the chain from three to two dimensions. The decrease in τ for σ^*_{SCD} greater than 0.1 can be due to the decrease in size of polyion which will facilitate the rotation of the polyion on the surface leading to a smaller τ.

IV. Conclusions

The static and dynamic properties of a single charged chain near a planar charged surface are studied using molecular dynamics simulation with explicit solvent. The properties of an adsorbed chain can be very different depending on the solvent quality. The chain is adsorbed parallel to the surface when the solvent is poor enough, independent of the surface charge, unlike in good solvents where the surface charge plays a dominant role. Interestingly the dynamics of the chain become faster as the solvent quality is decreased, which could be either due to the lower solvent density at the surface or due to the confinement of the chain to two dimensions.

References

(1) Forster, S.; Schmidt, M. *Adv. Polym. Sci.* **1995**, *13*, 51.

(2) Holm, C.; Joanny, J.; Kremer, K.; Netz, R.; Reineker, P.; Seidel, C.; Vilgis, T.; Winkler, R. *Adv. Polym. Sci.* **2004**, *166*, 67.

(3) Barrat, J.; Joanny, J. *Adv. Chem. Phys.* **1996**, *94*, 1.

(4) Netz, R.; Andelman, D. *Phys. Rep.* **2003**, *380*, 1.

(5) Wiegel, F. *J. Phys. A: Math. Gen.* **1977**, *10*, 299.

(6) Muthukumar, M. *J. Chem. Phys.* **1987**, *86*, 7230.

(7) Varoqui, R. J. *Phys. II Fr.* **1993**, *3*, 1097.

(8) Borisov, O.; Zhulina, E.; Birshtein, T. *J. Phys. II Fr.* **1994**, *4*, 913.

(9) Chatellier, X.; Senden, T.; Joanny, J.; di Meglio, J. *Europhys. Lett.* **1998**, *41*, 303.

(10) Netz, R.; Joanny, J. *Macromolecules* **1999**, *32*, 9013.

(11) Joanny, J. *Euro. Phys. J. B* **1999**, *9*, 117.

(12) Dobrynin, A.; Deshkovski, A.; Rubinstein, M. *Phys. Rev. Lett.* **2000**, *84*, 3101.

(13) Dobrynin, A.; Deshkovski, A.; Rubinstein, M. *Macromolecules* **2001**, *34*, 3421.

(14) Dobrynin, A. *J. Chem. Phys.* **2001**, *114*, 8145.

(15) Borisov, O.; Hakem, F.; Vilgis, T.; Joanny, J.; Johner, A. *Euro. Phys. J. E* **2001**, *6*, 37.

(16) Dobrynin, A.; Rubinstein, M. *Macromolecules* **2002**, *35*, 2754.

(17) Dobrynin, A.; Rubinstein, M. *J. Phys. Chem. B* **2003**, *107*, 8260.

(18) Linse, P. *Macromolecules* **1996**, *29*, 326.

(19) Chatellier, X.; Joanny, J. *J. Phys. II Fr.* **1996**, *6*, 1669.

(20) Borukhov, I.; Andelman, D. *Macromolecules* **1998**, *31*, 1665.

(21) Park, S.; Barrett, C.; Rubner, M.; Mayes, A. *Macromolecules* **2001**, *34*, 3384.

(22) Shafir, A.; Andelman, D.; Netz, R. *J. Chem. Phys.* **2003**, *119*, 2355.

(23) Cheng, H.; Olvera de la Cruz, M. *J. Chem. Phys.* **2003**, *119*, 12635.

(24) Patra, C.; Chang, R.; Yethiraj, A. *J. Phys. Chem. B* **2004**, *108*, 9126.

(25) Shafir, A.; Andelman, D. *Phys. Rev. E* **2004**, *70*, 061804.

(26) Beltran, S.; Hooper, H.; Blanch, H.; Prausnitz, J. *Macromolecules* **1991**, *24*, 3178.

(27) Kong, C.; Muthukumar, M. *J. Chem. Phys.* **1998**, *109*, 1522.

(28) Ellis, M.; Kong, C.; Muthukumar, M. *J. Chem. Phys.* **2000**, *112*, 8723.

(29) McNamara, J.; Kong, C.; Muthukumar, M. *J. Chem. Phys.* **2002**, *117*, 5354.

(30) Yamakov, V.; Milchev, A.; Borisov, O.; Dunweg, B. *J. Phys. Condens. Matter* **1999**, *11*, 9907.

(31) Messina, R. *Phys. Rev. E* **2004**, *70*, 051802.

(32) Chodanowski, P.; Stoll, S. *J. Chem. Phys.* **2001**, *115*, 4951.

(33) Chodanowski, P.; Stoll, S. *Macromolecules* **2001**, *34*, 2320.

(34) Stoll, S.; Chodanowski, P. *Macromolecules* **2002**, *35*, 9556.

(35) Messina, R.; Holm, C.; Kremer, K. *Langmuir* **2003**, *19*, 4473.

(36) Messina, R.; Holm, C.; Kremer, K. *Phys. Rev. E* **2002**, *65*, 041805.

(37) Messina, R.; Holm, C.; Kremer, K. *J. Chem. Phys.* **2002**, *117*, 2947.

(38) Messina, R. *J. Chem. Phys.* **2003**, *119*, 8133.

(39) Messina, R.; Holm, C.; Kremer, K. *J. Polym. Sci. Part B: Polym. Phys.* **2004**, *42*, 3557.

(40) Messina, R. *Macromolecules* **2004**, *37*, 621.

(41) Panchagnula, V.; Jeon, J.; Dobrynin, A. *Phys. Rev. Lett.* **2004**, *93*, 037801.

(42) Panchagnula, V.; Jeon, J.; Rusling, J.; Dobrynin, A. *Langmuir* **2005**, *21*, 1118.

(43) Panwar, A.; Kumar, S. *J. Chem. Phys.* **2005**, *122*, 154902.

(44) Wang, X. L.; Lu, Z. Y.; Li, Z. S.; Sun, C. C. *J. Phys. Chem. B* **2005**, *109*, 17644.

(45) Dubas, S. T.; Schlenoff, J. B. *Macromolecules* **1999**, *32*, 8153.

(46) Poptoshev, E.; Schoeler, B.; Caruso, F. *Langmuir* **2004**, *20*, 829.

(47) Kotov, N. *Nanostruct. Mater.* **1999**, *12*, 789.

(48) Hill, T. L. *An Introduction to Statistical Thermodynamics*; Dover: New York, 1960.

(49) Chang, R.; Yethiraj, A. *J. Chem. Phys.* **2001**, *114*, 7688.

(50) Chang, R.; Yethiraj, A. *J. Chem. Phys.* **2003**, *118*, 6634.

(51) Bitsanis, I.; Hadziioannou, G. *J. Chem. Phys.* **1990**, *92*, 3827.

(52) Feynman, R. P.; Leighton, R. B.; Sands, M. *The Feynman Lectures on Physics*; Addison-Wesley: Reading, 1964; Vol. 2.

(53) Grest, G.; Kremer, K. *Phys. Rev. A* **1986**, *33*, 3628.

(54) Tuckerman, M.; Berne, B. J.; Martyna, G. J. *J. Chem. Phys.* **1992**, *97*, 1990.

(55) Nose, S. *J. Chem. Phys.* **1984**, *81*, 511.

(56) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695.

(57) Yethiraj, A. *Adv. Chem. Phys.* **2002**, *121*, 89.

(58) Milchev, A.; Binder, K. *Macromolecules* **1996**, *29*, 343.

(59) Lai, P.-Y. *Phys. Rev. E* **1994**, *49*, 5420.

(60) Mansfield, K. F.; Theodorou, D. N. *Macromolecules* **1989**, *22*, 3143.

(61) Chakraborty, A. K.; Sha_er, J. S. Adriani, P. M. *Macromolecules* **1991**, *24*, 5226.

(62) Chakraborty, A. K.; Adriani, P. M. *Macromolecules* **1992**, *25*, 2470.

(63) Smith, G. D.; Bedrov, D.; Borodin, O. *Phys. Rev. Lett.* **2004**, *70*, 051802.

CT050267U

JCTC Journal of Chemical Theory and Computation

Multiscale Coarse-Graining of Mixed Phospholipid/ Cholesterol Bilayers

Sergei Izvekov and Gregory A. Voth*

Department of Chemistry and Center for Biophysical Modeling and Simulation, University of Utah, 315 South 1400 East Room 2020, Salt Lake City, Utah 84112-0850

Received November 30, 2005

Abstract: Coarse-grained (CG) models for mixed dimyristoylphosphatidylcholine (DMPC)/ cholesterol lipid bilayers are constructed using the recently developed multiscale coarse-graining (MS-CG) method. The MS-CG method permits a systematic fit of the bonded and nonbonded interactions and system pressure to trajectory and force data derived from an underlying reference all-atom molecular dynamics (MD) simulation. The CG sites for lipid and cholesterol molecules are associated with the centers-of-mass of atomic groups because of the simplicity in the evaluation of the forces acting on them from the atomistic MD data. Corresponding models with four-site and seven-site representations of the cholesterol molecule were also developed. The latter CG models differed by the bonding scheme of CG sites to represent intramolecular interactions. A one-site MS-CG model based on the TIP3P potential was used for water, with the interaction site placed at the molecular geometrical center, and the analytical fit of the model is presented. The MS-CG models were then used to conduct simulations in the constant NPT ensemble which reproduce accurately the structural properties as seen in the full all-atom MD simulation.

I. Introduction

Cell membranes are remarkably adaptive and durable molecule aggregates which are directly involved in a large variety of cecullar processes.[1,2] Due to the complexity of such membranes, most information on realistic biological membrane properties comes from experiment when available. However, with the continual increases in computer power, atomic-scale computer simulations of increasingly realistic membranes based on empirical force fields can provide valuable insight in the membrane structure and dynamics.[3−13]

At the present time, however, the all-atom molecular dynamics (MD) simulations of the membranes are limited to systems having several hundred lipid molecules on a time scale of about 100 ns. The extension of computer simulation of biological membrane systems to more biologically relevant time scales and system sizes requires a simpler treatment of the biological molecules of interest and the surrounding solvent molecules. The coarse-graining method is seen as one possible way to accomplish this task.[14−22] Numerous methods for coarse-graining and coarse-grained (CG) models have been reported in the past (for a recent review see, e.g., ref 23). In the framework of united atom-like CG models, which are next to all-atom models in length scale, small groups of atoms are represented as single interaction sites, thus preserving the molecular frame but with less resolution.

The most difficult stage in the implementation of such approaches is the construction of an effective force field to describe the interactions between the CG structural units. Most existing approaches for parametrization of effective CG potentials are not directly based on the underlying atomistic-scale forces which are available from, for example, all-atom empirical potential MD simulations. These approaches instead typically target the reproduction of the average structural properties seen in atomistic simulations, for example using an iterative adjustment of potential parameters starting from an approximation based on potentials of mean force[14,24] or the Inverse Monte Carlo technique.[25] They may also be parametrized to match thermodynamic properties.[14−16,21,22] These latter approaches can be computationally expensive

10.1021/ct050300c CCC: $33.50
Published on Web 02/18/2006

and may become rapidly less useful with an increase in the number of different types of interaction centers in the system.

The coarse-graining of potentials is further complicated due to the fact that the effective interaction between structural units is expected to be more dependent on the thermodynamic state compared to atomistic force fields. This is because the effective interaction between the CG structural units is defined by the average structure (e.g., average orientations, distances) within the complexes formed by those units in a particular phase. Rigorously speaking, the best CG potential should be a potential of mean force (PMF), which is a configurational *free energy* in a reduced phase space. Therefore, the structural properties, and thus the CG potentials, can in principle be sensitive to variations in temperature and other thermodynamic conditions due to both entropic and (anharmonic) energetic effects. A poor transferability of "coarser-grained" potentials may undermine the reliability of simulations in which the same CG potential is used to simulate different system conditions. Intuitively, the construction of a more transferable CG model may be improved by using a finer scale coarse-graining as has been observed in the modeling of proteins.[26,27] Of course, this comes at an increased computational cost. Another solution is to fit (or refit) CG potentials of a system under the same thermodynamic conditions for which the CG simulation is intended. This approach requires the availability of computationally efficient methods for the development of CG force fields as discussed below.

Recently, a significantly different approach for developing CG potentials from an underlying explicit atomistic molecular dynamics (MD) simulation has been presented.[28,29] This approach is called "multiscale coarse-graining" (MS-CG) because it provides a systematic way for coarse-graining the interactions from a reference all-atom MD simulation. This methodology is an extension of the force-matching (FM) approach developed in ref 30. The FM methodology enables one to obtain effective pairwise interactions, and a simpler variant of it was also successfully applied to construct all-atom models for several liquids from ab initio MD simulation data.[30,31]

The MS-CG method implements a FM procedure for the CG images from the reference atomistic trajectory/force data. It is natural to map atomic groups into a CG site through association of the CG site with the centers-of-mass (CM) because the force acting on the CM of the atomic groups can be straightforwardly evaluated from the atomistic MD data. The FM procedure, when applied to these data, should yield an approximation to the effective interactions between the CG sites. These interactions are a variational least-squares approximation to the average atomistic forces.[28,29]

The MS-CG method was first successfully applied to construct a CG model of the pure dimyristoylphosphatidylcholine (DMPC) lipid bilayer.[28] This model was able to accurately reproduce the structural properties of the system in constant NVT all-atom MD simulations. The FM method can also be extended to include a fit of the pressure in the all-atom simulation, with the resulting force field then being appropriate for MS-CG simulations under constant NPT conditions.[29]

In the present article, we report the application of the MS-CG method to a mixed DMPC/cholesterol lipid bilayer. Several MS-CG models are developed here that are different in the level of coarse-graining of the cholesterol sterol molecule and the underlying bonding scheme. The models are shown to capture many properties of the solvated bilayer and represent an important milestone in the coarse-graining of complex biological membrane systems. The applicability of the present MS-CG models is limited to preassembled bilayers. However, the MS-CG methodology permits an efficient parametrization of the interactions for simulation of other phases of phospholipid/cholesterol mixtures by fitting to an all-atom MD simulation of the phase of interest, which is generally expected to be better than a construction of a "universal" coarse-grained model. In general, the quality of an MS-CG model is determined by the length of the reference all-atom simulation as well as the accuracy of the underlying force field used in the reference MD simulation. Consequently, the present MS-CG force fields can be improved upon with the availability of more accurate atomistic mixed bilayer simulations.

This article is organized as follows: section 2 describes the MS-CG method used, and then section 3 reports applications of the method to the mixed lipid bilayer. This latter section starts with a presentation of general details of the FM fitting procedure from the underlying all-atom MD simulations and then presents the MS-CG models. The accuracy of the MS-CG models is also analyzed. In section 4, conclusions from this study are presented.

II. Coarse-Graining Methodology

The algorithmic development of the MS-CG method is presented elsewhere.[28,29] In the framework of the MS-CG approach, the atom−atom force $\mathbf{f}_i^{\mathrm{p}}(r_{ij})$ (acting from particle j on particle i) can be partitioned into a short-ranged part and a long-ranged Coulomb part as

$$\mathbf{f}_i^{\mathrm{p}}(r_{ij}) = -\left(f(r_{ij}) + \frac{q_i q_j}{r_{ij}^2}\right)\mathbf{n}_{ij} \tag{1}$$

where r_{ij} is the modulus of the vector $\mathbf{r}_{ij} = \mathbf{r}_j - \mathbf{r}_i$ connecting two atoms, q_i is the partial atomic charge (subject to fit), and $\mathbf{n}_{ij} = \mathbf{r}_{ij}/r_{ij}$. The effective short-ranged force is fitted by cubic splines connecting a set of points $\{r_k\}$ ($r_{k_{max}}$ defines a cutoff radius), thus preserving continuity of the functions and their first two derivatives across the junction, such that

$$f(r, \{r_k\}, \{f_k\}, \{f_k''\}) = A(r, \{r_k\})\, f_i + B(r, \{r_k\})\, f_{i+1} + C(r, \{r_k\})\, f_i'' + D(r, \{r_k\})\, f_{i+1}''; \; r \in [r_i, r_{i+1}] \tag{2}$$

where A, B, C, and D are known functions of r, $\{r_k\}$, and $\{f_k\}$, and $\{f_k''\}$ are tabulations of $f(r)$ and its second derivative on a radial mesh $\{r_k\}$. A spline representation depends linearly on its parameters $\{f_k, f_k''\}$, which permits the use of force averaging over trajectories as a part of the least-squares fit.

In the first step of the MS-CG procedure, all sets of system configurations recorded along the trajectories of the reference MD simulation are partitioned into blocks, each containing L configurations. For each block, the reference total force

$\mathbf{F}^{\mathrm{ref}}_{\alpha il}$, which acts on the ith interaction center of kind α in the lth configuration set of the block, and the same force predicted using the representation in eq 2 are required to be equal, which leads to the following equation

$$-\sum_{\gamma=nb,b}\sum_{\beta=1,K}\sum_{j=1,N_\beta}\left(f(r_{\alpha il,\beta jl},\{r_{\alpha\beta,\gamma,k}\},\{f_{\alpha\beta,\gamma,k}\},\{f''_{\alpha\beta,\gamma,k}\})+\frac{q_{\alpha\beta}}{r^2_{\alpha il,\beta jl}}\delta_{\gamma,nb}\right)\mathbf{n}_{\alpha il,\beta jl}=\mathbf{F}^{\mathrm{ref}}_{\alpha il},\ i=1,\ldots,N_\alpha,\ \alpha=1,\ldots,K,\ l=1,L \quad (3)$$

with respect to $\{f_{\alpha\beta,\gamma,k}, f''_{\alpha\beta,\gamma,k}, q_{\alpha\beta}\}$, which are the force parameters subject to fit. In eqs 3, $\{\alpha il\}$ labels the ith site of kind α in the lth configuration of the block; $r_{\alpha il,\beta jl}$ is the distance between sites $\{\alpha i\}$ and $\{\beta j\}$ in the lth configuration; $q_{\alpha\beta}$ is a product of partial charges q_α, q_β of sites of kind α and β; and N_β and K are, respectively, the number of sites of kind β and the total number of site kinds in the system.

The MS-CG method also allows for a systematic separation of bonded and nonbonded forces.[29] This is important for the CG sites which have an overlap in regions of intra- and intermolecular motions, causing the FM force field in those regions to be a mixture of nonbonded and bonded components, which are otherwise impossible to separate. To aid in fitting explicitly the bonded forces, the additional index $\gamma=\{nb, b\}$, which indicates whether the site pair $\alpha il,\beta jl$ is bonded (i.e., $\gamma=b$), is introduced. This index distinguishes the f, f'' parameters for bonded and nonbonded interactions in eq 3. For bonded pairs, the Coulomb term is absent, which is enforced in eq 3 by the $\delta_{\gamma,nb}$ term which is unity if $\gamma=nb$ and zero otherwise. It is permissible to choose the spline mesh $\{r_{\alpha\beta,\gamma,k}\}$ to be different for different pairs $\{\alpha\beta\}$ of site kinds.

If L is large enough, eq 3 forms an overdetermined system of linear equations with respect to the fitting parameters. Standard equations which are linear with respect to $\{f_{\alpha\beta,\gamma,k}, f''_{\alpha\beta,\gamma,k}\}$ must also be included into eq 3 to ensure that the $f(r)$'s first derivative, $f'(r)$, is continuous across the boundary between two intervals.[29,32] The number of equations in eq 3 can also be increased with an increment equal to the number of interaction sites in each configuration set.

The method described above is general and can be straightforwardly applied to derive an effective force field for coarse-grained models. To do this, the reference atomistic trajectories and forces have to be reduced to the trajectories and forces for the structural units subject to coarse-graining. The most natural way to coarse-grain a set of atomic groups is to place the CG site at each group's center-of-mass (CM) since the force acting on it can be straightforwardly evaluated from the atomistic data. Consequently, the FM algorithm can be applied to determine coarse-grained interactions from fully atomistic MD trajectory data. The CG force field so obtained is then an approximation to the effective interaction, or pairwise PMF, between the CG interaction sites.

Generally, a CG site can be mapped into an atomic group in other ways with the reference force evaluated in the atomistic MD simulation. For example, a CG site may instead be associated with the geometrical center (GC) of the structure (i.e., the CM of the system assuming that all atoms have the same mass). The latter choice may be more preferable for atomic groups with highly uneven mass distribution (e.g., the water molecule). Of course, the MS-CG force field will be dependent on where the CG sites are placed, and the same choices of that placement may be better than others. The optimal choice of MS-CG sites for a given problem is the topic of ongoing research in our group.

As is shown elsewhere,[29] CG force fields can perform well in the reproduction of the bulk phase structural properties; however, they may fail to maintain the proper internal pressure in the system, and as a result the density may also be incorrect (typically too low) in constant pressure CG MD simulations. This behavior can be corrected in the MS-CG approach through the virial equations used to evaluate the pressure in the MD simulation.[29] Because the virial depends linearly on the atomic forces and the kinetic energy does not rely on the forces at all, the MS-CG force field can also be constrained to produce the correct pressure.[29] This task can be accomplished by adding to eq 3 the proper constraints which include the system virial.[29]

The FM force field derived by including the virial constraint depends explicitly on the instantaneous kinetic energy and therefore on the temperature in the reference atomistic simulation. This may define the transferability (or lack thereof) of the present MS-CG models to other temperatures. Thermodynamic properties which rely on the derivatives of the temperature (e.g., thermal expansion coefficient) may also be less accurate.

Other CG approaches have developed a CG force parametrization by fitting a few preselected macroscopic and average structural properties. Our method is fundamentally different as it systematically maps the underlying atomistic-level interactions into effective interactions between CG structural units. Hence, it is a "multiscale" coarse-graining approach. Importantly, for systems such as mixed phospholipid/cholesterol bilayers (described below) for which fitting CG potentials to preselected average properties is difficult at best, the MS-CG approach appears to offer some real advantages.

III. Results and Discussion

A. Molecular Dynamics Simulation Details. In the present work, the MS-CG method was employed to obtain a CG model of a mixed DMPC/Chol bilayer at 50 mol % concentration of the Chol. The effective force field describing interactions between sites in the CG representation of the bilayer was derived using trajectory and force data from a single all-atom MD simulation. The DMPC molecules were modeled using a united atom force field,[12] and Chol molecules were modeled with a modified AMBER force field.[8] For water, the rigid TIP3P model[33] was used. The temperature of the system was kept constant at $T=308$ K using the Nosé-Hoover thermostat with a relaxation time of 0.2 ps. The electrostatic interactions were calculated via the particle mesh Ewald (PME) summation, and the van der Waals interactions were cut off at 0.7 nm. The system of 6336 total atoms (32 lipids, 32 cholesterol molecules, and 1312 water molecules) was integrated with a time step of 2

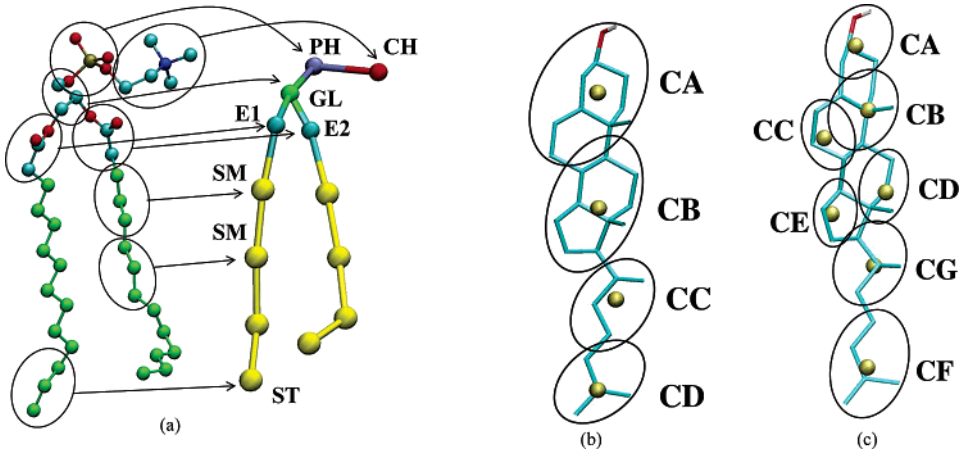


Figure 1. Coarse-grained representation of lipid and cholesterol molecules: panel (a) DMPC; panel (b) four-site cholesterol; and panel (c) seven-site cholesterol.

fs, and the initial structure of the DMPC/Chol bilayer taken from ref 34 was equilibrated for 4 ns. The reference properties of the atomistic system were obtained from the MD simulations under both constant NPT and NVT conditions. The equilibrium volume of the supercell was 92.029 nm^3 with an area per DMPC molecule of 0.440 nm^2 under the assumption of an effective area per Chol molecule of $\sim$0.336 nm^2.[35]

B. Coarse-Graining and Details of the Force-Matching. The MS-CG water molecules were mapped into a single CG interaction site associated with their GC instead of CM. As shown elsewhere,[28] with such a choice the FM force field produces a water structure which compares better to the atomistic reference MD system. The CG representation of a DMPC molecule is shown in Figure 1 (a) (also used in ref 28) and is similar to the models of Shelley et al.[14] and Marrink et al.[16] A coarse-graining of the acyl tail into a smaller set of spherical sites has the potential problem that could not fully describe trans−gauche isomerizations of the acyl chains. The trans−gauche conformations may result in a sharp change in the geometries of SM and ST CG groups in the all-atom MD simulation and thus may have an adverse impact on the quality of the MS-CG force field. On the other hand, the presence of cholesterol reduces the probability of gauche defects.[36]

For the Chol molecule, two CG representations were developed. In the first representation, the molecule was coarse-grained into four interaction sites. In the second representation, seven sites were used. The sites associated with each CG model of Chol are shown in Figure 1b,c. In the four-site representation, the tetracyclic ring structure is split between CA and CB sites. The CA site includes the polar hydroxyl group, while the CC−CD site pair represents the aliphatic side chain attached to the ring structure. For the model with four-site CG sterols, Table 1 summarizes the average atomistic radii of the groups forming the CG sites as evaluated from the same MD trajectory data used in the FM fit as well as the sum of the charges on the atomistic sites from which the CG site is formed. (It should be noted that in the final MS-CG models none of the CG groups have charges, i.e., they interact only through effective short-ranged forces.) In the four-site Chol representation, the arrangement

Table 1. Radii, $R_{CG} = \langle (r - r_{CM})_{max} \rangle$, and Total Charges, q_{CG}, of the underlying CG Sites in the All-Atom Model

site	R_{CG}, nm	q_{CG}, au	site	R_{CG}, nm	q_{CG}, au
CH	0.237	+1.4	ST	0.144	0
PH	0.165	−1.4	CA	0.392	−0.012
GL	0.138	+1.2	CB	0.266	0.009
E2	0.156	−0.7	CC	0.201	0.004
E1	0.153	−0.5	CD	0.160	−0.001
SM	0.140	0	WC	0.078	0

of the CG sites is close to linear. In the seven-site representation, three more sites were added to better reflect the planar geometry of the ring structure. The total number of DMPC plus Chol site pairs is 78 for the model with the four-site representation of the Chol molecule and 120 for the seven-site representation. Such a large variety of interaction sites makes it very difficult to utilize existing methods[14,16,25] for parametrization of the coarse-grained interactions. However, the MS-CG method is capable of treating such a complex system.

The intramolecular CG sites were linked by effective bonds, and two bonding schemes were used. In the first bonding scheme, which is referred to as "minimal", only consecutive sites were linked by bonds. In such a model, the force term describing bond bending was implicitly present in the FM nonbonded forces. In the second scheme, the next-nearest-neighbors were connected by bonds (e.g., CH-GL, PH-E1,2, etc., for DMPC and CA-CC, CB-CD, etc., for Chol). Such interactions are similar to a bond angle potential. The use of the bonds instead of angle potentials was necessary because the FM method can handle explicitly only site−site (either bonded or not) interactions. Therefore, the second bond scheme is referred to as the "bond angle" scheme. The use of the explicit separation of bonds in the FM procedure increased the number of interactions for the force-match. For example, in the model with the four-site Chol molecules and the "bond angle" scheme, the total number of interactions to match was 101.

For the purpose of force-matching, the atomistic MD simulation was carried out for 5 ns in a constant NVT ensemble using the equilibrium volume obtained from a constant NPT MD simulation. The trajectory, velocity, and

Table 2. Coefficients A_n of the Least-Squares Fit of TIP3P MS-CG One-Site Force Field Using the Expansion $f(r) = \sum_{n=2}^{16} A_n/r^n$ [a]

n	A_n	n	A_n
2	1121.994059416	10	1461833177752.
3	−100003.0777743	11	−2619871642943.
4	3956632.300647	12	1698105340916.
5	−91592923.27445	13	1534516677398.
6	1375454692.338	14	−940850812653.1
7	−14021182651.00	15	−2041415905133.
8	98418431541.42	16	−1742793289718.
9	−470296232562.7		

[a] Atomic units were used. A cut-off of 0.8 nm must be applied to this expansion.

force data were sampled with an interval of 1 ps so that the number of stored configurations was 5000. The instantaneous virial for each sampled configuration was also recorded. The velocity and virial data were necessary to implement the virial constraint.[29] Next, trajectories, velocity, and forces were transformed into corresponding data for the CG sites and then were used as an input into the FM algorithm with an explicit separation of bonds and virial constraint.[29] The FM force field was represented by a spline over a mesh with a grid spacing of approximately 0.01 nm for nonbonded and 0.0025 nm for bonded interactions. For nonbonded forces, the mesh was extended to 0.13 nm. The number of configurations used to build each of the overestimated systems of equations in eq 3 was $L = 3$.

As mentioned earlier, despite the fact that some composite CG sites in the underlying atomistic model are charged (see Table 1), it is possible to omit the Coulomb term in eq 3 if a large enough cutoff is used for the short-ranged interactions. For example, if a sufficiently large cutoff is used, the FM procedure with the Coulomb term explicitly present yielded effective charges on the CG sites to be much smaller compared to those inferred from the reference atomistic charges. This is a result of the inherent screening of Coulomb forces and the short-ranged forces then incorporating effectively the Coulomb interactions within their cutoff radius. The screening effects come from the water as well as the polar DMPC headgroups which are captured by the MS-CG force field. Therefore, the FM procedure was used with the Coulomb term removed in eq 3.

C. MS-CG Interactions. *1. Water Model.* The one-site CG representation of the TIP3P model as obtained from the FM to the all-atom MD simulation of the bilayer was tested on a bulk water system. The structural properties from the all-atom and MS-CG MD simulations of the 256 TIP3P molecules in the constant NVT ensemble and at ambient conditions were identical. However, the density of the CG water in the NPT MD simulation was slightly lower. By contrast, the FM (using the virial constraint) of the one-site water force field to the MD simulation of 256 TIP3P molecules at ambient conditions produced a force field which showed the same structural properties as one from FM to the bilayer system; however, the ambient density was correct. A polynomial representation of the MS-CG potential for this model is given in Table 2. A comparison of the water

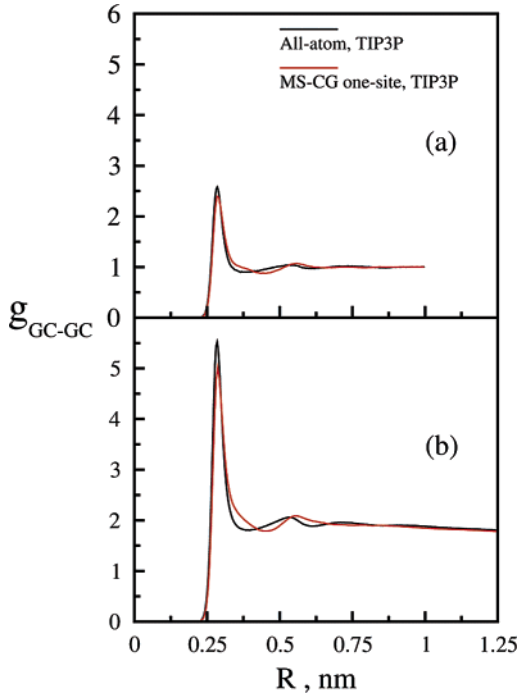


Figure 2. Comparison of the water GC-GC RDFs from the all-atom (black line) and MS-CG (red line) MD simulations of bulk water [panel (a)] and the bilayer system [panel (b)] using the MS-CG TIP3P model fitted to the bulk water data.

Table 3. Properties of TIP3P One-Site MS-CG CG Water Model[a]

property	TIP3P 1-site	TIP3P[b]	expt[c]
ρ (kg/m^3)	1004	1002	997.05
U^{pot} (kJ/mol)	−8.3	−41.1	−41.5
α_P (10^{-4} 1/K)	14.7	6.4	2.53
κ_T (10^{-5} 1/bar)	37.1	9.2	4.52
D_{diff} (10^{-9} m^2/s)	18.1	5.19	2.3

[a] Shown are ρ, density; U^{pot}, average configuration energy; α_P, thermal expansion coefficient; κ_T, isothermal compressibility; D_{diff}, self-diffusion coefficient. [b] References 33 and 44. [c] Reference 45 (25 °C).

geometrical center-geometrical center (GC-GC) RDFs from the bulk water and bilayer MD simulations is given in Figure 2.

Some other thermodynamic properties of the MS-CG water model are compiled in Table 3. The larger isothermal compressibility, κ_T, from the CG model is likely a consequence of the use of the virial constraint in the FM procedure which explicitly depends on the temperature in the reference MD simulation. A similar behavior has been reported for a one-site water model matched to ab initio MD simulation.[29]

The water−water (W−W) force field from the FM to the bulk water system with the virial constraint was used in all models of the bilayer reported here. However, in contrast to the bulk geometry, the all-atom water in the vicinity of the bilayer surface is "orientationally polarized", orienting its molecular dipoles toward the bilayer surface.[7,37] On the other hand, the correct pressure of the water seems to be a more influential factor in the formation of the bilayer properties even in MD simulations in the constant NVT ensemble. For example, the bilayer structural properties from the constant NVT MD simulation using the bulk MS-CG water force field

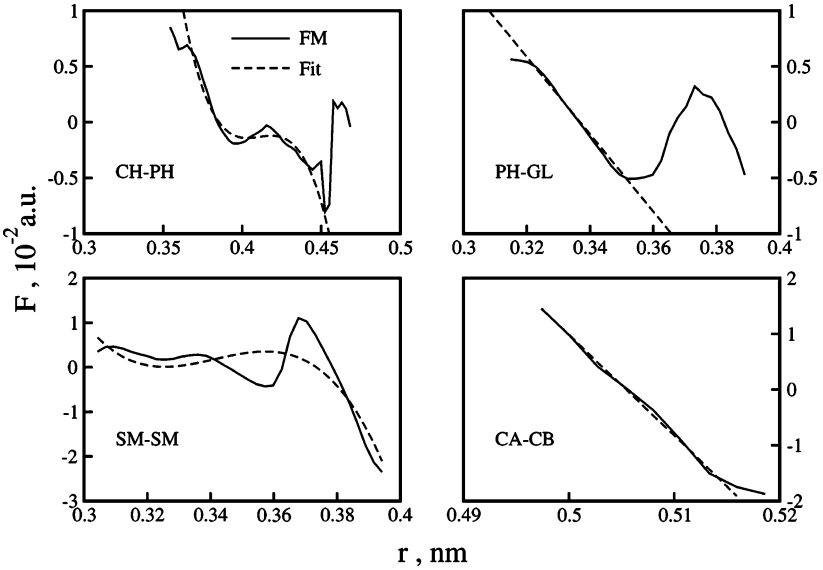


Figure 3. Effective pairwise bonded forces between selected coarse-grained interaction sites as a function of intersite separation (solid curve) and their analytical fit (dashed curve) used in the MS-CG MD simulations.

appeared to be slightly better compared to those properties if the MS-CG water force field from explicit FM to the bilayer system was used. It seems this behavior was due to a more correct pressure of the water on the bilayer surface from the bulk MS-CG water model.

2. DMPC and Chol MS-CG Models. Selected force profiles for bonded interactions for the models with the four-site representation of a Chol molecule are shown in Figure 3. The bonded forces of the model with the minimal bonding scheme were identical to the corresponding forces in the model with the bond angle scheme. As seen in a majority of the cases, the force deviates substantially from being harmonic (i.e., linear). In the MS-CG MD simulation, a least-squares approximation of the bonded forces either by harmonic or harmonic+quartic forces was used. A harmonic approximation was used, for example, to a lipid PH-GL bonded interaction, as it was found to be accurate enough. In those FM bonded forces for which the harmonic approximation obviously was not accurate, the quartic potential was instead used. A comparison of FM forces with their analytical fits is shown in Figure 3. An accurate description of the bonded forces is important for those parts of the molecules that have greater flexibility, such as the hydrocarbon chains in the DMPC molecule. To give an indication of the rigidity of the various bonds in the CG representations of the DMPC and Chol molecules, Table 4 shows parameters for the harmonic approximations to the bonded forces and associated frequencies (estimated as if the bonded pair were an isolated diatomic molecule). The low-lying spectra of bond frequencies also allows one to increase the time step in the MS-CG MD simulation compared to an all-atom simulation. The model with the minimal bonding scheme was deemed to be more accurate as the interactions between nearest-neighbor sites are included into nonbonded forces and, therefore, may be less influenced by errors introduced from analytical fits of the bonded forces.

Selected profiles of the MS-CG nonbonded forces for the model with the four-site representation of Chol are shown in Figure 4. The effective force field often shows substantial variations with distance in the vicinity of separations of the order of the sum of average radii of the CG sites (see Table 1). This is because at these distances the interaction in the all-atom MD simulation is most sensitive toward mutual orientations of the atomic groups representing the CG sites. For some of the site pairs, which were at large separations in the all-atom MD simulation, the unsampled core region was also significant. For example, for the CH−CD pair in the model with the four-site Chol, the r_{core} was ∼1.7 nm. Those DMPC-W and Chol-W interactions which have a large r_{core} value or are underconvergent/noisy might be improved by matching to the simulation of single DMPC or Chol molecules in the solvent. However, in the present study this approach was not used. The interaction within the r_{core} region

Table 4. Parameters of Harmonic Approximation to Bonded Forces and Corresponding Frequencies for the Model with the CG Chol Molecule of CG Four Sites[a]

bond	k_b, [kJ/(mol·Å^2)]	r_0, nm	ω, cm^{-1}
CH−PH (b)	0.184	0.385	132.5
CH-GL (a)	0.007	0.610	31.9
PH-GL (b)	0.187	0.337	168.5
PH-E2 (a)	0.022	0.435	51.1
PH-E1 (a)	0.038	0.661	67.4
GL-E2 (b)	0.554	0.282	315.9
GL-E1 (b)	0.237	0.326	206.4
GL-SM (a)	0.002	0.595	19.4
E2-SM (b)	0.365	0.374	254.7
E2-SM (a)	0.005	0.692	28.1
SM-SM (b)	0.203	0.345	204.3
SM-SM (a)	0.020	0.654	64.0
CA-CB (b)	0.871	0.506	235.0
CA-CC (a)	0.096	0.924	102.2
CB−CC (b)	0.648	0.426	266.1
CB−CD (a)	0.130	0.831	118.7
CC−CD (b)	0.278	0.424	206.5

[a] Labels (b) and (a) label "consecutive" and "angle" bonds, respectively.

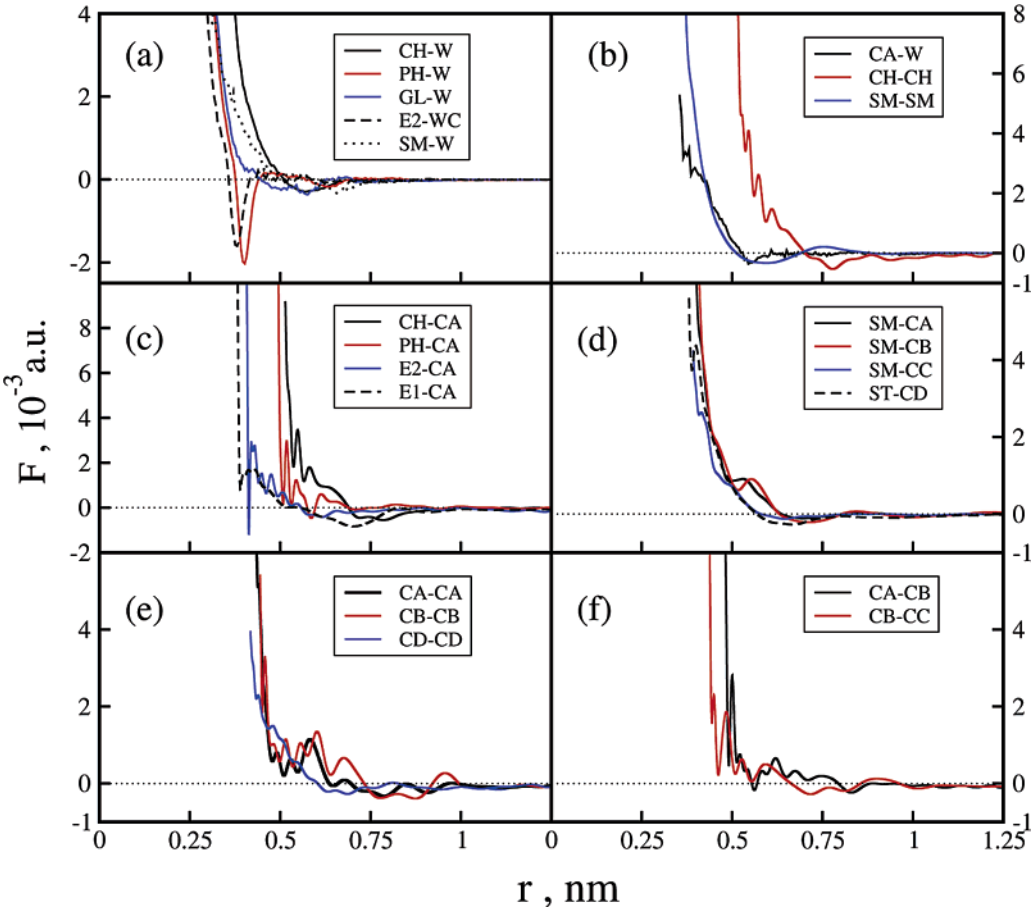


Figure 4. Effective pairwise nonbonded forces between selected coarse-grained interaction sites smoothed by a polynomial fit as a function of intersite separation. The approximate error bars in these curves are about 5−10% of the force value.

was therefore assumed to be weakly repulsive. Obviously, large r_{core} for certain pairs of CG sites limits the transferability of the model to other than bilayer geometries. Consequently, the models reported here can be only safely used to simulate preassembled bilayers.

The force profiles for some of the bilayer site−site pairs sometimes converged slowly in the FM procedure with some statistical noise. Such noise is not generally a source of difficulty in the MS-CG MD simulation because the force profiles can be further smoothed out using a least-squares polynomial fit of the FM spline data. However, such a fit in the present case gave somewhat worse properties. Therefore, the original spline fit was used to conduct the reported MS-CG MD simulations.

An inspection of the MS-CG effective forces also provides important insight into the nature of the interactions inside the bilayer. As seen from Figure 4(c), the most attractive interaction between lipid and Chol species occurs between E1 and CA sites. This is an interesting fact because there is experimental evidence of strong hydrogen bonding of the Chol hydroxyl group (belonging to the CA site), which acts as a proton donor with the carbonyl group of the sn-2 chain[38] (of which the E2 site is a part). MD simulations of DMPC/Chol bilayer at low sterol content have also shown that bilayer Chol can bind to lipid molecules via either hydrogen bonding or water bridges.[11,35] At a low Chol concentration (X_{Chol} = 12.5 mol %), the largest number of hydrogen bonds (both direct and through water bridges) occurs for carbonyl oxygen in the sn-2 chain;[35] however, as the Chol concentration increases (to X_{Chol} = 22 mol %), the sn-1 oxygen (which is in the E1 site) becomes a preferable place for the OH group of Chol to bind.[11] A similar behavior was suggested by NMR studies.[39] With an increase of the Chol concentration due to the "condensing effect", flexible acyl chains pack closely around the Chol molecules, limiting the ability of the latter to form H bonds with the sn-2 oxygen. The tendency of Chol molecules to orient their hydroxyl groups close to the carbonyl groups of the phospholipid and then to remain there was also observed in an MD simulation of DPPC/Chol bilayer.[40]

There is also a pronounced attractive part in the CH−CA force profile, while the PH−CA interaction is mostly repulsive. The CH is the second lipid CG site with which the CA site interacts most significantly. The attractive part of the CH−CA force is centered around 0.77 nm and is relatively broad. This correlates with MD simulation studies which pointed toward complexation of the choline moiety with the Chol hydroxyl group (the CA site) via formation of charged pairs.[11] The OH of Chol can bind to the oxygens of the phosphate group (PH site); however, the probability of formation of the hydrogen bonds at this location is much smaller compared to that of carbonyl groups (E1,2 sites).[11,35] The small negative dip in the PH−CA force profile might be a manifestation of the hydrogen bonding of the phosphate

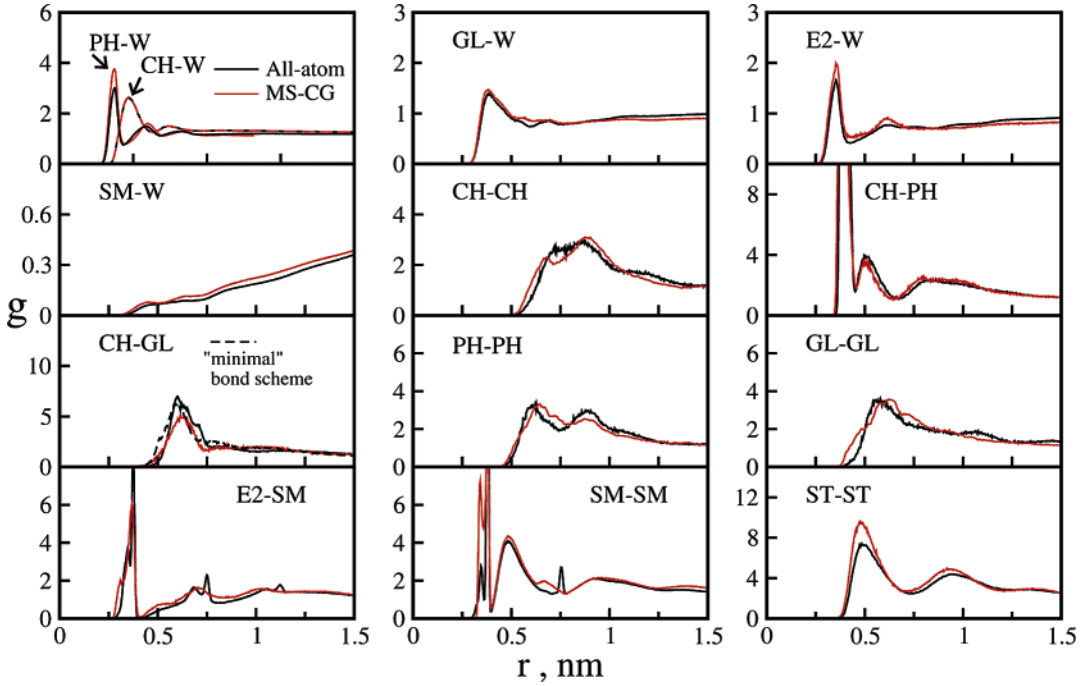


Figure 5. Selected DMPC−water and DMPC−DMPC site−site RDFs from the all-atom MD simulation (black lines) compared to those from the MS-CG MD simulation (red lines) using the four-site representation of the Chol molecule and the "bond angle" bonding scheme.

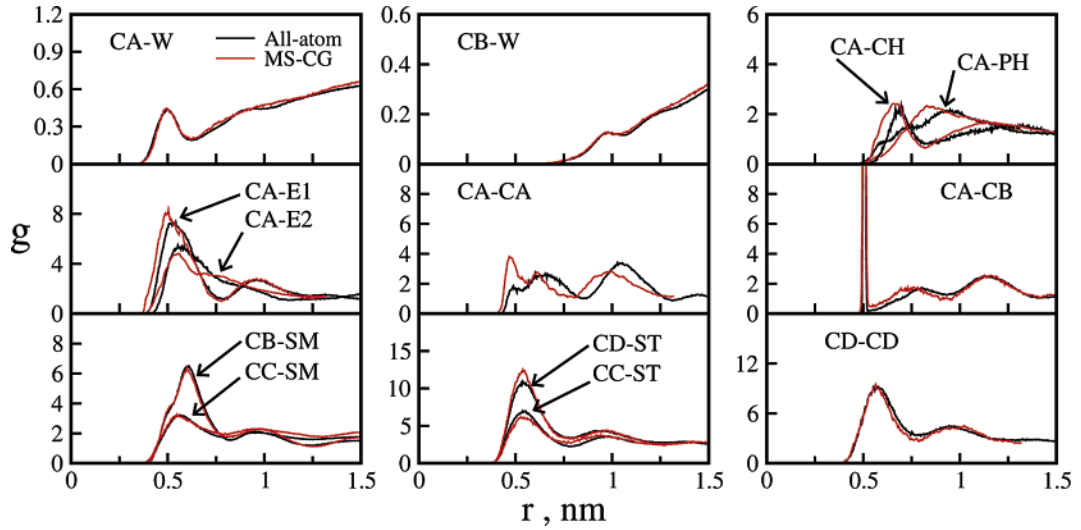


Figure 6. Selected Chol−water, Chol−DMPC, and Chol−Chol site−site RDFs from the all-atom MD simulation (black lines) compared to those from the MS-CG MD simulation (red lines) using the four-site representation of the Chol molecule and the "bond angle" bonding scheme.

oxygens and OH of Chol. The interactions of the Chol sites with acyl chain sites (SM, ST sites), which arise mostly from van der Waals forces, are only weakly attractive.

D. MS-CG Simulations of the Bilayer. To help validate the models, the system in the MS-CG MD simulation of the mixed bilayer was initially taken to be the same size as that of the all-atom MD simulation. At first, a 5 ns MD simulation for each MS-CG model was carried out using a constant NVT ensemble with time step of 0.005 ps. The MS-CG MD simulation for this system size was about 15 times faster. However, the gain in computational efficiency increases rapidly with an increase in system size, the latter (larger system sizes) being one purpose of CG simulations.

A comparison of the MS-CG RDFs from the model with the four-site CG Chol molecules with the bond angle scheme with their atomistic MD counterparts is given in Figures 5 and 6. The agreement is generally quite good. Some discrepancies in the bilayer site−site RDFs are attributed primarily to two factors. One of them is the analytical representation of the bonded forces, which as demonstrated in Figure 3 may deviate from the exact FM forces, thus influencing the subtle balance of the interactions in the MS-CG MD simulation. The second factor is the presence of sometimes large, unsampled regions of intersite separations. The most affected RDFs are those which involve solvated sites, and ways to improve this sampling are presently being

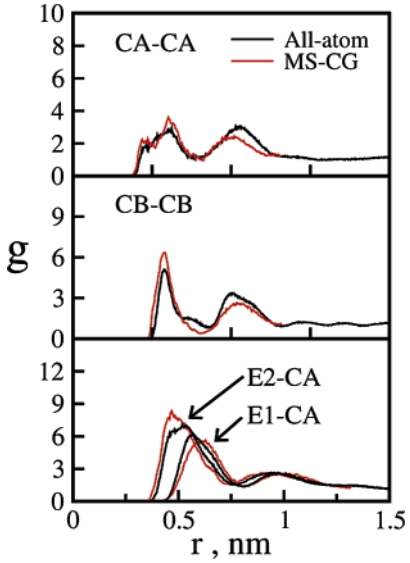


Figure 7. Selected site−site RDFs from the all-atom MD simulation (black lines) compared to those from the MS-CG MD simulation (red lines) using the seven-site representation of the Chol molecule and the "bond angle" bonding scheme.

explored. The use of the minimal bond scheme did not result in significant improvements or changes in the structure. For the pairs which were bonded by "angle" bonds in the bond angle scheme, the structure was slightly improved (e.g., the CH-GL pair, in Figure 5, panel CH-GL, dashed line).

An analysis of the CG structure of the bilayer is a convenient way to scrutinize the ordering in the bilayer because of the smaller number of degrees of freedom. The presence of a well pronounced first peak in DMPC-W and Chol-W RDFs (Figures 5 and 6) reflects the effective polarity of the respective CG sites. DMPC headgroup sites CH, PH, GL and ester groups E1,2 are solvated, and, as can be judged from a height of the first peak in the RDFs, the most strongly solvated site is PH followed by the CH, E2, GL, and E1 (RDF not shown) sites, respectively. The strong first peak in the PH-W RDF is a reflection of the fact that two phosphate oxygens are the most probable sites for the water to form hydrogen bonds with the DMPC.[35]

The good agreement of the CG and all-atom DMPC-W and Chol-W RDFs suggests that despite being nonpolar, the one-site CG water model adequately incorporates its interactions with the charged bilayer groups which arise from a polarity of the water molecules, e.g., hydrogen bonding with the PH and E1 sites. Nonpolarity of the SM and ST sites was manifested in an absence of the first peak in the SM-W (see Figure 5) and ST-W RDFs (the ST-W RDF is similar in shape to the SM-W RDF). As seen from Figure 6, the CA site is hydrophilic; however, it was solvated less compared to any of the lipid headgroup sites. The rest of the Chol CG sites are nonpolar.

As can be seen from Figure 7, the use of the seven-site representation of the Chol molecule improved the structure, presumably due to a better CG representation of the planar geometry of the Chol molecule. However, these improvements generally were not sufficient to consider the seven-site model to be universally superior compared to the four-site model, as the latter is computationally less expensive. The MS-CG models having a rodlike CG Chol molecular geometry of four interaction sites were already able to reproduce the bilayer structure at a sufficiently accurate level. Below is an analysis of the bilayer properties using the MS-CG model with the four-site Chol molecules with the bond angle bonding.

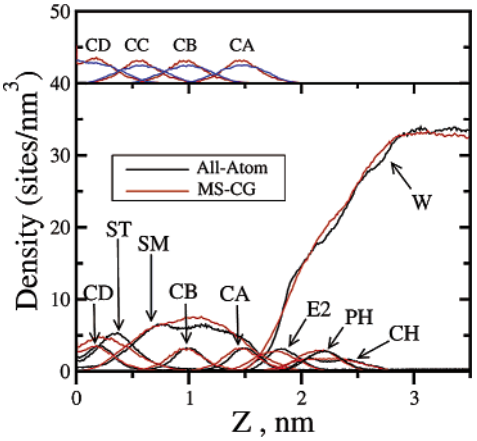


Figure 8. Comparison of the all-atom (black lines) and MS-CG (red lines) density profiles for several CG sites perpendicular to the bilayer. The insert compares z-density of Chol from smaller (red lines) and larger (blue lines) systems.

MS-CG MD simulations were next conducted using the constant NPT ensemble on the same size system. The simulation was 20 ns long. In Figure 8, the densities of the selected bilayer CG sites and water are shown, and the MS-CG model is again seen to capture the key structural properties of the bilayer accurately. The membrane thickness as estimated from the average distance between PH sites was 0.435 nm from the all-atom MD simulation compared to 0.422 nm from the MS-CG MD simulation. The area per DMPC molecule was 0.444 nm^2 (it was 0.440 nm^2 in the all-atom MD simulation). The z-density of Chol sites stretches to almost the same distance from the bilayer center as that of the acyl (SM, ST) sites. The front slope of the CA density follows closely the density profile of the SM sites.

Mutual arrangements of the CH and PH z-densities in Figure 8 and RDFs in Figure 6 indicate a tilt of the PH−CH bond toward the bilayer surface. The CH−CA and PH−CA RDFs suggest that the PH−CH bond points toward a neighboring CA group, forming the CH−CA charged pair. The first coordination number is 0.85 from the all-atom and 0.80 from the CG CH−CA RDFs. The CH−CA pair was most likely a solvent-separated pair as the position of the first peak in the CH−CA RDF at 0.69 nm exceeds the sum of 0.62 nm of the smallest CH−WC and CA-WC distances in the MS-CG MD simulation, as estimated from the corresponding RDFs. However, the formation of contact pairs is possible because the CH−CA separation reached a minimum at 0.51 nm. The Chol molecules tend to minimize exposure of their nonpolar sites to water molecules.[41] The CH sites shield the nonpolar parts of Chol (CA site represents a part of the first Chol ring) from the water that produced an additional driving force for the formation of the CH−CA pairs.

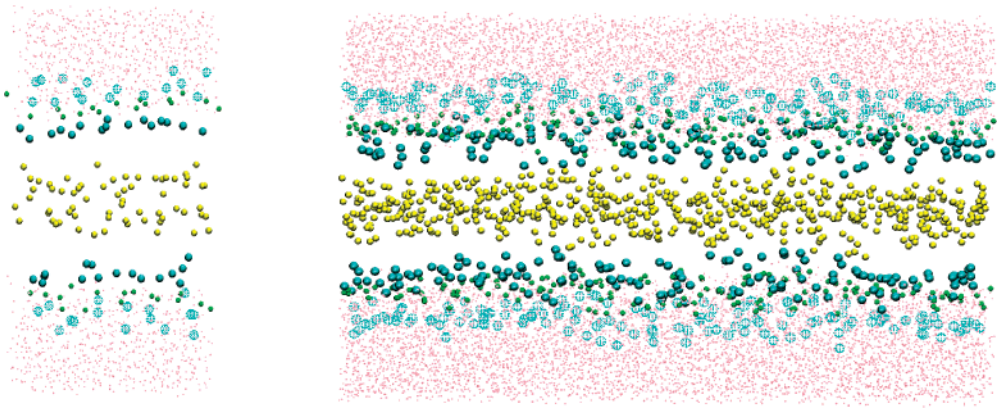


Figure 9. Side views of the bilayer with selected CG sites shown from the all-atom MD simulation [panel (a)] and the larger MS-CG MD simulation as described in the text (a segment of the supercell is shown) [panel (b)]. Referring to the CG sites in Figure 1, W: pink dots; CH: dotted spheres; GL: small green spheres; CA: large cyan spheres; and ST: small yellow spheres.

The MS-CG model was tested also on a system which was larger than the reference atomistic system by 4 times in the x and y directions (i.e., it contained 512 lipids, 512 cholesterol molecules, and 20 992 water molecules or 29 696 total sites). The simulation was initiated from the configuration obtained by a replication of the smaller system in the x and y directions four times and was integrated for a duration of 20 ns. Figure 9 presents snapshot images of the supercell with the selected sites shown from the all-atom MD simulation [panel (a)] and a segment of the simulation cell from the MS-CG MD simulation of the larger system. The z-density profiles from that simulation are similar to those of the smaller system; however, they exhibit a slight broadening which is most noticeable for Chol sites. Such a broadening might be partially attributed to undulations which are less pronounced in the smaller system due to finite size effects. The insert in Figure 8 compares the Chol z-density profiles systems of different sizes. The larger MS-CG mixed bilayer system was 50−100 times faster to integrate than the equivalent all-atom MD system of the same size depending on the choice of short-ranged cutoff and time step in the MD program for the MS-CG simulation.

The ordering in the bilayer was demonstrated by presenting the distributions of cosines of the angle between the PH−CH vector and the bilayer normal. In the all-atom MD simulation, the distribution of the PH−CH orientations has two well-defined maxima corresponding to orientations parallel and normal to the membrane. In pure lipid bilayer, the orientational distribution of the P−N vector is more flat.[8] A first maximum in the $\langle P_{\mathrm{PH-CH}}(\cos\theta)\rangle$ is attributed to the lipid molecules in close contact with Chol and corresponds to a charged CH−CA pair. The maximum at $\cos(\theta) = 1$ is due to the lipid molecules which are not associated (or associated less) with Chol molecules. These molecules prefer to orient their PH−CH bond outward. Several mechanisms have been suggested for such behavior,[8] depending on the concentration and packing of Chol molecules in the bilayer. For a Chol concentration of 50 mol %, probable mechanisms include a decrease in the area per lipid headgroup and the change in the solvation structure around the lipid headgroups as the average distance between lipid molecules in the Chol-contained bilayer increases. The distribution of PH−CH orientations from the MS-CG model is more flat with less structure from the parallel and perpendicular orientations to the bilayer. The distributions of the angles of acyl chain bonds with the bilayer normal are shown in Figure 10 (a).The MS-CG model shows broader distributions, suggesting more disorder.

In Figure 10(b), bilayer order parameters for bond orientations from the all-atom and MS-CG MD simulations are compared. The second-order parameter $P_2 = {}^1/_2\,(3\langle\cos^2(\theta)\rangle - 1)$ where θ is an angle between the bond and the bilayer normal, is shown. The value of $P_2 = 1$ indicates a perfect aligning of the bond pointing outward with the bilayer normal; $P_2 = -0.5$ means the bond pointing inward, and $P_2 = 0$ means a complete disorder in bond orientations. The PH-GL points predominantly into the bilayer interior. The linkage SM-E2 is in a more disordered state compared to the SM-E1 bond, likely as a result of a intensive sterol association with the E1 site. The SM-SM (II) bond, which is central in acyl chains, is in the most ordered state. The P_2 distribution is similar to those obtained by Marrink et al.[16] using their CG model of pure DPPC bilayer. The CB−CA vectors are well aligned along the bilayer normal, with an average angle $\theta = 9.8°$ from all-atom and $\theta = 13.0°$ from the MS-CG MD simulations.

The lateral diffusion coefficient, D, of the DMPC molecules was calculated from the slope of the diffusive part of the mean square displacement (MSD) of the molecular center-of-mass. The MSD from the MS-CG simulation is shown in Figure 11. The value $D = 1.1 \times 10^{-9}$ m^2/s was obtained for the DMPC diffusion, which is about four times larger than the experimental value,[42,43] $D = 3.0 \times 10^{-10}$ m^2/s for the ordered phase of DMPC/Chol at $X_{\mathrm{Chol}} = 33$ mol %. The diffusion of the CG water was $D = 11.2 \times 10^{-9}$ m^2/s, compared to the diffusion of the water in the all-atom MD simulation of $D = 2.74 \times 10^{-9}$ m^2/s. The dynamics of the MS-CG system is seen to be significantly (by about four times) faster than all-atom MD simulation, as is expected from the reduced number of degrees of freedom.[29] An advantage of this is a faster sampling of the configurational

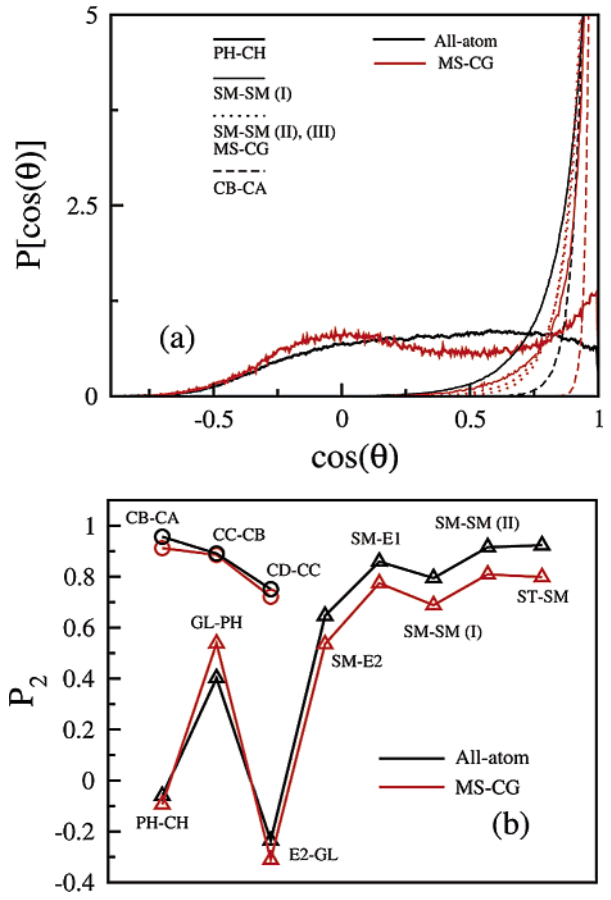


Figure 10. Panel (a): probability distribution of the angle between PH-CH, ST-SM, and CB-CA vectors and the bilayer normal from the reference all-atom (black lines) and MS-CG (red lines) MD simulations. Panel (b): P_2 order parameter of bonds with respect to bilayer normal from the all-atom (black line) and MS-CG (red lines) MD simulations.

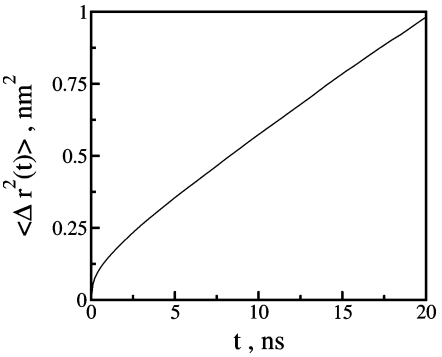


Figure 11. Center-of-mass mean squared displacement of DMPC molecules in the mixed bilayer from MS-CG MD simulation using the four-site CG Chol molecule.

space. However, the interpretation of the time scale in the MS-CG simulation is rather arbitrary.[29]

IV. Conclusions

Following our recent work on the application of the new MS-CG method to develop CG models of pure lipid bilayers,[28] we have applied the MS-CG method to construct a family of effective coarse-grained models for the mixed DMPC/Chol bilayer. The force-matching included a fit of the virial data, giving models that can be used for MS-CG MD simulations under constant NPT conditions. The models also explored different levels of coarse-graining of the cholesterol molecules and the intramolecular CG bonding scheme. The resulting MS-CG models were seen to predict the correct mixed bilayer structure, further demonstrating the effectiveness of the MS-CG method. Applications to more complex biomolecular systems are currently underway.

Acknowledgment. The research was supported by The National Institutes of Health (ROI GM63791). The computations in this project were provided in part by the United States National Science Foundation (NSF) cooperative agreement No. ACI-9619020 for the computing resources provided by the National Partnership for an Advanced Computational Infrastructure at the San Diego Supercomputer Center.

References

(1) Gennis, R. B. *Biomembranes: Molecular Structure and Function*; Springer: New York, 1989.

(2) *Lipid Bilayers: Structure and Interactions*; Katsaras, J., Gutberlet, T., Eds.; Springer: Berlin, 2001.

(3) Feller, S. E.; Pastor, R. W. *Biophys. J.* **1996**, *71*, 1350.

(4) Tieleman, D. P.; Marrink, S. J.; Berendsen, H. J. C. *Biophys. J.* **1997**, *1331*, 235.

(5) Feller, S. E.; Yin, D.; Pastor, R. W.; MacKerell, J. A. D. *Biophys. J.* **1997**, *73*, 2269.

(6) Feller, S. E.; Venable, R. M.; Pastor, R. W. *Langmuir* **1997**, *13*, 6555.

(7) Tu, K.; Klein, M. L.; Tobias, D. J. *Biophys. J.* **1998**, *75*, 2147.

(8) Smondyrev, A. M.; Berkowitz, M. L. *Biophys. J.* **1999**, *77*, 2075.

(9) Smondyrev, A. M.; Berkowitz, M. L. *J. Chem. Phys.* **1999**, *111*, 9864.

(10) Feller, S. E.; Pastor, R. W. *J. Chem. Phys.* **1999**, *111*, 1281.

(11) Pasenkiewicz-Gierula, M.; Róg, T.; Kitamura, K.; Kusumi, A. *Biophys. J.* **2000**, *78*, 1376.

(12) Smondyrev, A. M.; Berkowitz, M. L. *J. Comput. Chem.* **1999**, *20*, 531.

(13) Feller, S. *Curr. Opin. Colloid Interface Sci.* **2002**, *5*, 217.

(14) Shelley, J. C.; Shelley, M. Y.; Reeder, R. C.; Bandyopadhyay, S.; Klein, M. L. *J. Phys. Chem. B* **2001**, *105*, 4464.

(15) Marrink, S.; Mark, A. *J. Am. Chem. Soc.* **2003**, *125*, 15233.

(16) Marrink, S. J.; de Vries, A. H.; Mark, A. E. *J. Phys. Chem. B* **2004**, *108*, 750.

(17) Faller, R.; Marrink, S. J. *Langmuir* **2004**, *20*, 7686.

(18) Marrink, S. J.; Mark, A. E. *Biophys. J.* **2004**, *87*, 3894.

(19) Marrink, S. J.; Mark, A. E. *Biophys. J.* **2005**, *88*, 384A.

(20) Marrink, S. J.; Risselada, J.; Mark, A. E. *Chem. Phys. Lipids* **2005**, *135*, 223.

(21) Brannigan, G.; Brown, F. L. H. *J. Chem. Phys.* **2004**, *120* (2), 1059.

(22) Brannigan, G.; Philips, P. F.; Brown, F. L. H. *Phys. Rev. E* **2005**, *72* (1), 11915.

(23) Tozzini, V. *Curr. Opin. Struct. Biol.* **2005**, *15*, 144.

(24) Meyer, H.; Biermann, O.; Faller, R.; Reith, D.; Müller-Plathe, F. *J. Chem. Phys.* **2000**, *113*, 6264.

(25) Murtola, T.; Falck, E.; Patra, M.; Karttunen, M.; Vattulainena, I. *J. Chem. Phys.* **2004**, *121*, 9156.

(26) Brown, S.; Fawzi, N. J.; Head-Gordon, T. *Proc. Natl. Acad. Sci.* **2003**, *100*, 10712.

(27) Zacharias, M. *Protein Sci.* **2003**, *12*, 1271.

(28) Izvekov, S.; Voth, G. A. *J. Phys. Chem. B* **2005**, *109*, 2469.

(29) Izvekov, S.; Voth, G. A. *J. Chem. Phys.* **2005**, *123*, 134105.

(30) Izvekov, S.; Parrinello, M.; Burnham, C. J.; Voth, G. A. *J. Chem. Phys.* **2004**, *120*, 10896.

(31) Izvekov, S.; Voth, G. A. *J. Phys. Chem. B* **2005**, *109*, 6573.

(32) De Boor, C. *Practical Guide to Splines*; Springer-Verlag: New York, 1978.

(33) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.

(34) Ayton, G.; Smondyrev, A. M.; Bardenhagen, S. G.; McMurtry, P.; Voth, G. A. *Biophys. J.* **2002**, *83*, 1026.

(35) Smondyrev, A. M.; Berkowitz, M. L. *Biophys. J.* **2001**, *80*, 1649.

(36) Scott, H. L. *Biophys. J.* **1991**, *59*, 445.

(37) Tobias, D. J.; Tu, K.; Klein, M. L. *Curr. Opin. Colloid Interface Sci.* **1997**, *2*, 15.

(38) Wong, P.; Capes, S.; Mantsch, H. H. *Biochim. Biophys. Acta* **1989**, *980*, 37.

(39) Guo, W.; Kurze, V.; Huber, T.; Afdhal, N. H.; Beyer, K.; Hamilton, J. A. *Biophys. J.* **2002**, *83*, 1465.

(40) Hofsäss, C.; Lindahl, E.; Edholm, O. *Biophys. J.* **2003**, *84*, 2192.

(41) Levy, R. M.; Gallicchio, E. *Annu. Rev. Phys. Chem.* **1998**, *49*, 531.

(42) Filippov, A.; Orädd, G.; Lindblom, G. *Biophys. J.* **2003**, *84*, 3079.

(43) Almeida, P. F. F.; Vaz, W. L. C.; Thompson, T. E. *Biophys. J.* **2005**, *88*, 4434.

(44) Mahoney, M. W.; Jorgensen, W. L. *J. Chem. Phys.* **2000**, *112*, 8910.

(45) *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*, 20899; Linstrom, P. J., Mallard, W. G., Eds.; NIST: Gaithersburg, MD, 2003.

CT050300C

JCTC Journal of Chemical Theory and Computation

Probing Membrane Insertion Activity of Antimicrobial Polymers via Coarse-Grain Molecular Dynamics

Carlos F. Lopez,[†] Steven O. Nielsen,[†] Goundla Srinivas,[†] William F. DeGrado,[‡] and Michael L. Klein*[,†]

Center for Molecular Modeling, Department of Chemistry, University of Pennsylvania, 231 South 34th Street, Philadelphia, Pennsylvania 19104-6323, and Department of Biochemistry and Biophysics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104-6059

Received November 30, 2005

Abstract: Knowledge of the mechanism of action of antimicrobial agents is crucial for the development of new compounds to combat microbial pathogens. To this end, computational studies on the interaction of known membrane-active antimicrobial polymers with phospholipid bilayers reveal spontaneous membrane insertion and cooperative action at low and high concentrations, respectively. In late-stage attack, antimicrobials cross the membrane core and occasionally align to provide a stepping-stone pathway for water permeation; this suggests a possible new mode of action that does not depend on pore formation for transport to and across the inner leaflet. The computations rationalize the observed activity of a new class of antimicrobial compounds.

Introduction

The emergence of new bacteria and viruses,[1,2] increasing pathogen resistance to treatment,[3] and the threat of bioterrorism[4] underscore the vital need to understand the mechanisms of action of antimicrobial molecules. For these reasons, renewed interest in generic classes of antigenic therapeutic agents[5−11] has led to synthetic designs which show great potential for combating infectious pathogens. Indeed, nonpeptidic synthetic antimicrobials show promise as effective alternatives to their natural peptide counterparts.[12] Some advantages include lower cost of manufacturing and the fact that synthetic molecules are more resistant to degradation by the host organism. Various mechanisms of action have been proposed for membrane-targeting antimicrobials. Although no general mechanism of action has been established even for cationic peptides, the commonly accepted Shai-Matsuzaki-Huang model provides a reasonable explanation for the observed antimicrobial activity of such compounds.[6,13,14] A number of molecules exist, however, that do not seem to completely follow this model in their mechanism of action.[15,16]

Motivating the work presented here is a desire to understand the nature of poration caused by synthetic antimicrobial (AM) molecules interacting with a phospholipid membrane. The specific AM molecule considered is the amphipathic aryl amide dimer (Figure 1), which has been shown to have potent antimicrobial activity and was inspired by natural and synthetic peptides.[6,10,12,17] It was shown experimentally that arylamide polymers, at different concentrations, act as bactericidal agents.[12] Their interaction with the membrane was established by the measurement of AM induced calcein leakage from large unilamellar vesicles. Here, we report the results of coarse grained (CG) molecular dynamics (MD) simulations of a fully hydrated 256 dimyristoyl-phosphatidylcholine (DMPC) lipid bilayer in the presence of 0, 1, 2, 8, or 18 AM molecules, to assess concentration dependent effects on the membrane. Although the bulk concentration is known experimentally, the two-dimensional concentration of antimicrobials at the surface is not known. For this reason we probe the antimicrobial-membrane relation at different concentrations. These concentrations were chosen to be

* Corresponding author e-mail: klein@lrsm.upenn.edu.
[†] University of Pennsylvania.
[‡] University of Pennsylvania School of Medicine.

10.1021/ct050298p CCC: $33.50
Published on Web 03/24/2006

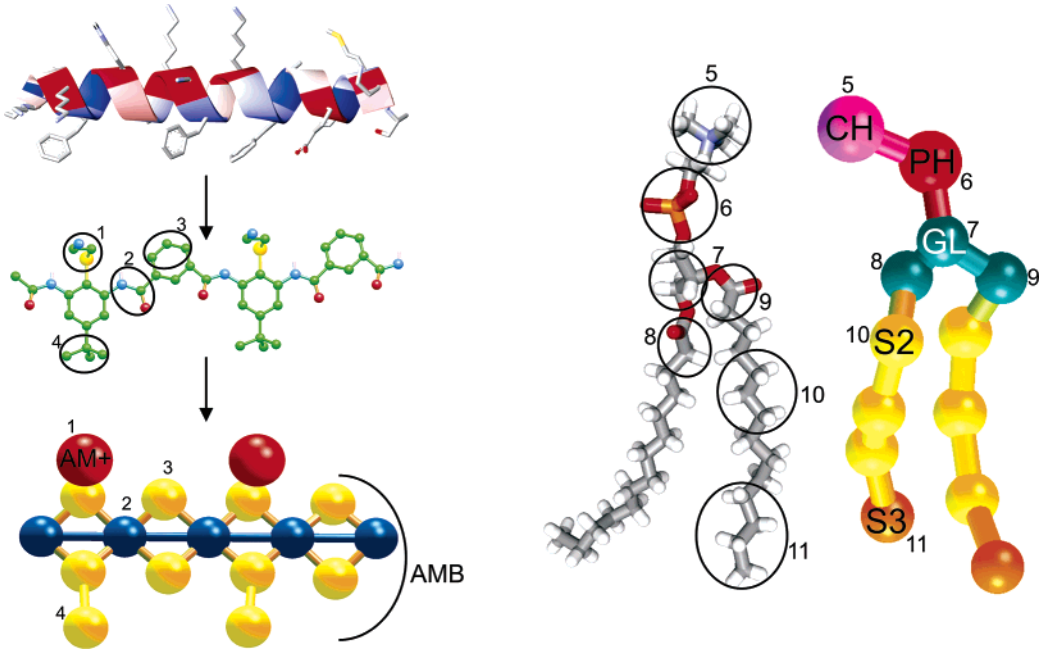


Figure 1. Inspiration for the design of synthetic antimicrobial targets and the coarse grain model used in the present simulations. Left: Ribbon representation of magainin, a potent natural antimicrobial. The α-helix ribbon is colored by hydrophobicity (red=hydrophobic, blue=hydrophilic), and selected individual amino acids are shown as sticks. Middle: The aryl amide dimeric polymer, a purely synthetic molecule inspired by natural antimicrobials. Color code: carbon, green; sulfur, yellow; oxygen, red; nitrogen, blue; hydrogen, white. Bottom: A coarse grain representation of the aryl amide dimeric polymer. Color code: cationic, red; peptidic, blue; hydrophobic, yellow. Right: The corresponding units between an all atom DMPC molecule (left) and its CG representation (right). Color code: nitrogen, blue; carbon, gray; oxygen, red; phosphate, orange; hydrogen, white. CG color code: choline (CH), purple; phosphate (PH), red; glycerol (GL), blue-gray; alkane (S2), yellow; terminal alkane (S3), orange. Circles in the atomistic representations correspond to the numbered units in the coarse grain representations.

similar to those observed by Huang in experiments with membrane-active peptides.[7]

Molecular dynamics (MD) computer simulations provide a powerful complement to experiments on biological systems,[18,19] but to date attempts to explore the interactions of antimicrobials with membranes have probed only relatively short-time behavior[20] (10 ns or so) and few have included more than one membrane-active molecule.[21] There is a lack of quantitative information due to the difficulty in simulating events that span the relevant range of time scales for the host membrane, which is a sluggish partially ordered liquid−crystalline system. Modern coarse grain (CG) work, inspired by the pioneering efforts of Smit and co-workers,[22] has resulted in the development of several models for membrane simulations in the past years.[23,24] Such advances in methodology have allowed access to unprecedented time scales in MD simulations, extending them toward biologically relevant time scales.[25] The longest simulation in this work, effectively 2 μs, would be difficult to run using all-atom interaction potentials, with current algorithms and computational resources.[26] Using CG methods it is now possible to systematically explore events in this temporal regime.

Methods

CG Model. Reduced representations for water, alkane, phospholipids, and the arylamide based antimicrobial are used in the present work. The model is shown in detail in Figure 1, with labels which will be referenced throughout the manuscript. The model consists of units that represent collections of atoms which are tuned to mimic some key physical or structural features known from experiment or atomistic simulation as shown in Figure 1a,b. The basic model and current developments are described in detail in our earlier papers.[23,25,27] Briefly, single spherical sites represent triplets of alkane carbon atoms and their accompanying hydrogen atoms (S2 and S3 units). The hydrocarbon sites are linked together to form chains using stretching and bending potentials. Single spherical sites also represent triplets of water molecules (WS, not shown). The headgroup units are chosen so that the choline (CH) unit is represented by one site, the phosphate (PH) unit by one site, and the glycerol (GL), along with the accompanying carboxylic acids (EST, not labeled in Figure 1 for clarity), by three sites. Several existing units were used to build the arylamide antimicrobial.[12] The positively charged section was approximated using an existing choline site (AM+), which is similar to the group we are trying to emulate. The remainder of the antimicrobial was modeled using generic hydrophobic sites (S2, S3 where appropriate), and carboxylic sites (EST) from the original model in an effort to preserve both the hydrophobic and slightly hydrophilic character of the antimicrobial bonds and the benzene ring. Special care was taken to preserve the overall size and shape of the arylamide molecules in the coarse grain construct.

One of the most important advantages of this approach is the effective increase in simulation times when compared to an equivalent all-atom molecular dynamics simulation. With the present model we observe an increase in time

efficiency of about 4 orders of magnitude,[25] currently allowing access to microsecond phenomena employing modest computational resources. In the present paper we report effective simulation times for each simulation. We have used the lateral lipid diffusion of the CG membrane and compared this to the CG lateral diffusion of an all-atom membrane. The two-dimensional lateral diffusion constant for the L_α phase of DMPC in the plane of the bilayer is 6.5×10^{-8} cm^2 s^{-1} for an all-atom simulation and 6.3×10^{-6} cm^2 s^{-1} for the CG model. The CG lipid motions are therefore 2 orders of magnitude faster than their all-atom counterparts. We use a factor of 1×10^2 to recalibrate the simulation time evolution and account for the molecular diffusion speedup. The effective simulation times reported in the paper therefore reflect this correction.

Simulation. All the simulations presented in this paper were run using our in-house CM3D molecular dynamics program. The program can perform calculations with full electrostatics, multiple time-step integrators following the RESPA[28] technique, and a wide variety of ensembles ranging from microcanonical (NVE) to isobaric-isothermic (NPT). The simulations contained in the present report were all run using the NPT ensemble at 303 K temperature and 1 atm pressure in an orthorhombic simulation cell with flexible sides. The temperature and pressure were controlled with the Nosé-Hoover chain methodology.[29] The outer simulation time step was chosen at 10 fs using a multiple time step algorithm.

Five simulations are reported in this work containing 0, 1, 2, 8, and 18 antimicrobial molecules, placed in close proximity (but outside) of a 256 DMPC hydrated lipid bilayer, which was constructed as described previously.[25] The simulations were run for 204 ns, 86 ns, 160 ns, 160 ns, and 1.9 μs effective simulation time (referred to throughout the text as AM0, AM1, AM2, AM8, and AM18, respectively). With the present model these simulations, for the longest run system, can be run on a high end Intel based single-processor workstation within a few weeks.

Analysis. The atom weighted density profiles were calculated by constructing a histogram along the axis perpendicular to the bilayer plane of each CG unit. Each unit was then weighted by the number of electrons of its corresponding all-atom counterpart (e.g. 42 electrons for the S2 site representing $(-CH_2-)_3$). The electron weight was then distributed along the histogram using a Gaussian function with a width corresponding to the Lennard-Jones radius of the CG unit. Given the nonequilibrium nature of the present simulations, the first quarter of the simulations was not included in the calculation of the atom weighted density profiles for simulations with 1, 2, and 8 AM, as shown in Figure 4. The first 50 ns of the 18 AM simulation were skipped. The next 150 ns were processed to obtain the data presented in Figure 4. The last 200 ns of the 18 AM simulation were used to generate the data presented in Figure 5.

Water transport across the bilayer was assessed by an algorithm written to monitor water molecules that cross the plane defined by the glycerol units on either bilayer leaflet. A water unit was counted once it crossed both glycerol planes passing across the bilayer core. Despite the length of the simulation water crossing is still a rare event. We simply divide the number of crossing waters by the simulation time after half of the AM molecules flip-flop across the core of the membrane to obtain a rough estimate of the qualitative passage rate.

The molecule insertion coordinate was calculated by obtaining the average *z*-axis coordinate between the end peptidic units (labeled "3" in Figure 1). The *z*-axis was defined as the axis perpendicular to the membrane plane, with the zero of the axis set to the center of the bilayer. As a visual aid, lines representing the average position of the choline groups and the glycerol groups were placed on the graphs for reference. Similarly the insertion angles were determined by obtaining the angle between the end peptidic units and the bilayer normal, which was approximated, in this case, as the *z*-axis.

Results and Discussion

Insertion Mechanism. We observe spontaneous insertion of antimicrobial polymers in all of the antimicrobial containing simulations reported here (Figure 2). It appears that cooperative activity occurs at high AM concentrations, as one molecule possesses the capacity to interact with and assist penetration of other AMs into the lipid bilayer. In the early stages of the simulation the AM molecules accumulate at the headgroup region of the outer leaflet lipid−water interface (Figure 2a) with their hydrophilic amide groups in the water. As the simulation evolves the AM molecules dive into the membrane to reside under the lipid headgroups (Figure 2b). At still longer times, as observed in the 18 AM simulation, some AM molecules cross the lipid bilayer core to reside under the headgroups of the inner leaflet (Figure 2c).

We have rationalized our observations by postulating the existence of two AM insertion mechanisms. The first consists of the spontaneous penetration of single, isolated AM molecules (1 AM simulation), which is relevant at lower concentrations. Data for the distance from the AM molecule to the center of the bilayer as well as the orientation of the AM relative to the bilayer plane is shown in Figure 3. The dashed lines show the average position of the lipid CH (∼18 Å) and GL (∼13 Å) groups throughout the simulation. The insertion distance decreases as the simulation evolves, with the molecule center crossing the CH-GL slab in about 15 ns of simulation time (Figure 3a). At this point, once the molecule has crossed the GL region there is further accommodation and reorientation. The AM+ groups, as will be shown below, remain near the headgroup region, while the body of the AM intercalates between the lipid tails. As shown in Figure 3b, the orientation of the molecule throughout the insertion is almost parallel to the membrane plane. This mechanism of insertion seems to indicate a simple spontaneous insertion at low concentrations for the antimicrobial. We are currently working on ways to characterize the insertion time scale for this event.

In addition to the previously described mechanism of insertion, a second mechanism is seen when AM molecules interact. To probe the latter, several simulations (2 AM, 8

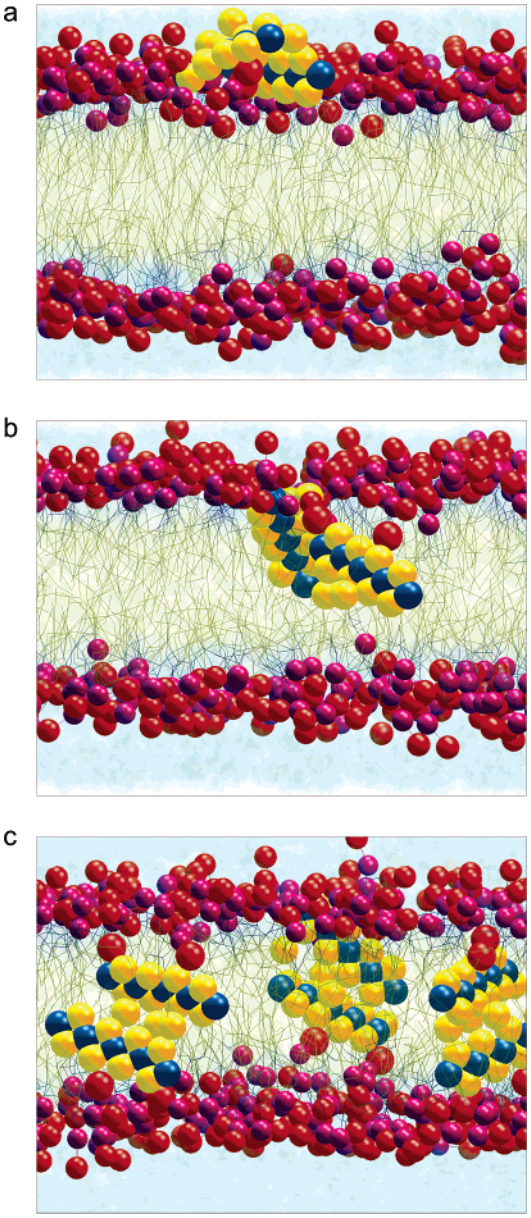


Figure 2. Depiction of the observed two-step - accommodation (a) and penetration (b) mechanism for membrane insertion of the antimicrobial (AM) aryl amide dimer molecules into the outer leaflet of a DMPC lipid bilayer (DMPC: choline, red; phosphate, purple; glycerol group, blue sticks; hydrophobic tails, green sticks). The initial interfacial insertion stage involves adsorption and "snorkeling" of the AM at the bilayer surface (a). In this example of the penetration stage, one AM rotates to become perpendicular to the bilayer plane and drags the accompanying AM into the membrane core (b). Subsequent population of both leaflets occurs at much longer times (c).

AM, and 18 AM) were initiated with AM molecules in close proximity relative to each other and wholly within the water subphase. The AM long axes were positioned parallel to the lipid membrane, and all the AM molecules were initially placed completely in the water subphase. We present the insertion distance and orientation angle for both AM molecules during insertion observed in the 2 AM simulation, shown in Figure 3c,d. Both antimicrobials initially approach the membrane and accommodate near the headgroup region as shown in Figure 2a. Figure 3c shows how the centers of the AM molecules approach the bilayer at about the same distance and time. It is important to note here that there is "contact" between the AMs throughout the insertion process. Within a short time of accommodation within the lipid headgroup region, one of the antimicrobials (labeled 2AM1 in Figure 3c) very rapidly changes orientation (Figure 3d). This antimicrobial will conserve this steep angle and its interaction with the second antimicrobial, which remains relatively parallel to the membrane throughout the insertion process. After this change in orientation, 2AM1 penetrates the membrane rather quickly. During this process it continues to interact with the second antimicrobial (2AM2) and "drags" it through the CH-GL region. After the groups have crossed the headgroup region dissociation between the two AMs occur and a shallower angle relative to the bilayer plane is recovered.

Eventually all of the arylamide monomers penetrate the membrane at sufficient time scales. Interfacial insertion times, throughout the simulations, were about 12 ns, and penetration times ranged from 10 to 20 ns throughout the different concentration simulations. About 50% of the AMs insert into the membrane via an accommodation/penetration mechanism as described here. The remaining 50% of the molecules are dragged into the bilayer core by those that experience a two stage mechanism. Further simulation might be required to extract proper time scales related to the accommodation and insertion events. The insertion and penetration mechanism observed in this study expands and adds pictorial detail to the traditional view of aggregation and oligomerization attributed to natural AM peptides such as magainins and melittins.[6,7,13,14] We are currently working on understanding how the cooperative interaction between AM molecules could be related to experimentally observed activity.

Membrane Structure. Figure 4 shows the results depicting the positions of the different protein components along the *z*-axis immediately after penetration of the membrane leaflet, nearest the side of attack, has taken place. The data analyzed here corresponds to the total simulation time for 0 AM, 1 AM, 2 AM, and 8 AM simulations and to about 200 ns of simulation time for the 18 AM simulation. As can be seen (Figure 4a), the membrane density on the side of attack decreases and broadens as the concentration increases. This is indicative of a displacement of the lipid headgroups due to the insertion of the AM polymers into the protein. There seems to be a decrease in the water density on the side of attack as the concentration increases as well. This is due, at least in part, to the antimicrobial molecules displacing water molecules at the lipid headgroup region. As shown by the dashed lines (Figure 4b), corresponding to the AM polymers, the AM+ units reside near the GL group of the lipid and between the GL and CH groups of the lipid. The body of the AM polymers however sits well past the GL groups. In the low concentration case a peak appears near the center of the bilayer. At higher concentrations, the distribution is broader, and the AM groups span several Ångstroms between the GL groups and parts of the opposite side of the lipid membrane. After insertion/penetration, the AM+ groups

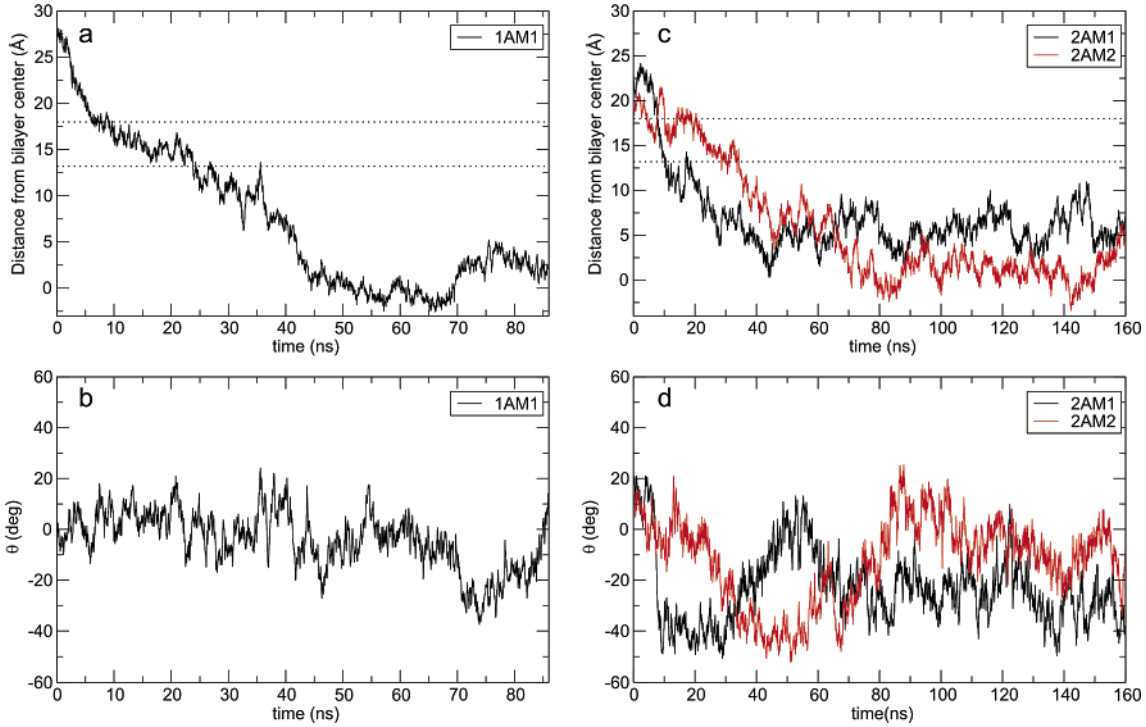


Figure 3. Insertion distance to bilayer center (top graphs) and orientation angle (bottom graphs) for the 1AM simulation (left) and the 2AM simulation (right). Dashed lines represent the average positions of the CH group (top line) and the GL group (bottom line) within the attacked leaflet.

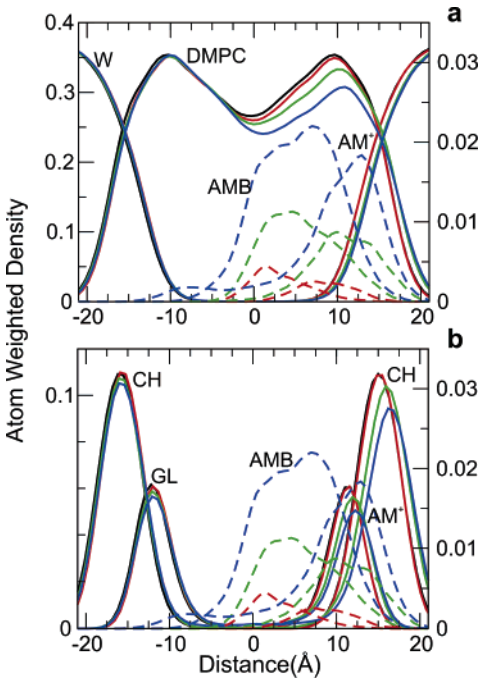


Figure 4. Atom weighted densities for the lipid bilayer (solid lines) and the AM molecules (dashed lines) for simulations with 0 (black), 2 (red), 8 (green), and 18 (blue) antimicrobial molecules. The scale on the left corresponds to the lipid molecules, while the scale on the right corresponds to the AM polymers.

reside close to the glycerol moieties, while the remainder of the molecule resides in the hydrophobic core.

At sufficiently long times all the AM molecules eventually insert into the membrane core and become oriented with their long axes parallel to the membrane surface. AM associations during the initial insertion process are lost via diffusion after further penetration into the membrane. After about 1 μs of simulation time, half of the AM polymers spontaneously cross the bilayer core and sit opposite the attack leaflet (18 AM simulation). The atom weighted densities showing the membrane structure are shown in Figure 5. At long times, there seems to be a preference for the AM body (labeled AMB in Figure 5a,b) to sit wholly near the center of the lipid membrane. The AM+ groups seem to prefer a position slightly below the GL groups (Figure 5b). This is in contrast to that observed initially after insertion where the AM+ group sits between the GL and CH groups. We believe this is due to transbilayer interactions with both lipids and other antimicrobial molecules. Indeed, as we will discuss below, AMs in one leaflet can interact and align with AMs from the opposite leaflet.

Typically, thinning of the system occurs when AM molecules are added to a hydrated lipid bilayer.[30] In the present simulations the membrane surface area (simulation cell cross section) increases slightly (6% at the highest concentration) and the overall system thickness (simulation cell height) decreases slightly (1% at the highest concentration) with increasing number of AMs.

Water Penetration and Transport. A decrease in water penetration at the attacked bilayer leaflet and an increase in water penetration at the leaflet opposite the AM attack are observed in the simulations. As AMs snorkel and dive through the membrane headgroup region they exclude water that normally resides at the interface. The leaflet opposite the AM attack expands in response to the swelling of the attacked leaflet, thereby enabling water penetration. This

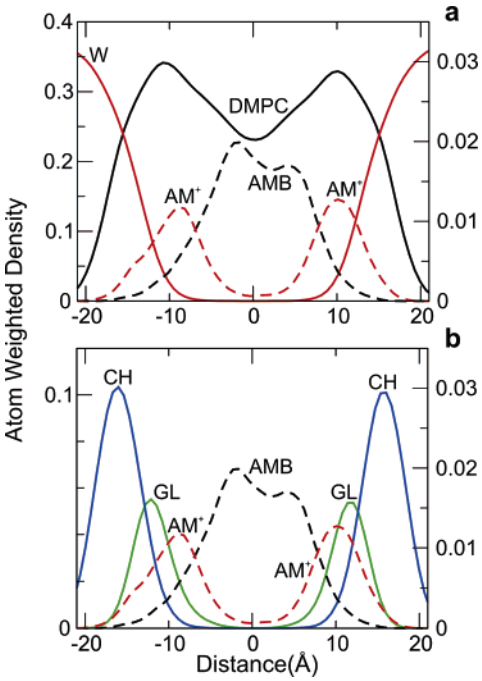


Figure 5. Atom weighted densities for the lipid bilayer (solid lines) and the AM molecules (dashed lines) for simulations with 0 (black), 2 (red), 8 (green), and 18 (blue) antimicrobial molecules. The scale on the left corresponds to the lipid molecules, while the scale on the right corresponds to the AM polymers.

effect is exacerbated with increasing AM concentration and has been noted previously for melittin penetration into a lipid bilayer.[31] However, the asymmetry between the pure inner leaflet and the AM containing outer leaflet may be exaggerated due to finite size effects, which in turn can enhance water penetration into the leaflet opposite to the AM attack.

In the absence of AM molecules in the bilayer, water can spontaneously diffuse past the lipid headgroups but is normally expelled back into the bulk region. The presence of AM molecules within both bilayer leaflets allows for occasional AM alignment perpendicular to the bilayer plane, thereby providing a stepping-stone pathway for water molecules to cross the membrane, as shown schematically in Figure 6. In this figure, snapshots of a single water molecule are taken during a passage across the membrane. The lipid bilayer and the surrounding water molecules have been abstracted into color panels for clarity. The water approaches from the water subphase and crosses the lipid headgroup region, where it finds an antimicrobial molecule. At this stage, the water unit interacts with the hydrophilic site until the bottom AM molecule aligns properly. At this time a quick jump occurs between both molecules, and the water now resides in the lower leaflet. The water finally crosses the lower leaflet headgroup region and returns to the water subphase.

In effect, water or ions could become transiently bound within the core of the bilayer by the hydrophilic groups of the AM. Thermal fluctuations can drive this bound water deeper into the membrane whereupon it can jump to an appropriately positioned AM in the opposite leaflet and eventually exit the membrane after crossing the bilayer core.

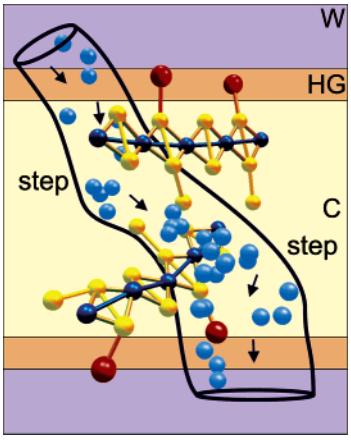


Figure 6. Representative water trajectories (blue spheres are CG waters) showing passage from the bulk water phase, W1 (purple), through the headgroup region HG1 (orange), and hydrophobic membrane core, C (yellow) to the HG2 and W2 regions via a pathway employing the aryl amide molecules as stepping stones.

The observed water transport is a rare event for a system of this size and even for a simulation of this length. In the highest concentration simulation, the water molecule passage rate is roughly 1/ns, which is about three times that of the pure CG membrane. This type of ion or water transport has been suggested as a viable mechanism for the function of hydraphile channels.[32] It is important to note that our water unit represents three loosely packed water molecules. As such, it is also comparable to a hydrated ion. This suggests that hydrated ions could gain passage through the bilayer by means of this mechanism. It is important to note that these water-passage results across the membrane are to be taken in a slightly qualitative manner as the coarse grain representation of the units interacting here does not allow for a more quantitative description of the passage event. It is clear, however, that the presence of the antimicrobials, mediating the water passage, results in an increase in the overall water-passage rate.

Conclusion

We see no evidence for the formation of pores when these synthetic AMs oligomerize during attack on the membrane and subsequently undergo trans-bilayer flips. This lack of pore formation could be due, at least in part, to the fact that the AMs are too short to span the bilayer.[12] However, interactions between AMs are fundamental for the disruption of the lipid bilayer. As such, these results emphasize AM cooperativity which is integral to the traditional Shai-Matsuzaki-Huang model and provides an alternative mechanism for transport across the membrane. This molecular level simulation of the interfacial insertion and penetration process suggests a possible new approach to generate targets drugs for pathogen eradication. The present results are not aimed to dispute the compelling evidence that exists regarding nanopore formation by natural AMs, rather they provide a basis for the design of low-cost synthetic AM molecules which could disrupt a cellular membrane through an alternative pathway.

The traditionally accepted mechanism for water transport involves the formation of pores by amphiphilic peptides. Here, however, the AMs, rather than forming a pore, provide a local hydrophilic environment to trap water within the globally hydrophobic membrane core. At high AM concentrations the probability for the alignment of AMs in opposite leaflets to form a bridge across the bilayer is increased. This suggests a previously unexplored pathway for membrane leakage and poration. Furthermore, in a fully atomistic environment, the formation of water wires across the bilayer seems plausible.[33,34] From our results we see that pore formation may not be necessary and that if AM bilayer flips can occur, water leakage is possible.

The efficiency of the aryl amide class of AMs, their ease of synthesis, and, as shown in this paper, their specific interaction with the membrane provides thought-provoking insight into possible AM design. The presented observations also provide an explanation of the fundamental interactions between generic amphiphilic molecules and lipid membranes. The present study presents researchers with a qualitative picture of the mode of action of a purely synthetic AM polymer which can be generalized to generic models and thus should be able to impact in the design of new and more potent antimicrobial agents.

Acknowledgment. This research was supported by generous grants from the National Institutes of Health (NIH). The authors thank Preston Moore, Bin Chen, Robert Doerksen, and Dahui Liu (PolyMedix) for discussions and important insights. C.F.L. would like to thank Richard Darst for assistance during the writing of this manuscript.

Note Added after ASAP Publication. This paper was released ASAP on March 24, 2006, with the incorrect Received Date. The correct version was posted on April 12, 2006.

References

(1) Brower, V. *EMBO Rep.* **2003,** *4*, 649−651.

(2) Cohen, M. L. *Nature* **2000**, *406*, 762−767.

(3) Anderson, R. M. *Nature Med.* **1999**, *5*, 147−149.

(4) Niller, E. *Nature Biotechnol.* **2002**, *20*, 21−25.

(5) Fernandez-Lopez, S.; Kim, H. S.; Choi, E. C.; Delgado, M.; Granja, J. R.; Khasanov, A..; Kraehenbuehl, K.; Long, G.; Weinberger, D. A.; Wilcoxen, K. M.; Ghadiri, M. R. *Nature* **2001**, *412*, 452−455.

(6) Zasloff, M. *Nature* **2002**, *415*, 389−395.

(7) Yang, L.; Harroun, T. A.; Weiss, M.; Ding, L.; Huang, H. W. *Biophys. J.* **2001**, *81*, 1475−1485.

(8) Bailey, H. *Curr. Opin. Biotech.* **1999**, *10*, 94−103.

(9) Lear, J. D.; Gratkowski, H. L.; DeGrado, W. F. *Biochem. Soc. T.* **2001**, *29*, 559−564.

(10) Raguse, T. L.; Porter, E. A.; Weisblum, B.; Gellman, S. H. *J. Am. Chem. Soc.* **2002**, *124*, 12774−12785.

(11) Patch, J. A.; Barron, A. E. *J. Am. Chem. Soc.* **2003**, *25*, 12092−12093.

(12) Tew, G. N.; Liu, D. H.; Chen, B.; Doerksen, R. J.; Kaplan, J.; Carroll, P. J.; Klein, M. L.; DeGrado, W. F. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5110−5114.

(13) Shai, Y. *Biopolymers* **2002**, *66*, 236−248.

(14) Matsuzaki, K. *Biochim. Biophys. Acta*, **1999**, *1462*, 1−10.

(15) Allen, N. E.; Nicas, T. I. *FEMS Microbiol. Rev.* **2003**, *26*, 511−532.

(16) Although glycopeptide antibiotics such as vancomycin have membrane activity, their efficacy relies on the inhibition of the production of bacterial peptidoglycan for cell wall maintenance.

(17) Porter, E. A.; Wang, X. F.; Lee, H. S.; Weisblum, B.; Gellman, S. H. *Nature* **2000**, *404*, 565−565.

(18) Daggett V.; Fersht, A. *Nat. Rev. Mol. Cell Biol.* **2003**, *4*, 497−502.

(19) Jensen, M. Ø.; Park, S.; Tajkhorshid, E.; Schulten, K. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 6731−6736.

(20) La Rocca, P.; Shai, Y.; Sansom, M. S. P. *Biophys. Chem.* **1999**, *76*, 145−159.

(21) Aliste, M. P.; MacCallum, J. L.; Tieleman, D. P. *Biochemistry* **2003**, *42*, 8976−8987.

(22) Karaborni, S.; Esselink, K.; Hilbers, P.; Smit, B.; Karthauser, J.; Vanos, N.; Zana, R. *Science* **1994**, *266*, 254−256.

(23) Shelley, J. C.; Shelley, M. Y.; Reeder, R. C.; Bandyopadhyay, S.; Klein, M. L. *J. Phys. Chem. B* **2001**, *105*, 4464−4470.

(24) Izvekov S.; Voth G. A. *J. Phys. Chem. B* **2005**, *109*, 2469−2473.

(25) Lopez, C. F.; Moore, P. B.; Shelley, J. C.; Shelley, M. Y.; Klein, M. L. *Comput. Phys. Commun.* **2002**, *147*, 1−6.

(26) Kale, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K. *J. Comput. Phys.* **1999**, *151*, 283−312.

(27) Nielsen, S. O.; Lopez, C. F.; Srinivas, G.; Klein, M. L. *J. Phys.: Condens. Matter* **2004**, *16*, R481−R512.

(28) Tuckerman, M. J.; Berne, B. J.; Martyna, G. J. *J. Chem. Phys.* **1992**, *97*, 1990−2001.

(29) Moore, P. B.; Klein, M. L. *Implementation of a General Integration for Extended System Molecular Dynamics*; Technol. Report; University of Pennsylvania, 1997. http://www.cmm.upenn.edu/moore/code/code.html.

(30) Chen, F. Y.; Lee, M. T.; Huang, H. W. *Biophys. J.* **2003**, *84*, 3751−3758.

(31) Berneche, S.; Nina, M.; Roux, B. *Biophys. J.* **1998**, *75*, 1603−1618.

(32) Leevy W. M.; Donato, G. M.; Ferdani, R.; Goldman, W. E.; Schlesinger, P. H.; Gokel, G. W. *J. Am. Chem. Soc.* **2002**, *124*, 9022−9023.

(33) Tepper, H. L.; Voth, G. A. *Biophys J.* **2005**, *88*, 3095−3108.

(34) Decoursey, T. E. *Physiol. Rev.* **2003**, *83*, 475−579.

CT050298P

JCTC Journal of Chemical Theory and Computation

Resolution Exchange Simulation with Incremental Coarsening

Edward Lyman[†] and Daniel M. Zuckerman*[,†]

Department of Computational Biology, School of Medicine, and Department of Environmental and Occupational Health, Graduate School of Public Health, BST W1041, 200 Lothrop Street, University of Pittsburgh, Pittsburgh, Pennsylvania 15261

Received March 9, 2006

Abstract: We previously developed an algorithm, called "resolution exchange", which improves canonical sampling of atomic resolution models by swapping conformations between high- and low-resolution simulations. Here, we demonstrate a generally applicable incremental coarsening procedure and apply the algorithm to a larger peptide, met-enkephalin. In addition, we demonstrate a combination of resolution and temperature exchange, in which the coarser simulations are also at elevated temperatures. Both simulations are implemented in a "top-down" mode, to allow efficient allocation of CPU time among the different replicas.

Atomic resolution simulations of proteins are currently limited to short durations (less than 1 μs)[2] or small systems (less than 100 residues).[3,4] Furthermore, accurate calculations involving large conformational changes are not possible for any system, as the cost of calculating entropic contributions is too great. Indeed, the cost of such calculations is only going to increase, as empirical force fields are improved by including polarization effects, either in a classical way[5,6] or in a semiclassical way.[7]

Thoroughly sampling the space of conformations is essential for a number of problems. From a purely biological perspective, there is a growing awareness of the importance of protein fluctuations—over and above the static picture—in the function of most proteins.[8] Allostery and conformational changes dramatic enough to be captured experimentally are just two examples of the existence of such fluctuations.[9,10] In a computational context, careful validation of empirical force fields requires confidence in the quality of conformational sampling, so that error may be attributed to the force field rather than undersampling. The calculation of free energy differences—as required for evaluation of binding affinities of small molecules[11] or the strength of protein−protein interactions[12]—also requires reliable sampling.[13,14]

The undersampling (or "quasi-ergodicity") problem is widely recognized, and consequently there have been many attempts to improve upon standard simulation protocols. Methods which aim to generate a canonical distribution of conformations include multiple time-step methods,[15,16] nonlinear variable transformations,[17] J-walking,[18] inverse renormalization group approaches,[19] and adaptive resolution methods.[20] The most widely used class of methods, however, comprises the generalized ensemble approaches.[21−23] Of the generalized ensemble approaches, perhaps the simplest and most popular is parallel tempering, in which a number of copies of the system are evolved in parallel at different temperatures.[24−27] Occasionally, configurations are swapped between neighboring replicas, presumably allowing the low-temperature replica to access more configuration space via high-temperature conformations.

Numerous[28−33] as well as extensive[34,35] parallel tempering simulations have been published, including some which claim to demonstrate the superior efficiency of the method over standard molecular dynamics (MD) simulation.[36−38] Regardless of the validity of those claims, there appears to be a limit to the utility of parallel tempering for the study of large proteins, nucleic acids, and macromolecular complexes: the number of replicas required to bridge a specified temperature gap increases as the square root of the number of degrees of freedom of the solution.[39,40] In other words, if atomic resolution information is desired, then very many atomic resolution simulations are required. Recent work by Berne and co-workers partly addresses this problem for explicitly

* Corresponding author e-mail: elyman@ccbb.pitt.edu.
† Present address: 3088 BST3, 3501 Fifth Avenue, University of Pittsburgh, Pittsburgh, PA 15213.

10.1021/ct050337x CCC: $33.50
Published on Web 04/22/2006

solvated systems, so that the number of replicas scales with the number of degrees of freedom of the solute only.[41] Solutes such as proteins, of course, can be quite large.

In this paper, we address the problem of insufficient sampling of implicitly solvated biomolecules using a different approach. We recently developed an algorithm, called resolution exchange, which uses a distribution generated by a coarse-grained model to improve the sampling of a higher-resolution simulation.[1] The resolution exchange (ResEx) algorithm guarantees canonical sampling for each level of resolution, so that the coarse-grained simulation introduces no bias into the high-resolution simulation. The algorithm is similiar in spirit to other exchange simulations, in that conformations are swapped between otherwise independent simulations. However, by employing replicas of reduced resolution, ResEx has the potential for significant efficiency gains. Other authors have recognized the value of improving sampling with reduced resolution representations. For example, Iftimie et al. used a classical potential as an importance function to improve sampling of an ab initio potential.[42] Here, our goal is to sample a *classical* atomic resolution potential, which leads us to a different algorithm. Also, Liu and Sabatti formally introduced a Markov chain Monte Carlo method which allows jumps between spaces of different dimensions.[43] Their algorithm has apparently not been applied to the simulation of macromolecules. Lwin and Luo recently introduced an algorithm similiar to ours, but it does not generate canonical sampling.[44]

We have also employed a modification of the usual parallel protocol used to carry out exchange simulations,[1] generalizing the "J-walking" approach previously introduced by Frantz et al.[18] The J-walking (or as we call it "top-down" exchange) method allows an unequal distribution of CPU time among the various replicas. We emphasize that any exchange simulation may be run in top-down mode. In contrast with other exchange methods, top-down exchange allows very little simulation time to be spent on the computationally expensive, high-resolution (or low temperature) replicas. Substantial improvement in sampling efficiency is therefore possible, in principle.

We previously applied the resolution exchange algorithm to butane and dileucine peptide.[1] Here, we confront issues which arise in the study of larger molecules. We show that a molecule may be coarsened incrementally, so that the overlap between models of neighboring resolution may be adjusted for improved sampling efficiency. We also demonstrate that resolution and temperature exchange are easily combined in a single simulation, so that sampling may be improved by both increasing temperature and decreasing resolution simultaneously. The incremental coarsening procedure is first demonstrated on dileucine, where we check that the correct conformational distribution is attained. We then demonstrate successful exchange between an all-atom model and an united-atom model of met-enkephalin, using two different exchange ladders: a ladder of varying resolution only and a ladder which combines resolution and temperature changes.

We will finish with a discussion of the next logical steps toward larger peptides and proteins.

1. Theory and Methods

The results presented in this paper concern two distinct, recently introduced simulation methods.[1] The first is resolution exchange, which allows exchange between simulations at different resolutions, and preserves canonical sampling. The second is top-down exchange, which allows unequal distribution of CPU time, maximizing the efficiency of an exchange simulation. In addition, we describe a general incremental coarsening strategy for building a ladder of models which improves exchange efficiency.

1.1. Resolution Exchange. Resolution exchange (ResEx) is motivated by the effectiveness of coarse-grained models for sampling of protein conformations[45,46] and by the need for atomic-level resolution for many calculations. ResEx uses coarse-grained simulation to accelerate basin-hopping in more detailed models. In contrast with *ad hoc* methods, ResEx guarantees canonical sampling of the atomic-resolution model.

The basic idea, as in any exchange simulation, is to exchange conformations between two simulations. How are trial configurations constructed for an exchange between models with different numbers of degrees of freedom? Consider a pair of models: a coarse-grained model, a configuration of which is described by a set of coordinates, Φ, and an atomic-resolution model described by a larger set, $\{\Phi,\mathbf{x}\}$. Note that the coarse model is built from a *subset* of the coordinates of the detailed model. Before the exchange, let the coarse-grained configuration be Φ_a, and let the atomic-resolution coordinates be $\{\Phi_b,\mathbf{x}_b\}$. By swapping only coarse variables, the trial configuration for the coarse-grained model is simply Φ_b and for the atomic-resolution model is $\{\Phi_a,\mathbf{x}_b\}$.

The exchange criterion is derived by considering the two simulations to consitute a single system characterized by the combined coordinates $\{\Phi_a,(\Phi_b,\mathbf{x}_b)\}$. Because the simulations—aside from the exchanges—run independently, the probability distribution of the combined system is the simple product of the individual distributions. Let the potential functions of the high- and low-resolution simulations be $U_H(\Phi,\mathbf{x})$ and $U_L(\Phi)$, respectively, and denote the Boltzmann factors as $\pi_H(\Phi,\mathbf{x};\beta_H) = e^{-\beta_H U_H(\Phi,\mathbf{x})}/Z_H$ and $\pi_L(\Phi;\beta_L) = e^{-\beta_L U_L(\Phi)}/Z_L$, where Z_H and Z_L are the partition functions. Then the exchange attempt is accepted with the Metropolis rate:

$$\min\left[1, \frac{\pi_H(\Phi_a,\mathbf{x}_b;\beta_H)}{\pi_H(\Phi_b,\mathbf{x}_b;\beta_H)} \frac{\pi_L(\Phi_b;\beta_L)}{\pi_L(\Phi_a;\beta_L)}\right] \quad (1)$$

The dependence on inverse temperature (β) is made explicit, as a reminder that the method is naturally combined with temperature exchange, though this of course extends to any type of exchange, such as Hamiltonian exchange.[47] Note that in the case of ordinary (temeprature-based) replica exchange, in which all the coordinates are swapped, eq 1 reduces to the familiar expression $\min[1,\exp(-\Delta\beta\Delta U)]$.

In a parallel implementation, eq 1, together with the protocol for trial move construction, ensures that the algorithm satisifies the detailed balance condition. To see this, consider "old" (o) and trial/"new" (n) configurations of the combined system. In the construction of any Boltzmann-preserving Monte Carlo move, two transition proba-

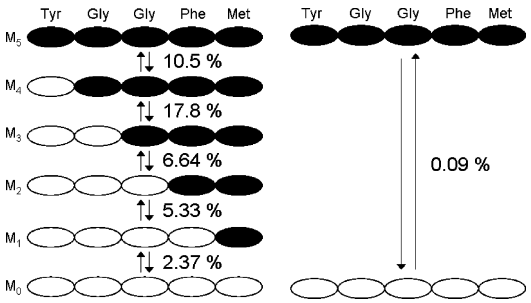


Figure 1. Two different ladders used to exchange all-atom with united-atom met-enkephalin. Residues are depicted with ovals—open corresponds to an all-atom representation, filled to united atom. The ratios of successful to attempted exchanges between each level are indicated by the percentages.

bilites must be considered: the conditional probability $\alpha(o \rightarrow n)$ of *generating* the move from configuration o to n and the conditional probability $w(o \rightarrow n)$ of *accepting* the move.[48] Detailed balance insists that $p(o)\alpha(o \rightarrow n)w(o \rightarrow n) = p(n)\alpha(n \rightarrow o)w(n \rightarrow o)$, where $p(j)$ is the equilibrium probability of configuration j. Yet the acceptance criterion (1) has the form

$$\frac{w(o \rightarrow n)}{w(n \rightarrow o)} = \frac{p(n)}{p(o)} \tag{2}$$

implying that the generating probabilities α are identical. This is indeed the case: given a predefined division into coarse and detailed coordinates, the conditional probability for the move $o = \{\Phi_a,(\Phi_b,\mathbf{x}_b)\} \rightarrow n = \{\Phi_b,(\Phi_a,\mathbf{x}_b)\}$ and its inverse are both one. That is, given the old configuration of the combined system, there is a unique trial configuration.

Lwin and Luo have constructed a similiar algorithm, except that *before* checking acceptance via eq 1, the high-resolution trial configuration is *minimized.*[44] Such minimization (even subject to constraints on the coarse coordinates Φ, as in ref 44) violates the detailed balance condition by biasing the generating probability, $\alpha(o \rightarrow n)$, without any compensating correction in the acceptance criterion. Put more simply, reverse moves into unminimized configurations are impossible. The consequences of the violation are readily seen, as shown in section 2.1.

1.2. Incremental Coarsening. An important practical issue is raised, however, by the construction of trial moves *without* minimization. The problem is that the degrees of freedom in the high-resolution simulation $\{\Phi,\mathbf{x}\}$ are strongly coupled—for a protein, think of Φ as backbone degrees of freedom (DoF) and $\mathbf{x}$ as side-chain DoF. Then it is clear that construction of trial moves by our method may lead to high rejection rates. We have solved this problem by noting that the rejection rate depends on the both the *number and type* of DoF in the set $\{\mathbf{x}\}$. Employing a ladder of incremental models at intermediate resolutions allows the acceptance rates to be tuned to reasonable values, as shown in Figure 1.

A ladder of incrementally coarsened models is straightforward to construct. Consider coarsening from an all-atom representation of a protein to a united-atom representation. In the first model above the all-atom level, only one residue is described by the united-atom representation—the protein is described by a "mixed model", with one united-atom residue and the rest all-atom. Then, in the next level up, there are two united-atom residues, and so on, until the entire protein is described by the united-atom force field. A similiar procedure may then be used to go beyond the united-atom level to a united-*residue* level. Notice that it may be desirable to coarsen more than one residue at a time, since some residues have fewer degrees of freedom than others.

Of course, implementation of the incremental ladder just described requires the construction of a potential function which has both united- and all-atom groups. In this work, we have built this mixed potential by combining the parameters for united and all-atom force fields into a single file. In other words, we created a larger parameter file, which contains both all-atom and united-atom types. This file also includes all of the interactions for both united- and all-atom types, with the united-atom interactions modified as described in section 1.1. Adding some parameters (taken from the all-atom potential) for the interfaces, where united and all-atom residues link, the mixed potential describes the whole molecule. The parameters (formatted for use in TINKER) are included as Supporting Information.

The incremental coarsening approach just described is rather general and not restricted to implementing exchange ladders spanning united- to all-atom resolutions. Lower resolution models could also be considered, for which it may be desirable to coarsen several residues at once. A first quantitative analysis of the incremental coarsening procedure, suggesting how efficiency can be improved, is given in sections 2.2 and 2.3.

1.3. Top-Down Exchange. In many exchange simulations, whether they are temperature-based,[38] Hamiltonian-based,[47] or use some other extended ensemble,[49,50] the goal is to improve the sampling of a hard-to-sample model (such as an all-atom protein model at native conditions) by sampling a related model, which is *presumed* easier-to-sample (such as the same all-atom model at higher temperature). [For a discussion of these issues from a statistical perspective, see ref 51.] Information is swapped between the simulations by occasionally exchanging configurations, in a way which preserves canonical sampling of each distribution. Usually there is little overlap between the hard-to-sample (henceforth, "bottom level") and the easy-to-sample (henceforth, "top-level") models, and therefore a ladder of intermediate models is required.

A critical observation is that the accuracy which is ultimately attained in the hard-to-sample, bottom-level model is effectively limited by that which is obtained in the easy-to-sample, top-level model.[18] In many cases, the top level is still quite difficult to sample well and will require considerable CPU time to reach an acceptable accuracy—much more than it would usually be allotted in a parallel implementation. It is this observation which motivates the top-down method. The top-down algorithm shown schematically in Figure 2 was developed previously for temperature-based simulation,[18] though was not widely recognized as such. The procedure is as follows:

(i) Run and store a simulation at the top level (model M_N) until it is judged to be sufficiently converged. This trajectory

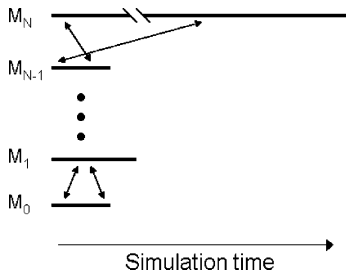


Figure 2. Schematic representation of top-down exchange. Thick horizontal lines are simulation trajectories (labeled "M_i" for model "*i*"), and arrows represent exchanges. The M_i may be differ in resolution, temperature, or both. Notice that the top-level simulation may be considerably longer than the others.

is a sample of the distribution $\pi_N(\mathbf{r})$ of the top level, where N labels the level and $\mathbf{r}$ labels the configurations. In the case of ResEx, $\mathbf{r} = (\Phi,\mathbf{x})$. Let n be a running index, with $n = N$ at this top level.

(ii) Start a simulation at the $n - 1$ level—for example, at the next lower temperature. Configurations $\mathbf{r}_{n-1}$ will be sampled according to π_{n-1} for model M_{n-1}.

(iii) Whenever an exchange is to be attempted, pull a random trial configuration $\mathbf{r}_n$ from the M_n trajectory. In the case of ResEx, one requires only the subset Φ.

(iv) Accept the trial configuration according to

$$\min\left[1, \frac{\pi_{n-1}(\mathbf{r}_n)}{\pi_{n-1}(\mathbf{r}_{n-1})}\frac{\pi_n(\mathbf{r}_{n-1})}{\pi_n(\mathbf{r}_n)}\right]$$

where $\pi_i(\mathbf{r}) = e^{-\beta_i U_i(\mathbf{r})}/Z_i$, Z_i is the partition function, U_i is the potential function, and β_i is the inverse temperature. Notice the partition functions need not be known, as they cancel between numerator and denominator.

(v) Continue with steps (iii) and (iv) until the sampling of the $n - 1$ level is judged sufficient. Store the $n - 1$ trajectory.

(vi) Continue with steps (ii)−(v) for $n = N - 2$, $N - 3$, ... until the bottom-level simulation is complete.

First, note that canonical sampling is maintained by eq 1.3,[18] just as in an ordinary parallel exchange simulation. On the other hand, detailed balance is not satisfied, as the trial configuration for the level n simulation ($\mathbf{r}_{n-1}$ above) is discarded—making reverse moves effectively impossible. The error is one of practice, not of principle, arising from the fact that the samples of π_n and π_{n-1} are finite, just as in any simulation.

To see intuitively that canonical sampling is maintained by top-down exchange, imagine a pair of simulations undergoing ordinary parallel exchange. Unbeknownst to the investigator, however, the top-level simulation is running on a very fast processor, while the other is running on an old, slow processor. Between neighboring exchange points, the trial conformations from the fast processor will be far more decorrelated than those of the slow simulation—which mimics the effect of the top-down protocol. However, these exchanges still satisfy detailed balance. In the limit of an infinitely fast top-level simulation, trial configurations are completely decorrelated, and one could equally well choose randomly from π_n as in step (iii).

Second, notice that because trial configurations are pulled *at random* from the sample of π_n in step (iii), transitions which are slow in the actual M_n trajectory occur rapidly among the trial configurations. Maximum benefit is thus obtained from successful exchanges—in contrast with a parallel exchange simulation, where trial configurations are typically separated by only a few picoseconds and remain highly correlated.

Third, good results may be obtained expending very little CPU time on all levels except the top one. This may be understood from an energy landscape perspective. The top level has been used to thoroughly sample the space—high barriers are crossed, and major sub-basins equilibrated. At lower levels, only local equilibration need occur. For example, let τ_{nonloc} be the time to cross high barriers, τ_{local} be the time to equilibrate locally, and say that m successful exchanges are needed to sample the space well. Then $\tau_{\text{local}} \times m$ CPU time is needed to sample the lower level. The required condition to save time over a parallel simulation is that $\tau_{\text{local}} \ll \tau_{\text{nonloc}}$. The degree to which this condition is satisfied will depend on the system studied, but the top-down approach allows the flexibility to take advantage of a separation in time scales. This idea is reminiscient of the "dragging" of fast degrees of freedom, suggested by Neal,[52] and the multiple time step approaches developed by Berne and co-workers.[15,16]

Finally, a major advantage of top-down simulation over parallel exchange protocols is that exchange attempts are nearly "free", in the sense that no communication is required between processors.[18] This means that exchanges may be attempted very frequently, and therefore much lower exchange rates may be accommodated. In the case of temperature exchange, this allows either for the steps in the temperature ladder to be more widely spaced or for the treatment of larger systems with fewer replicas.

1.4. Simulation Details. In ideal circumstances, low-resolution models used in ResEx simulations would be specifically optimized for resolution exchange. They would have maximal conformational overlap for the common degrees of freedom. Here we work with an "off the shelf" low-resolution model (united atom) which leads to some difficulties. Consider, for example, a $C^{\alpha}-C'$ bond which is parametrized in the two models by two slightly different natural bond lengths. In an exchange attempt, the configurations are swapped, and in each trial configuration the $C^{\alpha}-C'$ bond is moved a bit from its preferred length. These small contributions add up for every harmonic term in the entire molecule and have a noticeable effect on the acceptance of exchange moves. We have solved this problem by simply changing the harmonic parameters of the coarse model to match those of the detailed model. This makes the coarse model more "exchangeable" with the detailed model. Since the coarse model is simply "suggesting" configurations for the atomic model, and since eq 1 guarantees that no bias is introduced by the coarse model, we need not worry that the coarse model parameters are changed from their original values. We now describe in detail the two molecular systems which were studied in the present work.

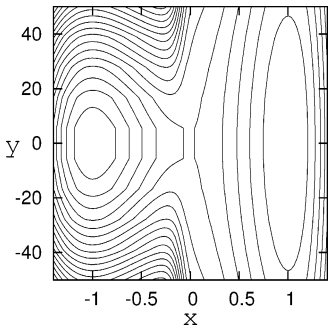


Figure 3. Contours of the potential surface $U(x, y)$, described by eq 3. Here we have reduced w to 10 so that both wells are visible in the figure.

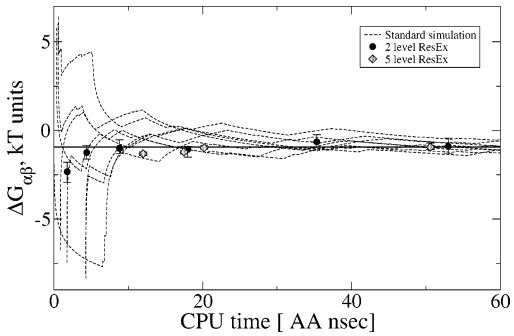


Figure 4. Comparison of different ResEx protocols for dileucine. Plotted are free energy difference estimates between the α and β states as a function of total CPU cost. The dashed lines are individual runs generated by standard Langevin dynamics (no exchange), and the solid horizontal line is the average of 4 independent 150 ns Langevin dynamics simulations. The solid circles are ResEx results from the two-level ladder from ref 1, and the diamonds are the ResEx results from the five-level ladder, averaged in each case over 8 independent runs. The error bars give the range of the 8 independent runs. The ResEx data points are displaced from the origin to accurately reflect the time invested in generating the top-level and all intermediate-level distributions. The exchange schedule for ResEx has not been optimized.

Dileucine. We first studied dileucine peptide (ACE-$(Leu)_2$-NME) using the same force field parameters as in ref 1. Here, we also carried out an incrementally coarsened ResEx simulation of dileucine in 5 levels. The coarsest level (M_4) was the same as in ref 1, namely a modified version of OPLSUA.[53] In lower levels, the molecule was rendered in finer detail beginning at the N-terminus: in M_3, the N-terminal methyl group, C^α, and C^β of the first residue were modeled in full atomistic detail; in M_2, C^γ and both C^δ's of the first residue were additionally modeled explicitly; in M_1, the C^α, C^β, C^γ, and one C^δ of the second residue were modeled explictly; and finally in M_0, the entire molecule was rendered in full atomic detail. The ladder of mixed models was chosen to keep approximately fixed the number of hydrogens by which neighboring levels differ, without splitting a methyl group.

The top level (M_4) was simulated first, for 25, 50, 100, or 200 ns. The different lengths of top-level simulation were used to generate the different data points in Figure 4. Then the higher resolution models were run, as per the top-down protocol (see section 1.3), attempting exchanges once per picosecond. A total of 2.5×10^3 exchanges were attempted between each level, for a total trajectory length per level of 2.5 ns. Frames were stored every 0.1 ps, for a total of 2.5×10^4 frames in the sample at each level below the top.

Met-Enkephalin. We next studied met-enkephalin (NH_3^+-Phe-Gly-Gly-Tyr-Met-COO^-). The united-atom force field was a modified version of OPLSUA.[53] The force field was modified so that the bond length and angle bending parameters matched those of the all-atom force field, which improves exchangeability (or conformational overlap) of the two models. The sample of the top-level model was constructed from two independent 100 ns trajectories, both started from the pdb structure 1plw (first NMR model), generated by Langevin dynamics as implemented in TINKER v. 4.2.[54] The friction coefficient was 5 ps^{-1}, and solvation was modeled with the GB/SA method.[55] The first 1 ns of each trajectory was discarded, and frames were stored every 10 ps for a total of 19 800 frames in the sample.

We then ran the next higher resolution simulation, as per the top-down algorithm (see section 1.3). This model was of mixed resolution, with the Tyr_1 residue represented by the OPLSAA all-atom force field,[56] and the remaining residues described by the OPLSUA force field. Every 1 ps, a random configuration was pulled from the top-level (M_5) trajectory, and a resolution exchange was attempted, with acceptance governed by eq 1. Since the acceptance ratio for the M_5 to M_4 exchanges was approximately 10%, the average length of M_4 trajectory between exchanges was 10 ps. A total of 10^4 ResEx moves were attempted, for a total M_4 trajectory length of 10 ns. Frames were stored every 0.1 ps for a sample of 10^5 frames.

This procedure was then repeated for each level shown in Figure 1, with the exception that the attempt frequency of ResEx moves was adjusted for the acceptance ratios, so that the segments of the simulations between exchanges were kept approximately constant at 10 ps. Also, the total number of attempted exchanges was adjusted so that approximately 10^3 successful exchanges were observed between each level, for a total trajectory length of 10 ns at each level. Given that the top level is presumed to be well-sampled, 10^3 exchanges should sample a large number of basins.

2. Results

In a previous short paper, we tested the ResEx algorithm on two small molecules: butane and dileucine peptide.[1] It was shown that the method produced results in agreement with those obtained by standard simulation methods. For the sake of clarity, here we first demonstrate our approach on a two-dimensional toy model, consisting of two basins which differ only *entropically*. We also extend the method to two peptides, dileucine and met-enkephalin, to demonstrate the viability of incremental coarsening.

2.1. Results: Two-Dimensional Model. An important consideration in designing any sampling method is whether it will correctly account for entropic differences. We therefore designed the potential surface shown in Figure 3 to compare three different sampling methods: a "standard" Monte Carlo

simulation, the same Monte Carlo with resolution exchange, and the same Monte Carlo with the "dual REM" method of Lwin and Luo.[44]

The surface $U(x, y)$ in Figure 3 is described by the function

$$U(x, y) = E_b(x^2 - 1)^2 + \frac{E_0 y^2}{1 + w(\tanh(x/0.1) + 1)/2} \quad (3)$$

where $E_0 \equiv k_B T$, $E_b = 10\ k_B T$ is the barrier height, and w controls the width of the right well in the figure. Notice that the profile of the surface at $y = 0$ is symmetric about $x = 0$: $U_x(x) \equiv U(x, y = 0) = E_b(x^2 - 1)^2$, i.e., the two minima are of equal depth. The parameters were chosen so that the equilibrium populations of the two wells differ greatly—we used $w = 500$, so that the right well holds 95% of the population, as measured by standard techniques.

For both the ResEx and the dual REM simulations, the "coarse-grained" potential was simply the one-dimensional potential U_x, i.e., a symmetric double well.

To describe the exchange moves explicitly, we denote 2*D* configurations by (x, y) and 1*D* configurations by $\tilde{x}$. For both algorithms, an exchange move consists of two parts: the construction of a 1*D* trial configuration ($\tilde{x}_{new}$) from a 2*D* configuration (x_{old}, y_{old}) and vice versa: the construction of a 2*D* trial configuration (x_{new}, y_{new}) from a 1*D* configuration ($\tilde{x}_{old}$). The construction of a 1*D* configuration in each case is simple—the "extra" (y) coordinate is dropped, i.e., $\tilde{x}_{new} = x_{old}$.

The only difference between the two simulations is in the construction of trial configurations for the 2*D* model from the 1*D* model. In ResEx, the trial configuration is the $\tilde{x}$ coordinate from the 1*D* model, with the (old) y coordinate from the 2*D* model, i.e., $x_{new} = \tilde{x}_{old}$ and $y_{new} = y_{old}$. In dual REM, on the other hand, the trial y coordinate is chosen randomly and then minimized. For the potential $U(x, y)$, this means that $y_{new} = 0$ *always*, i.e., $x_{new} = \tilde{x}_{old}$ and $y_{new} = 0$.

The ResEx simulation correctly samples the two wells, giving a population in the right well of 96.4 ± 1.6%. The dual REM simulation, on the other hand, yields a population of 51.0 ± 1.6% for the right well. What causes the error in the dual REM simulation? The answer is that the construction of dual REM trial moves violates detailed balance. More specifically, the minimization of the y coordinate means that the difference in *width* between the two wells is not accounted for correctly, since in dual REM $y_{new} = 0$ *always*. Notice that it is not the random selection of the y coordinate which intrinsically violates detailed balance, only the subsequent minimization.

What is the analogous situation in molecular simulations? In this case, both ResEx and dual REM construct trial moves in internal coordinates—the coarse-grained model is built from a subset of the degrees of freedom of the atomic model. For example, the coarse-grained model (the x coordinate above) may be the backbone coordinates of a protein, and the remainder (the y coordinate above) may be the side-chain degrees of freedom. In dual REM construction of trial moves, the side chains are minimized prior to exchange, and therefore differences in entropy between different side-chain conformations are neglected. In ResEx, there is no minimization prior to exchange, and canonical sampling is maintained.

2.2. Results: Incremental Coarsening of Dileucine. We previously reported on ResEx results for dileucine, demonstrating successful exhange and significant speedup from a direct exchange between all-atom and united-atom models. Here, it is shown that dileucine may be coarsened incrementally and that (i) the correct distribution is observed for the all-atom model and (ii) adding additional intermediate levels of resolution improves efficiency.

The additional levels boost the exchange acceptance by 2 orders of magnitude: exchanges were successful between M_4 and M_3 15.5% of the time, between M_3 and M_2 12.7% of the time, between M_2 and M_1 29.0% of the time, and between and M_1 and M_0 44.0% of the time. By comparison, exchanging AA and UA dileucine in a single step is successful only 0.16% of the time.[1] However, we need to ask whether it is really more efficient to introduce additional levels of simulation in order to boost the acceptance of exchange moves.

In fact, it appears to be substantially more efficient to use incremental coarsening rather than abrupt coarsening. The cost for a given ladder of N levels may be written

$$\text{total cost} = \text{top-level cost} + \sum_{i=0}^{N-1} \frac{m\tau_i}{r_i} \quad (4)$$

where the cost of the top level is fixed, m is the fixed number of successful exchanges which are desired, τ_i is the simulation cost for an interval between two exchange attempts at level i, and r_i is the acceptance rate between levels i and $i + 1$. We have assumed that the sampling of level i demands a fixed number of successful resolution exchanges, consistent with the motivation of the top-down protocol discussed in section 1.3.

Equation 4 implies that the effective exchange rate for an incremental ladder is a reciprocal sum of the individual rates. If we assume the τ_i are equal for all levels (which is exact for temperature exchange), then

$$1/r_{eff} = \sum_i 1/r_i \quad (5)$$

giving an effective rate for the five-level dileucine ladder of 5.1%. This result suggests an improvement in efficiency relative to the single stepladder, where the rate was 0.156%.[1]

In Figure 4, we compare the sampling of dileucine by three different simulation protocols: standard Langevin dynamics, resolution exchange with two levels, and resolution exchange with five levels. Sampling is assessed by examining the relative populations of the α and β states ($e^{-\Delta G_{\alpha\beta}/k_B T}$) considered in ref 1. The convergence of this relative population measure requires transitions between the two states, which occur infrequently in a standard simulation. The five-level ladder clearly outperforms the two-level ladder, as we are able to generate results both more accurately and more precisely with the five-level ladder in an equal amount of CPU time. Note that the total simulation time required for the entire ladder, including the top level, is included in the ResEx data points.

The efficacy of the ResEx approach is underscored by the fact that, at the top level (united atom), the sign of $\Delta G_{\alpha\beta}$ is

wrong. That is, the exchange process corrects for a substantial bias in the coarse model.

2.3. Results: Incremental Coarsening of Met-Enkephalin. Met-enkephalin is a flexible neurotransmitter which participates in immune responses and pain inhibition, among other roles.[57,58] By virtue of its small size and biological interest, it often is used to test new simulation methods[27,59,60] and compare existing protocols.[58,61]

Using met-enkephalin, we demonstrate the efficacy of the incremental coarsening procedure for a ladder of decreasing resolution at constant temperature and for a ladder of simultaneously decreasing resolution and increasing temperature. Because quantifying the quality of sampling for met-enkephalin is considerably more difficult than is commonly appreciated, we will not present a detailed efficiency analysis. More will be said on this second topic in the discussion.

We employed the ResEx algorithm in a top-down framework, as sketched in Figure 2. First, the top-level simulation (coarsest resolution—here, united atom) was run. We then ran an exchange simulation at the next highest resolution—here, one residue was represented at an all-atom resolution, and the rest of the peptide was united atom. This procedure was continued, "decoarsening" one residue at a time, until the entire peptide was represented at the all-atom level. Details are given in sections 1.2 and 1.4.

The incremental coarsening procedure substantially increases exchange acceptance. The five rates in the six-level ladder vary from 2.4% to 18%, as shown in Figure 1. For comparison, exchanging between all-atom and united-atom models of met-enkephalin, with no intermediate levels of resolution, results in an acceptance ratio of 0.09%. The acceptance ratios vary, in part, according to the number of degrees of freedom by which the two levels differ.

For met-enkephalin, the principal results are the acceptance rates shown in Figure 1, which are significant for several reasons. First, they demonstrate the first implementation of the incremental coarsening approach in a complex peptide. Second, because they are well within the practical range of the top-down protocol—see section 1.3 and the Discussion—they indicate that the ResEx algorithm could prove important for larger peptides. Last, by comparing the effective exchange rate suggested by eq 5, $r_{\mathrm{eff}} = 1.1\%$, with the rate of 0.09% for direct exchanges between united- and all-atom models, one sees that a substantial speedup has been achieved. Of course, the magnitude of the improvement is merely suggestive—without a rigorous quantification of the sampling quality, there can be no rigorous comparison of efficiency. Such a quantification is beyond the scope of this work, as noted in the Discussion.

It is useful to understand the intuitive reason behind the advantage of incremental coarsening. If one writes the acceptance criteria (1) and (1.3) in the form $\min[1, e^{-\epsilon}]$, then for exchanges between models of greatly differing resolution, one typically finds the dimensionless "energy" is large, i.e, $\epsilon \gg 1$. It seems to be roughly true that this energy is proportional to the difference in the number of degrees of freedom in the models being exchanged. However, if the change is made incrementally using many models M_i, then between levels i and $i + 1$ there is a relatively small cost $\Delta\epsilon_i$, with $\sum_i \Delta\epsilon_i \sim \epsilon$. It is clear that with enough increments, one can achieve $\Delta\epsilon_i \ll 1$ and thus create a high likelihood for exchange since the corresponding Boltzmann factors are much larger: $r_i \sim e^{-\Delta\epsilon_i} \gg e^{-\epsilon}$. This is exactly what is embodied in eq 4. The tradeoff is that one pays the cost for simulating the additional intermediate levels. However, as has been stressed in section 1.3, the intermediate-level simulations are very short compared to the top level. In the present context, for instance, the top-level met-enkephalin trajectory is 198 ns, while all other levels were simulated for only 10 ns. The net savings can be quite substantial, especially considering that good sampling is achieved by increasing the number of exchanges.

While we cannot yet rigorously measure sampling quality, we can show that the results obtained with ResEx are consistent with those obtained by standard methods, by comparing Ramachandran histograms (Figure 5) from the ResEx simulation, to those obtained by standard simulation (990 ns of simulation with the M_0 parameters). Overall, the agreement between the ResEx simulation and the 990 ns conventional simulation is quite good. However, a careful comparison reveals a region on the Phe_4 plot, labeled "A", which is noticeably undersampled by the ResEx simulation, as compared to the 990 ns Langevin dynamics trajectory. The explanation is provided by an inspection of the Phe_4 histogram of the united-atom simulation: region "A" was not sampled by the top-level simulation. The failure points to a potential weakness of the ResEx (or any exchange) method—regions which are not sampled by the top level will be difficult to sample in any of the other levels. This is a specific instance of a general problem that occurs whenever auxiliary distributions are used to enhance sampling of some "distribution of interest", namely the need to balance overlap with wider sampling via the auxiliary distribution.[51] In other words, it is a failure of the top-level simulation rather than the algorithm.

Interestingly, Figure 5 also presents two counterexamples to the foregoing discussion. Regions "B" of the Gly_3 and "C" of the Met_5 plots were both well-sampled by the ResEx simulation, despite being infrequently visited by the top level. That is, the ResEx acceptance criterion (1) correctly "reweights" the conformation space of the all-atom model by allowing normal dynamics to continue when appropriate. Nevertheless, we are in the process of experimenting with other "schedules" (combinations of attempt frequency and number of exchange attempts) to balance the normal and the exchange dynamics.

Ideally, the coarse model distribution would have better overlap with the high-resolution distribution, and the balance could be adjusted to favor exchanges over normal dynamics. This would allow the same quality of sampling with less simulation at each level below the top. In the long term, we hope to design coarse models constructed to *not* eliminate *any* regions of configuration space in more detailed models.

2.4. Resolution Exchange with Tempering. We have also explored the possibilty of combining resolution exchange with parallel tempering, so that the sampling of the reduced models is improved both by the reduction in resolution and

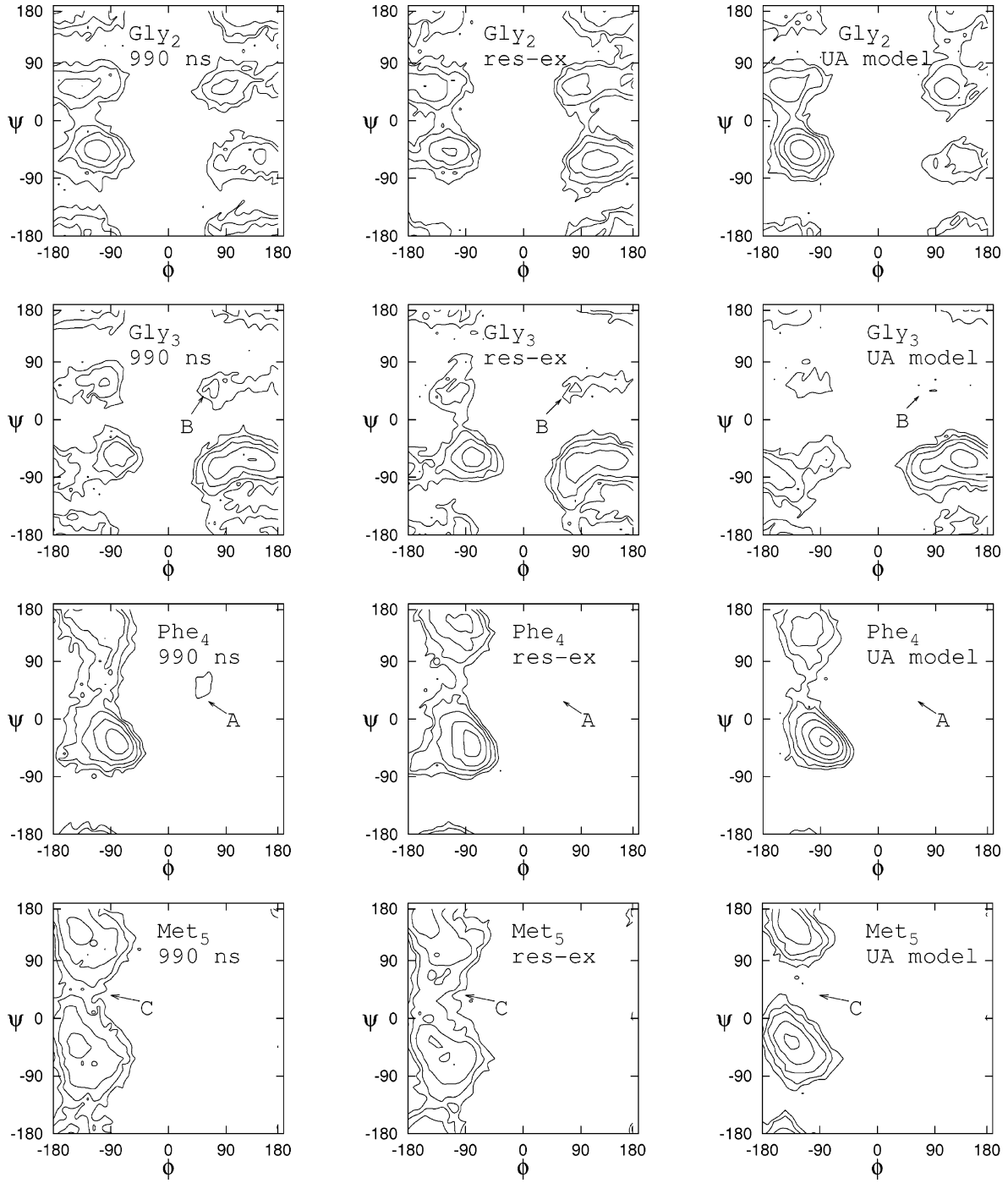


Figure 5. Ramachandran histograms for met-enkephalin. The left column is a 990 ns Langevin dynamics simulation at all-atom resolution, without resolution exchange; the middle column is the all-atom level (M_0) from resolution exchange as described in the text; the right column is the top-level united-atom simulation (level M_5) used for the resolution exchange simulation shown in the middle column. Note that since the peptide is unblocked, there are only 4 pairs of $\phi-\psi$ dihedrals. Res-ex fails to "find" one region (labeled "A") not present in the top-level simulation but finds two others (labeled "B" and "C").

by increased temperatures. In a standard parallel tempering simulation, the temperatures are roughly exponentially distributed, in order that the conformational overlap between neighboring temperatures is constant over the ladder. However, there is no simple relationship between the change in *resolution* and the acceptance of resolution exchanges. Some care must therefore be taken with the assignment of the temperature ladder.

We began with the ladder of models in Figure 1. The acceptance ratios give an idea of the temperature gap which may be tolerated between two levels—a higher acceptance ratio will tolerate a larger jump in temperature. However, compared to standard parallel tempering simulations,[26−35] it may seem that the acceptance ratios are already too low to accommodate tempering in addition to resolution exchange. After all, we may expect that any difference in temperature will lower the acceptance of exchange moves. In this regard, the top-down approach has an important advantage over a parallel implementation. Since exchange attempts are "free" (no commmunication between processors is required), they

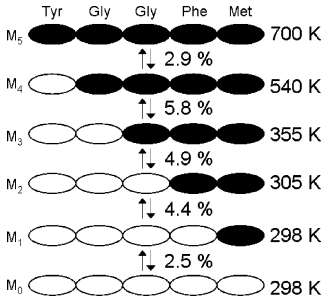


Figure 6. Ladder combining exchange between all-atom and united-atom met-enkephalin with tempering of reduced resolution simulations. Residues are depicted with ovals—open corresponds to an all-atom representation, filled to united atom. The ratios of successful to attempted exchanges between each level are indicated by the percentages. The temperature of each level is indicated on the right.

may be attempted much more frequently, and lower acceptance ratios may be tolerated.[18] Indeed, in our original study of dileucine peptide with top-down resolution exchange, the acceptance ratio was much less than 1%.[1] See also section 1.3.

The ladder combining temperature and resolution is shown in Figure 6. The temperature gaps were chosen by trial and error, aiming for an acceptance of attempted exchanges of a few percent between neighboring levels. Based upon this restriction, the top-level simulation was run at a temperature of 700 K, which is comparable to previously published parallel tempering studies of met-enkephalin.[27,36] We should expect, however, that fixed-CPU-cost sampling should be improved relative to ordinary replica exchange, by virtue of the reduction in resolution.

The reduction in resolution confers an additional benefit when combined with tempering. Since the overlap between neighboring levels in a parallel tempering simulation scales such as (number of DoF)$^{1/2}$, reducing resolution allows the temperature gaps to *increase* as the resolution is reduced. The overlap between neighboring levels in a combined resolution/tempering ladder is thus controlled both by the change in resolution and the change in temperature, with the two effects compensating one another in an unknown way. Indeed, we observed one puzzling case in our search for an appropriate resolution/temperature ladder. In one ladder (data not shown), exchange between levels M_2 and M_3 was successful about 7% of the time when both were at 298 K, while exchange occurred approximately 11% of the time when M_2 was thermostated to 305 K and M_3 to 320 K. We have not explained this result—though it should be remembered that different models have different landscapes, and therefore temperatures may not be directly compared.

3. Concluding Discussion

We have extended our resolution exchange (ResEx) method[1] using an incremental coarsening procedure for implicitly solvated peptides. After carefully testing the approach in the two-residue dileucine peptide, we applied it successfully to the five-residue met-enkephalin. Incremental coarsening allows tuning of the conformational overlap between models of differing resolution and therefore makes practical simulations which would otherwise be hampered by poor acceptance of exchange moves. We have also demonstrated that resolution exchange is naturally combined with parallel tempering, so that the reduced resolution models may be aided in their sampling of conformations by elevated temperatures.

Ramachandran histograms demonstrate that, for the most part, ResEx simulation is consistent with standard simulation techniques. In one case, however, they reveal a weakness of our method—a top-level simulation which eliminates important regions of conformation space will result in poor sampling at the bottom level. This weakness is shared by any simulation which relies upon auxiliary ensembles to sample among major sub-basins. In the future we hope to eliminate this problem by more careful construction of reduced models.

Of course, we hope to treat still larger molecules with the ResEx method. Since it is essential that the top-level be well-sampled, the treatment of larger molecules will require yet coarser top-level simulation. This will likely require incremental coarsening from the united-atom level to a model with one or two beads per residue. Suitable models are under development. It appears that the ResEx approach cannot be applied easily to explicitly solvated systems; however, given the difficulty and importance of sampling implicitly solvated systems, ResEx may prove very valuable for biomolecular simulation.

We have also developed an alternative rigorous algorithm which permits the use of coarse top-level simulations to generate atomically detailed canonical samples. It is essentially a "decorating" procedure, and it eliminates the potential issue of correlations between coarse coordinates Φ and detailed coordinates $\mathbf{x}$, which could reduce acceptance rates in resolution exchange. Specifically, after generating a low-resolution sample distributed according to $\pi_L(\Phi)$, one can independently sample detailed coordinates $\mathbf{x}$ according to an arbitrary distribution $\pi_x(\mathbf{x})$. (For example, π_x could be based on harmonic terms in the full force field.) Full configurations are thus generated according to the simple product $\pi_L(\Phi)\pi_x(\mathbf{x})$ and may be reweighted to generate a fully detailed, high-resolution distribution $\pi_H(\Phi,\mathbf{x})$ using standard methods.[62] In the long term, the decorating approach may prove useful for adding explicit solvent. It may also be implemented in an incremental fashion.

An "auxiliary" question which remains to be carefully addressed is the quantification of sampling efficiency. There are numerous proposals for judging whether a simulation is converged—some are based on principal components,[63] others on energy-based ergodic measures,[64] and our own work in progress uses structural histograms.[65] Which one provides an appropriate measure depends on what properties are of interest. For applications which depend on the relative populations of various conformations, such as calculation of binding affinities for small molecules, a measure which depends directly on the conformational distribution is needed. Such a method is under development—for now we only mention that structural histograms provide a much more sensitive signal of nonconvergence than energy-based methods.[65]

Acknowledgment. The authors would like to thank Marty Ytreberg, Bruce Berne, Michael Deem, and Angel García for interesting discussions and insightful comments. This work was supported by the Department of Environmental and Occupational Health and the Department of Computational Biology at the University of Pittsburgh and by the NIH (Grants ES007318 and GM070987).

Supporting Information Available: Parameter files (formatted for use with the Tinker molecular modeling simulation package) for the modified OPLSUA force field and for the mixed ua-aa force field and a sample "key" file for running the mixed force field simulations in Tinker. This information is available free of charge via the Internet at http:/pubs.acs.org.

References

(1) Lyman, E.; Ytreberg, F. M.; Zuckerman, D. M. *Phys. Rev. Lett.* **2006**, *96*, 028105.

(2) Aksimentiev, A.; Balabin, I. A.; Fillingame, R. H.; Schulten, K. *Biophys. J.* **2004**, *86*, 1332−1344.

(3) Duan, Y.; Kollman, P. A. *Science* **1998**, *282*, 740−744.

(4) Snow, C. D.; Nguyen, H.; Pande, V. S.; Gruebele, M. *Nature* **2002**, *420*, 102−106.

(5) Halgren, T. A.; Damm, W. *Curr. Opin. Struct. Biol.* **2001**, *11*, 236−242.

(6) Anisimov, V. M.; Lamoureux, G.; Vorobyov, Igor, V.; Huang, N.; Roux, B.; MacKerell, A. D., Jr. *J. Chem. Theor. Comput.* **2005**, *1*, 153−168.

(7) Iftimie, R.; Minary, P.; Tuckerman, M. E. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6654−6659.

(8) Huang, Y. J.; Montelione, G. T. *Nature* **2005**, *438*, 36−37.

(9) McCallum, S. A.; Hitchens, T. K.; Torborg, C.; Rule, G. S. *Biochemistry* **2000**, *39*, 7343−7356.

(10) Volkman, B. F.; Lipson, D.; Wemmer, D. E.; Kern, D. *Science* **2001**, *291*, 2429−2433.

(11) Shoichet, B. K. *Nature* **2004**, *432*, 862−865.

(12) Comeau, S.; Vajda, S.; Camacho, C. J. *Proteins* **2005**, *60*, 239−244.

(13) Kollman, P. *Chem. Rev.* **1993**, *93*, 2395−2417.

(14) Rodinger, T.; Pomès, R. *Curr. Opin. Struct. Biol.* **2005**, *15*, 164−170.

(15) Humphreys, D. D.; Friesner, R. A.; Berne, B. J. *J. Phys. Chem.* **1994**, *98*, 6885−6892.

(16) Hetényi, B.; Barnacki, K.; Berne, B. J. *J. Chem. Phys.* **2002**, *117*, 8203−8207.

(17) Zhu, Z.; Tuckerman, M. E.; Samuelson, S. O.; Martyna, G. J. *Phys. Rev. Lett.* **2002**, *88*, 100201−1−100201−4.

(18) Frantz, D. D.; Freeman, D. L.; Doll, J. D. *J. Chem. Phys.* **1990**, *93*, 2768−2783.

(19) Bai, D.; Brandt, A. *NATO science series III*; Technical Report 177; 2001.

(20) Praprotnik, M.; Site, L. D.; Kremer, K. *J. Chem. Phys.* **2005**, *123*, 224106.

(21) Mitsutake, A.; Sugita, Y.; Okamoto, Y. *Biopolymers* **2001**, *60*, 96−123.

(22) Gnanakaran, S.; Nymeyer, H.; Portman, J.; Sanbonmatsu, K. Y.; García, A. E. *Curr. Opin. Struct. Biol.* **2003**, *13*, 168−174.

(23) Earl, D. J.; Deem, M. W. *Phys. Chem. Chem. Phys.* **2005**, advance article.

(24) Swendsen, R. H.; Wang, J.-S. *Phys. Rev. Lett.* **1986**, *57*, 2607−2609.

(25) Hukushima, K.; Nemoto, K. *J. Phys. Soc. Jpn.* **1996**, *65*, 1604.

(26) Hansmann, U. H. E. *Chem. Phys. Lett.* **1997**, *281*, 140−150.

(27) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141−151.

(28) García, A. E.; Sanbonmatsu, K. Y. *Proteins* **2001**, *42*, 345−354.

(29) Zhou, R.; Berne, B. J. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12777−12782.

(30) Im, W.; Feig, M.; Brooks, C. L. I. *Biophys. J.* **2003**, *85*, 2900−2918.

(31) Kokubo, H.; Okamoto, Y. *J. Chem. Phys.* **2004**, *120*, 10837−10847.

(32) Im, W.; Brooks, C. L. I. *J. Mol. Biol.* **2004**, *337*, 513−519.

(33) Gnanakaran, S.; Hochstrasser, R. M.; García, A. E. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 9229−9234.

(34) García, A. E.; Onuchic, J. N. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13898−13903.

(35) Paschek, D.; Gnanakaran, S.; García, A. E. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6765−6770.

(36) Sanbonmatsu, K. Y.; García, A. E. *Proteins* **2002**, *46*, 225−234.

(37) Roe, D. R.; Hornak, V.; Simmerling, C. *J. Mol. Biol.* **2005**, *352*, 370−381.

(38) Zhang, W.; Wu, C.; Duan, Y. *J. Chem. Phys.* **2005**, *123*, 154105-1−154105-9.

(39) Kofke, D. A. *J. Chem. Phys.* **2002**, *117*, 6911−6914.

(40) Kone, A.; Kofke, D. A. *J. Chem. Phys.* **2005**, *122*, 206101-1−206101-2.

(41) Liu, P.; Kim, B.; Friesner, R. A.; Berne, B. J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13749−13754.

(42) Iftimie, R.; Salahub, D.; Wei, D.; Schofield, J. *J. Chem. Phys.* **2000**, *113*, 4852−4862.

(43) Liu, J. S.; Sabatti, C. Simulated sintering: Markov chain Monte Carlo with spaces of varying dimensions. In *Bayesian Statistics*; Bernardo, J., Berger, J., Dawid, A., Smith, A., Eds.; Oxford University Press: New York, 1998; Vol. 6.

(44) Lwin, T. Z.; Luo, R. *J. Chem. Phys.* **2005**, *123*, 194904-1−194901-9.

(45) Levitt, M.; Warshel, A. *Nature* **1975**, *253*, 694−698.

(46) Bradley, P.; Misura, K. M. S.; Baker, D. *Science* **2005**, *309*, 1868−1871.

(47) Fukunishi, H.; Watanabe, O.; Takada, S. *J. Chem. Phys.* **2002**, *116*, 9058−9067.

(48) Frenkel, D.; Smit, B. *Understanding Molecular Simulation*; Academic Press: San Diego, 1996.

(49) Yan, Q.; de Pablo, J. J. *J. Chem. Phys.* **2000**, *113*, 1276−1282.

(50) Sugita, Y.; Kitao, A.; Okamoto, Y. *J. Chem. Phys.* **2000**, *113*, 6042−6050.

(51) Neal, R. M. 2005. http://xxx.lanl.gov/abs/math.ST/0511216.

(52) Neal, R. M. *Taking bigger Metropolis steps by dragging fast variables*; Technical Report No. 0411; Dept of Statistics, University of Toronto: 2004. http://xxx.lanl.gov/abs/math.ST/0502099.

(53) Jorgensen, W. L.; Madura, J. D.; Swenson, C. J. *J. Am. Chem. Soc.* **1984**, *106*, 6638−6646.

(54) Ponder, J. W.; Richard, F. M. *J. Comput. Chem.* **1987**, *8*, 1016−1024. http://dasher.wustl.edu/tinker/.

(55) Still, W. C.; Tempczyk, A.; Hawley, R. C. *J. Am. Chem. Soc.* **1990**, *112*, 6127−6129.

(56) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *117*, 11225−11236.

(57) Plotnikoff, N. P.; Faith, R. E.; Murgo, A. J.; Good, R. A. *Enkephalins and endorphins: stress and the immune system*; Plenum: New York, 1986.

(58) Shen, M.-Y.; Freed, K. F. *Biophys. J.* **2002**, *82*, 1791−1808.

(59) Mitsutake, A.; Sugita, Y.; Okamoto, Y. *J. Chem. Phys.* **2003**, *118*, 6664−6675.

(60) Mitsutake, A.; Kinoshita, M.; Okamoto, Y.; Hirata, F. *J. Phys. Chem. B* **2004**, *108*, 19002−19012.

(61) Zaman, M. H.; Shen, M.-Y.; Berry, R. S.; Freed, K. F. *J. Phys. Chem.* **2003**, *107*, 1686−1691.

(62) Ferrenberg, A. M.; Swendsen, R. H. *Phys. Rev. Lett.* **1988**, *61*, 2635−2638.

(63) Hess, B. *Phys. Rev.* **2002**, *E65*, 031910-1−031910-10.

(64) Straub, J. E.; Rashkin, A. B.; Thirumalai, D. *J. Am. Chem. Soc.* **1994**, *116*, 2049−2063.

(65) Lyman, E.; Zuckerman, D. M. Preprint. http://xxx.lanl.gov/abs/physics/0601104.

CT050337X

JCTC Journal of Chemical Theory and Computation

Mapping All-Atom Models onto One-Bead Coarse-Grained Models: General Properties and Applications to a Minimal Polypeptide Model

Valentina Tozzini*

NEST − Scuola Normale Superiore, Piazza dei Cavalieri, 7 I-56126 Pisa, Italy

Walter Rocchia

Scuola Normale Superiore, Piazza dei Cavalieri, 7 I-56126 Pisa, Italy

J. Andrew McCammon

Department of Chemistry and Biochemistry, Center for Theoretical Biological Physics, Howard Hughes Medical Institute, Department of Pharmacology, University of California at San Diego, La Jolla, California 92093

Received November 29, 2005

Abstract: In the one- and two-bead coarse-grained (CG) models for proteins, the two conformational dihedrals ϕ and ψ that describe the backbone geometry are no longer present as explicit internal coordinates; thus, the information contained in the Ramachandran plot cannot be used directly. We derive an analytical mapping between these dihedrals and the internal variable describing the backbone conformation in the one- (two-) bead CG models, namely, the pseudo-bond angle and pseudo-dihedral between subsequent Cα's. This is used to derive a new density plot that contains the same information as the Ramachandran plot and can be used with the one- (two-) bead CG models. The use of this mapping is then illustrated with a new one-bead polypeptide model that accounts for transitions between α helices and β sheets.

1. Introduction

Coarse-grained (CG) models for proteins have become more and more popular in the past two decades, because of the necessity of simulating systems on the size scale of hundreds of nanometers and on the time scale of microseconds.[1,2] The seminal concept can be traced back to the 1970s:[3,4] these simplified models for proteins are based on united-atom representations using one to six interacting centers (beads) for each amino acid. Four- to six-bead models represent explicitly the backbone atoms [Cα, N(H), and C(O)] with additional "beads" for the side chain and for the backbone carbonylic oxygen and the backbone hydrogen,[5−7] and they use the ϕ and ψ dihedrals formed by the backbone atoms C−N−Cα−C and N−Cα−C−N, respectively, as internal variables. Conversely, in one- (two-) bead models, only the Cα (and additional beads for the side chain) is (are) explicitly present (see Figure 1).

The two-dimensional density plot of ϕ and ψ referring to a given Cα is called the Ramachandran map. This map is more dense in specific regions of the ϕ,ψ plane, each pertaining to a particular secondary structure. Other regions are "forbidden" because of the steric hindrance of the backbone atoms. Thus, the Ramachandran plot is a simple and immediate tool to check the reliability of a model protein. But, while this check can be directly done in four- (to six-) bead models, this is no longer possible in the one- (two-) bead models. In fact, in these models, the ϕ and ψ are no longer explicit internal variables and the description of the backbone geometry relies on different internal variables, that is, the Cα−Cα−Cα angle θ and the Cα−Cα−Cα−Cα dihedral α (see Figure 1). So far, the problem of how to

* Corresponding author fax: +39-050-509417; e-mail: tozzini@nest.sns.it.

10.1021/ct050294k CCC: $33.50
Published on Web 04/18/2006

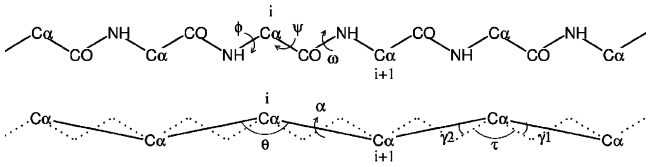


Figure 1. Schematic representation of the coarse-graining procedure. The internal coordinates (ϕ, ψ, ω) and θ,α and the angles defining the CG geometry (τ, $\gamma1$, $\gamma2$) are reported.

transfer the information contained in the Ramachandran plot to these simplified models has not been given any consideration. This is quite surprising, given the fact that one- and two-bead models have recently undergone rapid developments, becoming more accurate and sophisticated.[8−15]

In this paper, we derive an analytical correspondence of the all-atom internal backbone coordinates (ϕ,ψ) to the CG internal backbone coordinates (α,θ) that allows one to explicitly map the Ramachandran plot onto a new θ,α conformational density plot. The force field (FF) of the recent one- and two-bead CG models are reconsidered on the basis of the θ,α map. As an illustration of these concepts, a new one-bead model for polypeptides is presented that accounts for helix to sheet secondary structure transitions.

2. Coarse Graining and Mapping of the Internal Backbone Coordinates

The procedure leading from an all-atom model to a one-(two-) bead CG model is schematically described in Figure 1. The bond length and the backbone bond angle NH−Cα−CO display only small variations from the average values ($\tau = 111°$), and the peptide bond geometry is planar ($\omega = 180$), implying that the dihedral angles ϕ and ψ are basically the only free internal variables. [In this paper, we neglect the (rare) possibility of a cis conformation of the peptide bond ($\omega = 0$).] In the CG description, the backbone conformation is determined by the bond angle between three subsequent Cα's and the dihedral between four subsequent Cα's. The bond angle with vertex $C\alpha_i$, (θ_i) depends only on the two adjacent dihedrals ϕ_i and ψ_i, whereas the dihedral between $C\alpha_{i-1}$ and $C\alpha_{i+2}$, namely, $\alpha_{i,i+1}$, depends on the four dihedrals ϕ_i, ψ_i, ϕ_{i+1}, and ψ_{i+1}. Both θ and α depend parametrically on τ and on the two angles between the bonds Cα−Cα and Cα−CO or Cα−NH, which for the trans conformation assume the almost constant values of $\gamma1 = 20.7°$ and $\gamma1 = 14.7°$, respectively (see Figure 1 for the definition of angles and dihedrals). The angle θ as a function of ϕ,ψ is given by the following formula (valid for the trans conformation):

$$\begin{aligned}\cos(\theta_i) = \cos(\tau)[&\cos(\gamma1)\cos(\gamma2) - \\ &\sin(\gamma1)\sin(\gamma2)\cos(\phi_i)\cos(\psi_i)] - \\ &\sin(\gamma1)\sin(\gamma2)\sin(\phi_i)\sin(\psi_i) + \\ &\sin(\tau)[\cos(\psi_i)\sin(\gamma1)\cos(\gamma2) + \\ &\cos(\phi_i)\cos(\gamma1)\sin(\gamma2)] \quad (1)\end{aligned}$$

For (ϕ, ψ) = (180, 180), (0, 0), (180, 0), and (0, 180), one has $\theta = \tau + (\gamma1 + \gamma2)$, $\theta = \tau - (\gamma1 + \gamma2)$, $\theta = \tau + (\gamma1 - \gamma2)$, and $\theta = \tau - (\gamma1 - \gamma2)$, respectively. The first corresponds to the completely extended conformation; the others are planar but sterically forbidden or weakly allowed conformations (see below). The exact formula giving α explicitly as a function of the ϕ,ψ dihedrals is very complex; however, a precision of a few percent can already be obtained at the linear order in $\gamma1$ and $\gamma2$ (see the Supporting Information):

$$\alpha_{i,i+1} = 180 + \psi_i + \phi_{i+1} + \gamma1\sin(\psi_{i+1}) + \gamma2\sin(\phi_i) \quad (2)$$

This formula (with $\gamma1 = \gamma2$) was previously reported by Levitt.[3]

To make general considerations, we restrict ourselves to the case $\gamma1 = \gamma2$. Furthermore, we assume that the geometrical properties are uniform along the chain (or, equivalently, we consider cases of well-defined secondary structure), so that we can put $\phi_i = \phi_{i+1}$ and $\psi_i = \psi_{i+1}$ and drop the dependence on the index $i + 1$ in eq 2. In these conditions, the map (ϕ, ψ) → (α, θ) is symmetric under the exchange $\phi \leftrightarrow \psi$. Thus, the upper triangle in the (ϕ, ψ) plane is superimposed upon folding along the main diagonal (in magenta) over the lower triangle and mapped onto the same region in the α,θ plane. The periodicity and the additional symmetry $\theta(\phi, \psi) = \theta(-\phi, -\psi) = \theta(\phi + 180, \psi + 180)$ determine the peculiar "butterfly" shape of the image of the ϕ,ψ plane in the α,θ plane reported in Figure 2. In the general case when $\gamma1 \neq \gamma2$, the $\phi \leftrightarrow \psi$ symmetry is slightly broken and one has two slightly different "butterfly" images superimposed. It is to be noted that, as the whole α axis is spanned, the θ angle can assume only values ranging between $\tau - \gamma1 - \gamma2$ and $\tau + \gamma1 + \gamma2$, corresponding to the planar-contracted (forbidden) configuration ($\phi = \psi = 0$) and the planar extended configuration ($\phi = \psi = 180$). As the angle α approaches the value 0, the allowed range for θ becomes smaller. Ring structures correspond to α = 0: the maximum value for θ is $\tau = 111 + |\gamma2 - \gamma1|$, obtained for $\phi = 180°$ and $\psi = 0°$, corresponding roughly to a six-membered ring, while the minimum value for θ is ∼105°, obtained for $\phi = \psi \simeq \pm 75°$, corresponding roughly to a five-membered ring.

3. Conversion of the Ramachandran Plot into the α,θ Plot

We use the above relationships to convert the Ramachandran plot into a conformational density plot in the α,θ plane. The conversion for the generic Ramachandran plot and the Ramachandran plots of glycine, proline, and pre-proline residues are reported in Figure 3.

We observe that the symmetry-induced folding of the ϕ,ψ plane along the diagonal produces some peculiarities of the α,θ conformational maps. For instance, the forbidden region around $\phi = \psi = 0$ (simplified as a circle in Figure 3) is folded onto itself and mapped onto a flat region around $\alpha = \pm 180°$, $\theta = \tau - \gamma1 - \gamma2 = \simeq 75°$. Conversely, the forbidden region around $\phi = 0$, $\psi = \pm 180°$ is folded on a weakly allowed region ($\phi = \pm 180$, $\psi = 0$), implying that the corresponding region in the α,θ plane (represented as a small circle around α = 0, $\theta \simeq \tau$) is not strictly forbidden. The effect of the folding along the ϕ,ψ diagonal can also be seen on the core regions of the right-handed (green) and left-

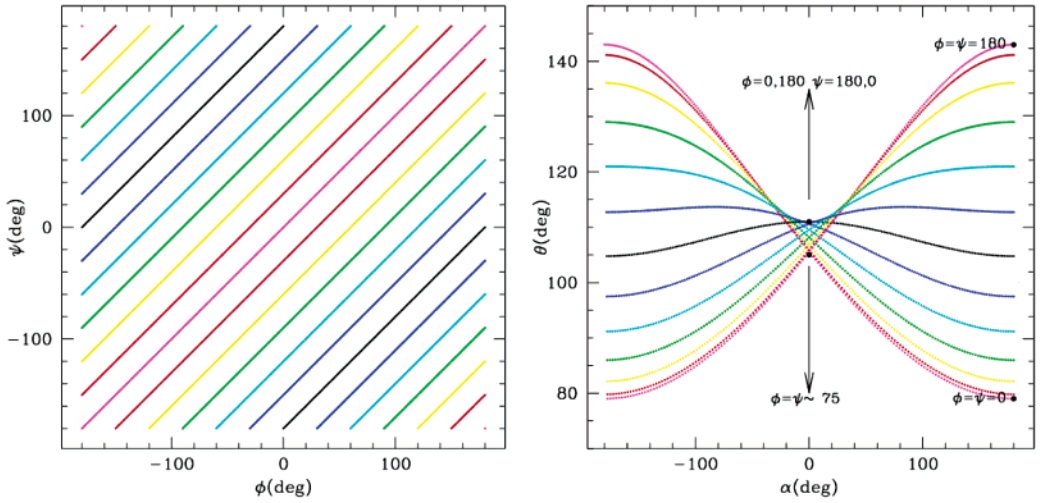


Figure 2. Mapping of the ϕ,ψ plane onto the θ,α plane. Left: lines $\phi - \psi =$ cost are represented in different colors. Right: The same lines are mapped onto the α,θ plane. The region that they define is the image of the whole ϕ,ψ plane onto the α,θ plane through the (ϕ, ψ) → (α, θ) map. Some relevant points are reported.

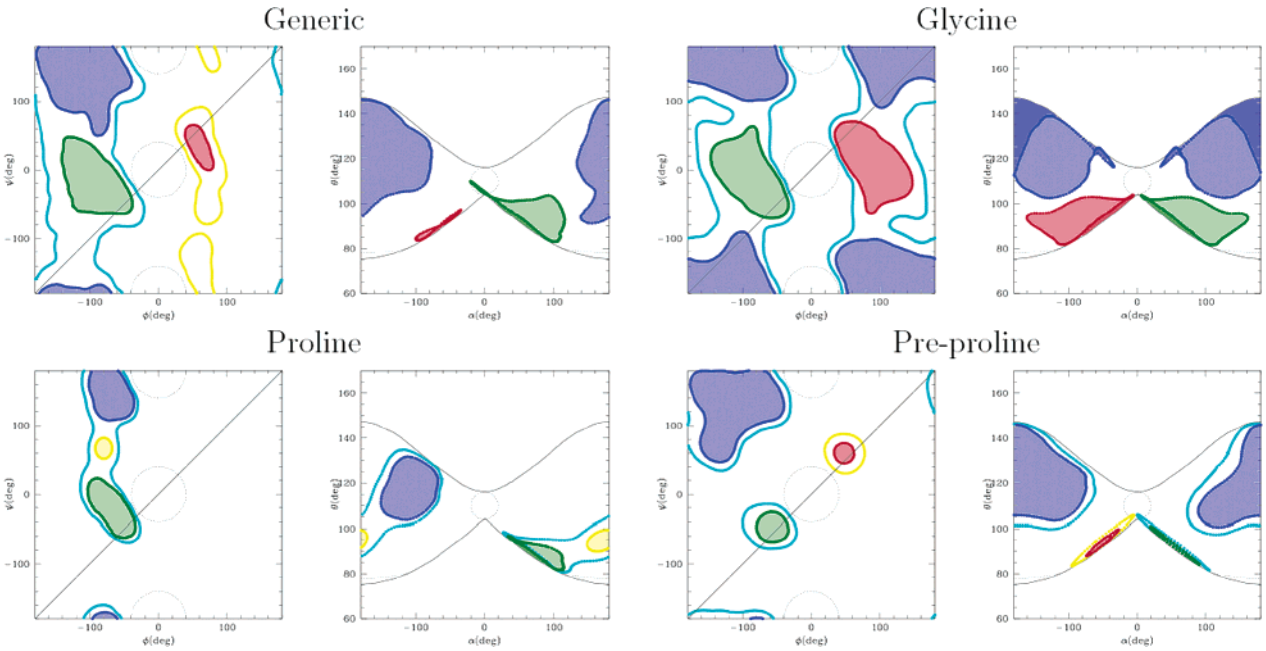


Figure 3. Mapping of the Ramachandran plot onto the α,θ plane for different amino acid types. The "core regions" are in green (right-handed helices), red (left-handed helices), and blue (sheets). The allowed regions (yellow and cyan) are omitted in the generic and glycine α,θ plots for clarity. Forbidden regions are enclosed in the dotted lines. The contours of the "butterfly" image are reported as black lines in the α,θ plane. The input data for the Ramachandran plot are taken from the Protein Data Bank.[18]

handed (red) regions, which are cut by the folding line. These assume peculiar folded and stretched shapes squeezed against the limits of the "butterfly" image. The same happens to the sheet region (blue) in the glycine plot.

However, despite the folding and deformation of the Ramachandran plot, the core regions of the different secondary structures remain separate in the α,θ plane. This is important, because it implies that passing from the ϕ,ψ to the α,θ system of internal coordinates still allows an unambiguous description of the backbone conformation in each secondary structure. This was previously assumed in one- and two-bead CG models but never directly investigated.

In Table 1, the typical ϕ,ψ values for the main secondary structures are reported, as well as the corresponding α,θ values. We observe that right- and left-handed helices have the same value of θ and opposite values of α. Conversely, the θ variable is more efficient in separating the helices from the sheet structures, especially in the glycine where both structures span almost the same α interval.

4. Using the α,θ Plot to Analyze the Force Fields

The potential of mean force $V(q_i)$ is related to the equilibrium probability distribution of the internal coordinate q_i through the relationship $P(q_i) \propto \exp[-V(q_i)/kT]$.

Table 1. Numerical Values of the Internal Angles and Dihedral Coordinates in the All-Atom and CG Representation[a]

structure	ϕ	ψ	θ	α
extended	180	180	146	180
β sheet antiparallel	−139	135	131	179
β sheet parallel ideal	−120	120	121	178
β sheet parallel	−120	113	119	177
flat ribbon	−78	59	92	163
α helix	−57	−47	92	52
3−10 helix	−49	−29	85	81
π helix	−57	−70	99	27
six-membered ring ideal	180	0	115	0
five-membered ring ideal	−75	−75	105	0
five-membered ring	−60	−105	108	0
α helix left-handed	57	47	92	−52
collagen triple helix	−51	153	117	−77
polypro−polygly left helix	−79	150	121	−109

[a] Typical values for secondary structures are reported. Data for ϕ and ψ were taken from textbooks;[19,20] the corresponding structures were built with the *biopolymer* module of Insight (MSI/Accelrys), and from those structures, the θ and α angles were measured.

Although it is well-known that $V(q_i)$ is only an approximation of the potential energy term $U(q_i)$,[21] comparing an approximate probability distribution $p(\alpha,\theta) = \exp[U(\theta,\alpha)/kT]$ to the α,θ plot can still give useful indications.

Very sophisticated $U(\theta,\alpha)$'s are included in the numerical FF derived by Bahar et al.[11,12] and in the UNRES FF by Sheraga et al.,[14,16] which is analytical but involves a very large number of parameters. Here, we analyze the FF with simpler analytical terms. Usually, these include a quite accurate dihedral energy term U^{α}, with terms in $\sin(n\alpha)$ and $\cos(n\alpha)$ with n up to 6,[3,10] although it was shown that the terms in $\cos(\alpha)$ and $\cos(3\alpha)$ are the most relevant.[8,9,13] In fact, these were included also in the most recent versions of the Gō models.[17] Conversely, the bond angle term $U(\theta)$ is usually treated as harmonic. To include the dependence of the equilibrium θ on the secondary structure, different strategies have been adopted. An explicit correlation $\theta(\alpha)$ was assumed by Levitt,[3] which reduced the dimensionality of the α,θ plot to a line. More realistic α,θ plots can be obtained with an approach like that of Head-Gordon et al.,[8] which uses an accurate dihedral term depending on the amino acid type, reproducing the pattern typical of the helices, sheets, or turns depending on the amino acid type (whose helix−sheet−turn propensity must be known a priori), while a harmonic bond angle term is used with an equilibrium $\theta_0 = 105°$ that is an average of the typical helix and sheet values (see Figure 4, left). Alternatively, angle and dihedral energy terms typical for the β sheet were used, and additional terms depending on the second and third neighbor distance along the chain were used to tune the α-helix propensity, such as in Mukherjee and Bagchi[13] (see Figure 4, right). As can be seen from the plots, these FFs may give a quite accurate pattern of the density plot (at least for the dihedral angle) for given secondary structures, but they are not very appropriate in describing situations when the propensity for a given secondary structure is not very well defined. For instance, they are inadequate to describe the α,θ plot for glycine.

5. A New Force Field with an Accurate Bond Angle Interaction

We recently proposed a different approach[15] including an accurate bond angle term U^{θ}. We use a quartic double-well potential for the bond angle potential:

$$U^{\theta} = \frac{1}{2}k_{\alpha}(\theta - \theta_{\alpha})^2 + \frac{1}{3}k'(\theta - \theta_{\alpha})^3 + \frac{1}{4}k''(\theta - \theta_{\alpha})^4$$

where $\theta_{\alpha} \simeq 90°$ corresponds to the first minimum (typical of the α helix). The other parameters are related to the position of the second minimum θ_{β}, to the relative stability $\Delta U^{\theta} = U^{\theta}_{\beta} - U^{\theta}_{\alpha}$, and to the (left and right) barriers $\Delta U^{\theta}_{\alpha}$ and $\Delta U^{\theta}_{\beta}$ via the following relations:

$$\Delta = \theta_{\beta} - \theta_{\alpha}$$

$$k' = -\frac{k_{\beta} + 2k_{\alpha}}{\Delta}$$

$$k'' = \frac{k_{\alpha} + k_{\beta}}{\Delta^2}$$

$$\Delta U^{\theta} = \frac{\Delta^2}{12}(k_{\beta} - k_{\alpha})$$

$$\Delta U^{\theta}_{\alpha} = \frac{\Delta^2}{12}\frac{k_{\alpha}^3}{(k_{\alpha} + k_{\beta})^3}(k_{\alpha} + 2k_{\beta})$$

$$\Delta U^{\theta}_{\beta} = \frac{\Delta^2}{12}\frac{k_{\beta}^3}{(k_{\alpha} + k_{\beta})^3}(k_{\beta} + 2k_{\alpha})$$

The dihedral term was taken as a harmonic (or harmonic cosine) with an equilibrium value depending on the a priori known secondary structure. The (amino acid dependent) parameters were derived through a statistical analysis of crystallographic structures, using the Boltzmann inversion.[15] The α,θ plots for helix-propense (green), sheet-propense (blue), and glycine-like amino acids are reported in Figure 5 (left). The accuracy of these maps is especially evident in the glycine-like pattern, where our force field reproduces the small separation between the helixlike and sheetlike regions, and in the sheet pattern (blue) where the peculiar shape with a small protuberance toward $\theta = 90°$ present in the "generic" map of Figure 3 is reproduced. This FF has proven successful in reproducing the flap opening dynamics in HIV-1 protease, which depends on a very peculiar motion of the glycine-rich flap tips, and its dependence on the mutations in this region.[15,22]

Here, we propose a further improvement of the potential. We combine our double-well U^{θ} with a cosine-sum U^{α} similar to the potential of Head-Gordon et al.[8]

$$U^{\alpha} = A[1 + \cos(\alpha)] + B[1 + \cos(3\alpha)] + C[1 + \cos(\alpha + \pi/4)] + D[1 + \cos(2\alpha)]$$

whose terms have the following meanings. The first has a minimum at $\alpha = 180°$ corresponding roughly to the sheet or extended structures. The second has additional minima at $\alpha = \pm 60°$, corresponding to helical structures. The third

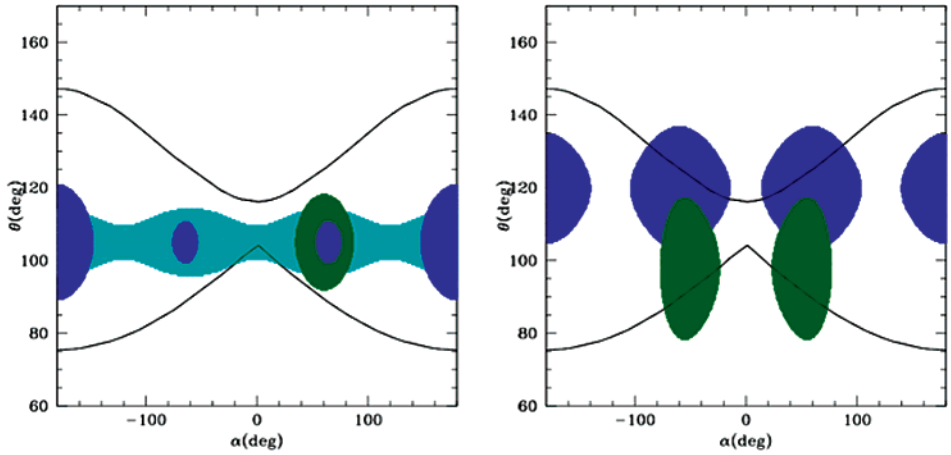


Figure 4. Typical α,θ density maps for two different kinds of FF. Left: Head-Gordon-like FF; right: Mukherjee-like FF. Green, helix; blue, sheet; cyan, turn.

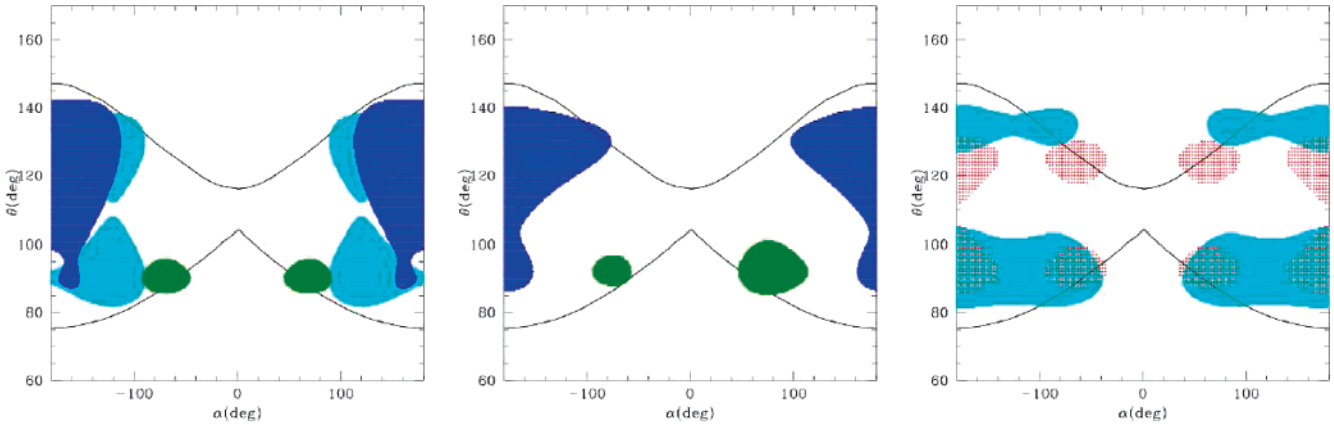


Figure 5. α,θ density maps for the FF in the present work. Left: version with harmonic cosine dihedral potential. Center and right: version with cosine sum dihedral potential. Different maps are obtained giving different weights to the terms, to reproduce the map for sheet (blue), helix (green), glycine (cyan), and generic (red) models.

has a single minimum at $\alpha = 135°$ and introduces an asymmetry that favors the right-handed helices with respect to the left-handed. The fourth has minima at $\alpha = \pm 90°$ and can enforce the helix propensity. By giving different weights to the helixlike or sheetlike terms in U^{α} and U^{θ} and tuning the asymmetry term, one can reproduce quite realistic density plots. In Figure 5, in the center, density plots for helixlike (green) and sheetlike (blue) structures are reported. In Figure 5, on the right, we report the density plot for glycine (cyan), obtained using approximately the same weight for the helix and sheet terms and no asymmetry. Finally, in red, we report a hypothetical "generic" potential as simply as possible, including only the first and second terms in U^{α}. This does not closely resemble any of the "experimental" density plots but contains all the secondary structures with approximately the same weight, also including the very uncommon polyproline helix (located at ∼100,∼120) and its right-handed counterpart.

6. Applications: α-Helix to β-Sheet Transition in a Minimal Polypeptide Model

We now illustrate the above concepts with a simple polypeptide model. We do not aim here to describe any specific poly amino acid type; an accurate amino acid type-dependent parametrization will be the matter of a future paper. However, this model could apply to a sheet former, like polyvaline. In addition to U^{α} and U^{θ}, we used a Morse nonbonded interaction potential

$$U^{nb} = \epsilon(\{1 - \exp[-\alpha(r - \sigma)]\}^2 - 1)$$

where $\sigma = 6.1$ Å is the bead diameter, $\alpha = 0.7$ is the well width parameter, and ϵ is taken as an energy unit. Because the helices, sheets, and random coils are characterized by a different number of nonbonded contacts, the relative stability of the different secondary structures depends on a delicate balance between ϵ and $\Delta U^{\alpha,\theta} = \Delta U^{\alpha} + \Delta U^{\theta}$. We chose the parameters in such a way that $\Delta U^{\alpha,\theta} = -5\epsilon$ favors the sheet conformation. The energy barriers separating the helix from the sheet were set at $\Delta U^{\alpha}_{\alpha} = 6\epsilon$ and $\Delta U^{\alpha}_{\theta} = 7\epsilon$, and a slight asymmetry toward the right-handed helix was added to the dihedral potential (see Figure 6 for the values of the parameters). The Cα−Cα bonds were constrained at the value 3.79 Å. We performed two simulated annealing runs on a 20-mer, starting from the α-helical and extended configurations, respectively. The DL_POLY code was used[23] for the simulation. The results are reported in Figures 6−8.

Figure 7 reports the representative angles and dihedrals along the simulation. In the simulation starting from the α

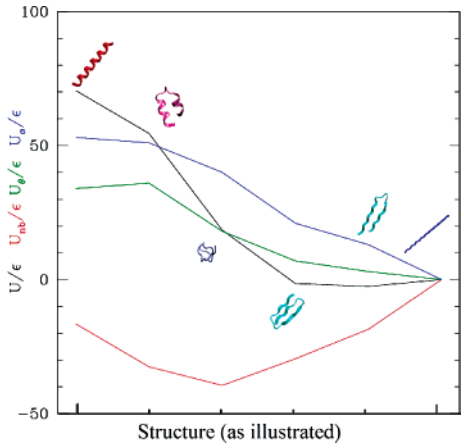


Figure 6. Relative stability and energy terms of the conformations explored during the simulations. Black, total potential energy; red, nonbonded energy; green, bond angle energy; blue, dihedral energy. All energies are expressed in units of ϵ. The extended conformation is taken as the zero level. Parameters of the dihedral and bond angle terms: $A = 2\epsilon$, $B = 4\epsilon$, $C = 0.5\epsilon$, $D = 0$, $\delta U_{\alpha}^{\theta} = 7\epsilon$, and $\delta U_{\alpha}^{\theta} = 9\epsilon$. Representative structures extracted from the simulations are reported: red, α helix; pink, "broken helix"; violet, random coil; cyan, β hairpin and 3-fold β sheet.

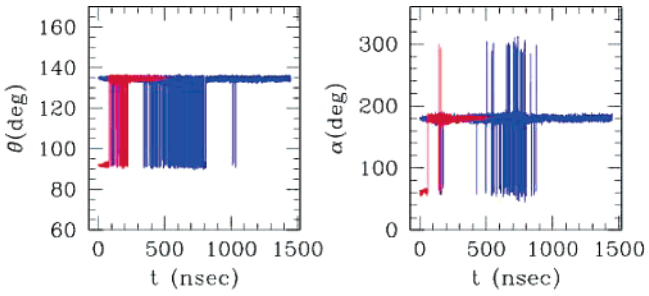


Figure 7. "Representative values" of bond angles and dihedrals evaluated during the simulations starting from the α helix (red) and from the extended structure (blue). The "representative value" is evaluated as follows. For θ, for each time step, all 18 θ angles along the chain are measured, and two average values are calculated, one for the "extended" and one for the "helical" (conventionally separated by the value 113°). The representative value for θ is then chosen as the (average) value of the most populated structure, for each time step. For α, the same procedure is followed, except that in this case three separate average values are calculated for right-handed (between 0 and 115°) or left-handed helices (between −115 and 0°) and for the extended conformation (otherwise).

helix (red), the transition to a globular random coil structure occurs after about 100 ns at the temperature $kT = 0.6\epsilon$, after passing through a metastable "broken helix" conformation (see Figure 6). As the temperature is raised to $kT = \sim\epsilon$, the system explores globular structures, and finally, upon cooling, it stabilizes in a 3-fold β sheet, that is, a β sheet with two β turns located approximately at the 8th and 14th residues along the chain (see Figure 6). In the simulation starting from the extended structure (blue), the system tends to contract to a globule after about 100 ns, then again stabilizes to a 3-fold β sheet. As the temperature is raised to

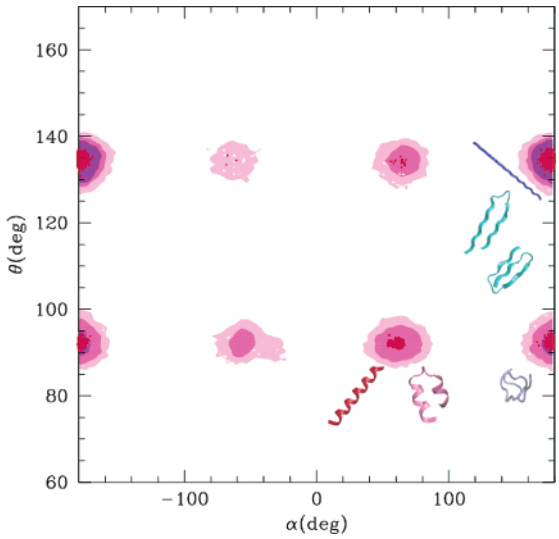


Figure 8. α,θ plot evaluated along the trajectories of the two simulations. Violet: simulation starting from the extended conformation. Pink−magenta−red: simulation starting from the α helix at increasing density values. The representative structures of the populated regions are reported.

$kT \simeq 2\epsilon$ (between 500 and 1000 ns), the system melts into a globular state. During cooling, it explores for long time intervals the β-hairpin structures (β sheets with one β turn approximately in the middle of the chain), before definitely stabilizing in the 3-fold β sheet. Remarkably, the structures we find as (meta)stable states were also observed in simulations performed with more sophisticated models.[24,25]

This dynamical behavior can be explained on the basis of the energy landscape of the system, reported in Figure 6. The energy terms combine in such a way that the helical structures are less stable, while the sheet structures (both β hairpin and 3-fold β sheet) have a very similar energy and are slightly more stable than the extended one. The definitive choice for the 3-fold β sheet may be due to the entropic factor.

The α,θ density plot for the two simulations is shown in Figure 8. The violet map is evaluated on the trajectory of the simulation starting from the extended conformation. Only the area corresponding to the β-sheet and extended structures is populated, with a small population of the globular region located at $\alpha \simeq 180$, $\theta \simeq 90$. The map in red corresponds to the simulation starting from the α helix. The transition to the β sheet passes through the molten globule region; however, as can be seen from the maps obtained with lower levels of populations (magenta and pink), other regions are also (slightly) populated. The pink map resembles the "generic" map (in red in Figure 5), although the effect of the asymmetry term that biases the system toward the "right-handed" structures is evident.

7. Conclusions

In this paper, we have reported two kinds of result. First, we have derived analytical relationships to convert the internal backbone coordinates of the all-atom representation of the polypeptide chain to the internal coordinates of the one- (two-) bead CG representation, namely, the bond angles

and dihedrals θ and α. This result has a general applicability to all cases where the geometry of the backbone must be described and only the coordinates of the Cα's are available, for instance, with low-resolution or CG models. We have used these relationships to convert the Ramachandran plot into new two-dimensional density plots in terms of the backbone variables θ,α. These α,θ density plots can be used as the Ramachandran plot, that is, to test the reliability of protein models, in low-resolution or CG models, when the original Ramachandran plot cannot be used.

Second, we have illustrated these ideas on a minimal one-bead polypeptide model that accounts for the transition from α helices to β sheets. With respect to the models available in the literature, our model contains a more sophisticated bond angle term, which allows one to reproduce quite accurately the α,θ plot. Furthermore, despite its simplicity, this model is shown to explore the conformational space as accurately as more sophisticated multiple-bead models. It is arguable that an optimization and fine-tuning of the parameters of the model can account for the helix or sheet propensities of the different amino acids and reproduce more complex structures and transitions.

Acknowledgment. We acknowledge the allocation of computer resources from the INFM-CNR parallel computing initiative. This work has been supported in part by NSF, NIH, CTBP, NBCR, Accelrys, and the NSF Supercomputing Centers.

Supporting Information Available: The derivation of the exact formula for the dihedral α as a function of ϕ, ψ. This information is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Head-Gordon, T.; Brown, S. *Curr. Opin. Struct. Biol.* **2005**, *13*, 160−167.

(2) Tozzini, V. *Curr. Opin. Struct. Biol.* **2005**, *15*, 144−150.

(3) Levitt, M. *J. Mol. Biol.* **1976**, *12*, 59−107.

(4) Ueeda, Y.; Taketomi, H.; Gō, N. *Biopolymers* **1978**, *17*, 1531−1548.

(5) Nguyen, H. D.; Hall, C. K. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 16180−18185.

(6) Fujisuka, Y.; Takada, S.; Luthey-Schulten, Z. A.; Wolynes, P. G. *Proteins* **2004**, *54*, 88−103.

(7) Favrin, G.; Irbäck, A.; Wallin, S. *Proteins* **2002**, *47*, 99−105.

(8) Brown, S.; Fawzi, N. J.; Head-Gordon, T. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 10712−10717.

(9) Friedel, M.; Shea, J.-E. *J. Chem. Phys.* **2004**, *120*, 5809−5823.

(10) McCammon, J. A.; Northrup, S. H.; Karplus, M.; Levy, R. M. *Biopolymers* **1980**, *19*, 2033−2045.

(11) Keskin, O.; Bahar, I. *Folding Des.* **1998**, *3*, 469−479.

(12) Bahar, I.; Jernigan, R. L. *J. Mol. Biol.* **1997**, *166*, 195−214.

(13) Mukherjee, A.; Bagchi, B. *J. Chem. Phys.* **2004**, *120*, 1602−1612.

(14) Pillardi, J.; Czaplewski, C.; Liwo, A.; Lee, J.; Ripoll, D. R.; Kazmierkiewicz, R.; Ołdziej, S.; Vedemeyer, W. J.; Gibson, K. D.; Arnautova, Y. A.; Saunders, J.; Ye, Y.-J.; Sheraga, H. A.; *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 2329−2333.

(15) Tozzini, V.; McCammon, J. A. *Chem. Phys. Lett.* **2005**, *413*, 123−128.

(16) Zacharias, M. *Protein Sci.* **2003**, *12*, 1271−1282.

(17) Koga, N.; Takada, S. *J. Mol. Biol.* **2001**, *313*, 171−180.

(18) RCSB Protein Data Bank. http://pdbbeta.rcsb.org/pdb/.

(19) Mathews, C. K.; van Holde, K. E.; Ahern, K. G. In *Biochemistry*; Addison-Wesley Longman Inc: Reading, MA, 2000; Benjamin−Cummings series.

(20) Voet, D.; Voet J. G. In *Biochimica*; Zanichelli: Bologna, Italy, 1997.

(21) Reith, D.; Pütz, M.; Müller-Plate, F. *J. Comput. Chem.* **2003**, *24*, 1624−1636.

(22) Chang, C.-E.; Shen, T.; Trylska, J.; Tozzini, V.; McCammon, J. A. *Biophys. J.* in press.

(23) Smith, W.; Forester, T. R. *J. Mol. Graphics* **1996**, *14*, 136. Smith, W.; Yong, C. M.; Rodger, P. M. *Mol. Simul.* **2002**, *28*, 385.

(24) Ding, F.; Borreguero, J. M.; Buldyrey, S. V.; Stanley, H. E.; Dokholyan, N. V. *Proteins* **2003**, *53*, 220−228.

(25) Nguyen, H. D.; Marchut, A. J.; Hall, C. K. *Protein Sci.* **2004**, *13*, 2909−2924.

CT050294K

JCTC Journal of Chemical Theory and Computation

Coarse-Grained Model of Coil-to-Helix Kinetics Demonstrates the Importance of Multiple Nucleation Sites in Helix Folding

Alan E. van Giessen and John E. Straub*

Department of Chemistry, Boston University, Boston, Massachusetts 02215

Received December 26, 2005

Abstract: An extension of a coarse-grained, implicit-solvent peptide model wherein each amino acid residue is represented by four interaction sites is presented and discussed. The model is used to study the coil-to-helix transition of five peptide sequences, ranging from all hydrophobic to all hydrophilic, for a 10-residue peptide. The thermodynamics of the folding transition are analyzed and discussed for each sequence, and the stability of the α-helix is correlated with the hydrophobic content of the sequence. In addition, for each sequence, the folding kinetics of the transition from random coil to full α-helix are analyzed, and the mean folding time is determined. Folding times vary from 59 ns for the most hydrophobic sequence to 132 ns for the most hydrophilic sequence. These folding times compare very well with those measured in experments. All sequences show single-exponential kinetics. A plot of the mean folding time versus the reciprocal of the Zimm−Bragg parameter σ—a measure of the free energy cost of nucleating a helix—is shown to be nonlinear, in contrast to the predictions of many theories of the coil-to-helix transition. It is proposed that the origin of this nonlinearity is due to multiple helix nucleation sites, indicating that even for short peptides such as those studied here, multiple folding pathways play an important role in the transition from random coil to native state.

1. Introduction

One of the fundamental problems in the study of protein folding is the coil-to-helix transition of proteins and peptides which form α-helices. That α-helices are common structural motifs in many biologically relevant proteins only underscores the importance of understanding the coil-to-helix transition. Many theoretical models for this transition have been developed, beginning with the pioneering work of Schellman,[1] Zimm and Bragg,[2] and Lifson and Roig[3] more than 40 years ago. For a recent survey of advances in the theory of the coil-to-helix transition, see the review by Doig.[4]

It is the model of Zimm and Bragg[2] which first defined the parameters by which the coil-to-helix transition is frequently described. In their model, which is isomorphic with the one-dimensional Ising model, the Hamiltonian is

$$H = -J\sum_{ij} s_i s_j - H\sum_{is_i} \tag{1}$$

where the spin−spin coupling constant J and the external field H are expressed in units of kT. Zimm and Bragg introduced the parameters s and σ, defined as

$$\sigma s = \exp(H) \qquad s = \exp(J + H) \tag{2}$$

which determine the free energy of helix propagation and nucleation, respectively. At the folding transition temperature, $s = 1$. Takano and co-workers[5] have shown recently how well the Zimm−Bragg model describes the thermodynamics of the coil-to-helix transition when compared to an all-atom molecular dynamics simulation.

The thermodynamics of the coil-to-helix transition is well understood, due largely to the use of the Zimm−Bragg model. The kinetics of the transition, however, are still poorly understood, despite much progress, both experimental and

* Corresponding author e-mail: straub@bu.edu.

10.1021/ct0503318 CCC: $33.50
Published on Web 04/22/2006

theoretical, over the past decade. Details of the folding rate and of the folding mechanism, or mechanisms, are still unclear. Early estimates for the folding rate were on the order of 1 μs.[6] More accurate temperature-jump experiments by Williams et al.[7] revealed a folding time of 160 ns for a 21 residue alanine-based peptide. Thompson et al.[8] and Lednev et al.[9] measured a slightly longer folding time of 220 ns and 240 ns, respectively. Using a stopped-flow CD measurement, Clarke[10] and co-workers measured a considerably longer folding time of miliseconds for a 16-residue peptide. Another set of T-jump experiments on a 153-residue globular protein by Woodruff and co-workers[11−13] showed a relaxation rate in the range of 10−160 ns. In the past few years, Gai and co-workers have performed both stopped-flow experiments[14,15] and temperature jump experiments[16−18] that have established a time scale on the order of 200 ns.

Theoretical predictions for the folding time of an α-helix tend to be considerably shorter than the 200 ns time scale established by experiment, though some are in agreement. An early MD simulation by Daggett and Levitt[19] suggested the time scale for helix propagation to be 100 ps. Coarse-grained simulations by Thirumalai and co-workers indicated the folding time at the folding temperature to be roughly 20 ns.[20] A different coarse-grained model developed by Takada and co-workers,[21] which will be discussed in more detail below, gives a folding time of approximately 15 ns. Margulis et al.[22] measured a folding time of 1 ns for a molecular dynamics simulation of an alanine pentapeptide in explicit solvent. One calculation that does agree with experiment is the nucleation-elongation theory of Doshi and Muñoz.[23] They establish a folding time of 150−300 ns.

Many models of the coil-to-helix transition have sought to relate the folding time, τ, to the Zimm−Bragg parameters s and σ. In a seminal work, Schwartz[24] estimated the folding time at the midpoint of the coil-to-helix transition (when $s=1$) to be

$$\tau = \frac{1}{4\sigma k_F} \tag{3}$$

where k_F is the rate of adding an additional helical residue at the helix end. Brooks[25] proposed a model based on a sequential formation of helical residues and demonstrated that the mean time for the folding/unfolding process scaled as $\tau \sim 1/\sigma$. In all these models, it is found that the mean folding time, τ, is inversely proportional to the Zimm−Bragg parameter σ. In addition, we note that all these models include the assumption that the helix propagates from a single nucleation site.

One recent theoretical model that relaxes this assumption is that due to Buchete and Straub.[26] This model, referred to as the active helix Ising model, is also based on the Zimm−Bragg model and allows one to solve the mean-first passage time equation. Buchete and Straub numerically determined the mean first passage time for a range of $1/\sigma$ values at fixed s and found significant nonlinear behavior. They observed linear behavior for small values of σ and strongly nonlinear behavior for large values of σ ($\sigma \geq 0.005$). This nonlinear behavior is increasingly important for longer polypeptide chains and for smaller values of the propagation constant s.

In contrast to almost all theoretical predictions, Gai[17] found evidence that the folding time for the coil-to-helix transition does *not* scale linearly with $1/\sigma$. The only model which predicts this behavior is the active helix Ising model of Buchete and Straub.[26] As noted above, one of the key assumptions in many models, but not that of Buchete and Straub, is that a helix is formed from a single nucleation site. It is suggested[26] that the origin of the nonlinearity lies in the ability of a peptide, even one as short as 10 residues in length, to have multiple nucleation sites. In this work, we propose to study the kinetics of the coil-to-helix transition via computer simulation for several model peptides in order to investigate the dependence of the folding time on the Zimm−Bragg parameter σ and to determine the relevance of multiple nucleation sites.

A natural choice for this investigation is the use of molecular dynamics (MD) simulations. The most accurate approaches employ all-atom MD simulations using an explicit molecular representation of the solvent. At the present time, for studies of the thermodynamics and kinetics of large-scale conformational transitions, and for phenomena that occur on time scales of hundreds of nanoseconds, such approaches are computationally too demanding in applications involving all but small peptides and proteins. Consequently, there is an ongoing effort to develop coarse-grained models of proteins using a reduced number of degrees of freedom. The most appealing approach is to include solvent effects implicitly in the interaction potentials and to replace the atoms in each amino acid residue by a small number of interaction sites, thereby drastically reducing the number of particles and interactions necessary for the calculation.

In this work, we present an extension of one such reduced model, originally developed by Takada et al.[21] In section 2, the peptide model is presented, while in section 3, the Langevin dynamics used to propagate the motion of the peptide forward in time is described. In section 4, the thermodynamcis of the coil-to-helix transition is discussed as well as the kinetics of the folding transition. Finally, section 5 presents some conclusions.

2. Peptide Model

The coarse-grained model used for the peptide in this work is a refinement of that proposed by Takada et al.[21] Similar structural models have been used by Hall[27] to study peptide aggregation. The structural model consists of four particles or "united atoms" per amino acid residue, shown schematically in Figure 1. Three of these united atoms represent the peptide backbone: one represents the amide nitrogen and its hydrogen, another the α-carbon and its hydrogen, and the third the carbonyl carbon and its oxygen. This high level of backbone representation is essential for reproducing correct secondary structure in the folded peptide.[28] The fourth united atom represents the amino acid side chain.

The model as presented here includes two types of side-chain interaction sites: hydrophobic and hydrophilic. While it is possible to introduce more detailed interaction potentials that more closely mimic the chemical identity of all the amino acids, for the present purpose it is sufficient to limit ourselves to this "two-letter" amino acid model. It has long

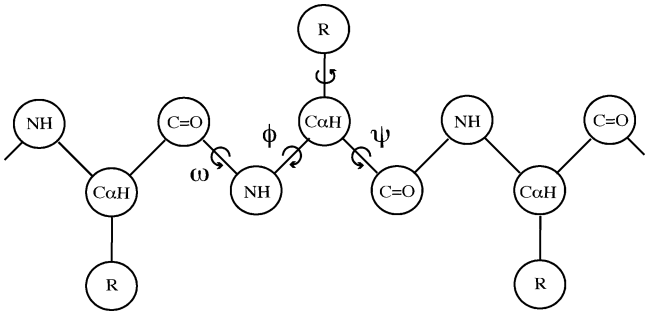


Figure 1. Schematic representation of the peptide model showing the dihedral angles ϕ, ψ, and ω. All bond lengths are fixed.

Table 1. Amino Acid Sequences

label	sequence[a]	T_f	τ (ns)
A	*PPPPPPPPPP*	411.3	59.4
B	*HPPPHPPPHP*	366.6	78.6
C	*PHPHPHPHPH*	349.4	118.7
D	*PHHHPHHHPH*	344.0	120.2
E	*HHHHHHHHHH*	323.8	131.7

[a] "P" denotes hydrophobic and "H" denotes hydrophilic amino acid residues.

Table 2. Structural Parameters

van der Waals diameters	σ (Å)	σ_{local} (Å)
C_α	3.30	2.64
C'	3.56	2.94
N	2.94	2.36
C_β	4.50	4.50

bond lengths	r (Å)
$C_\alpha-C'$	1.52
$C_\alpha-N$	1.45
$C'-N$	1.33
$C_\alpha-C_\beta$	1.80

bond angles	degrees	k_θ
$N-C_\alpha-C'$	111.6	200.0
$C_\alpha-C'-N$	117.5	200.0
$C'-N-C_\alpha$	120.0	200.0
$C'-C_\alpha-C_\beta$	110.0	200.0
$N-C_\alpha-C_\beta$	110.0	200.0

been realized that such a coarse-grained approximation captures many fundamental aspects of protein folding and can adequately be used to study the effect of amino acid sequence on equilibrium and dynamic properties.

In this work, we study five different 10-residue amino acid chains in order to determine the effect of amino acid sequence on the equilibrium properties of the coil-to-helix transition and on the mean folding time for helix formation. We use the letter *P* to denote a hydrophobic residue and the letter *H* to denote a hydrophilic reside. The five sequences are given in Table 1. Sequence A consists of 10 hydrophobic residues, while sequence E consists of 10 hydrophilic residues. Sequences B, C, and D are mixtures of hydrophobic and hydrophilic residues, with, in order, increasing hydrophilic content.

Table 3. Energetic Parameters

torsion potential	kcal/mol	
$v_{2,\phi}$	0.00	$-\pi < \phi < 0$
$v_{2,\phi}$	0.20	$0 < \phi < \pi$
$v_{3,\phi}$	0.45	$-\pi < \phi < -\pi/3$
$v_{3,\phi}$	4.00	$-\pi/3 < \phi < \pi/3$
$v_{3,\phi}$	0.45	$\pi/3 < \phi < \pi$
$v_{2,\psi}$	0.00	$-\pi < \psi < 0$
$v_{2,\psi}$	0.00	$0 < \psi < \pi$
$v_{3,\psi}$	1.50	$-\pi < \psi < -\pi/3$
$v_{3,\psi}$	4.00	$-\pi/3 < \psi < \pi/3$
$v_{3,\psi}$	0.45	$\pi/3 < \psi < \pi$
v_ω	40.0	

chiral potential	kcal/mol	degrees
k_χ	100.0	
$\chi_{0,i}$		52.52

van der Waals potential	kcal/mol
ϵ	0.060
ϵ_{local}	0.033

The interaction potentials can be divided into two types: local and nonlocal

$$V = V_{local} + V_{nonlocal} \tag{4}$$

The local interaction potentials consist of the bond angle, dihedral angle, 1−4 van der Waals, and an improper dihedral potential to maintain the chirality of the side chain

$$V_{local} = V_{BA} + V_\phi + V_\psi + V_\omega + V_{vdW-local} + V_\chi \tag{5}$$

The values for the various structural and energetic parameters are given in Tables 2 and 3. The bond angle potential is harmonic about the equilibrium bond angle and is given by

$$V_\theta = \sum \frac{1}{2} k_\theta (\theta_i - \theta_{0,i}) \tag{6}$$

The dihedral angle energy is given by

$$V_{TOR} = V_\phi + V_\psi + V_\omega \tag{7}$$

with

$$V_\phi = \sum_i \frac{1}{2}[v_{2,\phi}(1 - \cos 2\phi_i) + v_{3,\phi}(1 + \cos 3\phi_i)] \tag{8}$$

$$V_\psi = \sum_i \frac{1}{2}[v_{2,\psi}(1 - \cos 2\psi_i) + v_{3,\psi}(1 + \cos 3\psi_i)] \tag{9}$$

$$V_\omega = \sum_i \frac{1}{2} v_\omega (1 + \cos \omega_i) \tag{10}$$

The values for $v_{2,\phi}$, $v_{3,\phi}$, $v_{2,\psi}$, $v_{3,\psi}$, and v_ω were carefully chosen to produce, in conjunction with the van der Waals potential, a Ramachandran plot with realistic energy barriers for the alanine dipeptide,[29] shown in Figure 2. The barrier height between the α-helix and β-sheet regions is 4.2 kcal/mol. These values are also given in Table 3. The box in Figure 2 defines the α-helical region and is centered on the

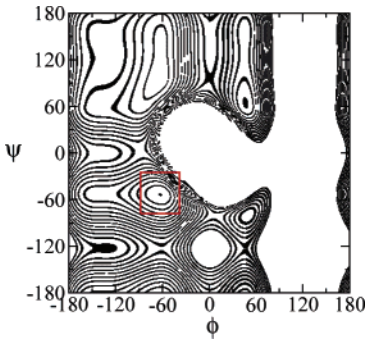


Figure 2. Ramachandran plot for the alanine dipeptide. Contour lines are 0.5 kcal/mol apart. The global minimum is at $\phi = -63$, $\psi = -54$. The "box" defines the α-helical region (see text for details).

average values of $\langle\phi\rangle_\alpha = -63°$ and $\langle\psi\rangle_\alpha = -54°$, where the subscript α indicates the averages are over helical configurations. The box in Figure 2 defines the α-helical region and marks out the area 25 degrees on either side of the minimum. Finally, the chirality is preserved via a harmonic potential for the improper dihedral angle formed by the vectors connecting the α-carbon and the nitrogen, the carbonyl carbon and the nitrogen, and the β-carbon and the α-carbon. The potential has the form

$$V_\chi = \sum \frac{1}{2} k_\chi (\chi_i - \chi_{0,i}) \tag{11}$$

Values for k_χ and $\chi_{0,i}$ are given in Table 3.

The nonlocal interaction has three contributions: the van der Waals interaction, hydrogen bonding, and the hydrophobic effect

$$V_{\text{nonlocal}} = V_{\text{vdW}} + V_{\text{HB}} + V_{\text{HP}} \tag{12}$$

The novel feature of the model developed by Takada et al.[21] is the dependence of the hydrogen bonding scheme and the hydrophobic interaction on the local peptide density, here referred to as ρ_I for residue I. The strength of both interactions is scaled by a function $0 \le S(\rho) \le 1$, which depends on the local density. For example, in the hydrogen bond potential, S is small for low peptide densities and equal to 1 for high densities. In this way, the competition for hydrogen bond formation between the (implicit) solvent and other hydrogen bond donors on the peptide is mimicked.

In full, the hydrogen bond potential is

$$V_{\text{HB}} = \epsilon_{\text{HB}} \sum_{ij(I \ge J+3)} S_{\text{HB},IJ} u_{\text{HB}}^{(\text{a,r})}(r_{ij}) + \frac{1}{2} \epsilon_{\text{HB}} \sum_I S_{\text{HB,c}}(\rho_I) \tag{13}$$

where i is the ith atom and is located in residue I. The strength of the interaction is given by ϵ_{HB}. The scaling function $S_{\text{HB},IJ}$ is given by $S_{\text{HB},IJ} = [S_{\text{HB}}(\rho_I) + S_{\text{HB}}(\rho_J)]/2$. The distance-dependent interaction between i and j is

$$u_{\text{HB}}^{(\text{a})}(r_{ij}) = 5\left(\frac{\sigma_{\text{HB}}}{r_{ij} - r_{\text{HB}}}\right)^{12} - 6\left(\frac{\sigma_{\text{HB}}}{r_{ij} - r_{\text{HB}}}\right)^{10} \tag{14}$$

and

$$u_{\text{HB}}^{(\text{r})}(r_{ij}) = 3\left(\frac{\sigma_{\text{HB}}}{r_{ij} - r_{\text{HB}}}\right)^{10} \tag{15}$$

The superscript (a) stands for attractive and (r) for repulsive. The attractive and repulsive forces in the hydrogen bond potential were introduced by Takada et al.[21] to model the anisotropy of the hydrogen bond. The attractive potential, eq 14, is used for the interaction between a carbonyl carbon and an amide nitrogen, while the repulsive potential, eq 15, is used for the interaction between an amide nitrogen with another amide nitrogen, between an amide nitrogen and an α-carbon, between a carbonyl carbon and another carbonyl carbon, and between a carbonyl carbon and an α-carbon. For a more indepth description of the hydrogen bond potential, readers are referred to the original work of Takada et al.[21]

The scaling function $S_{\text{HB}}(\rho_I)$, where $\rho_I = \sum_K u_{\text{HP}}(r_{\text{IK}})$, is defined as

$$S_{\text{HB}}(x, x_{\min}, x_{\max}) = \begin{bmatrix} 0 & \text{if } x < x_{\min} \\ \frac{1}{2}\left(1 + \cos\left(\pi \frac{x_{\max} - \boldsymbol{x}}{x_{\max} - x_{\min}}\right)\right) & \text{if } x_{\min} \le x \le x_{\max} \\ 1 & \text{if } x > x_{\max} \end{bmatrix} \tag{16}$$

The second term in eq 13 is a penalty term that accounts for buried non-hydrogen-bonded pairs. The function $u_{\text{HP}}(r)$ is defined below.

There are two contributions to the hydrophobic interaction, one from the side chains and one from the α-carbons. The interaction is given by

$$V_{\text{HP}} = \sum_I \delta_I \epsilon_{\text{HP},I}^{(\alpha)} S_{\text{HP}}(\rho_I) + \sum_\mu \delta_\mu \epsilon_{\text{HP},\mu}^{(\beta)} S_{\text{HP}}(\rho_\mu) \tag{17}$$

where I represents α-carbons and μ represents side-chain interaction sites. The parameter $\delta_{I,\mu}$ is equal to 1 for hydrophobic residues and -1 for hydrophilic residues. Similar to that used in the hydrogen bonding scheme, the scaling function $S_{\text{HP}}(\rho_I)$, where $\rho_I = \sum_K u_{\text{HP}}(r_{IK})$, is defined as

$$S_{\text{HP}}(x, x_{\max}) = \begin{bmatrix} \cos\left(\frac{\pi}{2} \frac{x_{\max} - x}{x_{\max}}\right) & \text{if } x \le x_{\max} \\ 1 & \text{if } x > x_{\max} \end{bmatrix} \tag{18}$$

The switching function u_{HP} used in both the hydrogen bonding scheme and the hydrophobic interaction is defined by

$$u_{\text{HP}} = \begin{bmatrix} 1 & \text{if } r < \sigma_{\text{HP1}} \\ \frac{1}{2}\left(1 + \cos\left(\pi \frac{r - \sigma_{\text{HP1}}}{\sigma_{\text{HP2}} - \sigma_{\text{HP1}}}\right)\right) & \text{if } \sigma_{\text{HP1}} < r < \sigma_{\text{HP2}} \\ 0 & \text{if } r > \sigma_{\text{HP2}} \end{bmatrix} \tag{19}$$

Finally, the van der Waals interaction is given by

$$V_{\text{vdW}} = \sum_{i,j>i} \phi_{ij}(r) \tag{20}$$

where

$$\phi_{ij}(r) = 4\epsilon\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \quad (21)$$

The cross-diameter σ_{ij} is given by $\sigma_{ij} = [\sigma_i + \sigma_j]/2$. For interactions between particles connected by three covalent bonds (1−4 pairs), the interaction strength ϵ and the diameter σ_{ij} are replaced by their reduced or "local" counterparts ϵ_{local} and $\sigma_{ij,\text{local}}$. The reduced parameters are introduced because these short-range interactions are better modeled by using the atomic parameters instead of the united atom parameters. Values for these parameters are given in Tables 2 and 3.

3. Langevin Dynamics

The motion of the model peptide is described using Langevin dynamics. The physical interaction between the solute molecule and solvent is mimicked by a random force, Γ, and solvent viscosity is modeled by a damping term with a coefficient ζ. The equation of motion for a generalized coordinate x_i is

$$m\ddot{x}_i = F_i + \Gamma_i - \zeta_i\dot{x}_i \quad (22)$$

The force on particle i due to the molecular configuration is represented by F_i. The friction constant ζ_i is related to the viscosity of water, η, by Stoke's Law

$$\zeta_i = 6\pi a_i\eta \quad (23)$$

where a_i is the effective radius of each particle and is equivalent to the sum of the van der Waals radius and the radius of a water molecule, 1.4 Å. As is usual in Langevin dynamics, the random force Γ_i has a mean of zero and a variance of

$$\langle\Gamma_i(t)\Gamma_i(t')\rangle = 2\zeta_i k_B T\delta(t - t') \quad (24)$$

Equation 22 is solved using the Velocity Verlet algorithm. The position at time $t + h$ is given by[30]

$$x_i(t + h) = x_i(t) + h\dot{x}_i(t) + \frac{h^2}{2m}[F_i + \Gamma_i - \zeta_i\dot{x}_i(t)] \quad (25)$$

Similarly, the velocity at time $t + h$ is given by

$$\dot{x}_i(t + h) = (1 - \alpha)(1 - \alpha + \alpha^2)\dot{x}_i(t) + \frac{h}{2m}(1 - \alpha + \alpha^2)[F_i(t) + \Gamma_i(t) + F_i(t + h) + \Gamma_i(t + h)] \quad (26)$$

where we have used the shorthand $\alpha = h\zeta_i/2m$. Bond lengths are held fixed via the RATTLE algorithm.

To improve the sampling of phase space, the Replica Exchange Method[31−34] is used. In this method, several *noninteracting* replicas are simulated in parallel, each at a different temperature. At regular intervals, a Monte Carlo exchange step is attempted between two replicas, say i and j, at neighboring temperatures, T_i and T_j. The transition probability of this replica exchange is given by

$$W(X \rightarrow X') = \begin{bmatrix} 1 & \text{if } \Delta \le 0 \\ \exp\{-\Delta\} & \text{if } \Delta > 0 \end{bmatrix} \quad (27)$$

where

$$\Delta = (\beta_i - \beta_j)[E_j - E_i] \quad (28)$$

Here, E_i is the potential energy for replica i at temperature $\beta_i = 1/kT_i$. The temperatures are chosen to be equally spaced on a logarithmic temperature scale. Exchanges are attempted every 10 ps, and the replica exchange acceptance ratios vary from 15% to 40%.

4. Results and Analysis

4.1. Thermodynamics. The thermodynamic and structural properties of all five peptides were studied using replica exchange MD. The fluctuations in the total energy and in the molecular configuration were measured. The first, given by the heat capacity C_v, is commonly used to determine the location of the collapse transition T_θ. The heat capacity is defined by

$$C_v(T) = \frac{\partial E_{\text{total}}}{\partial T} = \frac{\langle E^2\rangle - \langle E\rangle^2}{k_B T^2} \quad (29)$$

where the first equality is from thermodynamics, and the second is from statistical mechanics. Figure 3 shows the heat capacity for all five peptides as a function of the temperature. The peak in C_v is the collapse transition temperature. The data shown in Figure 3 were determined via the second equality and were subjected to the weighted histogram analysis method (WHAM). The heat capacity was also determined via the first equality and found to be in agreement with the statistical mechanical definition.

The peak in the conformational fluctuations of the peptide is used to determine the folding temperature, T_f, below which the polypeptide is predominantly in the native configuration. A measure of how much a given conformation differs from the native state is given by the parameter χ, called the "overlap function". There is no unique way of defining such a parameter, though all reasonable definitions lead to similar results. We follow Vietshans et al.[30] in defining χ as

$$\chi = \frac{1}{N_\alpha{}^2 - 5N_\alpha + 6}\sum_{i=1}^{N_\alpha-3}\sum_{j=i+3}^{N_\alpha}\Theta(\epsilon - |r_{ij} - r_{ij}^N|) \quad (30)$$

Here, N_α corresponds to the number of α-carbons, r_{ij} is the distance between α-carbons i and j, and r_{ij}^N is the same distance in the native state. Θ is the Heaviside function and is equal to 1 when its argument is positive and is equal to zero otherwise. Specifically, the Heaviside function is 1 when the difference between the pair distance r_{ij} and distance between the same pair in the native state, r_{ij}^N, is less than some tolerance ϵ. Thus, only "nativelike" pair-distances contribute to the sum in eq 30. The parameter ϵ is set to 0.5 Å. Note that χ is equal to 1 in the native state. We define the native state as a helix with ϕ and ψ angles of 63° and 54°, respectively. The fluctuations in χ are measured by

$$\Delta\chi = \langle\chi^2\rangle - \langle\chi\rangle^2 \quad (31)$$

The behavior of $\Delta\chi$ is shown in Figure 4. The temperature of the peak in $\Delta\chi$ for each sequence is the same as that for the heat capacity. Values for the folding temperature are given in Table 1. Note that the location of the peaks for both

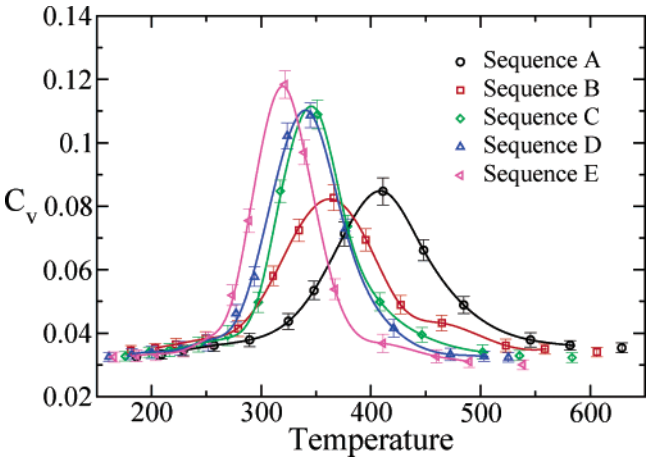


Figure 3. A plot of the heat capacity C_v versus the temperature. The peak in C_v is the collapse temperature, T_θ. The circles are sequence A (black), the squares are sequence B (red), the diamonds are sequence C (green), the up triangles are sequence D (blue), and the left triangles are sequence E (magenta). The colors and symbols are consistent throughout.

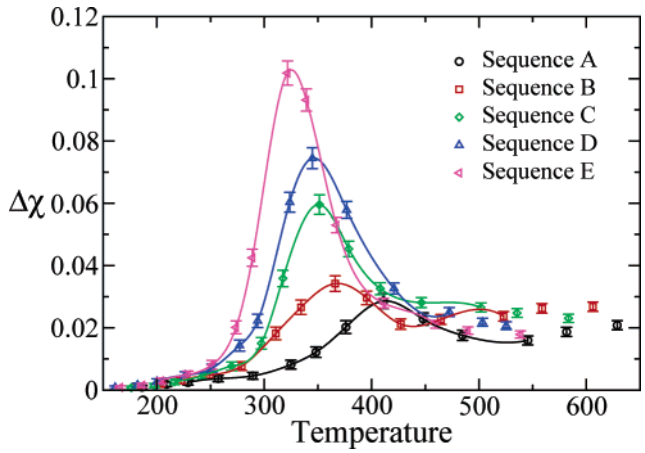


Figure 4. The fluctuations $\Delta\chi$ plotted versus the temperature. The peak in $\Delta\chi$ is the folding temperature, T_f. In this model, $T_f = T_\theta$. Note that the magnitude of the fluctuations increases, and the width of the transition region decreases as the hydrophilic content increases. The symbols are the same as in Figure 3.

C_v and $\Delta\chi$ shift to higher temperatures with increasing hydrophobicity. The more hydrophilic peptides are more easily solvated and consequently have a less stable α-helix and a folding transition at a lower temperature. The relative heights of the peaks are also consistent with the peaks in the heat capacity: the more hydrophobic sequences have smaller fluctuations at the folding transition than the more hydrophilic sequences. This is to be expected, as the more hydrophilic the peptide, the more readily it will be solvated. Note that while sequences C and D have similar transition temperatures, the width of the transition region for sequence C is narrower than that for sequence D.

Comparison with a plot of the radius of gyration, R_g, as a function of temperature, shown in Figure 5, clearly shows that R_g changes dramatically within the transition region centered on T_θ. Above T_θ, the radius of gyration increases with increasing temperature with the rate of increase greater

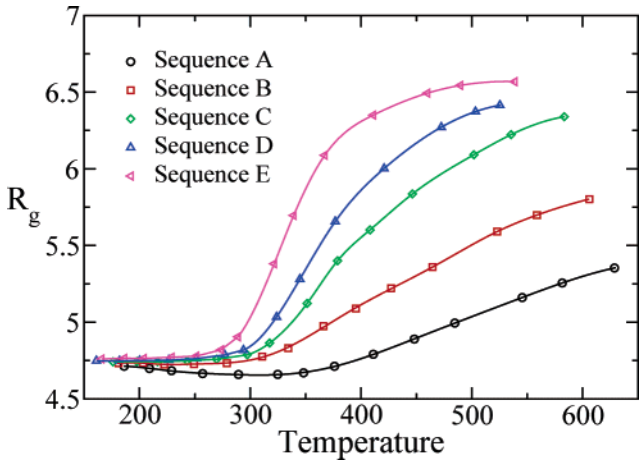


Figure 5. The radius of gyration R_g as a function of temperature. Note that R_g for sequence A has a mimumim. The symbols are the same as in Figure 3.

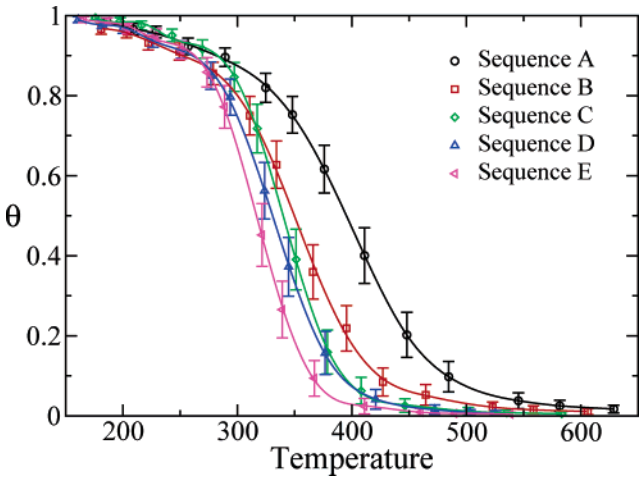


Figure 6. A plot of the fractional helicity θ versus temperature. The width of the transition region decreases with increasing hydrophilic content. The symbols and colors are the same as in Figure 3.

the more hydrophilic the peptide sequence. Below T_θ, the peptides all have a constant radius of gyration equal to that of a full α-helix, while above T_θ, the radius of gyration increases with increasing T. One interesting feature of the radius of gyration for sequence A is that it passes through a minimum at 305 K. This is due to the strong hydrophobic attraction of the P side chains which cause the α-helix to compress slightly. At very low temperatures, this compression is canceled by the bond- and dihedral angle potentials, which become increasingly important as the temperature decreases.

Figure 6 shows the behavior of the average helicity θ as a function of temperature. The helicity is defined as

$$\theta = \frac{N_H}{N_{res} - 4} \tag{32}$$

where N_H is the number of helical hydrogen bonds and N_{res} is the number of residues in the polypeptide chain. For each peptide sequence, $\theta = 1$ at low temperatures, indicating a full α-helix, and $\theta \approx 0$ at high temperatures, where each

peptide is fully unfolded and has no α-helical hydrogen bonds. The width of the transition region decreases with increasing hydrophobicity. Due to the short length of the peptides, sequence effects play a role in the behavior of θ. For example, sequence B has the lowest value for θ at low temperatures when it would be expected, as a more hydrophilic sequence, to have a value comparable to sequence A. This is due to fraying of the helix at the N-terminus.

One of the most useful models to describe the coil-to-helix transition is due to Zimm and Bragg.[2,5] Two parameters are central to their analysis: s, which is related to the free energy of helix propagation, and σ, which is related to the free energy of helix nucleation.

The fractional helicity for a peptide of N residues is given by

$$\theta = \frac{1}{N}\frac{\partial \ln Q}{\partial \ln s} \tag{33}$$

where Q is the partition function and can be written as a sum over eigenfunctions of the transition matrix

$$Q = \sum_n \lambda_n^N \tag{34}$$

In the case of the Zimm−Bragg model, $n = 2$, the partition function can be written as

$$Q = \lambda_+^N + \lambda_-^N \tag{35}$$

where

$$\lambda_\pm = \frac{1}{2}(1+s) \pm \frac{1}{2}\sqrt{(1-s)^2 + 4\sigma s} \tag{36}$$

To determine s and σ, eq 33 was fit to the fractional helicity determined via simulation and shown in Figure 6. Following the analysis of Takano et al.,[5] we have used the Levenburg-Marquardt nonlinear least-squares algorithm[35] to fit eq 33.

The behaviors of s and σ as a function of temperature are shown in Figures 7 and 8, respectively. In agreement with Ohkubo and Brooks,[36] we find that the large-N approximation for s and σ used in previous work[37−39] is not suitable for chain lengths of 10 residues. The behavior of s is qualitatively similar to that seen by Ohkubo and Brooks.[36] For a given T, the more hydrophobic sequence generally has a higher value of s. The exception is for sequence B, since s was determined from the fractional helicity, and the fractional helicity was low at low temperatures due to fraying of the helix ends. At the lowest temperatures sampled, s is between 1.5 and 2.0 for all sequences. As the temperature increases, s decreases, passing through 1.0 at a lower temperature than the folding temperature as determined by the peak in $\Delta\chi$. At the highest temperatures sampled, s is approximately 0.5, which is higher than expected, though the more hydrophilic the sequence, the smaller the value of s.

The behavior of σ is very interesting. At low and high temperatures, σ is appoximately 0.005 for the most hydrophobic sequence and 0.001 for the most hydrophilic sequence. However, at the folding temperature, σ has a maximum. The most hydrophobic sequence has a maximum

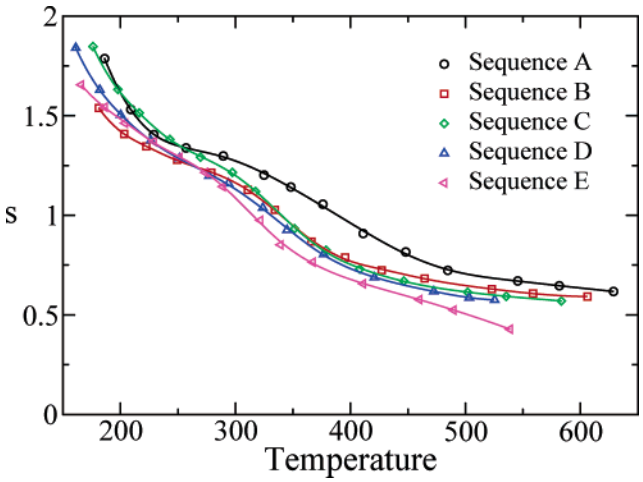


Figure 7. The Zimm−Bragg paramater s plotted as a function of temperature. A measure of the folding temperature is where $s = 1$. In this case, $T(s = 1)$ is generally lower than the folding temperature as determined by the peak in the heat capacity. The symbols and colors are the same as in Figure 3.

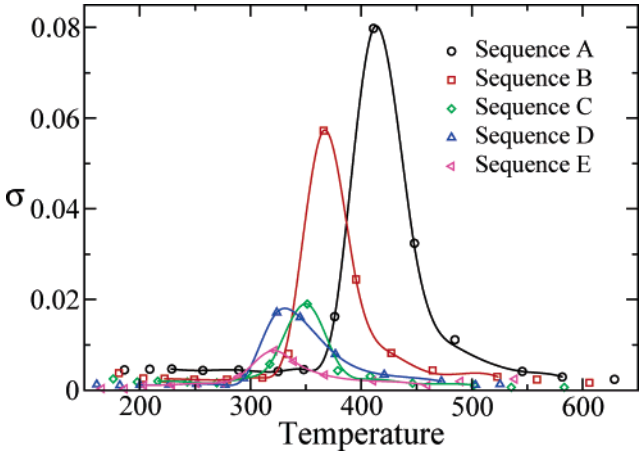


Figure 8. The Zimm−Bragg parameter σ shows a strong peak at the folding temperature. The magnitude of the peak increases as the hydrophobic content increases. The symbols and colors are the same as in Figure 3.

at $\sigma = 0.08$, while the most hydrophilic sequence has a peak of only $\sigma = 0.009$. The behavior for σ of sequence A is similar to that seen by Ohkubo and Brooks.[36] However, the value of σ at high temperatures determined here is considerably smaller than seen in that work.

4.2. Kinetics. To determine the sequence dependence of the kinetics of the coil-to-helix transition, a procedure similar to that developed by Veitshans et al.[30] was used. For each sequence, $M = 400$ independent initial configurations, generated from a high-temperature simulation run, were quenched to the folding temperature and were allowed to propagate forward via eqs 25 and 26 until the overlap function χ was equal to 1.0, whereupon the simulation was stopped. The fraction of unfolded peptides as a function of time, $P_u(t)$, was then used to characterize the folding kinetics. $P_u(t)$ is defined as

$$P_u(t) = 1 - \int_0^t P_{fp}(s)ds \tag{37}$$

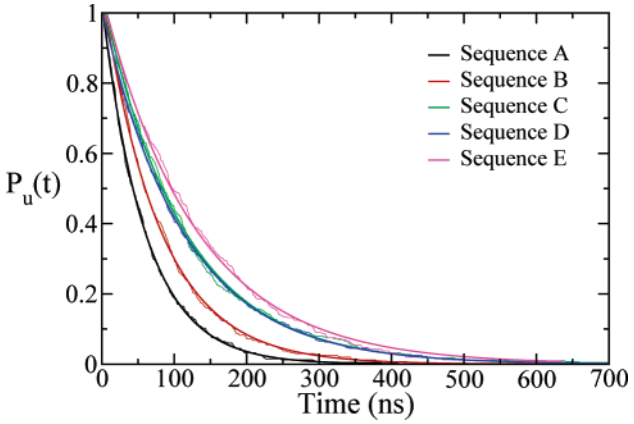


Figure 9. A plot of the fraction of unfolded peptides $P_u(t)$ as a function of simulation time (in nanoseconds). Each sequence displays single-exponential kinetics. Black is sequence A, red is sequence B, green is sequence C, blue is sequence D, and magenta is sequence E.

where $P_{fp}(t)$ is the distribution of first passage times

$$P_{fp}(s) = \frac{1}{M}\sum_{i=1}^{M}\delta(s - \tau_i) \tag{38}$$

and where τ_i is the first passage time of trajectory i. $P_u(t)$ for each sequence was then fit to a single exponential

$$P(t) = A_0\exp(-t/\tau_0) \tag{39}$$

where we shall refer to τ_0 as the mean folding time. Unlike Veitshans et al.,[30] $P_u(t)$ was best fit by a *single* exponential. Single-exponential folding kinetics were also seen by Bredenbeck et al.[40] in the folding of a 16-residue helix-forming peptide. Both $P_u(t)$ and the best-fit exponential for each sequence are shown in Figure 9. That the kinetics is best described by a single exponential is in line with what would be predicted from the value of σ_f, a measure of the "foldability" of the peptide. The parameter σ_f introduced by Thirumalai is defined by

$$\sigma_f = \frac{T_\theta - T_f}{T_\theta} \tag{40}$$

where T_θ and T_f are the collapse and folding temperatures determined in the previous section. For these model peptides, $T_\theta = T_f$, and $\sigma_f = 0$. Peptides with $\sigma_f \approx 1$ encounter misfolded structures, some of which can be very stable and which serve as kinetic traps.[30] Peptides with $\sigma_f \approx 0$ have folding kinetics that show two-state behavior,[30] which is best fit by a single exponential.

As expected, the most hydrophobic sequence has the fastest folding time, while the most hydrophilic sequence has the slowest. The folding times are given in Table 1 and vary from 59 ns for the fastest to 132 ns for the slowest. Sequences B, C, and D behave as expected, with the folding time increasing with increasing hydrophilic content. The data for sequences C and D lie on top of each other, though sequence C does fold slightly faster.

Two main folding pathways were observed for the coil-to-helix transition. The first, by which approximately 80% of the trajectories folded, began with a single helix nucleation site, followed by the growth of the helix to encompass the entire peptide. This mechanism is illustrated in Figure 10. In this figure, the unfolded peptide is shown in part A, a structure with a single helical hydrogen bond near the center of the peptide is shown in part B. In part C, the helix has grown in both directions and includes one of the end termini, and finally in part D, the full helix. The second pathway is the more interesting of the two, as helix formation begins with *two* helix nucleation sites. This pathway is shown in Figure 11. In part A of that figure, the unfolded peptide is shown. In part B, the first helix nucleation site is formed at the C-terminus. Before this helix can grow to include the entire peptide, a second helix is nucleated at the N-terminus, shown in part C. Given the small size of the peptide, multiple helix nucleation can only occur with a nucleation site at each terminus. Finally, in part D, the two helices meet to form a full helix. Multiple folding pathways have been seen in larger proteins,[41] and peptides with multiple helical structures were seen by Nymeyer and Garcia at low temperatures in simulations of the folding of a 21-residue helical peptide.[42]

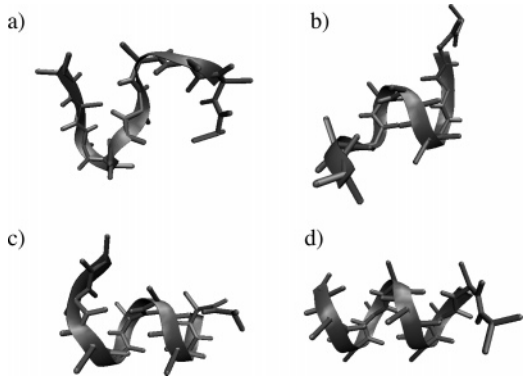


Figure 10. An illustration of a folding pathway that proceeds from a single nucleation site. Part (a) is the unfolded peptide, part (b) shows a single helical turn, in part (c) the peptide has grown in each direction, and part (d) shows the fully folded peptide.

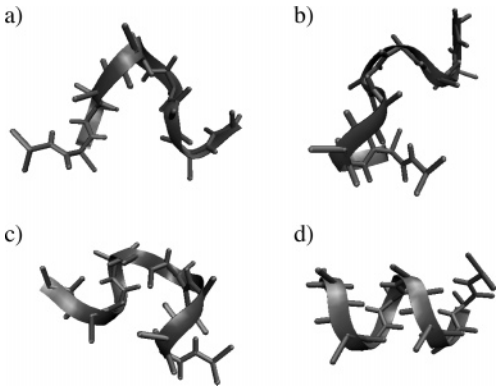


Figure 11. An illustration of a folding pathway that proceeds via multiple helix nucleation sites. Part (a) is the unfolded peptide, part (b) shows a single nucleated helix at the C-terminus, in part (c) the peptide has *two* helical segments, one at each terminus, and part (d) shows the fully folded peptide.

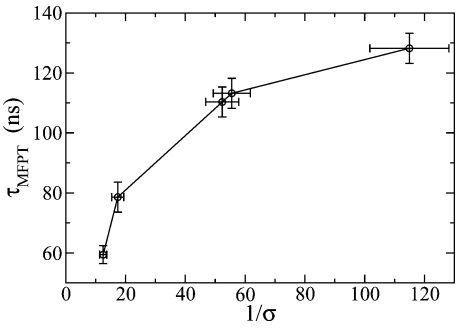


Figure 12. A plot of the mean folding time (in nanoseconds) versus the inverse of the Zimm−Bragg parameter σ. In agreement with the work of Buchete and Staub,[26] we find that τ is not simply proportional to $1/\sigma$.

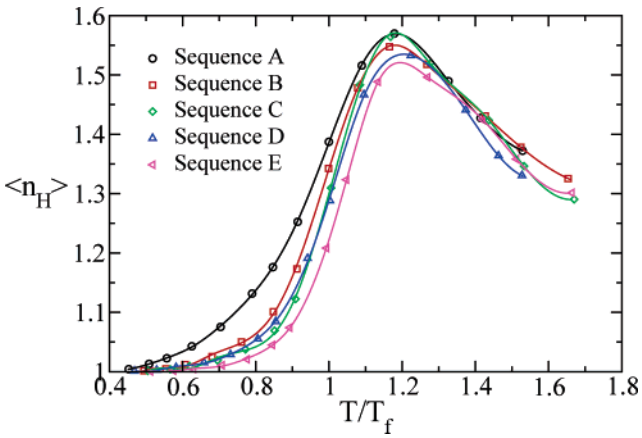


Figure 13. A plot of the average number of helical segments versus the reduced temperature T/T_f for each sequence. As the hydrophobic content of the sequence increases, so too does the average number of helices. The symbols and colors are the same as in Figure 3.

Finally, a plot of the mean folding time versus the Zimm−Bragg parameter $1/\sigma$ for all five sequences at their folding temperature is shown in Figure 12. In agreement with the results of Buchete and Straub,[26] we find that τ is not linear in $1/\sigma$. Sequences C and D lie very close to each other in Figure 12. They have very similar folding temperatures, very similar values of σ at the folding temperature, and very similar mean folding times. The origin of this similarity lies in the details of their sequences and in the relatively short chain length of the peptides studied. Sequences C and D differ only by 2 residues, residues 3 and 7, which are hydrophobic in sequence C and hydrophilic in sequence D. Changing the identity of two internal residues (which is 20% of the sequence) does not have much of an effect. In contrast, the sequence B differs from sequence C by 8 residues, and sequence A from sequence B by 3 residues, one of which is the N-terminus. There are marked differences between the behavior of these sequence pairs. Using longer chain lengths would give a clearer indication of the importance of amino acid sequence versus hydrophobic/hydrophilic content. It is certain, however, that increasing hydrophobic content increases the helix stability and decreases the mean folding time.

In their work, Buchete and Straub[26] identify the inclusion of multiple nucleation sites as the origin of the nonlinear behavior. To test whether this is a possible mechanism in these simulations, we can look at the average number of helical segments, $\langle n_H \rangle$ (defined as three or more consecutive residues in a helical state), at the folding temperature for each sequence. This is plotted versus the reduced temperature T/T_f in Figure 13. The more hydrophobic sequences, which have higher values of σ, do indeed have a higher average number of helical segments at the folding temperature than the more hydrophilic sequences. For example, sequence A has $\langle n_H \rangle = 1.39$, while sequence E has $\langle n_H \rangle = 1.21$. That all sequences have $\langle n_H \rangle > 1.0$ at the folding temperature indicates that multiple helix nucleation sites are important even for short chains such as those studied here.

5. Conclusion

A coarse-grained model peptide was introduced and used in a series of molecular dynamics simulations. For five different peptide sequences, each 10 residues in length and of varying hydrophobic/hydrophilic content, the thermodynamics of the coil-to-helix transition was characterized and the folding temperature was determined. The folding temperature, and hence the helix stability, increased with increasing hydrophobic content of the peptide. For each sequence, 400 independent configurations were simulated at their folding temperature, and the time was measured for each conformation to go from a random coil to an α-helix. From this distribution of folding times, the kinetics of the coil-to-helix transition was characterized. Folding times varied from 59 to 132 ns, which is slightly faster than the 200 ns established by experiment. However, given the small size (10 residues) of the peptides in this work, and the longer (16−21 residues) peptides used in the experimental work, this faster folding time is not unreasonable. While the coarse-grained interaction potentials used in this work are clearly approximate, especially the directionality of the hydrogen bond, the thermodynamic and kinetic results presented here indicate that it offers a realistic description of the coil-to-helix transition. For all five peptide sequences studied, single-exponential kinetics were observed, indicating a "two-state" folding process. Finally, the mean folding time was plotted versus the inverse of the Zimm−Bragg parameter σ, and a nonlinear dependence was found. The origin of the nonlinearity was ascribed to multiple helix nucleation sites, and pathways proceeding from single and from multiple helix nucleation sites were discussed. It was further shown that increasing values of σ correlated with an increasing average number of helical segments, indicating that even for small peptides such as those studied here, multiple helix nucleation sites play an important role in the folding kinetics.

Acknowledgment. This work was funded by the NIH (R01 NS41356) and the Center for Computational Science at Boston University.

References

(1) Schellman, J. A. The factors affecting the stability of hydrogen-bonded polypeptide structures in solution. *J. Phys. Chem.* **1958**, *62*, 1485−1494.

(2) Zimm, B. H.; Bragg, J. K. Theory of the phase transition between helix and random coil in polypeptide chains. *J. Chem. Phys.* **1959**, *31*, 526−535.

(3) Lifson S.; Roig, A. On the theory of helix-coil transition in polypeptides. *J. Chem. Phys.* **1960**, *34*, 1963−1974.

(4) Doig, A. J. Recent advances in helix-coil theory. *Biophys. Chem.* **2002**, *101*−*102*, 281−293.

(5) Takano, M.; Nagayama, K.; Suyama, A. Investigating a link between all-atom model simulation and the Ising-based theory on the helix-coil transition: Equilibrium statistical mechanics. *J. Chem. Phys.* **2002**, *116*, 2219−2228.

(6) Zana, R. On the rate-determining step for helix propagation in the helix-coil transition of polypeptides in solution. *Biopolymers* **1975**, *14*, 2425−2428.

(7) Williams, S.; Causgrove, T. P.; Gilmanshin, R.; Fang, K. S.; Callender, R. H.; Woodruff, W. H.; Dyer, R. B. Fast events in protein folding: helix melting and formation in a small peptide. *Biochemistry* **1996**, *35*, 691−697.

(8) Thompson, P. A.; Eaton, W. A.; Hofrichter, J. Laser temperature jump study of the helix-coil kinetics of an Alanine peptide interpreted with a 'kinetic zipper' model. *Biochemistry* **1997**, *36*, 9200−9210.

(9) Lednev, I. K.; Karnoup, A. S.; Sparrow, M. C.; Asher, S. A. α-helix peptide folding and unfolding activation barriers: a nanosecond UV resonance Raman study. *J. Am. Chem. Soc.* **1999**, *121*, 8074−8086.

(10) Clarke, D. T.; Doig, A. J.; Stapley, B. J.; Jones, G. R. The α-helix folds on the millisecond time scale. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 722−7237.

(11) Gilmanshin, R.; Williams, S.; Callender, R.; Woodruff, W.; Dyer, R. Fast events in protein folding: Relaxation dynamics of secondary and tertiary structure in native apomyoglobin. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 3709−3713.

(12) Dyer, R.; Gai, F.; Woodruff, W.; Gilmanshin, R.; Callender, R. Infrared Studies of Fast Events in Protein Folding. *Acc. Chem. Res.* **1998**, *31*, 709−716.

(13) Callender, R.; Dyer, R.; Gilmanshin, R.; Woodruff, W. FAST EVENTS IN PROTEIN FOLDING: The Time Evolution of Primary Processes. *Annu. Rev. Phys. Chem.* **1998**, *49*, 173−202.

(14) Huang, C.-Y.; Klemke, J. W.; Getahun, Z.; DeGrado, W. F.; Gai F. Temperature-dependent helix-coil transition of an alanine based peptide. *J. Am. Chem. Soc.* **2001**, *123*, 9235−9238.

(15) Huang, C.-Y.; Getahun, Z.; Wang, T.; DeGrado, W. F.; Gai, F. Time-resolved infrared study of the helix-coil transition using ^{13}C-labeled helical peptides. *J. Am. Chem. Soc.* **2001**, *123*, 12111−12112.

(16) Huang, C.-Y.; Getahun, Z.; Zhu, Y.; Klemke, J. W.; DeGrado, W. F.; Gai, F. Helix formation via conformation diffusion search. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2788−2793.

(17) Wang, T.; Du, D.; Gai F. Helix-coil kinetics of two 14-residue peptides. *Chem. Phys. Lett.* **2003**, *370*, 842−848.

(18) Wang, T.; Zhu, Y.; Getahun, Z.; Du, D.; Huang, C.-Y.; DeGrado, W. F.; Gai, F. Length dependent helix-coil transition kinetics of nine alanine-based peptides. *J. Phys. Chem. B* **2004**, *108*, 15301−15310.

(19) Daggett, V.; Levitt, M. Molecular dynamics simulations of helix denaturation. *J. Mol. Biol.* **1992**, *223*, 1121−1138.

(20) Klimov, D.; Betancourt, M.; Thirumalai, D. Virtual atom representation of hydrogen bonds in minimal off-lattice models of alpha-helices: Effects on stability, cooperativity and kinetics. *Fold. Des.* **1998**, *3*, 481−496.

(21) Takada, S.; Luthey-Schulten, Z.; Wolynes, P. G. Folding dynamics with nonadditive forces: a simulation study of a designed helical protein and a random heteropolymer. *J. Chem. Phys.* **1999**, *110*, 11616−11629.

(22) Margulis, C. J.; Stern, H. A.; Berne, B. J. Helix unfolding and intramolecular hydrogen bond dynamics in small α-helices in explicit solvent. *J. Phys. Chem. B* **2002**, *106*, 10748−10752.

(23) Dohsi, U. R.; Muñoz, V. The principles of α-helix formation: explaining complex kinetics with nucleation-elongation theory. *J. Phys. Chem. B* **2004**, *108*, 8497−8506.

(24) Schwartz, G. Kinetics of the helix-coil transition. *J. Mol. Biol.* **1965**, *11*, 64−77.

(25) Brooks, C. Helix-Coil Kinetics: Folding Time Scales for Helical Peptides from a Sequential Kinetic Model. *J. Phys. Chem.* **1996**, *100*, 2546−2549.

(26) Buchete, N.-V.; Straub, J. E. Mean first-passage time calculations for the coil-to-helix transition: the active helix Ising model. *J. Phys. Chem.* **2001**, *105*, 6684−6697.

(27) Voegler Smith, A.; Hall, C. K. Protein refolding versus aggregation: computer simulations on an intermediate-resolution protein model. *J. Mol. Biol.* **2001**, *312*, 187−202.

(28) Honig, B.; Cohen, F. E. Adding backbone to protein folding: why proteins are polypeptides. *Fold. Des.* **1996**, *1*, R17−R20.

(29) Cantor, C. R.; Schimmel, P. R. *Biophysical Chemistry; Part I*; Freedman: New York, 1980.

(30) Veitshans, T.; Klimov, D.; Thirumalai, D. Protein folding kinetics: time scales, pathways and energy landscapes in terms of sequence-dependent properties. *Fold. Des.* **1996**, *2*, 1−22.

(31) Swenson, R. H.; Wang, J.-S. Replica Monte Carlo simulation of spin-glasses. *Phys. Rev. Lett.* **1986**, *57*, 2607−2609.

(32) Hukushima, K.; Nemoto, K. Exchange Monte Carlo method and application to spin glass simulations. *J. Phys. Soc. Jpn.* **1996**, *65*, 1604−1608.

(33) Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **1999**, *314*, 141−151.

(34) Sugita, Y.; Okamoto, Y. Replica-exchange multicanonical algorithm and multicanonical replica-exchange method for simulating systems with rough energy landscapes. *Chem. Phys. Lett.* **2000**, *329*, 261−270.

(35) Press: W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed.; Cambridge University Press: Cambridge, 2002.

(36) Ohkubo, Y. Z.; Brooks, C. L., III Exploring Flory's isolated-pair hypothesis: Statistical mechanics of helix-coil transitions in polyalanine and the C-peptide from RNase A. *Proc. Natl. Acad. Sci.* **2003**, *100*, 13916−13921.

(37) Hansmann, U. H. E.; Okamoto, Y. Finite-size scaling of helix-coil transitions in poly-alanine studied by multicanonical simulations. *J. Chem. Phys.* **1999**, *110*, 1267−1276.

(38) Mitsutake, A.; Okamoto, Y. Helix-coil transitions of amino-acid homo-oligomers in aqueous solution studied by multicanonical simulations. *J. Chem. Phys.* **1999**, *122*, 10638−10647.

(39) van Giessen, A. E.; J. E. Straub, J. E. Monte Carlo simulations of polyalanine using a reduced model and statistics-based interaction potentials. *J. Chem. Phys.* **2005**, *122*, 024904.

(40) Bredenbeck, J.; Helbing, J.; Kumita, J. R.; Woolley, G. A.; Hamm, P. α-Helix formation in a photoswitchable peptide tracked from picoseconds to microseconds by time-resolved IR spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 2379−2384.

(41) Goldbeck, R. A.; Thomas, Y. G.; Chen, E.; Esquerra, R. M.; D. S. Kliger, D. S. Multiple pathways on a protein-folding energy landscape: kinetic evidence. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 2782−2787.

(42) Nymeyer, H.; Garcia, A. E. Simulation of the folding equilibrium of α-helical peptides: a comparison of the generalized Born approximation with explicit solvent. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13934−13939.

CT0503318

JCTC Journal of Chemical Theory and Computation

Sequence-Dependent Effects in the Cyclization of Short DNA

Luke Czapla,[†] David Swigon,[†,‡] and Wilma K. Olson*[,†]

Department of Chemistry & Chemical Biology, Rutgers, the State University of New Jersey, Wright-Rieman Laboratories, 610 Taylor Road, Piscataway, New Jersey 08854, and Department of Mathematics, University of Pittsburgh, 519 Thackeray Hall, Pittsburgh, Pennsylvania 15260

Received January 17, 2006

Abstract: A new, computationally efficient Monte Carlo approach has been developed to estimate the ring-closure properties of short, realistically modeled DNA chains. The double helix is treated at the level of base-pair steps using an elastic potential that accounts for the sequence-dependent variability in the intrinsic structure and elastic moduli of the base-pair steps, including the known coupling of conformational variables. Rather than using traditional Metropolis-Monte Carlo techniques to generate representative configurations, a Gaussian sampling method is introduced to construct three-dimensional structures from linear combinations of the rigid-body parameters defining the relative orientation and displacement of successive base pairs. The computation of the *J* factor, the well-known ratio of the equilibrium constants for cyclization vs bimolecular association of a linear molecule, takes into account restrictions on the displacement and directions of the base pairs joined in ring closure, including the probability that the end-to-end vector is null and the terminal base pairs coincide. The increased sample sizes needed to assess the likelihood that very short chains satisfy these criteria are attained using the Alexandrowicz half-chain sampling enhancement technique in combination with selective linkage of the two-half-chain segments. The method is used to investigate the cyclization properties of arbitrary-length DNA with greatly enhanced sampling sizes, i.e., $O(10^{14})$ configurations, and to estimate *J* factors lower than 0.1 pM with high accuracy. The methodology has been checked against classic theoretical predictions of the cyclization properties of an ideal, inextensible, naturally straight, DNA elastic rod and then applied to investigate the extent to which one can account for the unexpectedly large *J* factors of short DNA chains without the need to invoke significant distortions of double helical structure. Several well-known structural features of DNA—including the presence of intrinsic curvature, roll-twist coupling, or enhanced pyrimidine-purine deformability—bring the computed *J* factors in line with the observed data. Moreover, periodically distributed roll-twist coupling reduces the magnitude of oscillations in *J*, seen in plots of *J* vs chain length, to the extent found experimentally.

Introduction

The representation of chain molecules with ends confined to a fixed separation and orientation is a long standing problem in polymer physical chemistry that can be attacked from several points of view. In one approach the configurations of unconstrained linear molecules which meet preassigned geometric criteria, e.g., the distance and relative orientation of chain ends, are collected through exhaustive simulation studies.[1] This method has been used extensively in analyses of the kinetics of chain cyclization, i.e., ring

* Corresponding author fax: (732) 445-5958; e-mail: olson@rutchem.rutgers.edu.

[†] Rutgers, the State University of New Jersey.

[‡] University of Pittsburgh.

10.1021/ct060025+ CCC: $33.50
Published on Web 04/25/2006

closure,[2−6] and the formation of closed loops.[7−9] An alternative approach to the ring-closure problem is to derive an analytical expression for the configurational partition function of a simplified polymer, such as a freely jointed (Gaussian) chain or a freely rotating (wormlike) chain, that satisfies the requisite end conditions.[10−13] Such theories have provided important insights into experimental findings and serve as valuable benchmarks for simulations of the cyclization propensities of polymer chains with similar molecular properties.

The probability of polymer ring closure is typically described in terms of the Jacobson-Stockmayer J factor, or cyclization constant, defined as the ratio of the equilibrium constants for cyclization and bimolecular association of a linear molecule.[10] If the chain is sufficiently long and exhibits ideal Gaussian (random-coil) behavior, the J factor decreases with chain length N as $N^{-3/2}$.[10] At short chain lengths near the rigid-rod limit, where there is very little variation in the overall molecular structure, the likelihood of ring closure drops off sharply with a decrease in N. Thus the J factor exhibits a maximum at intermediate chain lengths comparable to 2−3 times the persistence length.[12,14,15]

The unexpected, spontaneous cyclization of DNA molecules much shorter than the persistence length[16,17] has renewed interest in ring-closure measurements and the intrinsic structure and deformability of double helical DNA. DNA chain segments less than 100 bp in length are cyclized up to 5 orders of magnitude more efficiently than expected from theoretical predictions based on the conventional representation of DNA as an ideal, inextensible, naturally straight elastic rod, i.e., helical wormlike chain model with a persistence length of ∼500 Å.[15] New theories account for these discrepancies in terms of the spontanous melting of isolated base pairs,[18] the presence of occasional sharp kinks in the chain backbone,[19,20] and/or the reduction of the torsional modulus.[17] The published models, however, do not consider the unique sequence-dependent structural features of the short minicircles and the possibility that the wide range of J factors which have been reported[6,16,17,20] may be a natural consequence of the differences in intrinsic structure and deformability of the selected sequences. For example, the DNA molecules found to close most easily into tight minicircles contain a well-known nucleosome-positioning sequence[21] with periodically repeating chemical features. This sequence includes two notable patterns: (i) TA base-pair steps, which recur at increments of ∼10 bp, i.e., roughly a complete double helical turn, and (ii) AT-containing dimers (AA, TA, or TT), which alternate at half helical turns with GC base-pair steps. The TA steps are among the most easily deformed of all base-pair steps,[22] and the separation of tracts of AA and TT dimers by GC-rich steps is implicated in DNA intrinsic curvature.[23] Furthermore, complementary base pairs remain intact even in the most severely deformed protein-mediated DNA bends, such as the ∼80 bp wrapped a complete superhelical turn around the histone proteins in the nucleosome core particle,[24] and even the very sharpest (40−50°) known protein-induced DNA bends[25,26] are typically smaller than the (∼90°) values posited to account for the cyclization tendencies of short chains.

The occurrence of small rings thus presents new technical challenges to the simulation of DNA ring closure. It is difficult to accumulate a meaningful sample of closed configurations from the random sampling of a relatively stiff, naturally straight molecule. The approximation of DNA as an ideal elastic rod has only limited success in predicting the behavior of the molecule in the short-length regime, where configurational fluctuations do not result in the ends having sufficiently random alignment for the persistence length to account entirely for the cyclization efficiency. This is despite the fact that most mixed-sequence DNA is expected to have only negligible intrinsic curvature, and hence a naturally straight model would be considered appropriate. Some sequences, such as those containing certain motifs beginning with four to six adenines (A-tracts), however, possess significant intrinsic curvature that is known to facilitate ring closure.[12,27]

Here we present the details of a new computationally efficient technique to estimate the J factors of short, realistically modeled DNA chains. We treat the double helix at the level of base-pair steps, i.e., dimers, making use of an elastic potential which governs the fluctuations in the relative orientation and displacement of successive base pairs from their intrinsic values. We account for sequence-dependent variability in both the intrinsic structure and the elastic moduli of the base-pair steps, including the known coupling of conformational variables and anisotropy of bending. In place of conventional Metropolis-Monte Carlo methods[28] for sampling configurations of constrained molecules, we utilize the quadratic form of the DNA elastic energy and sample unconstrained configurations using a standard Gaussian random number generator. We increase sample size considerably by combining the Alexandrowicz half-chain sampling enhancement technique,[29] with a novel method of selectively linking the pairs of half-chain segments that are likely to satisfy the end-to-end ring-closure criteria. The method can be used to estimate J factors lower than 0.1 pM with high accuracy and hence is suited for the investigation of short DNA molecules. DNA polymers with different degrees of intrinsic curvature and different patterns of dimeric flexibility are investigated with the new technique. After checking the methodology against classic theoretical predictions of the cyclization properties of an ideal, elastic DNA rod, we use the approach to investigate the extent to which one can account for the observed ring-closure properties of short chains without the need to invoke significant distortions of double helical structure. We consider the effects of intrinsic curvature, anisotropic bending, roll-twist coupling, and enhanced pyrimidine-purine deformability and compare the computed J factors against values found experimentally for DNA chains of the same length and chemical content.

Methods

DNA Model. We make use of a dimeric representation of DNA which incorporates the known effects of base sequence on the intrinsic structure and deformability of the constituent dinucleotide steps.[22] The rest state of each dimer is described by six independent step parameters which specify the preferred orientation and displacement of neighboring base-

pair planes—three angular variables termed tilt, roll, and twist and three variables called shift, slide, and rise with dimensions of distance.[30] The configuration of a base-pair step is denoted by the vector Θ, with components θ_i ($i = 1...6$) corresponding respectively to the instantaneous values of tilt, roll, twist, shift, slide, and rise at the given step.

We use an elastic potential to describe the fluctuations in structure at each base-pair step. The potential is parametrized in terms of the differences between the values of the base-pair step parameters for the assumed geometry of the step and the intrinsic base-pair step parameter values assigned to the minimum-energy state of the step. Coupling is incorporated in the model through off-diagonal terms in the energy expression (see below).

Defining Θ^0 as the vector containing the intrinsic base-pair step parameters, the potential of a single XZ base-pair step of the DNA is expressed by the double summation:

$$\Psi = \frac{1}{2}\sum_{i=1}^{6}\sum_{j=1}^{6} f_{ij}\Delta\theta_i\Delta\theta_j$$

$$\Delta\theta_i = \theta_i - \theta_i^0 \qquad (1)$$

Here the f_{ij} are elements of the symmetric 6×6 elastic force constant matrix **F**, which contains the noncoupled elastic constants along its diagonal and the coupled terms in its off-diagonal elements. The θ_i^0 are the step parameters of the minimum-energy reference state stored in Θ^0 and the θ_i are the corresponding values of the instantaneous configurational state stored in the vector Θ. Hence, a set of intrinsic base-pair step parameters Θ^0 and an elastic constant matrix **F** are sufficient to monitor the fluctuations of a single base-pair step. The configuration of a DNA chain segment depends on the choice of parameters at each of the N base-pair steps, and the total energy U is a sum of the energies over all steps:

$$U = \sum_{n=1}^{N} \Psi_n \qquad (2)$$

Configurational Sampling. The sampled configurations are chosen to satisfy a Boltzmann distribution. Here, rather than using traditional Metropolis-Monte Carlo sampling techniques[28] to generate representative configurations, we take advantage of the quadratic form for the energy of DNA in eq 1 and the consequent simplification of the configurational partition function as a product of Boltzmann terms for individual base-pair steps. Hence, the probability that a base-pair step will adopt configuration Θ is proportional to the Boltzmann factor of the dimer deformation energy

$$p(\Theta) \propto \exp(-\beta\Psi) = \exp\left[-\frac{\beta}{2}\sum_{i=1}^{6}\sum_{j=1}^{6} f_{ij}(\theta_i - \theta_i^0)(\theta_j - \theta_j^0)\right] \qquad (3)$$

where $\beta = 1/k_BT$, k_B is the Boltzmann constant, and T is the absolute temperature. Here, for simplicity, we omit indices that denote the sequential location n and chemical identity XZ of the base-pair step.

To express the probability density function as a product of independent terms, we write the dimeric energy Ψ in matrix form, i.e., $\Psi = 1/2\Delta\Theta^T\mathbf{F}\Delta\Theta$, diagonalize the force-constant matrix **F**, and rewrite the energy in terms of a diagonal matrix **D** and a basis variable Ω, with elements ω_i ($i = 1...6$) given by linear combinations of the base-pair step parameters:

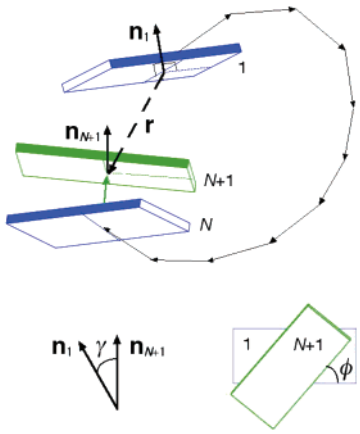


Figure 1. Linear DNA segment of N base pairs (blue blocks) in a configuration approaching the requirements for cyclization. The end-to-end vector **r** (thick dashed arrow) joins the first base pair to a hypothetical (green) base pair N+1, which coincides with the first base pair in a perfectly closed chain. The net bending angle γ (lower left image) is defined by the normals $\mathbf{n}_1$ and $\mathbf{n}_{N+1}$ of base pairs 1 and N+1 (thick solid arrows) and the end-to-end twist ϕ (lower right image) by the long axes and normals of the same base pairs.

$$\Psi(\Theta) = \frac{1}{2}\Delta\Theta^T\mathbf{F}\Delta\Theta = \frac{1}{2}\Delta\Theta^T\mathbf{Q}(\mathbf{Q}^T\mathbf{F}\mathbf{Q})\mathbf{Q}^T\Delta\Theta$$

$$= \frac{1}{2}(\Delta\Theta^T\mathbf{Q})\mathbf{D}(\mathbf{Q}^T\Delta\Theta) = \frac{1}{2}\Omega^T\mathbf{D}\Omega \qquad (4)$$

Here **Q** is the eigenvector matrix specifying the directions of the principal axes of deformation. The superscript ^{T} is used to denote the transpose of a matrix or vector.

Elimination of the cross terms in the energy expression allows us to write the probability density function for a single base-pair step, including normalization, as a product of Gaussians:

$$p(\Theta) = \prod_{i=1}^{6}\left(\frac{1}{2\pi\beta D_{ii}}\right)^{1/2}\exp\left(-\frac{\beta}{2}D_{ii}\omega_i^2\right) \qquad (5)$$

To sample this function, we modify a standard Gaussian random number generator[31] and collect a Boltzmann distribution of dimeric states without the necessity of using the Metropolis method. This approach, which we term Gaussian sampling, is superior to the Metropolis method in that it is computationally more efficient and does not suffer from correlations between sample points or incomplete coverage of phase space.

Calculation of the *J* Factor. To monitor the closure of a DNA chain of N base pairs, we add a virtual (N+1)th base pair (of the same type as the first base pair) and consider the DNA to be closed when the (N+1)th base pair coincides with the first base pair (see Figure 1). Following Flory et al.,[32] we express the J factor as a product of probabilities which describe the contribution of the spatial configuration

to the cyclization equilibrium constant:

$$J = \frac{4\pi}{N_A} W(\mathbf{r} = \mathbf{0})\, \Gamma_r(\cos\gamma = 1)\, \Phi_{r,\cos\gamma}(\phi = 0) \qquad (6)$$

Here $W(\mathbf{r} = \mathbf{0})$ is the probability that the configuration is closed, i.e., the end-to-end vector $\mathbf{r}$, which connects the first and $(N+1)$th base pairs of the DNA, is zero. The factor $\Gamma_r(\cos\gamma = 1)$ is the conditional probability that the normals of terminal base pairs are aligned when the chain ends coincide, i.e., the cosine of the net bending angle is unity ($\cos\gamma = 1$) when the end-to-end separation $r = |\mathbf{r}|$ is zero ($r = 0$). The term $\Phi_{r,\cos\gamma}(\phi = 0)$ is the conditional probability that the first and last, i.e., $(N+1)$th, base pairs coincide, i.e., the end-to-end twist ϕ is zero when both $r = 0$ and terminal base pairs are coplanar ($\cos\gamma = 1$). The quotient $N_A/4\pi$, where N_A is Avogadro's number, is the normalization constant associated with the uniformly distributed probability density of bimolecular association

For a DNA configuration specified by the angular parameters Θ_n, $n = 1...N$, we calculate the end-to-end vector and orientation angles from the serial product of 4×4 generator matrices $\mathbf{A}_n$ that incorporate the 3×1 displacement vector $\mathbf{r}_n$ and the 3×3 transformation matrix $\mathbf{T}_{n,n+1}$, which relates coordinate frames on successive base pairs $(n, n+1)$:

$$\mathbf{A}_{1:N} = \mathbf{A}_1\mathbf{A}_2...\mathbf{A}_{N-1}\mathbf{A}_N$$

$$\mathbf{A}_n = \begin{bmatrix} \mathbf{T}_{n:n+1} & \mathbf{r}_n \\ \mathbf{0} & 1 \end{bmatrix} \qquad (7)$$

The dependence of $\mathbf{T}_{n,n+1}$ and $\mathbf{r}_n$ on Θ_n follows the formulation introduced by Zhurkin et al.[33] and subsequently developed by El Hassan and Calladine,[34] in which the angular step parameters are defined in terms of a sequence of symmetric Euler rotations and the translational components are expressed in the "middle" base-pair frame corresponding to the axis positions generated by half the rotational operation that brings adjacent base-pair frames into coincidence. See the literature[33−35] for further details.

The end-to-end vector $\mathbf{r} = \mathbf{r}_{1:N}$ is accumulated in the far right column of $\mathbf{A}_{1:N}$:

$$\mathbf{r} = [\mathbf{I}_3 \;\; \mathbf{0}]\, \mathbf{A}_{1:N} \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix} \qquad (8)$$

Here $\mathbf{I}_3$ is the identity matrix of order three, and the $\mathbf{0}$'s are null matrices of orders necessary to fill the 3×4 premultiplication and 4×1 postmultiplication vectors.

The cosine of the angle between the normals of terminal base pairs is the (3,3) element of $\mathbf{A}_{1:N}$:

$$\cos\gamma = [0 \;\; 0 \;\; 1 \;\; 0]\, \mathbf{A}_{1:N} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \qquad (9)$$

The end condition imposed on the twisting of terminal base pairs is extracted from the trace of the end-to-end transformation matrix $\mathbf{T}_{1:N+1}$ accumulated in $\mathbf{A}_{1:N}$.

$$\mathrm{Tr}(\mathbf{T}_{1:N+1}) = \cos\phi(1 + \cos\gamma) + \cos\gamma \qquad (10)$$

Sampling Criteria. Because there is negligible likelihood that all three constraints on base-pair juxtaposition and orientation will be met exactly in numerical calculations, we relax the end conditions and consider as closed those configurations which meet the following criteria: (i) the length of the end-to-end vector $\mathbf{r}$ is less than r_ϵ; (ii) the cosine of the angle γ between the normals of terminal base pairs is larger than $1 - \nu_\epsilon$; and (iii) the magnitude of the end-to-end twist ϕ is less than τ_ϵ. Each generated configurational state is checked against these criteria, and each satisfaction is recorded. The probability densities are obtained as Bernoulli trials maximum-likelihood estimates, with each probability normalized by the volume of the enclosed phase space (see below).

The radial probability density $W(\mathbf{r} = \mathbf{0})$ is calculated as the number of configurations M_r obeying criterion (i), normalized by the sample size M and the phase space volume given as $4/3\ \pi r_\epsilon^3$. The contribution $\Gamma_r(\cos\gamma = 1)$ from the alignment of normals is calculated as the number of configurations $M_{r,\cos\gamma}$ for which both criteria (i) and (ii) hold, normalized by M_r and the phase space volume ν_ϵ. Finally, the twist alignment density $\Phi_{r,\cos\gamma}(\phi = 0)$ is calculated as the number of configurations $M_{r,\cos\gamma,\tau}$ satisfying all three criteria, normalized by $M_{r,\cos\gamma}$ and its phase space $2\tau_\epsilon$. The estimated J factor, obtained by substitution of these values in eq 6, is thus given by

$$J = \frac{M_{r,\cos\gamma,\tau}}{QM} \qquad (11)$$

where $Q = 4\pi N_A r_\epsilon^3 \nu_\epsilon \tau_\epsilon/3$.

The standard deviation of this estimate, assuming sampled configurations are uncorrelated, is given by

$$\sigma = \frac{1}{Q}\sqrt{\frac{M_{r,\cos\gamma,\tau}(M - M_{r,\cos\gamma,\tau})}{M^3}} \qquad (12)$$

When $M_{r,\cos\gamma,\tau}$ is much smaller than M, as is the case for closure probability calculations, the relative error is given by the approximate formula $\sigma/J \cong (M_{r,\cos\gamma,\tau})^{-1/2}$. Because the half-chain sampling method (see below) introduces slight correlation between sampled configurations, the true relative error (estimated as the standard deviation of multiple runs) is about 2.4 times larger than σ/J. Thus, the relative error is significant if $M_{r,\cos\gamma,\tau}$ is less than 100. For reported values of J greater than 0.1 pM, i.e., $\log J > -13$, the relative error is consistently below 10% (around 3% whenever feasible). The largest relative error occurs for smaller J factors (below 0.1 pM); data shown in this molar range are generally obtained with less accuracy, with as much as 32% relative error when the ends of short, naturally straight chains are out of phase, e.g., 10 accepted configurations for a 90 bp ideal, elastic B-DNA chain.

Sampling Enhancement. It is computationally expensive to generate ensembles of more than $O(10^7)$ configurations. Since DNA segments shorter than the persistence length are stiff and the probability that a randomly generated configuration satisfies all three ring-closure criteria is very small, the computed J factors are subject to large error if straightforward sampling approaches are used. Therefore, for such segments we use a sampling enhancement technique intro-

duced by Alexandrowicz[29] and subsequently applied by Levene and Crothers[4] in Monte Carlo simulations of DNA ring closure. Rather than generate the configurations of the entire DNA with N base pairs, we divide the chain into two equal (or nearly equal) pieces and sample L configurations of each half-chain segment separately. By taking all pairwise combinations of both halves we can theoretically achieve L^2 configurations of the full-length chain. It should be noted that, even though the configurations of half-chain segments are uncorrelated, the multiplicative combination of a finite number of states introduces some bias in the full-chain ensemble.

If we let m be a base pair close to the midpoint of the chain, the end-to-end matrix $\mathbf{A}_{1:N}$ can be factored into two submatrices containing the structural details of the two-half-segments of the molecule:

$$\mathbf{A}_{1:N} = \mathbf{A}_{1:m}\mathbf{A}_{m:N} \tag{13}$$

Elements of these two submatrices are stored during the simulation and multiplied when half-segments combine successfully to form a closed, full-length chain. It follows from eq 13 and the definition of $\mathbf{A}_n$ in eq 7 that the end-to-end vector can similarly be divided into two vectors that connect the chain ends via a central base pair.

$$\mathbf{r}_{1:N} = \mathbf{r}_{1:m} + \mathbf{T}_{1:m}\mathbf{r}_{m:N} \tag{14}$$

The requirement of ring closure, i.e., $\mathbf{r}_{1:N} = \mathbf{0}$, imposes a constraint on the vectors in the two halves of the DNA, namely

$$\mathbf{r}_{m:N} = -\mathbf{T}_{1:m}^{-1}\mathbf{r}_{1:m} \tag{15}$$

To reduce the number of unnecessary half-chain combinations, we join only those pairs of segments that are likely to satisfy the end-to-end ring-closure criterion. The current algorithm keeps track of the values of $-\mathbf{T}_{1:m}^{-1}\mathbf{r}_{1:m}$ for the first half-segments and $\mathbf{r}_{m:N}$ for the second half-segments and combines only those for which eq 15 is approximately satisfied. This is implemented as follows: the three-dimensional Cartesian space of $\mathbf{r}$ is divided into cubes of volume V. As first and second half-configurations are being sampled we record cubes in which those configurations terminate in the sense that the values of $-\mathbf{T}_{1:m}^{-1}\mathbf{r}_{1:m}$ and $\mathbf{r}_{m:N}$ fall within the cube boundaries. Then, for each cube we combine all first half-configurations terminating in that cube with all second half-segments terminating in this same (primary) cube. We also combine those first half-configurations with the second half-segments terminating in all cubes that are less than r_ϵ away from any point in the primary cube, where r_ϵ is the radius of the sphere used to approximate the radial probability density $W(\mathbf{r} = \mathbf{0})$. First and second half-segments terminating in cubes that are farther than r_ϵ apart need not be combined because they cannot satisfy the initial closure criterion, i.e., $\mathbf{r} = \mathbf{0}$.

This method discards the majority of possible configurational combinations, especially when sampling chain segments with a very low J factor. Thus, the total sample size can be, effectively, as large as 10^{14}, when necessary, to achieve a good estimate of the contribution of the conditional twist probability density $\Phi_{r,\cos\gamma}(\phi = 0)$. The latter quantity is the most difficult to sample reliably and, except for the recent biased Monte Carlo calculations of the Vologodskii group,[6] which unfortunately are not applicable to the present study of sequence-dependent effects, has not been considered in previous Monte Carlo simulations of DNA ring closure. Published estimates of the J factors obtained from Brownian dynamics simulations have also omitted the contribution of terminal base-pair overlap to the likelihood of chain closure.[36]

The calculations were performed on a single 2.1 Ghz AMD Athlon(tm) processor with 2 gigabytes of available RAM. Numerical estimation of the J factor of a single molecule takes 1−10 CPU h depending upon DNA chain length and intrinsic structure. Simulations are longest for short, naturally straight chains where there is the least variation in overall molecular structure.

Results and Discussion

Intrinsically Straight DNA. We used the Gaussian sampling method in combination with the improved half-chain generation technique to investigate the cyclization probabilities of various models of double helical DNA. We tested the computational approach by first determining the ring-closure properties of an ideal, inextensible, naturally straight DNA molecule over the entire range of chain lengths between 90 and 450 bp. We chose a minimum-energy rest state with an intrinsic helical repeat of 10.5 bp/turn (∼34.3° twist at every base-pair step) and a pitch of 35.7 Å (3.4 Å rise at every base-pair step). All other step parameters were equated to zero.

In this simple *ideal* model, analogous to the Shimada-Yamakawa twisted wormlike chain representation of DNA[15] (where the molecule is naturally straight, inextensible, subject to isotropic bending, and able to undergo independent fluctuations in twist), we allow deformations in tilt, roll, and twist (θ_1, θ_2, θ_3) but fix the translational parameters shift, slide, and rise (θ_4, θ_5, θ_6) near their intrinsic values (by use of very large force constants). The root-mean-square (RMS) fluctuation of tilt is equated to that of roll (isotropic bending), i.e., $\langle\Delta\theta_1^2\rangle = \langle\Delta\theta_2^2\rangle$, and is assigned a value of 4.84°, corresponding to a persistence length $a = 2\Delta s/(\langle\Delta\theta_1^2\rangle + \langle\Delta\theta_2^2\rangle)$ of nearly 500 Å (if Δs, the per residue base-pair displacement, is taken as 3.4 Å). The assumed RMS fluctuation in twist $\langle\Delta\theta_3^2\rangle = 4.09°$ corresponds to a global twisting constant $C = k_BT/\langle\Delta\theta_3^2\rangle$ somewhat larger in magnitude than the global bending constant A, i.e., $C/A = 1.4$, where $A = ak_BT$. The choice of C is compatible with measurements of the equilibrium topoisomer distributions of DNA minicircles[37] and the fluorescence depolarization anisotropy of ethidium bromide molecules intercalated in DNA minicircles.[38]

To estimate J, we set a radial bound r_ϵ of 30 Å on the sphere used to monitor the end-to-end vector of the full-length chain, a limit of 30° or less on the angle between the normals of the first and last base pairs ($\cos\gamma > \sqrt{3}/2$), and a restriction on the magnitude of the end-to-end twist τ_ϵ to values less than 30°. The same values are adopted in all subsequent simulations unless otherwise noted. As is clear from Table 1, the computed J factors are generally insensitive to the magnitude of the selected bounds. More restrictive bounds, however, lead to smaller numbers of acceptable configurations and hence to larger errors.

Table 1. Effect of Boundary Conditions on the Computed *J* Factor of a 210 bp, Ideal DNA Molecule[a]

r_ϵ (Å)	γ_ϵ (deg)	τ_ϵ (deg)	*M*	*J* (M)	relative error[b] (%)
30	30	30	3.6×10^{13}	1.26×10^{-8}	0.41
20	30	30	3.6×10^{13}	1.25×10^{-8}	0.74
10	30	30	3.6×10^{13}	1.27×10^{-8}	2.11
10	10	10	3.6×10^{13}	1.03×10^{-8}	12.5
10	10	10	1.44×10^{14}	0.92×10^{-8}	6.38

[a] See legend to Figure 2 for details of model. [b] The relative error is estimated from multiple runs and equal to roughly $2.4\sigma/J$.

As is clear from Figure 2, the computed dependence of *J* on DNA chain length mimics the predictions of the Shimada-Yamakawa theory,[15] showing a close fit with the observed cyclization efficiencies of chains of 240 bp or more[39−42] and accounts satisfactorily for the reported ring-closure tendencies of other short (105−130 bp) DNA sequences.[20] The model, like the Shimada-Yamakawa theory, underestimates *J* by up to 3 orders of magnitude for some very short (89−105 bp) chains.[16,17] These discrepancies have prompted our extension of the calculations to more realistic models of DNA with concomitant analysis of the effects of chemical features on cyclization properties as described below.

Influence of Natural Curvature. Individual base-pair steps adopt characteristic spatial forms and show different deformational tendencies in high-resolution DNA structures.[22] These local turns and twists, if appropriately concatenated and then repeated in phase with the double helical repeat, can introduce a natural "static" curvature or superhelicity in the DNA.[27,43,44] Indeed, it is rare to find segments of intrinsically straight DNA,[45] particularly in fragments as short as those employed in recent DNA ring-closure measurements.[16,17,20]

Here we study an idealized, naturally curved DNA made up of two types of base-pair steps arranged in such a way that a molecule of 150 bp forms a stress-free, nearly circular configuration, or O-ring.[35] Half of the steps in the 10 bp repeating sequences are intrinsically straight ($\theta_1^0 = \theta_2^0 = 0$) with twist ($\theta_3^0 = 36°$) corresponding to a 10 bp/turn helical repeat. The remaining steps have a different rest state with a positive roll angle ($\theta_2^0 = 7.41°$) and an intrinsic twist ($\theta_3^0 = 35.57°$) slightly lower than that of the preceding steps. The elastic properties at all steps, however, mimic those of an ideal DNA rod, i.e., the force constants are the same as those applied above in the simulation of naturally straight DNA. The sequence is constructed such that the first five steps assume the ideal B-DNA rest state and the last five steps adopt the perturbed rest state, i.e., an X_5Z_5 repeating pattern with XX and XZ dimer steps assigned the ideal B-DNA step parameters and ZZ and ZX steps assigned the set of modified parameters. As above, the intrinsic shift and slide are fixed at 0 Å, the rise at 3.4 Å, and the tilt at 0° at all base-pair steps.

Not surprisingly, the computed *J* factors of chain segments of the 150 bp O-ring are many orders of magnitude greater than those of a naturally straight molecule with the same elastic properties, chain length, and double helical repeat (compare the results reported in Figure 3 for naturally curved molecules of 90−210 bp and ideal, straight DNA chains of the same chain length and with the same 10 bp helical repeat). The enhancement in *J* is greatest for curved segments containing an integral number of helical turns. The ends of such chains readily meet in perfect register for successful ring closure. Interestingly, the chain segment which closes most easily into a circle is somewhat shorter than the perfect O-ring (130 vs 150 bp), reflecting the effects of thermal fluctuations on the configuration of DNA. Moreover, the cyclization probabilty maximum at 130 bp and the accompanying enhancement of the *J* factor compared to that of ideal, naturally straight DNA, i.e., $\log [J/J_0] = 7.48$, where J_0 is the Jacobson-Stockmayer factor of the latter model, agree remarkably well with values predicted for a naturally curved DNA with the same ratio of intrinsic curvature and persistence length.[12] Finally, although the curvature enhances the likelihood of DNA ring closure, the greater amplitude of the sawtooth oscillations in log *J* with *N* in the naturally curved DNA compared to those of a naturally straight model is counter to the observed changes in *J* at short chain lengths, where the amplitude is lower than that expected for a naturally straight DNA.[17] The computed variation of log *J* vs *N* in Figure 3, however, is exaggerated by the lower radial bound imposed on the curved DNA sequences ($r_\epsilon = 10$ Å) compared to that imposed on the straight chain ($r_\epsilon = 30$ Å).

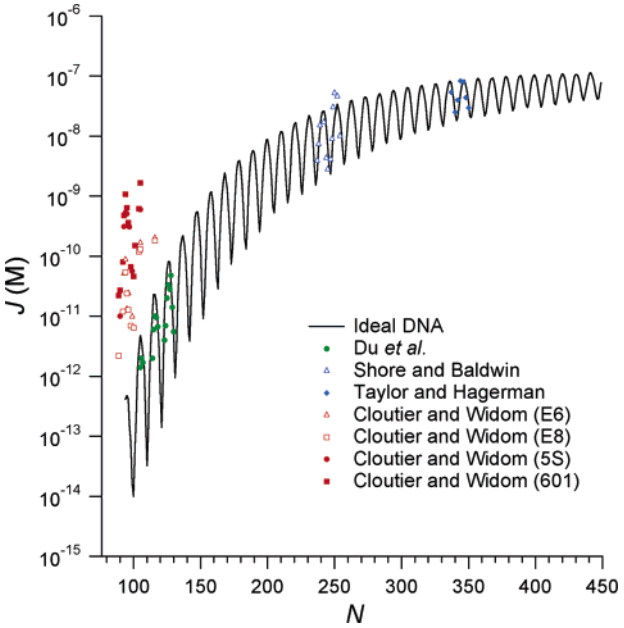


Figure 2. Monte Carlo estimates of the dependence on chain length of the *J* factors of ideal DNA, which has properties analogous to those of the Shimada-Yamakawa twisted worm-like chain:[15] the double helix is assumed to be naturally straight in its equilibrium rest state, inextensible, and capable of isotropic bending and independent twisting at the base-pair level. The intrinsic twist and bending fluctuations are consistent with experimental measurements, i.e., 10.5 bp per helical turn and a persistence length of approximately 500 Å. Observed data points are denoted by color-coded symbols: filled and open red symbols (short nucleosome-binding and control sequences, respectively);[16,17] filled green circles (other short sequences);[20] open blue triangles (*Eco*RI restriction fragments);[40,41] and filled blue diamonds (*Eco*RI (E)-ended constructs).[42] The curve is made up of a series of line segments connecting the computed *J* factors of molecules that differ in chain length by one base pair.

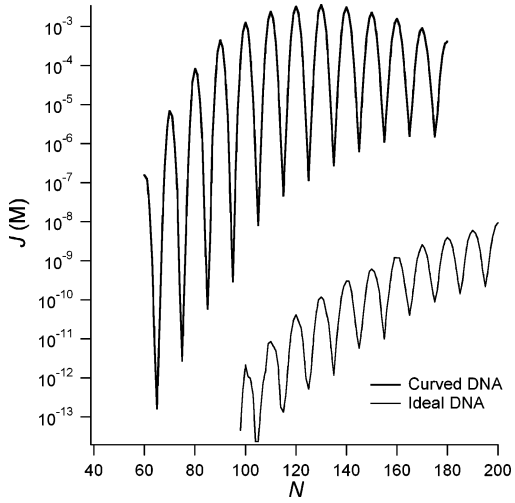


Figure 3. Predicted effects of intrinsic curvature on the *J* factors of DNA chains of 60−180 bp (thick solid curve). The intrinsically curved DNA is made up of two types of base-pair steps, each subject to isotropic bending and arranged such that a molecule of 150 bp forms a stress-free, nearly circular configuration (O-ring).[35] The corresponding variation of log *J* vs *N* for an ideal DNA with the same assumed helical repeat (10 bp per turn) is shown as the thin solid line.

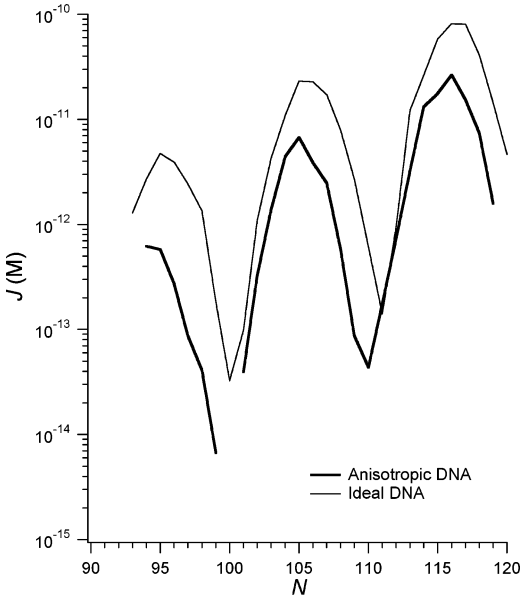


Figure 4. Influence of phased bending anisotropy on the computed likelihood of DNA ring closure in naturally straight chains of 94−119 bp (thick solid curve). Here the DNA is constituted from two types of base-pair steps: five successive steps are subject to ideal isotropic bending and the next five or six steps in the $(X_5Z_5X_5Z_6)_n$ conformational repeating pattern bend preferentially via roll in a hingelike manner. The data point at $N = 100$ bp is subject to large sampling error and thus omitted. The corresponding variation of log *J* vs *N* for an ideal DNA with the same assumed helical repeat (10.5 bp per turn) is shown as the thin solid line.

Table 2 reports the *J* factors for a series of short (95, 100, 105 bp) DNA molecules of different intrinsic curvature, all with a 10.5 bp helical repeat and all determined with the same radial bound ($r_\epsilon = 30$ Å). The chains are constructed, as described above, from two types of base-pair steps: one with $\theta_2^0 = 0°$ and $\theta_3^0 = 34.3°$ and the other with nonzero values of intrinsic roll and slightly reduced dimeric twist (see the legend to the table). Only a small intrinsic roll in the latter steps ($\theta_2^0 \approx 2°$) is required to bring the computed *J* factors in line with those reported for very short nucleosome-positioning sequences, e.g., $J_{obs} = 6.39 \times 10^{-10}$ M at $N = 95$ bp and 1.66×10^{-9} M at $N = 105$ bp.[16,17] The computed oscillation in the *J* factor between 95 and 105 bp, however, is still greater than that observed experimentally ($J_{obs} = 4.57 \times 10^{-11}$ M at $N = 100$ bp).

Phased Bending Anisotropy. Compared to the ideal isotropic rod model, in which the double helix is assumed to be equally likely to bend in all directions, the bending of real DNA is anisotropic, with deformability in the direction of the major and minor grooves typically exceeding that across the grooves. In other words, the variation in roll is generally greater than that in tilt.[22,33]

To assess the effects of anisotropic bending on ring closure, we consider a naturally straight, inextensible DNA model in which roughly half of the base-pair steps are assigned the bending properties of an ideal, isotropic rod, and the remainder are subject to preferential bending via roll. The force constants of the former steps are the same as those employed above ($f_{11} = f_{22} = 4.84^{-2}$), while those of the latter steps ($f_{11} \approx 0^{-2}$ and $f_{22} = 6.84^{-2}$) are assigned so that the flexibility in roll exceeds that in tilt to such an extent, that fluctuations via tilt are prevented. This deformational pattern, which preserves the ∼500 Å persistence length of DNA,[46] corresponds to Schellman's classic "hinge" model of DNA bending.[47] The fluctuations in twist are governed at all steps by the same elastic constant used above, i.e., $f_{33} = 4.09^{-2}$. A 21-bp repeating sequence is employed to preserve the assumed 10.5 bp helical repeat, i.e., $(X_5Z_5X_5Z_6)_n$, where the XX and XZ steps bend isotropically and the ZZ and ZX steps exhibit hinge bending. Other step parameters are assigned the intrinsic values listed above.

Surprisingly, the DNA modeled with phased anisotropic bending is stiffer than an ideal DNA elastic rod in terms of

Table 2. Effect of Intrinsic Curvature on the *J* Factors of Short DNA Duplexes of Chain Length *N*

	intrinsic curvature[a]				
r_ϵ	210 bp	420 bp	630 bp	840 bp	straight
95	8.42×10^{-6} M	3.83×10^{-9} M	1.52×10^{-10} M	4.48×10^{-11} M	4.75×10^{-12} M
100	3.21×10^{-10} M	1.62×10^{-12} M	3.94×10^{-14} M	2.74×10^{-14} M	3.25×10^{-14} M
105	3.77×10^{-5} M	2.37×10^{-8} M	1.18×10^{-9} M	2.18×10^{-10} M	2.31×10^{-11} M

[a] Intrinsic curvature is measured in terms of the number of base pairs of an idealized $X_5Z_5X_5Z_6$ repeating sequence designed to form an O-ring of the specified chain length. Intrinsic roll (θ_2^0) is fixed at zero and intrinsic twist (θ_3^0) at ∼34.3° (10.5 bp per helical turn) at XX and XZ steps and at the following values at ZZ and ZX steps: 210 bp (5.4°, 34.03°); 420 bp (2.7°, 34.21°); 630 bp (1.8°, 34.25°); 840 bp (1.35°, 34.26°); straight (0°, ∼34.3°).

its ring-closure propensities, even though its closed configuration has lower elastic energy. The computed *J* factors are as much as an order of magnitude lower than those of an ideal rod with the same helical repeat for particular chain lengths in the range 94−119 bp (Figure 4). While the bending anisotropy may enhance deformations into the grooves, the local restrictions on bending across the grooves seemingly contribute to more effective cyclization. A possible reason for these observations may be that, whereas cyclized, ideal DNA is free to rotate about its helical axis, the corresponding global torsional motion of a covalently closed DNA subject to anisotropic bending is hindered by an energy barrier. That is to say, the ideal DNA molecule can "turn inside-out" and thereby cyclize with a bend in any direction with respect to the first base pair, whereas the hinged DNA chain is restricted to global bending in a single plane.

Roll-Twist Coupling. The localized twisting of DNA is strongly correlated with its bending.[48] The unwinding of the double helix (via decrease in twist) is generally accompanied by increased roll, i.e., local bending into the major groove, and the overwinding of the duplex (via increase in twist) is usually accompanied by decreased roll, i.e., local bending into the minor groove. Furthermore, the degree of twist-roll coupling is sequence-dependent, being generally most pronounced for pyrimidine-purine dimers and weakest for purine-pyrimidine steps.[22,48]

Here we explore the effects of roll-twist coupling in an otherwise ideal model of B DNA, i.e., the chain is inextensible and naturally straight with a 10.5 bp helical repeat and the bending of successive base pairs is isotropic in the sense that $f_{11} = f_{22}$. The base-pair steps are subdivided into two categories and grouped, as above, into repeating $(X_5Z_5X_5Z_6)_n$ conformational blocks. The bending is independent of twisting in five successive (XX and XZ) base-pair steps, but roll is coupled to twist in the next five or six (ZZ and ZX) steps. The sign and magnitude of the twist-roll coupling modulus, $f_{23} = f_{32} = 5.41^{-2}$, are compatible with the interdependence of base-pair step parameters seen in high-resolution structures,[22] i.e., the increase in roll will tend to lower twist and vice versa. Other elastic constants are held at canonical B-DNA values ($f_{11} = f_{22} = 4.84^{-2}$, $f_{33} = 4.09^{-2}$). The translational parameters and tilt are assigned the standard intrinsic values listed above.

Remarkably the introduction of roll-twist coupling into an otherwise ideal DNA chain increases the *J* factor and concomitantly reduces the amplitude of oscillations in log *J* with *N* (Figure 5). The computed variation in *J* with chain length closely matches the measured *J* factors of the "random" E8 sequences used as a control in the ring-closure studies of Cloutier and Widom[16,17] (open symbols in Figure 5). The overall base-pair composition of the latter sequences is similar to that of the more easily cyclized "601" nucleosome-binding sequences (filled symbols in Figure 5), i.e., both sequence families are slightly GC-rich with 2−3 more pyrimidine-purine (YR) steps than random expectation, but there are distinct differences in the chemical content and distribution of the YR steps. Specifically, there is a greater proportion of TA dimers than CA·TG steps in the nucleosome-positioning sequences compared to the control sequences and the TA steps in the positioning sequence are spaced ~10 bp apart, in nearly perfect phase with the helical repeat. Notably, the energy required for melting the TA step, i.e., the TA·TA duplex fragment, is less than that of the CA·TG step,[49,50] suggesting that conformational deformations of the TA are more facile and could dominate the motions of the double helix.

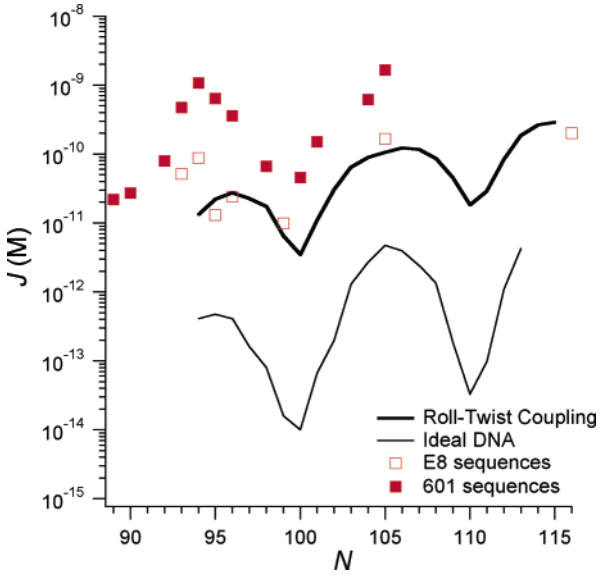


Figure 5. Predicted effects of phased roll-twist coupling on the *J* factors of naturally straight DNA chains of 94−115 bp (thick solid curve). Here the DNA is made up of two types of base-pair steps: five successive steps exhibit ideal elastic behavior, and the next five or six steps in the $(X_5Z_5X_5Z_6)_n$ conformational repeating pattern are subject to roll-twist coupling. The corresponding variation of log *J* vs *N* for an ideal DNA with the same assumed helical repeat (10.5 bp per turn) is shown as the thin solid line. The observed *J* factors of "601" nucleosome-positioning sequences[16,17] are denoted by filled symbols and those of E8 control sequences[16,17] by open symbols.

The natural changes in local twist, which accompany fluctuations in roll, clearly make it unnecessary to postulate[17] that the DNA which forms small rings is more easily twisted than the classic ideal elastic rod. The build-up of roll, which is required to bring the ends of the DNA in contact, automatically enhances the degree of twisting in the model. See ref 51 for a related discussion showing that the observable, effective twisting resistance of a closed DNA ring with the same pattern of roll-twist coupling is close to zero.[51]

Phased Deformability. Large protein-induced bends of DNA occur predominantly at pyrimidine-purine steps,[25,26] the dimers expected to be most easily bent on the basis of steric clash arguments[52] and potential energy calculations.[53,54] The enhanced deformabilty of such dimers contributes to the global plasticity of DNA without significant disruption of the double helical structure.

To test whether the regular spacing of flexible TA steps may enhance the ring closure of short nucleosome-positioning sequences, we constructed a hypothetical, naturally straight DNA model (with the same 10.5 bp helical repeat and elastic properties as ideal B-DNA) interspersed every 10 bp by a dimer with twice the RMS fluctuation of tilt and

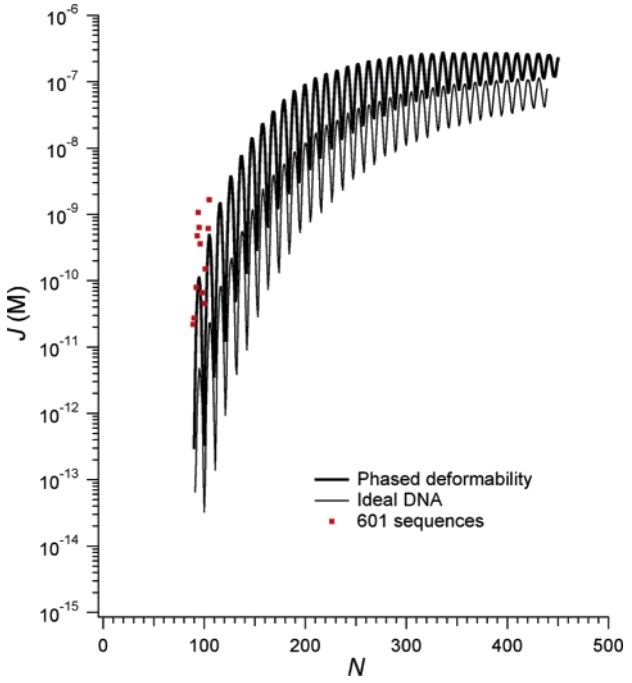


Figure 6. Influence of phased deformability on the computed likelihood of DNA ring closure in chains of 90−430 bp (thick solid curve). The ideal B-DNA model is interspersed every 10 bp by a deformable step subject to isotropic bending fluctuations twice those of the other steps. The corresponding behavior of an ideal DNA with the same assumed helical repeat (10.5 bp per turn) is shown as the thin solid line. The observed *J* factors of "601" nucleosome-positioning sequences[16,17] are denoted by filled symbols.

roll, i.e., $f_{11} = f_{22} = (2 \times 4.84)^{-2} = 9.68^{-2}$. The latter range of bending is comparable to the composite variation of tilt and roll in the TA steps found in well resolved X-ray structures. For example, the sum of the fluctuations of tilt and roll ($\langle\Delta\theta_1^2\rangle + \langle\Delta\theta_2^2\rangle$) for TA steps in a dataset of 239 nonredundant protein-bound DNA structures of 2.5 Å or better resolution (Y. Li and W. K. Olson, unpublished data) is 15.26^2 if all extant data are considered and 7.02^2 if the most extremely deformed TA steps are omitted, i.e., values that bound the $(2 \times 9.68^2) = 13.69^2$ range of deformations assumed here. Values of intrinsic tilt, shift, slide, and rise are identical to those used in the analysis of naturally straight DNA.

As seen in Figure 6, the softening of DNA with force constants that are consistent with the deformability of TA steps and their placement in nucleosome-positioning sequences increases the *J* factor by over 2 orders of magnitude to the range of values found[16,17] for short chains (filled symbols in the figure). The added fluctuations, however, do not dampen the oscillations of log *J* with *N* to the extent found experimentally. The amplitude of the variation in log *J* with *N* for the DNA chains with phased deformability is roughly 90% of that for the ideal DNA elastic rod. Although the imposed pattern of localized flexibility increases the *J* factor at all chain lengths, the effects are less pronounced at longer chain lengths. The enhancement in *J* levels off with increase in chain length to a factor ~2.5 times the value predicted for the ideal B-DNA duplex, the difference reflecting the lower persistence length of the locally softened DNA. The regularly spaced deformations also lower the chain length of maximum ring-closure probability, from a calculated value of 483 bp for the naturally straight DNA to 336 bp for the modified duplex. The cyclization propensities of long DNA multimers with regularly spaced, structural features such as those considered here are thus expected to differ from those of the mixed sequences traditionally used in the physical characterization of DNA.

Concluding Remarks. The present calculations clearly show that, in contrast to earlier suggestions,[17−20] it is not necessary to invoke significant distortions in the DNA double helix, e.g., melting, kinking, "free" twisting, to bring the computed *J* factors in line with the enhanced cyclization propensities of molecules much shorter than the persistence length.[16,17] A number of well-known features of DNA double helical structure—including the presence of intrinsic curvature, roll-twist coupling, or enhanced pyrimidine-purine deformability—can account for the unexpectedly large *J* factors of short chains. Moreover, chains with roll-twist coupling dampen the sawtooth oscillations in *J* with increase in chain length to the extent found experimentally. The internal coupling of dimeric variables also avoids the known detrimental effects of enhanced DNA deformability on nucleosome formation, i.e., the entropic penalty associated with the binding of a flexible sequence.[50]

These conclusions could not have been drawn without the improvements in chain sampling introduced herein. Many of the rare "closed" arrangements of DNA captured in the course of these calculations would be missed with conventional Metropolis-Monte Carlo configurational sampling. We enhance the efficiency of finding closed states of DNA through the selective addition of half-chain segments, combining only pairs of half-segments that are known in advance to match the end-to-end spacing of the full-length molecule. The Gaussian sampling technique ensures complete coverage of phase space and does not suffer from the correlations between sample points associated with Metropolis-Monte Carlo approaches. The generation of spatial configurations from random moves along the eigenvectors of the base-pair step parameters allows us to consider the coupling of conformational variables that accounts so well for the mesoscale looping of DNA.

The improved sampling of DNA configurations also provides a valuable guide for the development of theories that use expansion methods to estimate the configurational partition function of a closed ring. The reliable fitting of the highly irregular spatial density distributions of short polymers, including DNA, has long been known to require spatial moments higher than second order.[1,2,32,55] The extent to which approximate partition functions based on series expansions of the Boltzmann factor about the minimum energy configuration match the computed partition function will be reported elsewhere.

Finally, the present calculations omit consideration of the polyanionic character of DNA. Calculations of ring-closure probabilities that include long-range electrostatic interactions between anionic phosphate groups are expected to increase the chain length of maximum ring-closure probability.

Acknowledgment. We thank Drs. Jonathan Widom and Alexander Vologodskii for generously sharing experimental data, Dr. A. R. Srinivasan for assistance in modeling naturally curved DNA, and Dr. Edward W. Castner, Jr. for access to plotting software. Support of this work through U.S.P.H.S. Grant GM34809 is gratefully acknowledged. Computations were carried out at the Rutgers University Center for Computational Chemistry.

References

(1) Yoon, D. Y.; Flory, P. J. Moments and distribution functions for polymer chains of finite length. II. Polymethylene chains. *J. Chem. Phys.* **1974**, *61*, 5366−5380.

(2) Suter, U. W.; Mutter, M.; Flory, P. J. Macrocyclization equilibria. 2. Poly(dimethylsiloxane). *J. Am. Chem. Soc.* **1976**, *98*, 5740−5745.

(3) Hagerman, P. J. Analysis of the ring-closure probabilities of isotropic wormlike chains: application to duplex DNA. *Biopolymers* **1985**, *24*, 1881−1897.

(4) Levene, S. D.; Crothers, D. M. Ring closure probabilities for DNA fragments by Monte Carlo simulation. *J. Mol. Biol.* **1986**, *189*, 61−72.

(5) Kahn, J. D.; Crothers, D. M. Protein-induced bending and DNA cyclization. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 6343−6347.

(6) Podtelezhnikov, A. A.; Mao, C. D.; Seeman, N. C.; Vologodskii, A. Multimerization-cyclization of DNA fragments as a method of conformational analysis. *Biophys. J.* **2000**, *79*, 2692−2704.

(7) Marky, N. L.; Olson, W. K. Loop formation in polynucleotide chains. I. Theory of hairpin loop closure. *Biopolymers* **1982**, *21*, 2329−2344.

(8) Vologodskii, A. V.; Levene, S. D.; Frank-Kamenetskii, M. D.; Cozzarelli, N. R. Conformational and thermodynamic properties of supercoiled DNA. *J. Mol. Biol.* **1992**, *227*, 1224−1243.

(9) Rippe, K.; von Hippel, P. H.; Langowski, J. Action at a distance: DNA-looping and initiation of transcription. *TIBS* **1995**, *20*, 500−506.

(10) Jacobson, H.; Stockmayer, W. H. Intramolecular reaction in polycondensations. I. The theory of linear systems. *J. Chem. Phys.* **1950**, *18*, 1600−1606.

(11) Yamakawa, H.; Stockmayer, W. H. Statistical mechanics of wormlike chains. II. Excluded volume effects. *J. Chem. Phys.* **1972**, *57*, 2843−2854.

(12) Livshits, M. A. Calculating the probability for cyclization of a short intrinsically curved and twisted DNA fragment. *Mol. Biol.* **1996**, *30*, 85−90.

(13) Zhang, Y. L.; Crothers, D. M. Statistical mechanics of sequence-dependent circular DNA and its application for DNA cyclization. *Biophys. J.* **2003**, *84*, 136−153.

(14) Kratky, O.; Porod, G. Röntgenuntersuchung Gelöster Fadenmoleküle. *Rec. Trav. Chim.* **1949**, *68*, 1106−1122.

(15) Shimada, J.; Yamakawa, H. Ring-closure probabilities for twisted wormlike chains. Application to DNA. *Macromolecules* **1984**, *17*, 689.

(16) Cloutier, T. E.; Widom, J. Spontaneous sharp bending of double-stranded DNA. *Mol. Cell.* **2004**, *14*, 355−362.

(17) Cloutier, T. E.; Widom, J. DNA twisting flexibility and the formation of sharply looped protein-DNA complexes. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 3645−3650.

(18) Yan, J.; Marko, J. F. Localized single-stranded bubble mechanism for cyclization of short double helix DNA. *Phys. Rev. Lett.* **2004**, *93*, 108108.

(19) Wiggins, P. A.; Phillips, R.; Nelson, P. C. Exact theory of kinkable elastic polymers. *Phys. Rev. E* **2005**, *71*, 021909.

(20) Du, Q.; Smith, C.; Shiffeldrim, N.; Vologodskaia, M.; Vologodskii, A. Cyclization of short DNA fragments and bending fluctuations of the double helix. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 5397−5402.

(21) Lowary, P. T.; Widom, J. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* **1998**, *276*, 19−42.

(22) Olson, W. K.; Gorin, A. A.; Lu, X.-J.; Hock, L. M.; Zhurkin, V. B. DNA sequence-dependent deformability deduced from protein-DNA crystal complexes. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 11163−11168.

(23) Zhurkin, V. B.; Tolstorukov, M. Y.; Xu, F.; Colasanti, A. V.; Olson, W. K. Sequence-dependent variability of B-DNA: an update on bending and curvature. In *DNA Conformation and Transcription*; Ohyama, T., Ed.; Landes Bioscience/Eurekah.com, Georgetown, TX, 2005; pp 18−34.

(24) Davey, C. A.; Sargent, D. F.; Luger, K.; Mäder, A. W.; Richmond, T. J. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J. Mol. Biol.* **2002**, *319*, 1087−1113.

(25) Suzuki, M.; Yagi, N. Stereochemical basis of DNA bending by transcription factors. *Nucleic Acids Res.* **1995**, *23*, 2083−2091.

(26) Werner, M. H.; Gronenborn, A. M.; Clore, G. M. Intercalation, DNA kinking, and the control of transcription. *Science* **1996**, *271*, 778−784.

(27) Crothers, D. M.; Drak, J.; Kahn, J. D.; Levene, S. D. DNA bending, flexibility, and helical repeat by cyclization kinetics. *Methods Enzymol.* **1992**, *212*, 3−29.

(28) Metropolis, N. A.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, H.; Teller, E. Equation of state calculations by fast computing machines. *J. Chem. Phys.* **1953**, *21*, 1087−1092.

(29) Alexandrowicz, Z. Monte Carlo of chains with excluded volume: a way to evade sample attrition. *J. Chem. Phys.* **1969**, *51*, 561−565.

(30) Dickerson, R. E.; Bansal, M.; Calladine, C. R.; Diekmann, S.; Hunter, W. N.; Kennard, O.; von Kitzing, E.; Lavery, R.; Nelson, H. C. M.; Olson, W. K.; Saenger, W.; Shakked, Z.; Sklenar, H.; Soumpasis, D. M.; Tung, C.-S.; Wang, A. H.-J.; Zhurkin, V. B. Definitions and nomenclature of nucleic acid structure parameters. *J. Mol. Biol.* **1989**, *208*, 787−791.

(31) Press: W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical Recipes in C*; Cambridge University Press: New York, 1986.

(32) Flory, P. J.; Suter, U. W.; Mutter, M. Macrocyclization equilibria. 1. Theory. *J. Am. Chem. Soc.* **1976**, *98*, 5733−5739.

(33) Zhurkin, V. B.; Lysov, Y. P.; Ivanov, V. I. Anisotropic flexibility of DNA and the nucleosomal structure. *Nucleic Acids Res.* **1979**, *6*, 1081−1096.

(34) El Hassan, M. A.; Calladine, C. R. The assessment of the geometry of dinucleotide steps in double-helical DNA: a new local calculation scheme. *J. Mol. Biol.* **1995**, *251*, 648−664.

(35) Coleman, B. D.; Olson, W. K.; Swigon, D. Theory of sequence-dependent DNA elasticity. *J. Chem. Phys.* **2003**, *118*, 7127−7140.

(36) Merlitz, J.; Rippe, K.; Klenin, K. V.; Langowski, J. Looping dynamics of linear DNA molecules and the effect of DNA curvture: a study of Brownian dynamics simulation. *Biophys. J.* **1998**, *74*, 773−779.

(37) Horowitz, D. S.; Wang, J. C. Torsional rigidity of DNA and length dependence of the free energy of DNA supercoiling. *J. Mol. Biol.* **1984**, *173*, 75−91.

(38) Heath, P. J.; Clendenning, J. B.; Fujimoto, B. S.; Schurr, J. M. Effect of bending strain on the torsion elastic constant of DNA. *J. Mol. Biol.* **1996**, *260*, 718−730.

(39) Shore, D.; Langowski, J.; Baldwin, R. L. DNA flexibility studied by covalent closure of short fragments into circles. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 4833−4837.

(40) Shore, D.; Baldwin, R. L. Energetics of DNA twisting. I. Relation between twist and cyclization probability. *J. Mol. Biol.* **1983**, *170*, 957−981.

(41) Shore, D.; Baldwin, R. L. Energetics of DNA twisting. II. Topoisomer analysis. *J. Mol. Biol.* **1983**, *170*, 983−1007.

(42) Taylor, W. H.; Hagerman, P. J. Application of the method of phage T4 DNA ligase-catalyzed ring-closure to the study of DNA structure. II. NaCl-dependence of DNA flexibility and helical repeat. *J. Mol. Biol.* **1990**, *212*, 363−376.

(43) Trifonov, E. N. DNA in profile. *Trends Biochem. Sci.* **1991**, *16*, 467−470.

(44) Hagerman, P. J. Straightening out the bends in curved DNA. *Biochim. Biophys. Acta* **1992**, *1131*, 125−132.

(45) Anderson, J. N. Detection, sequence patterns and function of unusual DNA structures. *Nucleic Acids Res.* **1986**, *14*, 8513−8533.

(46) Olson, W. K.; Marky, N. L.; Jernigan, R. L.; Zhurkin, V. B. Influence of fluctuations on DNA curvature. A comparison of flexible and static wedge models of intrinsically bent DNA. *J. Mol. Biol.* **1993**, *232*, 530−554.

(47) Schellman, J. A. The flexibility of DNA. *Biopolymers* **1974**, *13*, 217−226.

(48) Gorin, A. A.; Zhurkin, V. B.; Olson, W. K. B-DNA twisting correlates with base pair morphology. *J. Mol. Biol.* **1995**, *247*, 34−48.

(49) Protozanova, E.; Yakovchuk, P.; Frank-Kamenetskii, M. D. Stacked-unstacked equilibrium at the nick site of DNA. *J. Mol. Biol.* **2005**, *342*, 775−785.

(50) Travers, A. DNA dynamics: bubble 'n flip for DNA cyclization. *Curr. Biol.* **2005**, *15*, R377−379.

(51) Olson, W. K.; Swigon, D.; Coleman, B. D. Implications of the dependence of the elastic properties of DNA on nucleotide sequence. *Philos. Trans. R. Soc.* **2004**, *362*, 1403−1422.

(52) Calladine, C. R. Mechanics of sequence-dependent stacking of bases in B-DNA. *J. Mol. Biol.* **1982**, *161*, 343−352.

(53) Ulyanov, N. B.; Zhurkin, V. B. Sequence-dependent anisotropic flexibility of B-DNA. A conformational study. *J. Biomol. Struct. Dynam.* **1984**, *2*, 361−385.

(54) Sarai, A.; Mazur, J.; Nussinov, R.; Jernigan, R. L. Sequence dependence of DNA conformational flexibility. *Biochemistry* **1989**, *28*, 7842−7849.

(55) Olson, W. K. The flexible DNA double helix. In *Stereodynamics of Molecular Systems*; Sarma, R. H., Ed.; Pergamon Press: New York, 1979; pp 297−314.

CT060025+

JCTC Journal of Chemical Theory and Computation

The Extent of Cooperativity of Protein Motions Observed with Elastic Network Models Is Similar for Atomic and Coarser-Grained Models

Taner Z. Sen,[†,‡] Yaping Feng,[‡,§] John V. Garcia,[†] Andrzej Kloczkowski,[†] and Robert L. Jernigan*[,†,‡]

L.H. Baker Center for Bioinformatics and Biological Statistics, Iowa State University, Ames, Iowa 50011-3020, and Department of Biochemistry, Biophysics, and Molecular Biology, and Bioinformatics and Computational Biology Program, Iowa State University, Ames, Iowa 50011

Received February 13, 2006

Abstract: Coarse-grained elastic network models have been successful in determining functionally relevant collective motions. The level of coarse-graining, however, has usually focused on the level of one point per residue. In this work, we compare the applicability of elastic network models over a broader range of representational scales. We apply normal mode analysis for multiple scales on a high-resolution protein data set using various cutoff radii to define the residues considered to be interacting or the extent of cooperativity of their motions. These scales include the residue, atomic, proton, and explicit solvent levels. Interestingly, atomic, proton, and explicit solvent level calculations all provide similar results at the same cutoff value, with the computed mean-square fluctuations showing only a slightly higher correlation (0.61) with the experimental temperature factors from crystallography than the results of the residue-level coarse-graining. The qualitative behavior of each level of coarse graining is similar at different cutoff values. The correlations between these fluctuations and the number of internal contacts improve with increased cutoff values. Our results demonstrate that atomic level elastic network models provide an improved representation for the collective motions of proteins compared to the coarse-grained models.

Introduction

Elastic Network Models[1−3] have been quite successful in predicting the large-scale motions of proteins and other biological structures, even for such large complexes as the ribosome.[4−6] These models originated from the theory of polymer networks[7,8] using the pioneering idea of Tirion,[3] who proposed a single uniform spring constant parameter for all atom−atom contacts used in a normal mode analysis. Elastic Network applications have usually focused on coarse-grained representations of proteins, using mostly C^{α}-atoms and relying upon $C^{\alpha}-C^{\alpha}$ proximity for placement of springs. The predicted position fluctuations of amino acids in proteins obtained from Elastic Network Models usually give quite good agreement with experimental B-factors measured by crystallographers, but as we will see here more detailed atomic models yield similar, if slightly better, results. This is an important finding that may be particularly important for developing mixed coarse-grained models wherein the functionally important part of the protein is represented by atoms, and the remainder of the structure is rendered in lesser detail. The only information utilized in Elastic Network Models is the structure of the protein, from the Protein Data Bank (PDB),[9] but this approach can also be applied to hypothetical protein models based on sequence similarities

* Corresponding author phone: (515)294-3833; fax: (515)294-3841; e-mail: jernigan@iastate.edu.
† L.H. Baker Center for Bioinformatics and Biological Statistics.
‡ Department of Biochemistry, Biophysics, and Molecular Biology.
§ Bioinformatics and Computational Biology Program.

10.1021/ct600060d CCC: $33.50
Published on Web 04/08/2006

or other techniques. The essential aspect of these models is a representation of proteins as highly interconnected structures, which represents well their cohesive and cooperative nature. It has been shown that fluctuations of residues in proteins depend mostly on the packing density and that the slowest modes corresponding to the motions of large domains depend essentially on the protein shape.[10,11] Elastic Network Models have been useful in studies of protein binding[12] and the analysis of the binding pocket flexibility.[13]

One of the strengths of the Gaussian Network Model is its success in the determination of functionally significant collective motions in proteins with an extremely simple model based only on packing density and geometry. However, does such a simple model, which does not differentiate between various bonded and nonbonded interactions (such as covalent and hydrogen bonds), produce physically meaningful results? There is strong evidence that it actually does. First, the accumulated normal mode analysis results demonstrate clearly that GNM produces experimentally verifiable results, e.g. for X-ray analysis,[2,14] NMR,[15] hydrogen-exchange,[16] and cryo-EM[4,17,18] experiments. Second, the normal mode results correlate well with results of molecular dynamics (MD) simulations[19] based on detailed atomistic force fields. These studies have proven that the normal mode analysis using coarse-grained models is extremely useful and that collective motions derived from the equilibrium structure depend largely on the shape of the protein, rather than on particular types of interactions.[10,11] A lack of any dependence on discriminating between bonded and nonbonded interactions is most likely due to the large number of interactions inside compact structures of biomolecules that leads to their cohesiveness and cooperativity. Essentially for large compact structures the number of covalent bonds is small compared to the number of nonbonded interactions. Note that this conclusion does not negate the differential importance of certain types of interactions for protein stability or for the folding process.

Although elastic network models have proven to provide a good description of protein collective motions, the effect of coarse-graining over the full range of scales has not been thoroughly explored. Jernigan and co-workers have mostly analyzed one end of the spectrum—coarser-grained models of proteins—and have observed that even when 40 residues of hemagglutinin A are represented by a single node, the global motions are only slightly affected[20,21] in comparison to more detailed models. Here we will explore the other end of the spectrum, and study the effect of more detailed representations of proteins for the elastic network models. We will analyze the effect of scaling in elastic network models by comparing results obtained at varying levels of coarse-graining. These levels will include one point per residue, one point per atom for heavy atoms alone, and the case when protons are also included. Additionally we will investigate the effect of explicit inclusion of solvent molecules insofar as they are reported for high-resolution protein structures. For the residue-level coarse-graining, a single node (located at C^{α}) is assigned to each residue. For the atomic-level representation, each heavy atom in the protein is assigned a node, and hydrogen atoms are neglected. For the proton-level additional nodes for each hydrogen atom in the protein are included. Finally, in the explicit solvent-level representation oxygen and hydrogen atoms of the water molecules reported in the crystallographic data are also taken into account, and each of these atoms is represented by a node. Our study will allow us to analyze the effects of scaling at various levels of accuracy and present a multiscale picture of the normal mode of protein dynamics.

Previously[22] we had observed a strong correlation between the entropies computed from the elastic network models with the number of internal contacts in the given protein. This corresponds to a simple view of protein stabilities, in which the number of contacts (stabilizing energy) compensates directly for the extent of motions within the structure (motional entropy). Conceivably such a simple relationship could also depend on the level of cooperativity in the model, i.e., the cutoff distance defining both the number of contacts and their restraining effects on the motions of the protein. We have investigated this correlation for the same set of proteins and at different levels of coarse-graining with the same elastic network models.

Although normal mode analyses provide a remarkable tool for probing protein dynamics, they have some limitations: every interaction is treated identically for all contacts regardless of the contact distance or type of interaction. We have observed however that the results obtained by using residue-type specific potentials[23,24] at the residue-level coarse-graining (unpublished results) or adjusted springs based on number of contacts[25] are not substantially different from those obtained by using a harmonic potential with a single uniform spring constant. Furthermore, elastic network model results are comparable to those of molecular dynamics based on AMBER potential.[19] Here, we take a different route and explore the effect of assigning a harmonic potential with a single uniform spring constant for each pair of nodes being in contact regardless of the type of the interaction, at all of the different scales of coarse-graining.

There are other important reasons to introduce more detailed atomic level elastic network models. For other types of studies such as enzyme mechanisms,[26] unraveling the details of molecular hinges or detailed investigations of residue conservation around hinges, further detail is likely to be important. One potential outcome from the atomic elastic networks could be the identification of specific conserved atomic groups, in more detail than residue conservation, relating to critical functional motions and flexibility, within molecular hinges, enzyme active sites, or other functional loci. This could be information of importance for protein design. One of the appealing aspects of the atomic models is that they can be conveniently combined with other more coarsely grained parts of the structure (mixed coarse-graining), as has been demonstrated previously.[20,21,27]

Methods

Data Set. We used search tools available on the PDB Web site to find proteins with resolution better than 0.8 Å and with less than 50% sequence similarity to one another. We narrowed our list for this initial study to only eight mostly single chain proteins whose lengths range from 64 to 158

amino acids. These proteins, listed in order by their increasing size are as follows: type III antifreeze protein rd1 (pdb id: 1ucs) (64 residues), syntenin Pdz2 domain (1r6j) (82 residues), high-potential iron−sulfur protein (1iua) (83 residues), Lys-49 phospholipase A2 homologue (lysine 49 PLA2) (1mc2) (122 residues), cobratoxin (1v6p) (2 chains 62 residues each), bacterial photoreceptor pyp (1nwz) (125 residues), carbohydrate binding domain Cbm36 (1w0n) (131 residues), and *E. coli* pyrophosphokinase HPPK (1f9y) (158 residues).

Multiscale Representations. Our defined models are as follows: "residue-level models" include only C^{α} atoms; "atomic-level models" include every atom in a protein except hydrogen atoms; "proton-level models" include every atom in a protein including hydrogen atoms; and finally, "explicit solvent-level models" include every protein atom and also every oxygen and hydrogen atom of water molecules in the crystallographic data provided in the protein PDB. If the positions of hydrogen atoms are not found in the pdb file, Accelrys DS ViewerPro is used to generate locations of missing hydrogen atoms. Ligands are removed from the protein structures and are not included in the present analyses.

Gaussian Network Models. The details of the Gaussian Network Model[2] (GNM) and its extension considering the directionalities of fluctuations—the Anisotropic Network Model[1] can be found elsewhere. The GNM originates from the theory of rubberlike elasticity[7,8] and Tirion's approach of using a uniform spring constant parameter in the harmonic analysis of protein motions.[3] The cohesiveness of the protein structure in the elastic network model is represented by assuming that all pairs of nodes separated by less than a certain cutoff distance are connected by uniform springs. In the standard coarse-grained version, each residue is represented by a single point (node) positioned at its C^{α} atom, but we will also use an atomic version here where the points represent atoms. There are two parameters in the model: the cutoff distance R_c and the spring constant γ. The cutoff distance R_c determines whether two residues are connected by a spring, i.e., are in contact, without differentiating between bonded and nonbonded interactions. These contacts are mathematically expressed as the contact (Kirchhoff) matrix, Γ, where the ijth element of the matrix is -1 if nodes i and j are connected by a spring, and zero otherwise, and the diagonal elements are the sums of nondiagonal elements in a given row (or column) taken with the negative sign. Because of this definition the matrix Γ is singular (its determinant is zero), and only the pseudoinverse of Γ can by calculated by using the singular value decomposition (SVD) method. It can be shown that the zero eigenvalues of Γ that are eliminated by using SVD correspond to the six external rigid body degrees of freedom. The equilibrium correlations $\langle \Delta \boldsymbol{R}_i \cdot \Delta \boldsymbol{R}_j \rangle$ between fluctuations of residues i and j are proportional to the ijth element of the inverse of Γ

$$\langle \Delta \boldsymbol{R}_i \cdot \Delta \boldsymbol{R}_j \rangle = \frac{3k_B T}{2\gamma} (\Gamma^{-1})_{ij} \tag{1}$$

where $\Delta \boldsymbol{R}_i$ and $\Delta \boldsymbol{R}_j$ are the vectors representing the instantaneous displacements of the ith and the jth nodes from their mean positions. Here k_B is Boltzmann's constant, T is temperature, and γ is is the spring constant. The mean-square fluctuation $\langle (\Delta R_i)^2 \rangle$ of the ith node is then given by the ith diagonal element $[\Gamma^{-1}]_{ii}$ of the matrix Γ^{-1}. The mean-square fluctuations may be compared directly with the experimental crystallographic Debye−Waller temperature factors (B-factors) usually available in the pdb files by the equation

$$B_i = 8\pi^2 \langle (\Delta R_i)^2 \rangle / 3 \tag{2}$$

The pseudoinverse matrix Γ^{-1} can be expanded in the series of eigenvalues λ_k and eigenvectors $\boldsymbol{u}_k$ of the contact matrix Γ as follows

$$\Gamma^{-1} = \sum_k {\lambda_k}^{-1} \mathbf{u}_k \mathbf{u}_k^T \tag{3}$$

where zero eigenvalues (that physically correspond to motions of the center of mass of the system) are excluded from the summation. This eigenexpansion has a direct physical meaning by showing contributions from individual modes associated with the eigenvalues of Γ.[28] The ith component of the eigenvector $\boldsymbol{u}_k$ (corresponding to the kth normal mode) specifies the magnitude of the mean-square fluctuations of the ith node in the kth mode. It can also be shown that all eigenvalues of Γ are non-negative. If we order eigenvalues according to their ascending values starting from zero, then the most important contributions in eq 3 are given by the smallest nonzero eigenvalues λ_k, that correspond to the large-scale, slow, collective modes. Slowest modes play a dominant role in the fluctuational dynamics of protein structures, because of their contributions to the mean-square fluctuations scale with ${\lambda_k}^{-1}$. It has been shown that the most important motions of proteins[29−31] or large biological structures (such as the ribosome)[4−6,32,33] that are associated with their biological function can be clearly identified with a few slowest modes of GNM. The large-scale changes of protein conformations between 'open' and 'closed' forms, or domain swapping in proteins, can be also well represented with elastic network models.[34] Reviews of elastic network applications can be found in refs 16 and 35.

Correlation Coefficients. The usual criterion for choosing parameters is based upon achieving the best agreement between the computed fluctuations and the experimental B-factors. For this purpose, here we use the linear correlation coefficient:

$$C = \frac{\sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N} (x_i - \bar{x})^2 \cdot \sum_{i=1}^{N} (y_i - \bar{y})^2}} \tag{4}$$

In this equation, N is the number of nodes, and x_i and $\bar{x}$ are the mean-square fluctuations of the ith node calculated by GNM and their mean over all nodes, respectively. Similarly, y_i and $\bar{y}$ are the experimentally determined B-factor for the ith node and the mean over all nodes. The linear correlation coefficient is a straightforward way to analyze the extent of linear dependence between any two quantities. Its value can range between 1 and -1, where the limiting

values 1 and −1 correspond to perfect correlation and perfect anticorrelation.

Overlaps. Absolute overlap between two eigenvectors, each representing specific motions, is defined as

$$|\cos\theta| = \frac{|\sum_{i}^{n} x_i y_i|}{|\boldsymbol{x}|\cdot|\boldsymbol{y}|} \quad (5)$$

In this equation, $\boldsymbol{x}$ and $\boldsymbol{y}$ are two eigenvectors, x_i and y_i denote their ith components, and θ is the angle between $\boldsymbol{x}$ and $\boldsymbol{y}$. If two eigenvectors are exactly collinear, then their absolute overlap equals 1. If they are orthogonal to each other, i.e. the angle between the two eigenvectors is 90°, then the absolute overalp will be equal to zero. This provides a measure of the extent of similarity in the directions of motions for different modes.

Entropy. In the Gaussian Network Model fluctuations of residues about their mean positions obey the Gaussian distribution

$$W(\Delta\boldsymbol{R}_i) = A\exp\{-3(\Delta R_i)^2/2\langle(\Delta R_i)^2\rangle\} \quad (6)$$

The conformational entropy change ΔS_i resulting from fluctuations in the position of the ith residue can be obtained from the equation

$$\Delta S_i = k_B \ln W(\Delta\boldsymbol{R}_i) = -\gamma(\Delta R_i)^2/(2T[\Gamma^{-1}]_{ii}) \quad (7)$$

Equation 1 for the case $i = j$ was applied in the above derivation. Equation 7 can be used to calculate the free energy increase of entropic origin contributed by the ith residue, upon distortion $\Delta\boldsymbol{R}_i$ of its coordinates

$$\Delta G_i = -T\Delta S_i = \frac{\gamma}{2}(\Delta\boldsymbol{R}_i)^2/[\Gamma^{-1}]_{ii} \quad (8)$$

This free energy change is inversely proportional to $\langle(\Delta\boldsymbol{R}_i)^2\rangle$. Physically, this signifies a stronger resistance to deformation, including unfolding, of residues subject to smaller amplitude fluctuations in the folded state.[16]

Results and Discussion

Choosing Spring Constants for Different Resolution Scales. The Gaussian Network Model requires specification of two parameters: the spring constant that defines the strength of interactions and the cutoff distance that defines whether two given nodes are in contact or not. The spring constant ultimately scales the amplitudes of motions calculated from the contact matrix. When comparing results obtained at different scales, the spring constant should be adjusted to reflect the scale at which the protein is modeled.[27] Here, the spring constants at each scale are calculated for each protein by comparing fluctuations predicted by GNM with experimentally determined B-factors, as this method has proven to be generally successful in the past.

Choosing Cutoff Radii for Different Resolution Scales. Correlations between the GNM-derived mean-square fluctuations and crystallographic B-factors calculated from eq 4 clearly show the extent to which GNM results represent actual protein motions. Phillips and co-workers[14] showed that

Table 1. Average Correlation Coefficients between Computed Mean-Square Fluctuations and Experimental B-Factors for Four Different Resolution Levels of Coarse-Graining as a Function of the Cutoff Distance[a]

cutoff (Å)	residue level	atomic level	proton level	solvent level
1				
2		0.17	0.37	0.27
3		0.46	0.59	0.51
4	0.17	**0.61**	0.58	0.42
5	0.38	0.59	0.59	0.48
6	0.38	0.57	0.59	0.52
7	0.51	0.60	0.60	0.56
8	0.55	0.60	0.61	0.56
9	0.52	0.60	**0.61**	0.57
10	0.56	0.60	0.60	0.58
11	**0.56**	0.61	0.60	0.59
12	0.54	0.60	0.60	0.59
13	0.55	0.59	0.59	0.59
14	0.55	0.58	0.59	**0.60**
15	0.54	0.57	0.58	0.59

[a] A correlation of 1 shows perfect correlation and 0 the lack of correlation (maxima are indicated in bold).

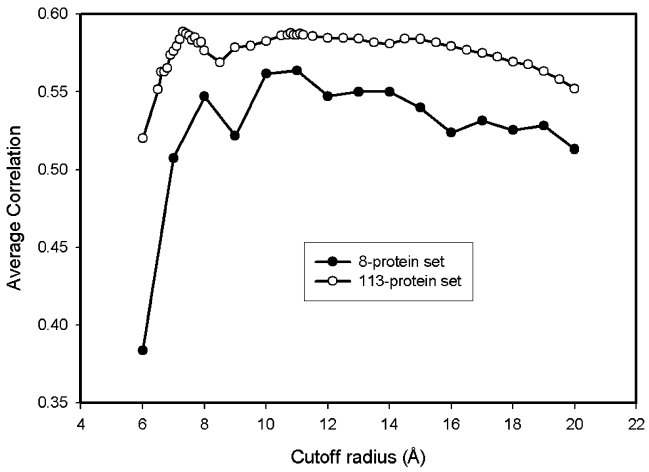


Figure 1. Average correlation coefficients as a function of the cutoff radius for the 8-protein set used in this study and the 113-protein set used by Phillips.[14] The correlation coefficients between the results of residue-level coarse-grained model and experimental B-factors for both data sets suggest two optimal cutoff radii around 7.3 and 11.1 Å.

GNM coarse-grained at the residue-level has a correlation of about 0.6 with the experimental data, depending on the cutoff radius and on the extent of inclusion of neighboring molecules packed in the crystal. Although 60% correlation at the residue-level is rather impressive, here we are studying the effect of including other atoms together with solvent molecules in the crystal on these correlations. Table 1 shows the correlation coefficients for C^{α}-atoms calculated at the residue, atomic, proton, and the explicit solvent levels for various cutoffs.

The results in Table 1 show that at the residue level, the correlation increases with the increasing cutoff radius reaching a peak around 11 Å as shown in Figure 1. However, the average correlation coefficient never exceeds 0.56. Although the value of this correlation is close to the result (∼0.6)

Table 2. Optimum Cutoff Radii (Å) for Eight Proteins in the Data Set for Four Different Resolution Level Models[a]

	1ucs	1iua	1r6j	1w0n	1mc2	1nwz	1v6p	1f9y
residue	8 (0.65)	13 (0.54)	14 (0.76)	12 (0.48)	7 (0.60)	10 (0.54)	6 (0.63)	19 (0.78)
atom	4 (0.67)	5 (0.56)	14 (0.72)	8 (0.58)	5 (0.73)	4 (0.67)	7 (0.64)	22 (0.78)
proton	3 (0.66)	5 (0.54)	15 (0.71)	8 (0.59)	5 (0.68)	3 (0.63)	7 (0.67)	23 (0.78)
solvent	15 (0.59)	15 (0.53)	18 (0.69)	9 (0.51)	5 (0.62)	10 (0.66)	7 (0.64)	23 (0.78)

[a] The correlation coefficients are given in parentheses.

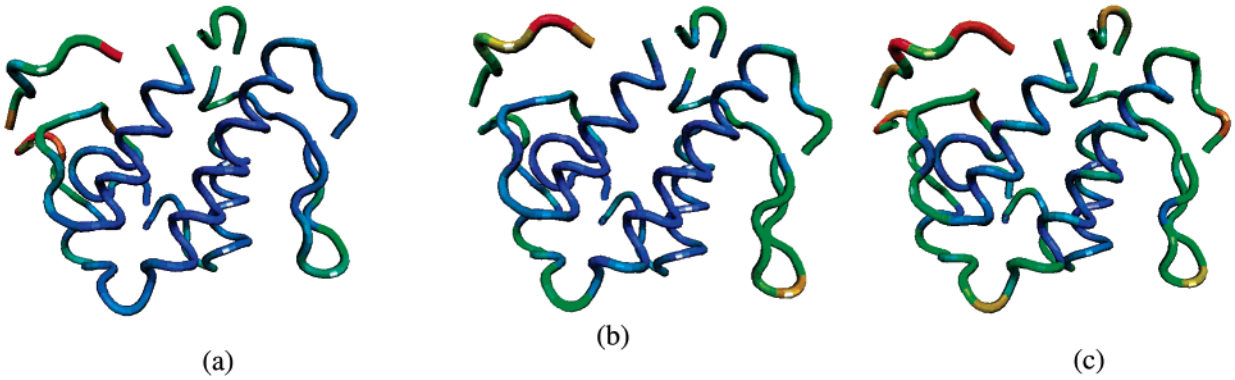


Figure 2. The schematic picture of lysine 49 PLA2 (PDB id: 1mc2). The backbone is colored according to the magnitude of mean-square fluctuations obtained (a) experimentally, (b) computed from the residue-level GNM, and (c) calculated from the atomic-level GNM. Most mobile regions are colored with red, less mobile regions with green, and, finally, almost immobile regions with blue.

obtained by Phillips,[14] the optimum cutoff radius (11 Å) found here is much larger than the Phillips' optimum cutoff of 7.3 Å. One major difference is that we have neglected intermolecular contacts due to packing in crystal. It is also important to note that the number of proteins in our data set is quite limited (8 proteins only). For further comparison with the Phillips group's results,[14] we repeated the average correlation coefficient calculations as a function of cutoff distance with their data set of 113 proteins. These results shown in Figure 1 indicate that for the 113-protein data set, another peak around 11.1 Å is also clearly visible. Figure 1 also demonstrates that although the 8-protein set consistently exhibits lower correlations than the 113-protein set, the average correlation coefficients of both sets have similar patterns; thus the 8-protein set seems to be sufficiently representative to make comparisons at various radii.

Table 2 lists the optimum cutoff distances for all eight proteins for each of the four different resolution level models studied here. The correlation coefficients are also given in Table 2 in parentheses. A real surprise comes upon examination of average correlation coefficients obtained at better resolution with more detailed scales. The inclusion of other atoms in the normal mode analysis increases the average correlation coefficient for the fluctuations of the C^{α}-atoms by 0.05−0.61. This is highly interesting, because although all interactions are treated similarly, a better correlation is obtained. The inclusion of all heavy atoms clearly provides a superior representation of protein structure and protein dynamics. Interestingly, the further inclusion of protons or even atoms of the solvent does not enhance these correlations and only shifts the optimum cutoff radius. The optimum cutoffs for various scales differ: for atomic and proton-level calculations, the optimum cutoff values are 4 and 9 Å, respectively, and for the explicit solvent level the optimum cutoff is 14 Å. It is worth emphasizing that the inclusion of atoms redefines the packing density critical for protein dynamics. While the consideration of protons in protein structure is associated with small uncertainties such as the ionization state of histidine, the inclusion of atoms of the explicit solvent is much more uncertain. At least it is encouraging that there is no visible loss of correlation when these possibly incomplete sets of solvent atoms are included.

Atomic and Proton Resolution Level Models Give Better Results than the Residue-Level Models. To analyze the effect of the resolution scale of the model, we have chosen one of the proteins from the data set lysine 49 PLA2 (pdb code: 1mc2) for a more detailed presentation of the results. A schematic representation of the protein backbone colored according to the magnitude of mean-square fluctuations of residues derived from the experimental data and from residue-level and explicit solvent-level models is shown in Figure 2 (parts a−c, respectively). The residue-level model computations were performed with the cutoff radius 7 Å and the atomic-level calculations with the cutoff 5 Å. Figure 3 shows the computed mean-square fluctuations of C^{α}-atoms for the residue-level and the atomic-level models. B-factors are also provided for comparison. The predicted fluctuations are calculated by summing over all internal normal modes. The mean-square fluctuations obtained for the residue-level model have a correlation of 0.60 with B-factors, whereas the atomic-level model calculations with 5 Å cutoff give a correlation 0.73 with the experimental data. Figure 3 shows that mean-square fluctuations predicted from the atomic-level model are significantly closer to the experimental B-factors, both qualitatively and quantitatively.

What is the source of the discrepancy between theoretical predictions and the experimental data? For further analysis, we focus on the PDZ2 domain of syntenin (1r6j). PDZ

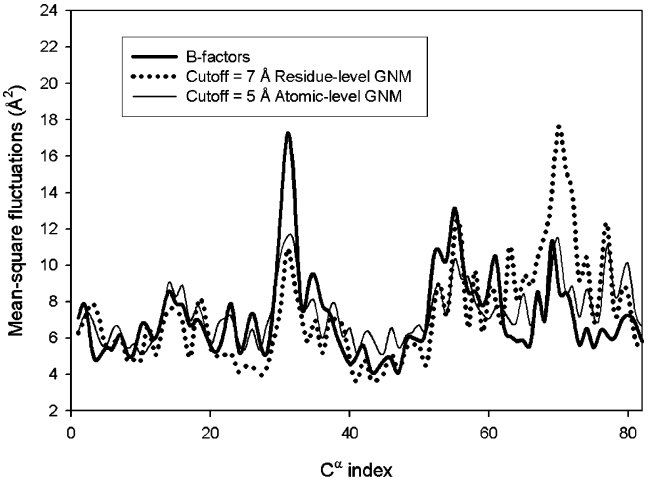


Figure 3. The mean-square fluctuations for lysine 49 PLA2 computed from the residue-level and the atomic-level models using optimal cutoffs. Results are shown for C^{α} atoms only.

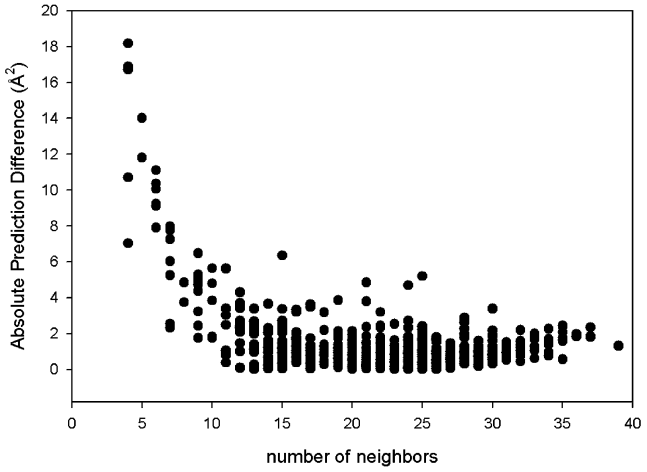


Figure 4. The absolute differences between atomic-level model predictions and experimental B-factors for the PDZ2 domain. The calculations are performed at the cutoff 5 Å as a function of the number of contacts (neighbors).

domains are mainly involved in the regulation of intracellular signaling and in the assembly of large protein complexes.[36] The structure of the PDZ2 domain of syntenin was resolved with a resolution 0.73 Å, allowing determination of coordinates of the hydrogen atoms in the crystal.[37] The PDZ2 domain contains 82 residues and 1867 atoms (including solvent atoms and hydrogen atoms). Figure 4 shows the dependence of the absolute value of the difference between predicted mean-square fluctuations and experimental B-factors as a function of the number of contacts in the protein structure. An inverse relationship can clearly be seen between this difference and the number of neighbors (contacts). Since nodes inside the protein core have more contacts, Figure 4 shows that the GNM predictions are generally less accurate on the protein surface. This implies that atoms on the protein surface should perhaps be treated in a more cooperative way than atoms of residues inside the core.

Since the GNM is mainly used to analyze cooperative global motions with functional relevance, a detailed analysis of slowest normal modes is of critical importance. For this purpose, we show in Figure 5 the overlaps of the eigenvectors computed for the residue-level and proton-level models. The overlap is defined by eq 5 as the absolute value of the cosine of the angle between these two eigenvectors. The absolute value of the overlap is used because the term $\mathbf{u}_k\mathbf{u}_k^T$ in eq 3 does not depend on the direction of the eigenvector $\mathbf{u}_k$, and the use of absolute cosine ensures that a 180° rotation still specifies the same type of motion. The overlap is calculated only for the eigenvector components corresponding to the C^{α}-atoms. Figure 5a−d illustrates these overlaps for four different proteins: (5a) 1ucs, (5b) 1r6j, (5c) 1w0n, and (5d) 1f9y.

Each point in Figure 5a−d shows a pair of eigenvectors, one computed from the residue-level model and the other from the proton-level model that have an absolute overlap of at least 0.4. The results were obtained by using optimum cutoff radii for each level of resolution for various proteins according to Table 2. For the case of syntenin, the eigenvectors corresponding to the first 10 slowest modes in both the residue-level and proton-level models have overlap higher than 0.4. However, this correspondence does not always hold; for example, for the case of pyrophosphokinase HPPK, this overlap is less good. More detailed studies are needed to conclude whether there may be certain regularities in the overlaps of modes in protein multiscale models.

Figure 5 shows scattered, sporadic, rather weak overlaps for 1ucs (Figure 5a) but not for other proteins (Figure 5b−d): The small (64 residues) type III antifreeze protein rd1 (pdb id: 1ucs) indeed shows very scattered overlaps, but for the larger proteins, there is a strong overlap between corresponding eigenvectors (around the diagonal of the plot) and very weak overlap between dissimilar eigenvectors (far from the diagonal). These high overlaps between these two different scales can be due to the protein size, which is indirectly related to packing density (the larger the protein, then the larger is its core having high packing density). Since the successes of Elastic Network Models depend on having an adequate representation of protein packing, larger proteins in general might be expected to exhibit better multiscale overlaps.

The Effect of Fluctuations in Elastic Network Models on Protein Entropy. We have calculated the correlation coefficient (defined by eq 4) between the free energy change of entropic origin given by eq 8 and the numbers of contacts for alpha-carbons of each residue at four different levels of coarse graining. The results have been averaged over the set of eight proteins and are shown in Table 3 as the function of the cutoff distance used for defining contacts. It is interesting that Table 3 strongly resembles Table 1. This resemblance originates from the fluctuational nature of these free energy changes.

Figure 6 shows plots of the absolute value of the entropy of fluctuations as a function of the total number of contacts for 3 different proteins: 1f9y, 1iua, and 1mc2. The calculations have been performed for the standard residue-level coarse-grained GNM. We used six different values of the cutoff radius defining contacts, ranging from 5 Å to 10 Å with increments of 1 Å. Each of these six cutoffs is represented by a marked point in Figure 6 starting from 5 Å

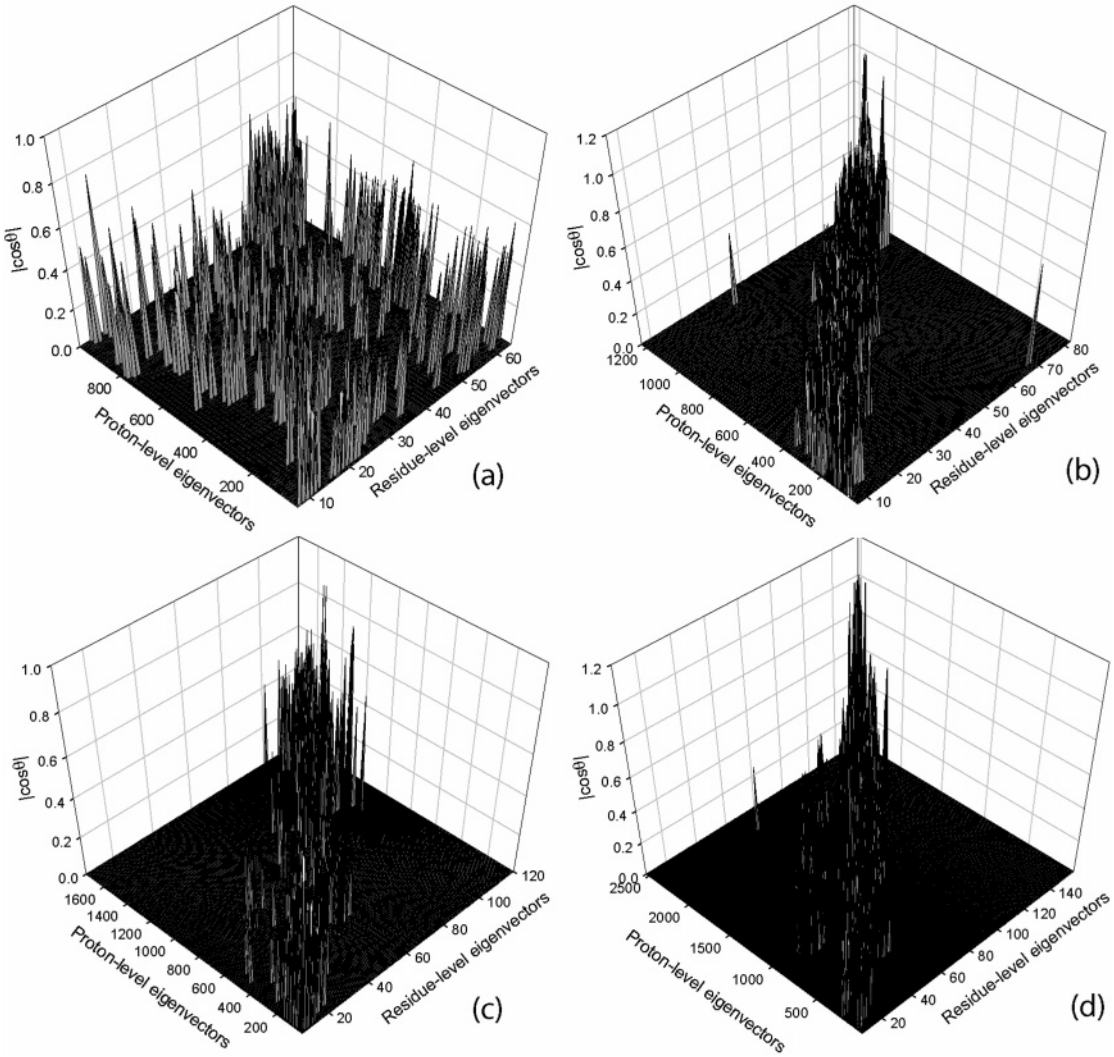


Figure 5. The absolute overlaps, $|\cos\theta|$, between eigenvectors obtained for the residue-level and the proton-level models for (a) type III antifreeze protein rd1 (1ucs), (b) syntenin Pdz2 domain (1r6j), (c) carbohydrate binding domain Cbm36 (1w0n), and (d) *E. coli* pyrophosphokinase HPPK (1f9y). The calculations were performed by using optimum cutoffs for each protein for a given model. Proteins are arranged from (a) to (d) according to increasing protein size.

Table 3. Average Correlation Coefficients between the Free Energy Change Due to Fluctuations (Entropy) and the Contact Number (Energy) as a Function of the Cutoff Distance for Four Different Resolution Level Models[a]

cutoff (Å)	residue level	atomic level	proton level	solvent level
1				
2		−0.32	−0.03	0.01
3		0.15	0.50	0.50
4	0.19	0.76	0.91	0.89
5	0.61	0.95	0.99	0.97
6	0.73	0.99	1.00	0.99
7	0.89	1.00	1.00	1.00
8	0.95	1.00	1.00	1.00
9	0.98	1.00	1.00	1.00
10	0.99	1.00	1.00	1.00

[a] Correlation coefficients have been averaged over the set of eight proteins. High values are achieved for the three more detailed models at lower cutoff values, as is also seen in Table 1.

on the left to 10 Å on the right. The linearity of the plots in Figure 6 re-emphasizes the dependence of entropy on packing density. A related study was also done by us[38] and by Halle,[39] where an inverse relationship between mean-square fluctuations and contact densities can be seen. It is also worth noting that entropy depends on the size of the protein. The largest of the three proteins 1f9y (158 residues) has the smallest entropies, and the smallest one 1iua (83 residues) has the largest entropies for the same number of contacts, as seen in Figure 6. This means that the fluctuation entropy *per contact* is smaller for larger proteins, i.e., large proteins exhibit more cooperative motions.

Conclusions

We have applied normal mode analysis with multiscale coarse-graining to high-resolution protein structures. The atomic, proton, and explicit solvent level models all provide quite similar results, showing significantly higher correlations of the predicted fluctuations of C^{α}-atoms with the experimental B-factors than the residue level GNM. At the residue-level coarse-graining, the optimum cutoff radius is ∼11 Å, which is significantly larger than the value 7.3 Å obtained

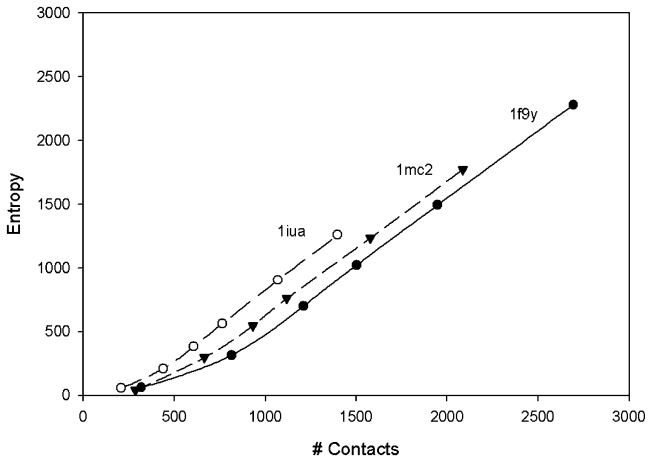


Figure 6. The absolute value of the entropy of fluctuations as a function of the total number of contacts for 3 different proteins. The residue-level coarse-grained model was used. For each protein, there are 6 points corresponding to 6 different cutoffs varying from 5 Å on the left to 10 Å on the right in increments of 1 Å.

by Phillips and co-workers.[14] This suggests that the optimum cutoff radius may depend on the specific protein structure, and the inclusion of intermolecular contacts in the crystal seems to be necessary at the residue level resolution. The absence of these intermolecular contacts in our model must be compensated by an increased cutoff that increases the number of springs and leads to better agreement with experimental data. The inclusion of atoms in our models significantly improves predictions of fluctuations of C^{α}-atoms and gives better correlations with experimental B-factors. Additionally better resolution atomic scale models require small cutoff radius (4 Å). More detailed atomic resolution level elastic network models are likely to provide a better representation of motions in proteins. Our results also show that small proteins may require atomic scale resolution models to achieve a good representation of their dynamics. However, the atomic level GNM computations for larger proteins require significantly larger computer resources than those for the residue-level GNM. An alternative that offers a compromise might be mixed coarse-grained modeling of proteins proposed by Doruker and Jernigan[20,21,27]—to include a high level of detail for the most important parts of the protein structure and less detail for other parts. Our analysis shows that the multiscale normal mode analysis can be useful for understanding and predicting the collective motions in proteins.

Acknowledgment. The authors thank Lei Yang for the critical reading of the manuscript. The authors acknowledge the financial support provided by the NIH Grants R01GM-072014 and R33GM066387. J.V.G. was a NIH-NSF BBSI Summer Institute in Bioinformatics and Computational Biology fellow.

References

(1) Atilgan, A. R.; Durell, S. R.; Jernigan, R. L.; Demirel, M. C.; Keskin, O.; Bahar, I. *Biophys. J.* **2001**, *80*, 505−515.

(2) Bahar, I.; Atilgan, A. R.; Erman, B. *Folding Des.* **1997**, *2*, 173−181.

(3) Tirion, M. M. *Phys. Rev. Lett.* **1996**, *77*, 1905−1908.

(4) Tama, F.; Valle, M.; Frank, J.; Brooks, C. L. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 9319−9323.

(5) Wang, Y. M.; Rader, A. J.; Bahar, I.; Jernigan, R. L. *J. Struct. Biol.* **2004**, *147*, 302−314.

(6) Wang, Y. M.; Jernigan, R. L. *Biophys. J.* **2005**, *89*, 3399−3409.

(7) Flory, P. J. *Proc. R. Soc. London, Ser. A* **1976**, *351*, 351−380.

(8) Kloczkowski, A.; Mark, J. E.; Erman, B. *Macromolecules* **1989**, *22*, 1423−1432.

(9) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235−242.

(10) Doruker, P.; Jernigan, R. L. *Proteins: Struct., Funct., Genet.* **2003**, *53*, 174−181.

(11) Lu, M. Y.; Ma, J. P. *Biophys. J.* **2005**, *89*, 2395−2401.

(12) Tobi, D.; Bahar, I. *PNAS* **2005**, 0507603102.

(13) Zheng, W.; Brooks, B. R. *Biophys. J.* **2005**, *89*, 167−178.

(14) Kundu, S.; Melton, J. S.; Sorensen, D. C.; Phillips, G. N. *Biophys. J.* **2002**, *83*, 723−732.

(15) Haliloglu, T.; Bahar, I. *Proteins: Struct., Funct., Genet.* **1999**, *37*, 654−667.

(16) Bahar, I.; Rader, A. J. *Curr. Opin. Struct. Biol.* **2005**, *15*, 586−592.

(17) Ming, D.; Kong, Y. F.; Lambert, M. A.; Huang, Z.; Ma, J. P. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 8620−8625.

(18) Beuron, F.; Flynn, T. C.; Ma, J. P.; Kondo, H.; Zhang, X. D.; Freemont, P. S. *J. Mol. Biol.* **2003**, *327*, 619−629.

(19) Micheletti, C.; Carloni, P.; Maritan, A. *Proteins: Struct., Funct., Bioinformatics* **2004**, *55*, 635−645.

(20) Doruker, P.; Jernigan, R. L.; Bahar, I. *J. Comput. Chem.* **2002**, *23*, 119−127.

(21) Kurkcuoglu, O.; Jernigan, R. L.; Doruker, P. *QSAR Comb. Sci.* **2005**, *24*, 443−448.

(22) Bahar, I.; Wallqvist, A.; Covell, D. G.; Jernigan, R. L. *Biochemistry* **1998**, *37*, 1067−1075.

(23) Miyazawa, S.; Jernigan, R. L. *Macromolecules* **1985**, *18*, 534−552.

(24) Miyazawa, S.; Jernigan, R. L. *J. Mol. Biol.* **1996**, *256*, 623−644.

(25) Sen, T. Z.; Jernigan, R. L. In *Normal-Mode Analysis: Theory and Applications to Biological and Chemical Systems*; Cui, Q., Bahar, I., Eds.; CRC: Boca Raton FL, 2005; Chapter 9, pp 171−186.

(26) Kurkcuoglu, O.; Jernigan, R. L.; Doruker, P. *Biochemistry* **2006**, *45*, 1173−1182.

(27) Doruker, P.; Jernigan, R. L.; Navizet, I.; Hernandez, R. *Int. J. Quantum Chem.* **2002**, *90*, 822−837.

(28) Haliloglu, T.; Bahar, I.; Erman, B. *Phys. Rev. Lett.* **1997**, *79*, 3090−3093.

(29) Keskin, O.; Durell, S. R.; Bahar, I.; Jernigan, R. L.; Covell, D. G. *Biophys. J.* **2002**, *83*, 663−680.

(30) Keskin, O.; Bahar, I.; Flatow, D.; Covell, D. G.; Jernigan, R. L. *Biochemistry* **2002**, *41*, 491−501.

(31) Navizet, I.; Lavery, R.; Jernigan, R. L. *Proteins: Struct., Funct., Genet.* **2004**, *54*, 384−393.

(32) Rader, A. J.; Wang, Y. M.; Bahar, I.; Jernigan, R. L. *Biophys. J.* **2004**, *86*, 190A.

(33) Trylska, J.; Konecny, R.; Tama, F.; Brooks, C. L.; McCammon, J. A. *Biopolymers* **2004**, *74*, 423−431.

(34) Kundu, S.; Jernigan, R. L. *Biophys. J.* **2004**, *86*, 3846−3854.

(35) Ma, J. P. *Structure* **2005**, *13*, 373−380.

(36) Sheng, M.; Sala, C. *Annu. Rev. Neurosci.* **2001**, *24*, 1−29.

(37) Kang, B. S.; Devedjiev, Y.; Derewenda, U.; Derewenda, Z. S. *J. Mol. Biol.* **2004**, *338*, 483−493.

(38) Liao, H.; Yeh, W.; Chiang, D.; Jernigan, R. L.; Lustig, B. *Protein Eng. Des. Select.* **2005**, *18*, 59−64.

(39) Halle, B. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 1274−1279.

CT600060D

JCTC Journal of Chemical Theory and Computation

Protein Structure Prediction: The Next Generation

Michael C. Prentiss,[†,‡] Corey Hardin,[||] Michael P. Eastwood,[‡] Chenghang Zong,[†,‡] and Peter G. Wolynes*[,†,‡,§]

Center for Theoretical Biological Physics, La Jolla, California 92093, Department of Chemistry and Biochemistry, University of California at San Diego, La Jolla, California 92093, Department of Physics, University of California, La Jolla, California 92093, and Department of Chemistry, University of Illinois, Urbana−Champaign, 600 South Mathews Avenue, Urbana, Illinois 61801

Received January 4, 2006

Abstract: Over the last 10−15 years a general understanding of the chemical reaction of protein folding has emerged from statistical mechanics. The lessons learned from protein folding kinetics based on energy landscape ideas have benefited protein structure prediction, in particular the development of coarse grained models. We survey results from blind structure prediction. We explore how second generation prediction energy functions can be developed by introducing information from an ensemble of previously simulated structures. This procedure relies on the assumption of a funneled energy landscape keeping with the principle of minimal frustration. First generation simulated structures provide an improved input for associative memory energy functions in comparison to the experimental protein structures chosen on the basis of sequence alignment.

Introduction

Every other summer, research groups compare their different protein structure prediction methods via the Critical Assessment of Techniques for Protein Structure Prediction (CASP) experiment. During the CASP experiment, sequences of experimentally determined protein structures that are not publicly available are placed on the Web. This exercise is double blind where neither the organizers nor the participants know the experimentally determined structure. Groups respond with up to 5 ranked predictions, before a predetermined date, such as the publication of the structures. Since the inception of CASP, a three-dimensional structure prediction category has expanded to address related prediction questions such as the sequence to structure alignment quality, amino acid side-chain placement, multidomain domain boundaries, and the ordered or disordered nature of a protein sequence.[1]

* Corresponding author e-mail: pwolynes@chem.ucsd.edu.
[†] Center for Theoretical Biological Physics.
[‡] University of California at San Diego.
[§] University of California, La Jolla.
[||] University of Illinois, Urbana−Champaign.

These different prediction questions can be examined from a common framework: the principle of minimal frustration. The principle of minimal frustration states that native contacts must be more favorable, in a strict statistical sense,[2] than non-native contacts in order for proteins to fold on physiologic time scales.[3] Without a sufficient energetic bias toward the native state, the multidimensional energy surface as a function of native structure possesses too many minima for an efficient stochastic search. Such an energy surface would lead to slow folding kinetics, even if the proteins never found a sufficiently stable native state. This is not true since we know most proteins fold without assistance.[4] The opposite of a rough energy surface is biased toward the native basin without any local minima is an absolute manifestation of the principle of minimal frustration. Funneled energy surfaces have no unfavorable energetic traps (i.e. Gō Models) and have been shown to reproduce most features of experimental folding kinetics.[5−7] These energy landscape concepts can richly be applied in several areas of chemistry and physics.[8] Apparently, evolution's energy function is minimally frustrated.

The correlation between a protein sequence and its three-dimensional structure can be described using similar land-

10.1021/ct0600058 CCC: $33.50
Published on Web 04/26/2006

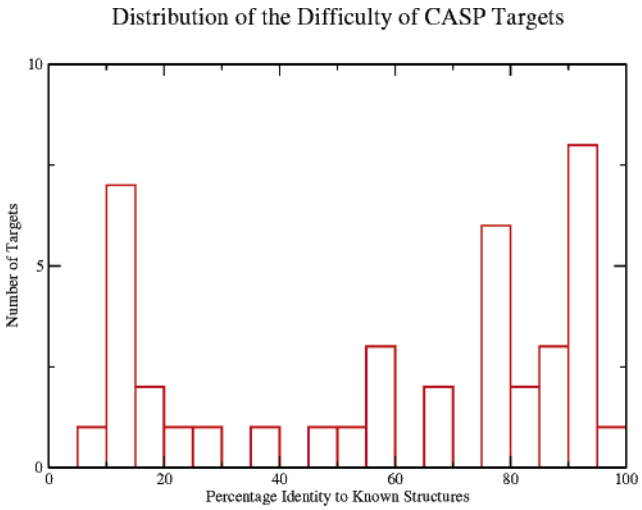


Figure 1. The difficulty of the prediction targets as defined by percent identity. Proteins below 25 sequence identity are usually considered ab initio or fold recognition targets.

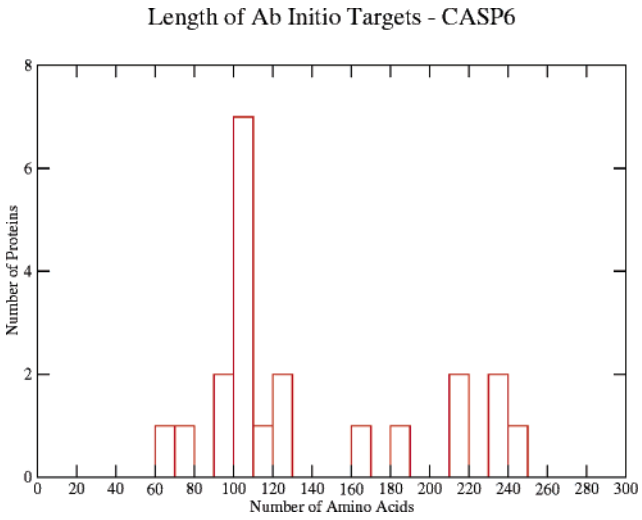


Figure 2. The ab initio prediction targets amino acid lengths for CASP6.

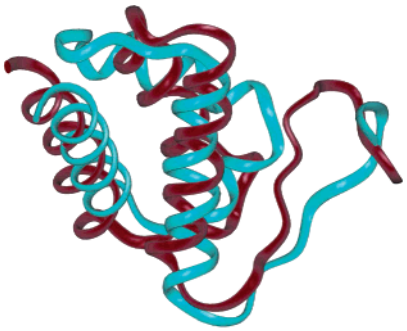


Figure 3. Sequence dependent superpositions of model 1 structure against the native state for CASP5 target T0170 (PDB ID 1UZC). Blue represents the prediction, and the native state is represented with red.

scape language. As a protein sequence diverges away from a consensus wild type sequence, the potential for energetically unfavorable interactions increases. The wild type sequence and its homologues will fold toward the same native basin. Only once enough frustrating contacts are added to wild type sequences will the sequence no longer correspond to the native state ensemble. Sequences with over 25% sequence identity to previously determined protein structures are called comparative modeling targets. The energy landscape underlying such a prediction is a Gō Model based on the structure of the known homologue. This heavily funneled energy surface yields high-resolution structures, with the discrepancies in the turns and residues which have poor sequence to structure alignments. Figure 1 demonstrates the distribution of homology of proteins sequence to known structures included in CASP6. Since proteins below 25% sequence identity are considered new fold recognition targets, 70% of the structures were comparative modeling targets. Recently sequenced genomes such as *E. coli* have the same ratio of ab initio to comparative modeling targets, which suggests the analysis of this ratio over time could be a useful measure of the progress of efforts to experimentally find examples of all of Nature's protein structures.

In contrast to comparative modeling, ab initio structure predictions do not have the advantage of creating Gō-like energy surfaces. While many ab initio targets contain less than 150 residues, and thus are candidates for standard techniques, there are several that are longer as shown in Figure 2. Most longer sequences will be multidomain proteins. This causes new problems. Folding a protein with two hydrophobic cores allows for new sources of frustration, beyond those present in single domain proteins. To obtain predictions for such problematic sequences, they usually must be divided into their constituent domains. Current methods for dividing the sequence into domains range from purely sequence based algorithms, which look for sequence patterns in multiple sequence alignments, to simulation techniques that look for hydrophobic core formation among multiple independent simulations.[9−11]

The case studies we highlight of difficult structure predictions were chosen from our participation in the CASP5 and CASP6 experiments. In CASP5, we utilized several improved techniques, such as a backbone hydrogen bond term for the proper formation of beta sheets and a liquid crystal-like term to ensure parallel or antiparallel sheet formation.[12] We also performed target sequence averaging which enhances the funneling of the prediction landscape[13] and assessed our ensemble of sampled structures with a 20 letter contact for submission.[14] Our most striking result from this round of blind prediction was a prediction for target T0170 protein databank[15] code (PDB ID IUZC). Figure 3 presents the sequence dependent overlay of our model 1 structure with the experimentally determined structure. The sequence dependent alignment quality of this structure is high as measured by a Q score of 0.38. Q is an order parameter defined in eq 1 that measures the sequence dependent structural complementarity of two structures, where Q is defined as a normalized summation of C-alpha pairwise contact differences.

$$Q = \frac{2}{(N-1)(N-2)} \sum_{i<j} \exp\left[-\frac{(r_{ij} - r_{ij}^{N})^2}{\sigma_{ij}^{2}} \right] \qquad (1)$$

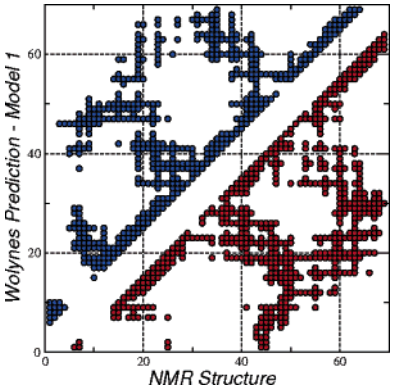


Figure 4. Contact map of target T0170 (PDB ID 1UZC) model 1 structure against the NMR structure.

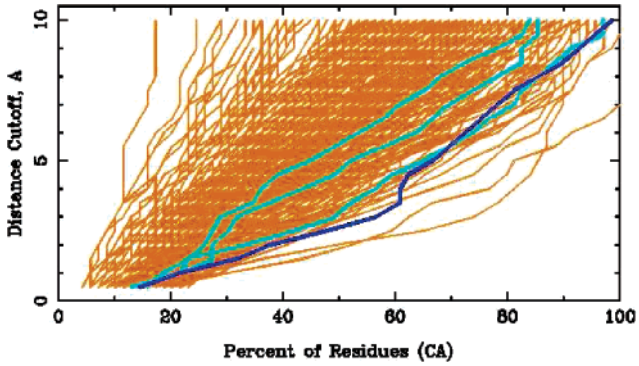


Figure 5. Percentage of residues under a RMSD limit (dark blue − model 1, light blue − model 2-5, orange − other groups prediction).

The resulting order parameter, Q, ranges from 0, when there is no similarity between structures at a pair level, to 1 which is an exact match. Q has been shown to be more sensitive in determining the quality of intermediate quality protein structure predictions.[13] Q scores of 0.4 for single domain proteins equals an RMSD of 5 Å. In most cases the reference state for the Q score is the native state, but often one wants to compare structural similarity between structures in a simulation. A sequence independent measure CE[16] also scores well (CE Z-score = 4.1). The CE Z-score measures structural complementarity without regard to sequence information and is parametrized such that structure between a Z-score greater than 4 belong to the same protein structure family. The contact map of the prediction, Figure 4, which identifies all of the C-alpha intermolecular interactions within 9 Å where the axes are the index of the protein, shows the correct packing of the helices. Figure 5 shows the size of partially correct continuous in sequence segments under an RMSD cutoff. When compared against the other predictions, our model 1 prediction (dark blue) was among the best of all submitted structures. Also the relative success of the prediction classifies this target as being of moderate difficulty. In this example CASP demonstrates that small (70 residues) all-alpha proteins are beginning to be successfully predicted by a variety of ab initio techniques.

Methods

Energy Functions and Sampling. We used an Associative Memory Hamiltonian (AMH) with optimized parameters to sample and predict structures.[17−19] The AMH uses a reduced description of the amino acid chain in order to gain the orders of magnitude computational acceleration over all atom models needed to fold moderate length proteins with ordinary computational resources and has been described in great detail before.[13] This is possible due to reducing the number of atoms per residue from over 10 to only three backbone atoms: the C_α, C_β, and O. The remaining backbone heavy atoms (N, C′) can be reconstituted using the ideal geometry of the peptide bond as a template. Also we reduced the complexity of the amino acid code from 20 letters to four. We chose the four letter code, which has the advantage of preserving a diversity of contacts, because it is still simple enough that the number of coefficients that need to be optimized does not create problems of inaccurate statistics due to limits of interactions encountered in the molten globule state. Specifically four amino acid classes are defined: hydrophilic (A, G, P, S, T), hydrophobic (C, I, L, M, F, W, Y, V), acidic (N,D,Q,E), and basic (R,H,K).[20] The optimization procedure produces an energy landscape that discriminates the native state from misfolded states, while avoiding kinetic traps reasonably well.[2,21] The AMH is an analogue to the neural networks designed by Hopfield to synthesize information from multiple previous experiences.[22] This energy function recalls structural patterns in a set of known protein structures. The Hamiltonian produces an energetically favorable minimum when there is sufficient coherence between a set of three-dimensional protein structures.

The AMH energy function, in its most general sense, contains a backbone term, E_{back}, and interaction term, E_{int}, defined by

$$E_{total} = E_{back} + E_{int} \quad (2)$$

The backbone energy term consists of several terms that reproduce the self-avoiding behavior of the polypeptide chain given by

$$E_{back} = -(E_{SHAKE} + E_{rama} + E_{ev} + E_{chain} + E_{chi}) \quad (3)$$

As in many molecular mechanics energy functions, covalent bonds are preserved by using the SHAKE algorithm[23] E_{SHAKE}, which enables an increase of the time step size and eliminates the need for a traditional harmonic calculation. The SHAKE algorithm preserves the distances between neighboring $C_\alpha-C_\beta$ and $C_\alpha-O$ atoms. The neighboring residues limit the variety of angles the backbone atoms can occupy, producing a Ramachadran plot.[24] This distribution of angles is reinforced by a potential, E_{rama},with low barriers to encourage rapid local backbone movements. Another term, E_{ev}, maintains a sequence specific excluded volume constraint between $C_\alpha-C_\alpha$, $C_\beta-C_\beta$, $O-O$, $C_\alpha-C_\beta$ atoms. The chain connectivity and planarity of the peptide bond due to resonance is ensured by means of a harmonic potential, E_{chain}. Also the chirality of the C_α, due to its four different bonding partners, is maintained using scalar product of neighboring unit vectors of carbon and nitrogen bonds, E_{chi}.

While E_{back} creates peptide-like stereochemistry, it does not introduce the majority of the attractive interactions that result in folding. Such interactions are supplied by the rest

of the potential E_{int}. The interactions described by E_{int} depend on the sequence separation $|i - j|$. Specifically, they are divided into three proximity classes $x(|i - j|)$: $x =$ short ($|i - j| < 5$), $x =$ medium ($5 \leq |i - j| \leq 12$) and $x =$ long ($|i - j| > 12$) as defined by eq 4

$$E_{int} = E_{short} + E_{med} + E_{long} \tag{4}$$

Also these distance classes are also referred to as local, supersecondary, and tertiary, respectively.

The AMH interaction potential E_{int} is based on correlations between a target's sequence signified by i, j, and the sequence-structure patterns in a set of memory proteins μ represented as i', j', and a pairwise contact potential. The pairs in the target and in the memory are first associated using a sequence-structure threading algorithm.[14] The database is assumed to contain a subset of pair distances, which may match the associated pair distances in the target structure. The general form of the associative memory interaction is

$$E_{int} = -\frac{\epsilon}{a}\sum_{\mu}^{N_{mem}} \sum_{j\text{-}12\leq i\leq j-3} \gamma(P_i, P_j, P^{\mu}_{i'}, P^{\mu}_{j'}) \exp\left[-\frac{(r_{ij} - r^{u}_{i'j'})}{2\sigma^2_{ij}}\right] + -\frac{\epsilon}{a}\sum_{k=1}^{3} C_k(N)\gamma(P_i, P_j, k)U_k(r_{ij}) \tag{5}$$

where the similarity between target pair distances r_{ij}, with aligned memory pair distances $r^{u}_{i'j'}$, is measured by Gaussian functions whose widths are given by $\sigma_{ij} = |i - j|^{0.15}$ Å. The set of parameters, γ, encode the similarity between residues i and j and the memories residues i' and j'. Favorable interactions occur during coherence in the distances achieved in the sequence to structure alignments. The encoding of the alignment information in eq 5 is only an example of what is used for the all-alpha energy functions. Other encodings have been used in the alpha−beta energy function[12] to improve the discrimination between helices and strands. While the first term in eq 5 is the superposition of interactions over a set of experimentally determined structures, it also shares a dependence on the sequence separation between the interacting residues. For residues separated by greater than 12 residues, a contact potential E_{long}, as described by the second term in eq 5, which does not depend on interaction information from the structures, is used to define local in sequence interactions. In this term $C_k(N)$ represents a sequence length dependence scaling to account for the variation in probability distributions based on sequence length. Five wells instead of the three defined here by $U_k(r_{ij})$ determine interactions in the alpha−beta energy function.[12] Energy units ϵ are defined excluding backbone contributions in terms of a native state energy in eq 6

$$\epsilon = \frac{|E^{N}_{amh}|}{4N} \tag{6}$$

where N is the number of residues. A distance class scaling a is constant in each of the energy classes because they are designed to be equal during the optimization.

The solvent in these energy functions is treated in a mean field manner, where the implicitly solvated native states of the proteins define the energy gap to the molten globule state. Solvent effects are also present in the sequence to structure alignment energy functions, but they are not explicitly represented in the molecular dynamics energy function. Water mediated contacts with an expanded 20 letter code in the contact potential were introduced,[25] based upon previous work which examined protein recognition.[26,27] The water mediated contacts along with a new one-dimensional burial term has shown promising results especially for long proteins.

Once the energy function is optimized, the minima of the energy function are probed via simulated annealing with molecular dynamics simulations. This minimization technique integrates Newton's equations of motions to determine the energy of the next time step. Simulated annealing slowly reduces the temperature from a high value as in the tempering of steel in metallurgy. This minimization algorithm allows for local searches, while allowing modest energy barriers to be overcome.

Energy landscape ideas have generated an optimization scheme for creating funneled energy surfaces. While funneled, the parametrization does not eliminate all non-native minima. The superposition of several energy surfaces reduces the likelihood of such trapping in local minima.[28,29] The flexibility of the AMH framework provides several ways of incorporating multiple sequence alignment information. Some of the options include creating a consensus sequence,[13] simulating different homologue sequences concurrently, and averaging the resulting forces and energies.[12] The averaged AMH energy function we used averages the forces and the energies of these simulation over a set of sequences, because it allows for more generalizable results than may occur with other techniques, and is described as in eqs 7 and 8

$$E_{short+med} = -\frac{1}{N_{seq}}\frac{\epsilon}{a}\sum_{1}^{seq}\sum_{\mu}^{N_{mem}} \sum_{j\text{-}12\leq i\leq j-3} \gamma(P_i, P_j, P^{\mu}_{i'}, P^{\mu}_{j'} \exp\left[-\frac{(r_{ij} - r^{u}_{i'j'})^2}{2\sigma^2_{ij}}\right] \tag{7}$$

$$E_{long} = -1/N_{seq}\frac{\epsilon}{a}\sum_{1}^{seq}\sum_{k=1}^{3} C_k(N)\gamma(P_i, P_j, k)U_k(r_{ij}) \tag{8}$$

To superimpose multiple energy landscapes, we need a multiple sequence alignment to a set of sequence homologues. Sequences homologous to the target sequence are first identified by using PSI-Blast with default parameters.[30] Each sequence above and below a certain sequence identity thresholds (70% and 30% in this work) is then aligned against each other, and proteins that have greater than 90% sequence identity to other identified sequence homologues are removed. The culling of the sequence homologues via open source bioinformatic libraries is necessary for two reasons.[31] Some classes of proteins have a large number of sequence homologues, and performing a multiple sequence alignment can be impractical. Also removing sequence homologues attempts to remove biases introduced when there are few homologues. The remaining sequences were aligned using

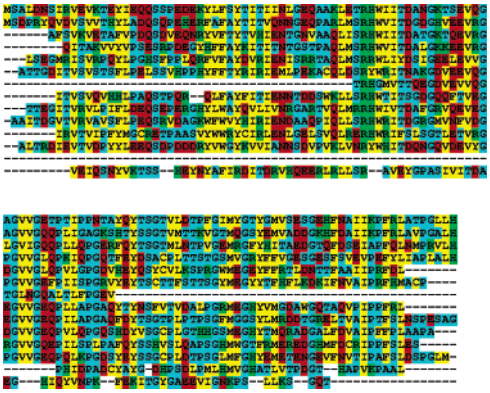


Figure 6. Multiple sequence alignment for target T0212 (PDB 1TZA) colored with respect to a four letter code, where red represents acidic residues, blue represents polar residues, yellow represents nonpolar residues, and green represents basic residues.

a multiple sequence alignment algorithm.[32] Within the AMH energy function, gaps occurring in a sequence alignment could be addressed in a variety of ways, in this work gaps in the target sequence are ignored, while gaps within homologues are completed with residues from the target protein. This strategy may introduce biases toward the target sequence, but this approach is preferred to ignoring interactions. Figure 6 shows a representative multiple sequence alignment for a target, colored with respect to the four letter code of the AMH. If one focuses on the hydrophobic yellow residues, the alternating hydrophobic hydrophilic patterns for beta strands formation are apparent.

Another way of introducing the characteristics of multiple funneled energy landscapes is using information derived from neural networks trained on multiple sequence alignments. Even with different architectures, neural networks typically achieve 75% accuracy when predicting secondary structure. Recently it has been shown that artful combinations of two different predictions can slightly improve the results.[33] This secondary structure information was added by a biasing energy function to either a helix or a strand via, $E_{Q_{ss}} = 10^5\epsilon(Q-Q_{ss})^4$,[13] where Q_{ss} is defined by eq 9

$$Q_{ss} = \sum_{k}^{n} \frac{2}{(N_k - 1)(N_k - 2)} \sum_{i<j-1} \exp\left[-\frac{(r_{ij} - r_{ij}^{ss})^2}{\sigma_{ij}^2}\right] \quad (9)$$

Q_{ss} takes the same form of the Q defined before in eq 1 except that potential acts over n independent secondary structures units derived from secondary structure prediction. The distances that define energy minimum, r_{ij}^{ss}, are determined from experimentally determined Cartesian distances. Previously in an effort to incorporate this secondary structure information, the Ramachandran potential has been altered to bias the backbone.[34] The local in sequence potential $E_{Q_{ss}}$ is preferred to the Ramachandran potential biasing because it avoids SHAKE violations when the strength of the bias is increased.

For most selected CASP6 targets, we followed the same protocol. We averaged the AMH potential over multiple sequence homologues when they were available. In most cases, information from secondary structure prediction was used to bias secondary structure units to their predicted structures. Molecular dynamics with simulated annealing sampled low-energy structures. Also constant temperatures slightly above the predicted glass temperature were used to generate candidate structures. We collected structures above T_K, which usually gives the fastest folding thereby compromising between the funneled and glassy behavior of the energy function. Once the kinetics of the structure slows, the diversity of structures encountered disappears. The slow kinetics regime typically predominates around a temperature of 0.9. While using a linear annealing schedule up to T_K, about 25 different collapsed structures were collected during each simulation. The amount of sampling performed for each structure varied from about 500 to 20 000 different structures. While this was roughly 50 times more sampling than we had previously performed in the CASP setting, it is dwarfed by the efforts of others who can sample in the millions of structures by using more powerful computational resources.[35] Subsequently, a smaller subset of structures was selected for submission by evaluating the size of the hydrophobic core and the hydrophilic surface area. Further selection criteria included visual inspection, agreement with the preliminary secondary structure prediction, and low energies predicted from a second optimized contact energy function.

Selection of Structures

To select candidate structures from independent simulated annealing or constant temperature trajectories, we calculated both the buried hydrophobic surface area and the exposed hydrophilic surface area along the trajectory. In an effort to calculate the buried or exposed surface area, we assigned residues which have greater than the mean total surface area as solvent exposed, and the converse as solvent buried. We scaled each surface area by a weight to represent the likelihood of amino acid burial. It was modeled to the free energy cost of transferring each amino acid from octanol to water[36] in an effort to introduce a sequence specificity as shown in eq 10

$$E_{\text{Burial}} = \sum_{i}^{N} \begin{cases} \gamma_i * \text{SA}_i, \text{ if SA}_i > \text{total surface}/N \\ 0, \text{ if SA}_i < \text{total surface}/N \end{cases} \quad (10)$$

This normalization is desirable because the surface accessibility is calculated from our minimal C_α, C_β, and O atoms, which produces amino acids of the same volume. Such an energy term would be more valuable if nonadditive interactions and a larger number of hydration layers were added. The unavoidable inaccuracies in atomistic force fields and the slow glassy kinetics of side-chain rearrangements prevented any completion of the backbone and side chains with all-atoms or minimization of putative structures.[37]

Another parameter we used after sampling to select and examine structures was based on sequence specific backbone probabilities. The specificity of local interactions has been fruitful for improving collapsed proteins structure predictions.[38] In a similar spirit sequence specific nearest neighbor probabilities were also used.[39] Local signals have also been

Table 1: Linear Regression of Hydrophobic Burial Energy

proteins	fold class	correlation coefficient
1R69	α	0.22
1BG8	α	0.33
1UTG	α	0.63
1MBA	α	0.40
2MHR	α	0.46
1IGD	α/β	−0.70
3IL8	α/β	−0.06
1TIG	α/β	0.02
1BFG	β	0.16
1CKA	β	−0.14
1JV5	β	0.11
1K0S	β	0.27

Table 2: Linear Regression of Mscore

proteins	fold class	correlation coefficient
1R69	α	0.29
1BG8	α	0.04
1UTG	α	0.26
1MBA	α	0.26
2MHR	α	0.10
1IGD	α/β	0.37
3IL8	α/β	0.13
1TIG	α/β	0.19
1BFG	β	0.08
1CKA	β	0.03
1JV5	β	−0.07
1K0S	β	−0.10

theoretically shown to contribute roughly a third of the total folding gap for α helical proteins.[40] Similarly we started looking at such probabilities to further improve the backbone potential of the AMH but without needing secondary structure prediction.

$$E_{\text{trimer}} = \sum_{i=2}^{N-1} \text{LogP}(i-1, i, i+1, \phi, \psi) \qquad (11)$$

Somewhat surprisingly, the summation of the resulting log probabilities from 4012 highly resolved protein structures could be used as an additional measure as part of a strategy for the selection of structures out of an ensemble. Table 1 shows the linear correlation coefficients between structures of varying Q-scores, sampled above T_K which is where the best predictions usually occur before glassy dynamics dominates the kinetics. For both proteins with all α and α/β compositions, the summed log probabilities provide discrimination but not within the all β folds. These results shown in Table 2 echo the previous findings in terms of the ϕ, ψ probability maps and also that all beta structures are less well predicted when a dihedral angle energy function is minimized. The weakness of nearest neighbor excluded volume effects to determine local structure is also demonstrated in the consistent weakness of secondary structure prediction with respect to beta strands. Alpha helices are correctly predicted to roughly 80% accuracy, while beta strands average 60% accuracy by such pure sequence based algorithms. The difficulty of predicting some circular dichroism spectroscopy results for beta to coil transitions can also be attributed to the weakness of the local backbone excluded volume interactions.

Table 3: CASP6 Results: Best Submitted and Sampled Structures

target	length	fold	sub Q	samp Q	temp	traj	CASP
T0281	70	α/β	0.34	0.48	0.85	986	NF
T0201	94	α/β	0.36	0.44	1.39	199	NF
T0212	123	β	0.26	0.42	1.30	97	FR/A
T0230	102	α/β	0.31	0.42	1.05	395	FR/A
T0207	76	α/β	–	0.39	0.98	297	–
T0224	87	α/β	0.30	0.38	1.20	501	FR/H
T0263	97	α/β	0.34	0.38	0.94	404	FR/H
T0272-a	85	α/β	0.30	0.37	0.94	30	FR/A
T0265	102	α/β	0.29	0.34	0.83	374	CM/H
T0213	103	α/β	0.26	0.32	0.98	448	FR/H
T0243	88	α/β	0.31	0.32	0.95	418	FR/H
T0239	98	α/β	0.25	0.32	0.99	424	FR/A
T0214	110	α/β	0.24	0.30	0.41	348	FR/H
T0242	115	α/β	0.27	0.30	0.89	358	NF
T0270-b	125	α/β	0.27	0.28	0.99	32	–
T0270-a	122	α/β	0.25	0.27	0.80	47	–
T0272-b	124	α/β	–	0.26	0.81	34	FR/A
T0273	186	α/β	0.22	0.24	0.98	189	NF

Results

Blind Simulations. For ab initio blind predictions in CASP6, we selected sequences if there were no experimentally determined homologous structures found by automated comparative modeling servers. The overall results for the ab initio structure prediction simulation are summarized in Table 3, where the abbreviations are length = the number of amino acids, temp = temperature where the best structure was encountered, sub Q or samp Q = the best sampled and submitted structures, respectively, as judged by a function of Q, and traj = number of independent trajectories simulated. The CASP6 targets are classified under the following categories (NF=new fold, FR/A=fold recognition analogue, FR/H=fold recognition homologue, CM/H=comparative modeling hard). Targets T0207 and T0270 were removed from the experiment so their CASP class are undefined. Structures for T0207 and T0272-b were not submitted. There are a few main points from these data. Using a Q of 0.4 as a measure of successful prediction, we were able to encounter high quality structures for 4 targets and nearly so for 4 others. The temperature at which the best structures were sampled was between the 1.2 and 0.8, which is the annealing regime we investigated most throughly. This suggests our annealing schedules were close to the behavior we sought a priori. The longer the length of the target sequence clearly reduced the quality of our predictions. Also the proteins where we had a greater number of trajectories naturally showed better structures. A final observation identifies the difference between the best submitted structure and the best sampled structure as disappointingly large for some of the targets. This can be attributed to our strategy of maximizing the number of simulations performed rather than more carefully studying our trajectories. This

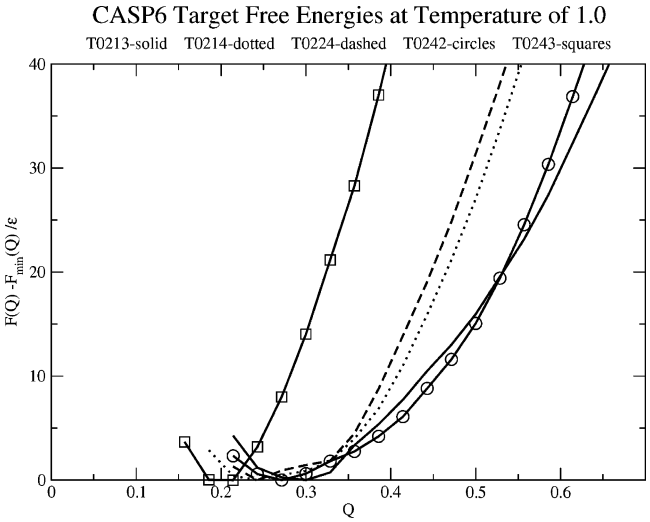


Figure 7. Free energy calculations for CASP6 targets T0213, T0214, T0224, T0242, and T0243.

Table 4: Likely Quality of Structure Seen at a Free Energy of 10 CASP6

target	PDB	length	probable *Q*	sampled *Q*
T0213	1TE7	103	0.43	0.32
T0214	1S04	110	0.40	0.30
T0224	1RHX	87	0.39	0.38
T0242	2BLK	123	0.45	0.30
T0243	–	88	0.28	0.32

difference would be smaller if greater care was taken in the selection of the structures, but the number of high quality structures would have been less.

Calculating the free energy of several randomly chosen CASP6 targets in Figure 7 provides us with probabilities of what we would have expected to see if more simulations had been performed during the CASP season. We can estimate how many independent structures need to be seen at this temperature to sample the region 10 k_BT greater than the minimum of the free energy. We see roughly $e^{10} \approx 2 \times 10^4$ independent sampled structures would be needed at a temperature of 1.0. Target T0242 (PDB ID 2BLK) illustrates why the best structure we encountered had a *Q* score of 0.3. For this target, we sampled roughly 7000 different structures. To achieve a *Q* of 0.45, according to the free energy analysis we would need to increase our sampling by a factor of 3.

When extrapolating to lower temperatures, we see lower barriers to the folded state, and thus if sampling were more complete one would see better structures at these temperatures. This further cooling would be a favorable strategy except that dynamic slowing due to the approach of the glass transition interferes, which occurs at a temperature of 0.9. Naturally, it is best to sample just above the glass transition temperature, which can be approximately found from $Q-Q$ correlation ($< Q(t)Q(t+\tau) >$),[41] and by using the Kolmogorov-Smirnov test to assess the independence of samples.[42] Table 4 indicates what was the best structure we would be likely to see under such sampling conditions. The differences between thermodynamically accessible structures and those that were sampled suggests that increased simulations would have improved the best structures sampled considerably. The free energy of target T0243 (PDB ID not available) is significantly different due to its unusual architecture that contains a buried helix.

As in Figure 4, we compare contact maps between the predictions and the experimentally resolved structure. Often contact maps give more insight than superimposed structures especially when viewing in 2 dimensions. We compare the submitted structures with the best structure encountered during our sampling to determine what aspects of folding are being captured by our energy functions. For a short target T0201 (PDB ID 1S12), we see that sometimes a small difference in the contact maps in Figure 8 can greatly improve the quality of the prediction even though a large number of contacts are already correct. There was a larger fraction of incorrect contacts in our best submitted structure for target T0230 (PDB ID 1WCJ) than we would have seen in the best generated structure as shown in Figure 9. The incorrect parallel docking of the first two helices is largely resolved in the best sampled structure, and the *Q* score improves considerably. Similar analysis for target T0281 (PDB ID 1WHZ) shows incorrect long range contacts between the two otherwise properly oriented helices and disordered intermediate interactions as in Figure 10. Again the best sampled structure has these problems largely resolved.

One amusing way to analyze predicted structures is to view the results of different structure prediction schemes as intermediates along a kinetic folding coordinate. How far did the simulated annealing get in the folding pathway? By mapping the likelihood of folding[43] against its location on a folding free energy surface, we can assess how close the model structure is to the folded state in a kinetic sense. The energy function for the kinetic modeling is a Gō model i.e. ideally nonfrustrated energy function. The difference between the Gō model and the structure prediction energy functions is a measure of the quality of those structure prediction schemes. A pairwise additive Gō model was created based on the native structure of the experimentally determined protein. As it has been discussed previously,[13] this Gō model has both polypeptide backbone energy terms that are the same as in the structure prediction energy function as described by eq 3 and an interaction potential where the Gaussian interaction potential distances r_{ij}^N are determined by the native state formally described in eq 12.

$$E_{G\bar{o}} = -\frac{\epsilon}{a}\sum_{i<j} - 3\gamma_{G\bar{o}}[x_{(|i-j|)}]\exp\left[-\frac{(r_{ij}-r_{ij}^N)^2}{\sigma_{ij}^2}\right] \quad (12)$$

The interactions are defined in this minimal model as residues with greater than three residues in sequence separation between $C^\alpha-C^\alpha$, $C^\alpha-C^\beta$, $C^\beta-C^\alpha$, $C^\beta-C^\beta$ atom pairs. The weights $\gamma_{G\bar{o}}$ or the depth of the Gaussian wells are set to (0.177,0.048,0.430) in order to approximately divided the interaction energy equally between the different distance classes as defined in the original structure prediction energy function. The width of the gaussians is defined by the sequence separation as before. Notice that the Gō Hamilto-

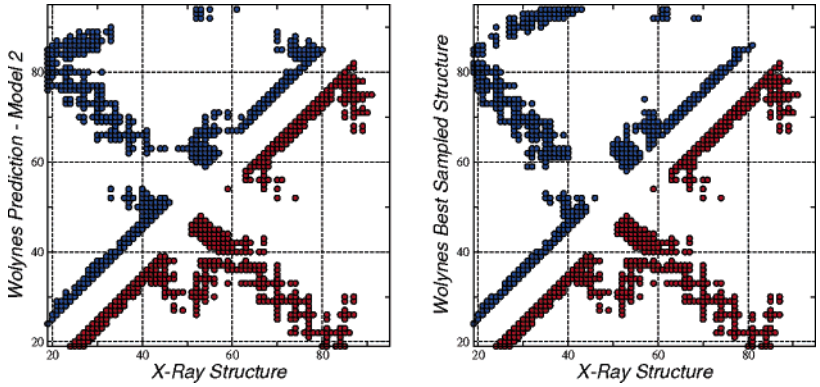


Figure 8. Contact maps for the best submitted (Q=0.36) and the best sampled (Q=0.44) structures for target T0201.

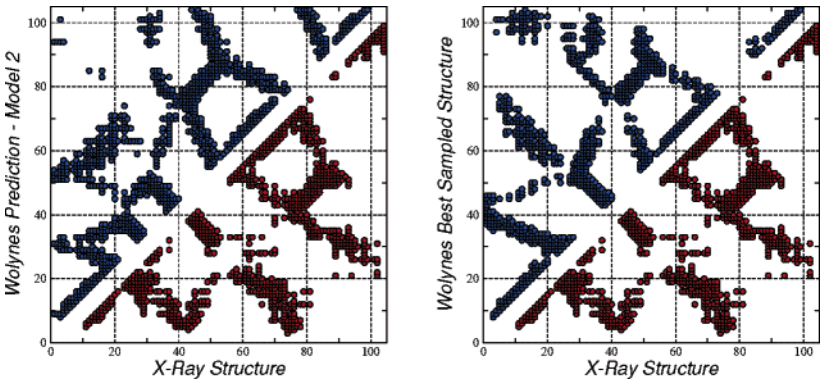


Figure 9. Contact maps for the best submitted (Q=0.31) and the best sampled (Q=0.42) structures for target T0230.

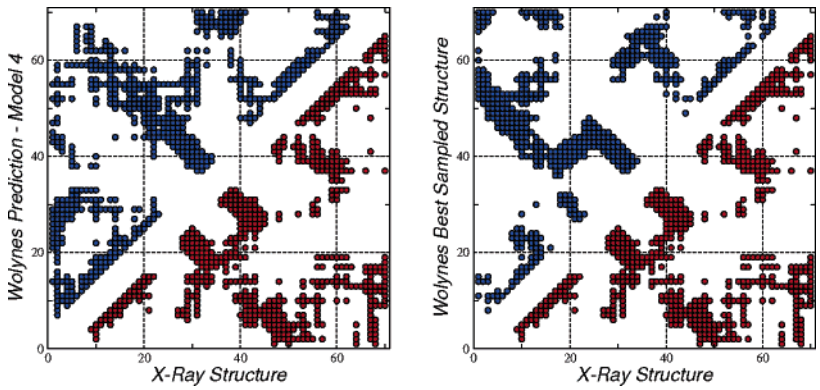


Figure 10. Contact maps for the best submitted (Q=0.34) and the best sampled (Q=0.48) structures for target T0281.

nian does not contain a summation over a set of memory structures as in the AMH; this is because all of the contacts in this definition of a Gō model use only the native state. One hundred independent simulations of this Gō energy function are performed starting with the best structure of three different structure prediction groups. Pfold is then calculated by simply determining whether the simulation started from the model structure folds to the native structure. The results in Figure 11 compare three minimalist models, one of which (the Baker Group) has undergone a further atomistic refinement. The minimalist models are only a few k_BT from the barrier's peak; they only infrequently cross it. It also suggests that a detailed less coarse grain sampling procedure may be necessary for correctly assigning hydrophobic packing and hydrogen bonding patterns.

The Next Generation in Structure Prediction

Examining the contact maps of structures encountered during the CASP experiment, we observed that contacts between residues with a large separation in sequence can be inaccurate, even when most of the contacts within a 12 residues sequence separation are nativelike. A different way of expressing this idea is that the amount of funneling is different within the different distance classes. While this was not used in the recent CASP exercise, we thought it would be interesting and straightforward to improve the prediction energy function by using these first generation results as better memory structures in the AMH. Sequence to structure alignments yield gapless identity alignments thereby eliminating any possibility of secondary structure registry shift irregularities.

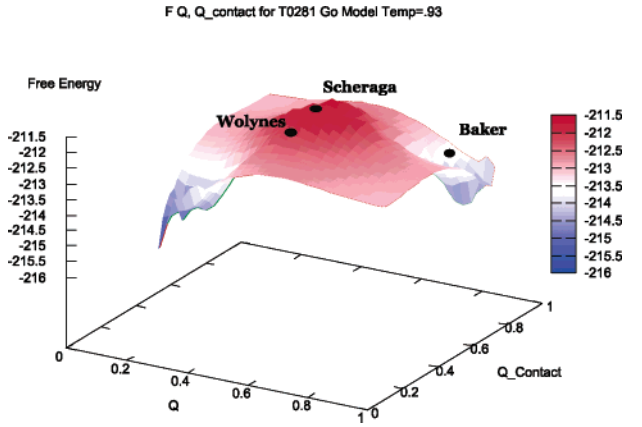


Figure 11. Gō Model free energy surface with final prediction structures shown. The Pfold values for the three proteins are the Wolynes Group 0.07, the Scheraga Group 0.02, and the Baker Group 0.97 with an error of ±0.1.

Different energy functions have been used to identify nativelike proteins from an ensemble of simulated structures. Alternatively, one can rely on energy landscape ideas and assume a mean field contact potential derived from the energy minima of the simulated energy function. This approach has the additional advantage that it does not rely on using a distinct energy function: one is simply seeing how close simulated annealing was to completely accessing the global minimum of the prediction energy function. To select structures a pairwise Q denoted by a lower case q is calculated between all of the ground-state structures encountered in 200 independent simulations.

By dividing the interchain interactions under the same definitions as used in the energy function, the potential for improvements from such second generation structures over the original memories is considerable for protein 256B. As seen in Figure 12, the low-temperature structure as identified by q has an increased amount of nativelike contacts in all distance classes. This style of analysis also suggests potential changes in the energy function. The long distance in sequence interactions are also improved over that original memory used in the energy function. To utilize this improvement the energy function in the distant interaction class was modified. The original function used a multiwell contact potential,

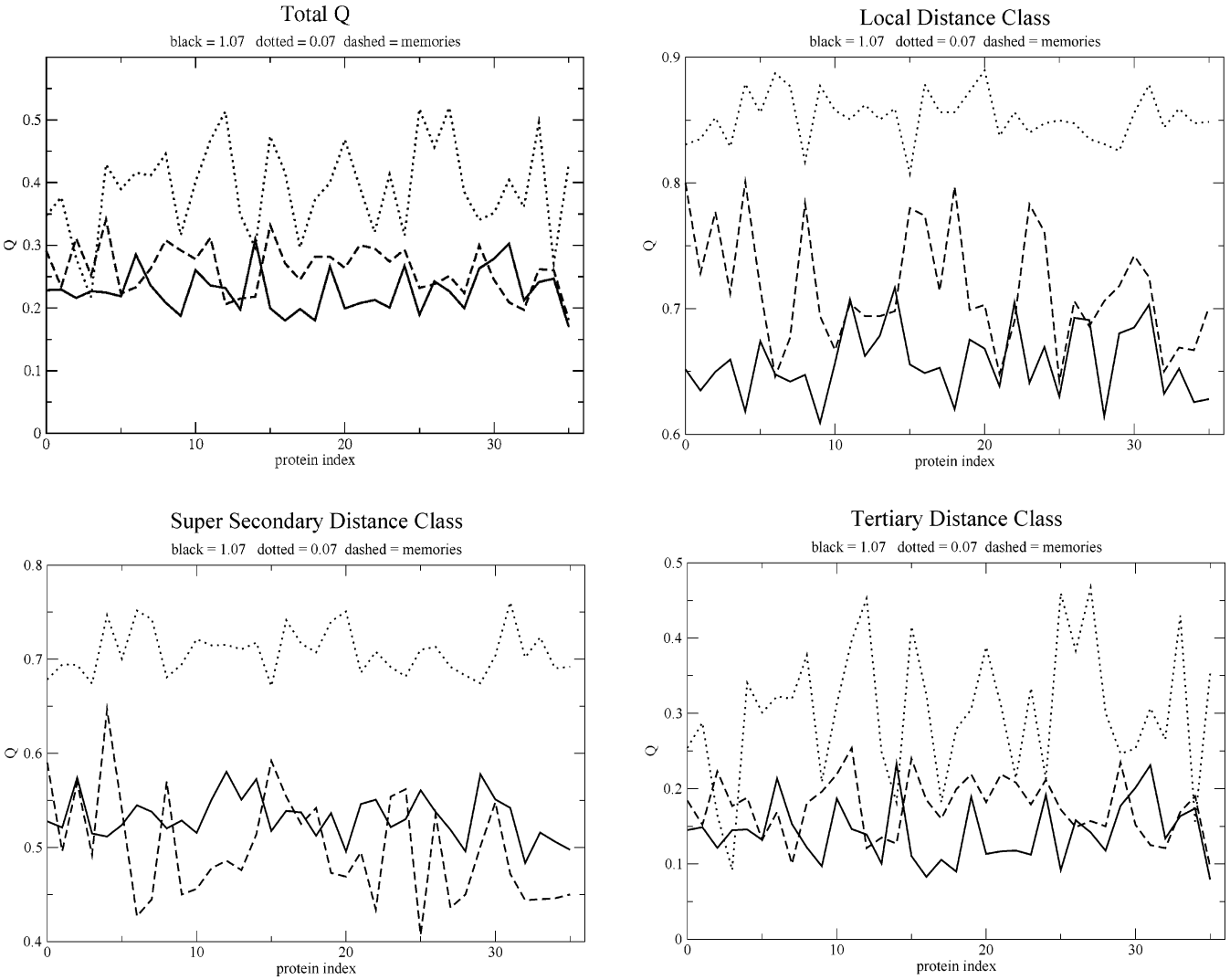


Figure 12. These figures show the total Q and the Q in the different distance classes between PDB structures, structures from a temperature of 1, and a temperature near zero for structures used as inputs to AMH simulations. The lowest temperature show the largest improvement because they are fully collapsed.

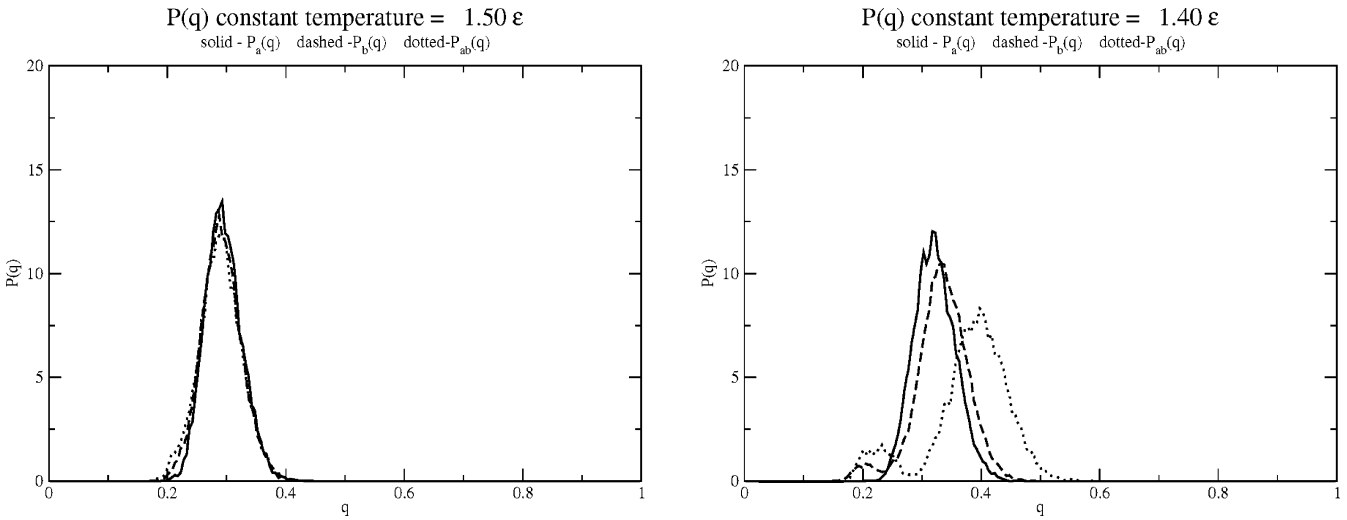


Figure 13. Kolmogorov-Smirnov test shows the constant temperature simulation falling out of equilibrium at a lower temperature of 1.4. The different probability distributions of structures between two independent simulations is no longer the same.

which does not use any information from the memory proteins. For this third distance class the next generation energy function uses associative memory contacts much as was done before for modeling with homologues.[44] The energy function now takes the form

$$E_{\text{int}} = -\sum_{3}^{c}\frac{\epsilon}{a_c}\sum_{\mu}^{n}\sum_{i<j}^{N}\gamma(P_i, P_j, P_{i'}^{\mu}, P_{j'}^{\mu})\Theta(r_{ij} - r_{i'j'}^{\mu}) \quad (13)$$

The parameters for this new distance class are taken from the second distance class. The total energy is defined over the set of memory structures as defined by eq 14

$$\epsilon = \frac{1}{36}\sum_{1}^{\mu}\frac{E_{\text{amh}}^{\text{model}}}{4N} \quad (14)$$

instead of using the values taken from the optimization. Some next generation memory structures are more collapsed than the memory structures used in the initial round of simulation. Furthermore the scaling is changed from the initial round of simulation's 1:1:1 scaling among the three different (local, super-secondary, tertiary) distance classes to 1.5:0.5:1 in an effort to approximate the equal division of energy in each distance class. To examine the equilibrium properties of this energy function, we need to estimate the glass transition temperature. As previously explored,[42] we use the Kolmogorov-Smirnov test to determine if two independent simulations have been sampled from the same equilibrium distribution. This test ensures that simulations are equilibrated.

Once the glass transition temperature (T_K) is estimated using the Kolmogorov-Smirnov test as seen in Figure 13, we can use now standard techniques to quantify the equilibrium proprieties of different energy functions. The proteins we used for study of the next generation AMH strategy are cytochrome B562 (PDB ID 256b) and HDEA (PDB ID 1BG8), because they are both of moderate size and one of them (1BG8) was not in the training set of proteins that optimized the original energy function. An additional advantage of this choice is these proteins have different fold types. According to CATH[45] HDEA belongs to the orthogonal bundle architecture, while cytochrome B562 represents an up−down bundle. Using umbrella sampling combined with the weighted histogramming method, we are able to sample parts of phase space that would rarely be encountered during a simulation.[46] When using memories with a larger number of native contacts, we see improved free energy and energy profiles as shown in Figure 14. This is even more impressive when we consider this energy function has not yet been properly optimized for this new Hamiltonian. For the other target, the results are also not surprising. In this case the next generation memories used to simulate this protein were not of greater structural quality than the initial set. Thus a very similar free energy profile was generated as seen in Figure 15. Our use of q as an order parameter successfully identified the high Q structures for the 256B test case. This is due to the highly funneled characteristic of the first generation energy function. The original energy function for 1BG8 is not as funneled, so therefore there is poorer enrichment by scanning with q. This limitation could be overcome by increasing the amount of sampling of structures in the first generation simulations. More simulations would guarantee better structure as was demonstrated during the CASP5 exercise summarized in Table 3. This difference in the enrichment could be anticipated by using the Kolmogorov-Smirnov measure to differentiate the distribution of the little q values encountered between the memories derived from simulation and the protein databank.

Conclusion

These case studies from our participation in the CASP experiment only provide a snapshot of our group's prediction schemes. It produces a series of lessons for us, and we hope for others. In the future, more balanced efforts between the sampling and selection of structures from that ensemble would appear to be desirable. More efforts in selection would have clearly improved the results submitted in CASP6. While it was computationally impractical to quench all of the structures simulated during the prediction season, the com-

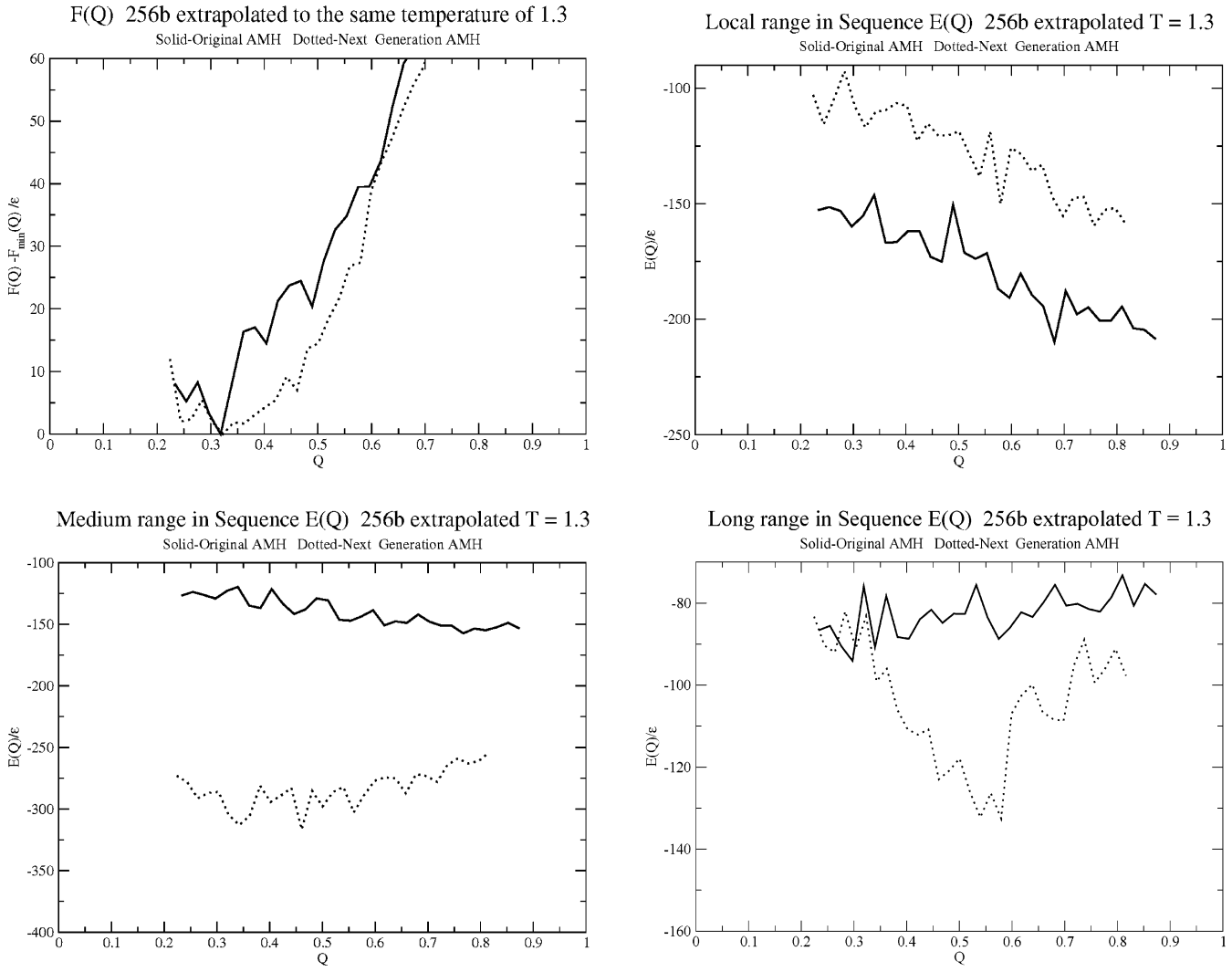


Figure 14. The free energy the two different energy functions for the protein 256B, shows roughly a 5−10 $k_B T$ improvement for this protein. The primary improvements are in the medium- and long-range nonlocal distance classes.

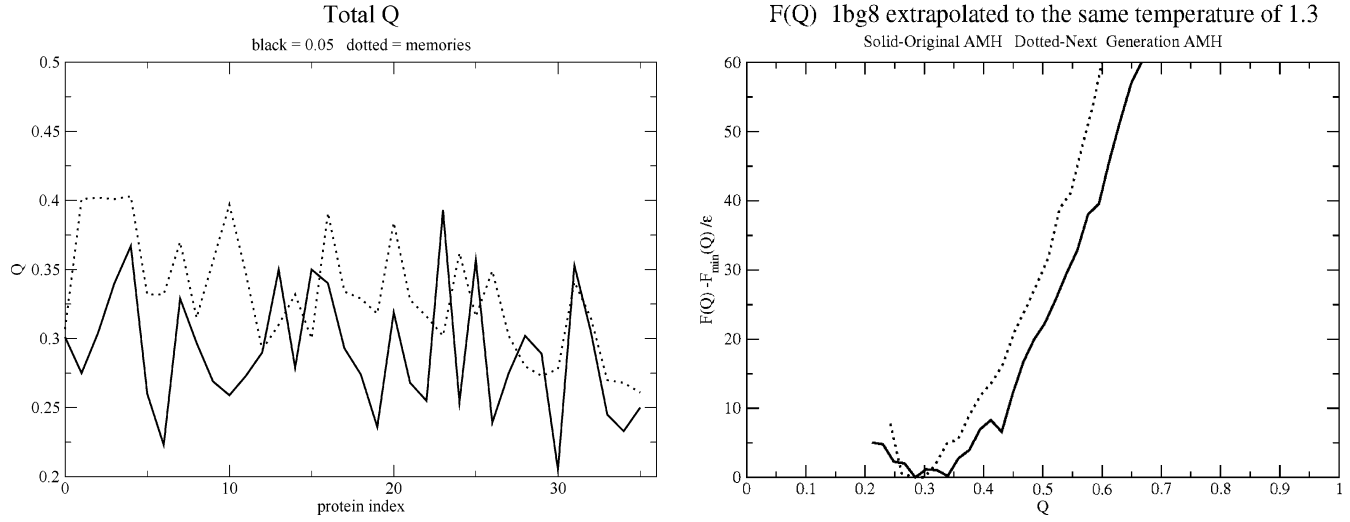


Figure 15. The free energy the two different energy functions for the protein 1BG8 show little improvement. The memories though show no enrichment in native contacts.

parison of the contact maps demonstrated further that tempering of the structure would have improved intermediate range ordering. Using preliminary structures as input to a next generation of AMH modeling improves the quality of the prediction results. While these results may initially appear to be a model or energy function specific, we feel that any algorithm that uses structures as an input would benefit from similar next generation approaches.

Acknowledgment. The authors thank Joe Hegler, Zaida Luthey-Schulten, Garegin Papoian, and Marcio Von Muhlen for their key roles in developing codes used in this study and for many helpful discussions over the years. The efforts of P.G.W. are supported through the National Institutes of Health Grant 5RO1GM44557. Computing resources were supplied by the Center for Theoretical Biological Physics through National Science Foundation Grants PHY0216576 and PHY0225630.

References

(1) Moult, J.; Fidelis, K.; Zemla, A.; Hubbard, T. *Proteins* **2003**, *53 Suppl 6*, 334−339.

(2) Goldstein, R. A.; Luthey-Schulten, Z. A.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 4918−4922.

(3) Bryngelson, J. D.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 7524−7528.

(4) Anfinsen, C. B. *Science* **1973**, *181*, 223−230.

(5) Gō, N. *Annu. Rev. Biophys. Bioeng.* **1983**, *12*, 183−210.

(6) Koga, N.; Takada, S. *J. Mol. Biol.* **2001**, *313*, 171−180.

(7) Portman, J. J.; Takada, S.; Wolynes, P. G. *Phys. Rev. Lett.* **1998**, *81*, 5237−5240.

(8) Wales, D. *Energy Landscapes;* Cambridge University Press: Cambridge, U.K., 2003.

(9) Wheelan, S. J.; Marchler-Bauer, A.; Bryant, S. H. *Bioinformatics* **2000**, *16*, 613−618.

(10) George, R. A.; Heringa, J. *J. Mol. Biol.* **2002**, *316*, 839−851.

(11) Rigden, D. J. *Protein Eng.* **2002**, *15*, 65−77.

(12) Hardin, C.; Eastwood, M.; Prentiss, M.; Luthey-Schulten, Z.; Wolynes, P. G. *Proc. Nat. Acad. Sci. U.S.A.* **2002**, *100*, 1679−1684.

(13) Eastwood, M. P.; Hardin, C.; Luthey-Schulten, Z.; Wolynes, P. G. *IBM Syst. Res.* **2001**, *45*, 475−497.

(14) Koretke, K. K.; Luthey-Schulten, Z.; Wolynes, P. G. *Protein Sci.* **1996**, *5*, 1043−1059.

(15) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucl. Acids Res.* **2000**, *28*, 235−242.

(16) Shindyalov, I.; Bourne, P. *Protein Eng.* **1998**, *11*, 739−747.

(17) Friedrichs, M. S.; Wolynes, P. G. *Science* **1989**, *246*, 371−373.

(18) Friedrichs, M.; Wolynes, P. G. *Tet. Comp. Meth.* **1990**, *3*, 175.

(19) Friedrichs, M. S.; Goldstein, R. A.; Wolynes, P. G. *J. Mol. Biol.* **1991**, *222*, 1013−1034.

(20) Hardin, C.; Eastwood, M.; Luthey-Schulten, Z.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 14235−14240.

(21) Goldstein, R.; Luthey-Schulten, Z. A.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 9029−9033.

(22) Hopfield, J. J. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 2554−2558.

(23) Ryckaert, J.; Ciccotti, G.; Berendsen, H. *J. Comput. Phys.* **1977**, *23*, 327−341.

(24) Ramachandran, G.; Sasisekharan, V. *Adv. Protein Chem.* **1968**, *23*, 283−438.

(25) Papoian, G. A.; Ulander, J.; Eastwood, M. P.; Luthey-Schulten, Z.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 3352−3357.

(26) Papoian, G. A.; Wolynes, P. G. *Biopolymers* **2003**, *68*, 333−349.

(27) Papoian, G. A.; Ulander, J.; Wolynes, P. G. *J. Am. Chem. Soc.* **2003**, *125*, 9170−9178.

(28) Maxfield, F. R.; Scheraga, H. A. *Biochemistry* **1979**, *18*, 697−704.

(29) Finkelstein, A. V. *Phys. Rev. Lett.* **1998**, *80*, 4823−4825.

(30) Altschul, S.; Madden, T.; Schaffer, A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. *Nucl. Acids Res.* **1997**, *25*, 3389−3402.

(31) Stajich, J. E. et al. *Genome Res.* **2002**, *12*, 1611−1618.

(32) Thompson, J.; Higgins, D.; Gibson, T. *Nucl. Acids Res.* **1994**, *22*, 4673−4680.

(33) Zhang, Y.; Kolinski, A.; Skolnick, J. *Biophys. J.* **2003**, *85*, 1145−1164.

(34) Hardin, C.; Eastwood, M.; Prentiss, M.; Luthey-Schulten, Z.; Wolynes, P. G. *J. Comput. Chem.* **2002**, *23*, 138−146.

(35) Bonneau, R.; Tsai, J.; Ruczinski, I.; Chivian, D.; Rohl, C.; Strauss, C. E. M.; Baker, D. *Proteins* **2001**, *Suppl 5*, 119−126.

(36) Zhou, H.; Zhou, Y. *Proteins* **2004**, *54*, 315−322.

(37) Kussell, E.; Shakhnovich, E. I. *Phys. Rev. Lett.* **2002**, *89*, 168101.

(38) Simons, K.; Kooperberg, C.; Huang, E.; Baker, D. *J. Mol. Biol.* **1997**, *268*, 209−225.

(39) Betancourt, M.; Skolnick, J. *J. Mol. Biol.* **2004**, *2*, 635−649.

(40) Saven, J. G.; Wolynes, P. G. *J. Mol. Biol.* **1996**, *257*, 199−216.

(41) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids;* Clarendon Press: New York, U.S.A., 1987.

(42) Eastwood, M.; Hardin, C.; Luthey-Schulten, Z.; Wolynes, P. G. *J. Chem. Phys.* **2003**, *118*, 8500−8512.

(43) Du, R.; Pande, V. S.; Grosberg, A. Y.; Tanaka, T.; Shakhnovich, E. I. *J. Chem. Phys.* **1997**, *108*, 334−350.

(44) Koretke, K. K.; Luthey-Schulten, Z.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 2932−2937.

(45) Pearl, F. et al. *Nucl. Acids Res.* **2005**, *33*, D247−251.

(46) Kong, X.; Brooks, C. L., III *J. Chem. Phys.* **1996**, *105*, 2414−2423.

CT0600058

JCTC Journal of Chemical Theory and Computation

Classical Trajectory-Based Tunneling Splittings: The Hydrogen Atom Transfer in the Hydroperoxyl Anion

Kai Giese and Oliver Kühn*

Institut für Chemie und Biochemie, Physikalische und Theoretische Chemie, Freie Universität Berlin, Takustr. 3, D-14195 Berlin

Received November 28, 2005

Abstract: The hydroperoxyl anion HO_2^- is one of the simplest examples for multidimensional hydrogen-atom transfer. This is reflected in the tunneling splittings, which have been theoretically predicted to be highly selective concerning the excitation of either the O−O stretching, the O−H stretching, or the H−O−O bending vibration. Using the available quantum mechanical spectrum up to 5000 cm^{-1}, we scrutinize the performance of a recently proposed trajectory-based method to calculate tunneling splittings (Giese, K.; Ushiyama, H.; Kühn, O. *Chem. Phys. Lett.* **2003**, *371*, 681). It is found that this new method is capable of reproducing the general behavior of the tunneling splittings rather well. Furthermore, for this particular system, the error shows a systematic trend, which suggests the applicability of a scaling correction that could be based on a small number of exact splittings.

1. Introduction

Tunneling is a ubiquitous phenomenon in reactions involving hydrogen atoms and its isotopes.[1,2] Being a fundamental quantum effect, it continues to provide an inspiration for developing approximate methods for the efficient calculation of tunneling splittings. The semiclassical Wentzel−Kramers−Brillouin (WKB) approximation is not only the oldest but also the most rigorous approach in this respect, and its one-dimensional formulation is well-established.[3] However, multidimensional WKB theory appears to be too demanding to be turned into a practical scheme,[4−7] and quasi-one-dimensional approximations are often used.[8] Degrees of freedom (DOF) orthogonal to a one-dimensional tunneling path are included, for instance, in instanton theory,[9] for which a number of successful implementations have been developed.[10−17] On the other hand, the revival of semiclassical methods in the time domain[18] triggered the development of trajectory-based approaches, for example, within the semiclassical initial value representation[19−21] or by using the concept of multiple spawning.[22]

An even simpler approach which has the advantage of being straightforwardly interfaced with ab initio on-the-fly molecular dynamics goes back to Makri and Miller (MM method;[23] for applications, see also refs 24−27). Here, the recipe for calculating a tunneling splitting consists of running an ensemble of classical trajectories in one of the minima of the potential. Whenever the projection of the momentum onto a *predefined* tunneling direction changes sign, the trajectory is continued as a straight line along the tunneling direction through the classically forbidden region.

Reasonable choices for the tunneling direction depend on the actual potential, for example, on the type of coupling between the reaction coordinate and the considered vibrational mode. Although several possibilities have been discussed,[23−25,28] in particular, straight-line tunneling paths can only perform reasonably if the coupling between the reaction coordinate and vibrational mode is small.[29] Following the classification of the configuration space according to the solutions of the Hamilton−Jacobi equations, the action S can be real (R), imaginary (I), or complex (C) (see, e.g., ref 6 and Figure 1), and considering a simple two-dimensional symmetric mode coupling potential, the MM method based on straight-line paths connecting the minima and parallel to the reaction coordinate will perform best if tunneling goes through the C region. When tunneling mainly goes through the I region, the MM method will fail.

Recently, an extension of the MM method (called EMM) has been proposed,[29] which is based on the concept of tunneling trajectories developed by Takatsuka and co-

* Corresponding author fax: +49083854792; e-mail: ok@chemie.fu-berlin.de.

10.1021/ct0502918 CCC: $33.50
Published on Web 03/18/2006

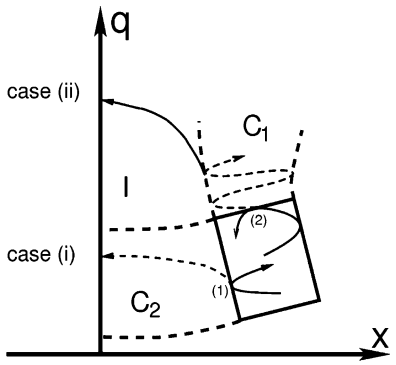


Figure 1. Illustration of the EMM method for a symmetric two-dimensional (x, q) potential and two turning points (1 and 2) leading to tunneling through the C (case i) and I (case ii) regions, respectively.

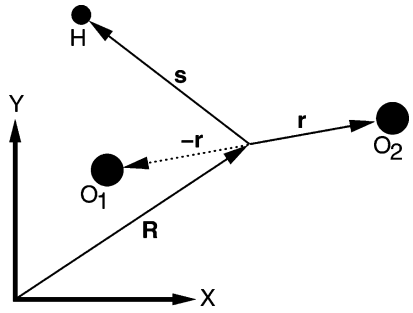


Figure 2. Definition of coordinates used for the hydroperoxyl anion (see text). The geometrical parameters as obtained by Chan and Hamilton using the QCISD(T)/6-311++G(2df,pd) method are $R_{HO} = 0.9619$ Å (1.1491 Å), $R_{OO} = 1.527$ Å (1.6991 Å), and $\gamma = 97.34°$ (42.3°) for the minimum (transition state)[32] (the figure is not to scale). The barrier height for hydrogen atom transfer is 71 kJ/mol (5941 cm^{-1}).

workers.[30] For the case of the above-mentioned two-dimensional symmetric mode coupling potential, it was shown that EMM outperforms MM in the case of strong couplings. This is possible because the tunneling is not restricted to occur along a predefined path. The aim of the present paper is to extend our previous model study and apply the EMM method to a real molecular system, that is, the hydroperoxyl anion, HO_2^-.

Among the nonrigid triatomics of the form HX_2, the hydroperoxyl anion has a rather low barrier for hydrogen-atom transfer and the distances of heavy atoms differ considerably between the C_s minimum and the C_{2v} saddle point geometry (see caption to Figure 2). Therefore, the minimum-energy path has a large curvature, which is an indication of a truly multidimensional transfer process. Chan and Hamilton[31,32] obtained an accurate potential energy surface using the QCISD(T)/6-311++G(2df,pd) level of quantum chemistry. The parameters of the relevant geometries are given in Figure 2. The lowest vibrational eigenstates were obtained by the same authors using a combination of a discrete variable representation and a Lanzcos diagonalization (in the following, denoted QM, see Table 1). The fundamental transitions are at 3575.1 cm^{-1} (O−H stretch), 1066.7 cm^{-1} (H−O−O bend), and 709.8 cm^{-1} (O−O stretch). The tunneling splittings are considerably mode-specific: For instance, for the O−O stretch fundamental transition, the splitting is 2×10^{-4} cm^{-1}, while for a combination excitation with one quanta, respectively, in the O−O stretch and H−O−O bend, it is 1×10^{-2} cm^{-1}. Thus, the hydroperoxyl anion appears to be very well-suited as a test case for the newly developed EMM method, because the tunneling splittings differ by orders of magnitude among different eigenstates.

Table 1. Mode Specific Tunneling Splittings ΔE of the Hydroperoxyl Anion[a]

	$E_n^{\pm}/hc$ (cm^{-1})		Δ_n/hc (cm^{-1})		
state	+	−	QM	MM	EMM
(0,0,0)	0	0		0.000 03	0.000 02
(0,0,1)	709.8	709.8	0.0002	0.000 04	0.0003
(0,1,0)	1066.7	1066.7	0.0009	0.001	0.004
(0,0,2)	1408.7	1408.7	0.001	0.0001	0.002
(0,1,1)	1753.4	1753.4	0.01	0.01	0.03
(0,2,0)	2084.3	2084.3	0.03	0.1	0.2
(0,0,3)	2092.5	2092.5	0.004	0.0007	0.007
(0,1,2)	2426.2	2426.3	0.08	0.1	0.1
(0,2,1)	2745.3	2745.6	0.3	1.2	1.1
(0,0,4)	2756.8	2756.8	0.02	0.005	0.03
(0,3,0)	3063.5	3064.1	0.6	3.4	2.8
(0,1,3)	3081.2	3081.5	0.3	0.9	0.7
(0,2,2)	3388.2	3390.2	2.0	7.1	5.6
(0,0,5)	3398.9	3399.1	0.2	0.06	0.2
(0,3,1)	3692.5	3698.3	5.8	10.3	9.9
(0,1,4)	3714.9	3716.1	1.2	3.9	3.2
(0,4,0)	4000.4	4011.0	10.6	14.2	13.5
(0,2,3)	4010.6	4013.5	2.9	10.4	9.2
(0,0,6)	4017.1	4020.0	2.9	0.3	0.6
(0,3,2)	4285.0	4316.4	31.4	10.1	9.5
(0,1,5)	4323.9	4328.1	4.3	8.9	6.7
(0,4,1)	4562.1	4606.9	44.8	9.9	9.2
(0,2,4)	4600.0	4620.9	20.9	10.8	8.6
(0,0,7)	4608.3	4624.3	16.0	1.6	2.4
(0,5,0)	4826.9	4916.3	89.4	7.7	7.4
(0,3,3)	4895.2	4919.6	24.4	6.1	5.0
(0,1,6)	4907.2	4931.4	24.2	9.5	8.9
(0,1,6)	4907.2	4931.4	24.2	.5	8.9

[a] The values $E_n^{\pm}$ and the quantum mechanical (QM) splittings Δ_n are taken from ref 32.

The paper is organized as follows: In section 2, we will first give a brief review of the EMM method. We continue by discussing the specific implementation of MM and EMM strategies for the case of HO_2^-. The results obtained for both methods are compared to the numerically exact quantum mechanical data in section 3. We summarize in section 4.

2. Theory

2.1. The EMM Method. Before we review the EMM approach, let us briefly recall the MM method.[23] Here, an ensemble of classical trajectories is propagated in the classically allowed region in the neighborhood of a minimum of the potential $V(\mathbf{q})$. If at time t_n the projection of the momentum $\mathbf{p}$ onto a tunneling direction $\mathbf{d}$ changes sign, a turning point is reached and the trajectory is continued as a straight line in the classically forbidden region. Denoting this turning point $\mathbf{q}_n^*$, the tunneling probability for an individual trajectory at time t is given by (in mass-weighted coordinates)

$$P(t) = \sum_{t \le t_n} \exp\left(-\frac{1}{\hbar} \int_0^{\xi_{max}} d\xi \sqrt{2[V(\mathbf{q}_n^* + \xi\mathbf{d}) - V(\mathbf{q}_n^*)]} \right) \quad (1)$$

where ξ_{max} is the maximum length of the straight-line path for which the square root is real. The tunneling splitting is then given by

$$\Delta = 2\hbar \frac{d}{dt}\langle P(t)\rangle \tag{2}$$

where $\langle ... \rangle$ denotes the ensemble average.

While this approach involves an a priori definition of a tunneling direction, which is usually taken as a straight-line path, the EMM method does not have such constraints. Here, the trajectories are propagated, and a parity of motion, $\sigma_j = \pm 1$, is assigned to each DOF such that real-valued momenta can always be defined via[30]

$$\bar{p}_j = p_j/\sqrt{\sigma_j} \tag{3}$$

For $\sigma_j = +1$ and $\sigma_j = -1$, the motion is classically allowed and forbidden, respectively. Note that transformation eq 3 is noncanonical if $\sigma_j = -1$. Inserting eq 3 into the usual Hamiltonian yields

$$\bar{H}(\bar{\mathbf{p}},\mathbf{q};\sigma) = \sum_j \frac{\sigma_j}{2}\bar{p}_j^{\,2} + V(\mathbf{q}) \tag{4}$$

where $\sigma = (\sigma_1,...,\sigma_N)$ is the vector of parities of motion. Trajectories in the $(\bar{\mathbf{p}},\mathbf{q})$ space can be generated by Hamilton's equation of motion

$$\dot{p}_j = -\frac{\partial \bar{H}}{\partial q_j} = -\frac{\partial V}{\partial q_j} \tag{5}$$

$$\dot{q}_j = \frac{\partial \bar{H}}{\partial \bar{p}_j} = \sigma_j \bar{p}_j \tag{6}$$

The first equation, eq 5, is the unchanged Newton's equation of motion. However, the second equation, eq 6, determines that, for $\sigma_j = -1$, the velocity $\dot{q}_j$ and momentum $\bar{p}_j$ have opposite directions. Notice that eqs 5 and 6 are equivalent to the formulation in ref 30 and have a canonically invariant form. This implies that the method of characteristics can be used to construct an action function based on a field of trajectories.[33] However, for the present purpose, we define the action along a trajectory on a σ sheet (for constant energy) by

$$S(\bar{\mathbf{p}}_0,\mathbf{q}_0,\sigma;t) = \sum_j \int_0^t \sqrt{\sigma_j}\,\bar{p}_j(\tau)\,\dot{q}_j(\tau)\,\mathrm{d}\tau \tag{7}$$

where $(\bar{\mathbf{p}}_0,\mathbf{q}_0)$ are the initial conditions of the trajectory. As was noted before, the transformation eq 3 is noncanonical; that is, trajectories generated for different sets of parities σ refer to different dynamical systems, and it is necessary to resort to intuitive arguments in order to connect solutions belonging to different sheets.

The calculation of tunneling splittings in the EMM approach proceeds as follows: (i) Initial conditions for an ensemble in either the left or the right well are generated, for example, from normal-mode sampling.[25] (ii) A trajectory is propagated in the R region of the potential (see Figure 1) until it hits a classical turning point $\mathbf{q}_n^{(cl)}$. In ref 29, it was suggested that these turning points can be defined by the condition $\bar{\mathbf{p}}\cdot\mathbf{d}_i = 0$, where $\{\mathbf{d}_i\}$ is a set of tunneling directions which is defined by the normal-mode vectors of the potential minimum. (iii) Whenever a trajectory encounters such a turning point, the parity σ_j of the corresponding direction $\mathbf{d}_j$ is flipped from $+1$ to -1 and the propagation is continued according to the eqs 5 and 6. The turning point condition is also used to determine *nonclassical* turning points outside the R region: When the turning point condition is fulfilled for a trajectory in the forbidden region, then another parity σ_k of the corresponding direction is flipped from $+1$ to -1. Both trajectories in the forbidden region, the one with $\sigma_j = -1$ and $\sigma_k = +1$ as well as the one with $\sigma_j = -1$ and $\sigma_k = -1$, will be integrated in this case (see Figure 1). The extension to more than two dimensions is straightforward. (iv) When a trajectory that is propagated in the forbidden region crosses the symmetry line Σ, the complex action of that trajectory is computed. The action is given by the integral eq 7 along the trajectory from $\mathbf{q}_n^{(cl)}$ to the crossing point on the symmetry line. There may be several trajectories that emanate from one classical turning point $\mathbf{q}_n^{(cl)}$. Only the action that has the smallest imaginary part is kept, and its absolute value is denoted W_n. (v) The contributions of all nonclassical trajectories with the smallest action that have emanated from classical turning points $\mathbf{q}_n^{(cl)}$ at time t_n are summed up according to

$$P(t) = \sum_{t_n \le t} \exp\left(-\frac{2}{\hbar}W_n\right) \tag{8}$$

from which the splitting is obtained via eq 2. For more details, we refer to the original publication[30] as well as to the earlier paper on the EMM method.[29]

2.2. Implementation for HO_2^-. Both methods, MM and EMM, are formulated in Cartesian coordinates. For MM, this is because the straight-line paths are defined as Cartesian vectors; for EMM, the parities of motion are associated with Cartesian momenta. Recently, Guo and Thompson[28] applied the MM approach to hydrogen peroxide (HOOH) and used the torsional angle as the tunneling coordinate instead of a straight line. This was possible because a single internal coordinate was (assumed to be) responsible for the tautomerization. However, a generalization seems to be hardly possible when more internal non-Cartesian coordinates have to be considered. To apply these methods to a case like the hydroperoxyl anion, it is therefore necessary to describe the molecule in a Cartesian coordinate system.

First, we notice that two rotational DOF can be removed by confining the motion to a plane. For the remaining DOF, the coordinate vectors **R**, **r**, and **s**, as defined in Figure 2, are employed. The positions of the two oxygen atoms, $\mathbf{r_1}$ and $\mathbf{r_2}$, and the position of the hydrogen atom, $\mathbf{r}_3$, can be expressed as

$$\mathbf{r_1} = \mathbf{R} - \mathbf{r} \tag{9}$$

$$\mathbf{r_2} = \mathbf{R} + \mathbf{r} \tag{10}$$

$$\mathbf{r_3} = \mathbf{R} + \mathbf{s} \tag{11}$$

Assuming (without restriction) that the center of mass of the

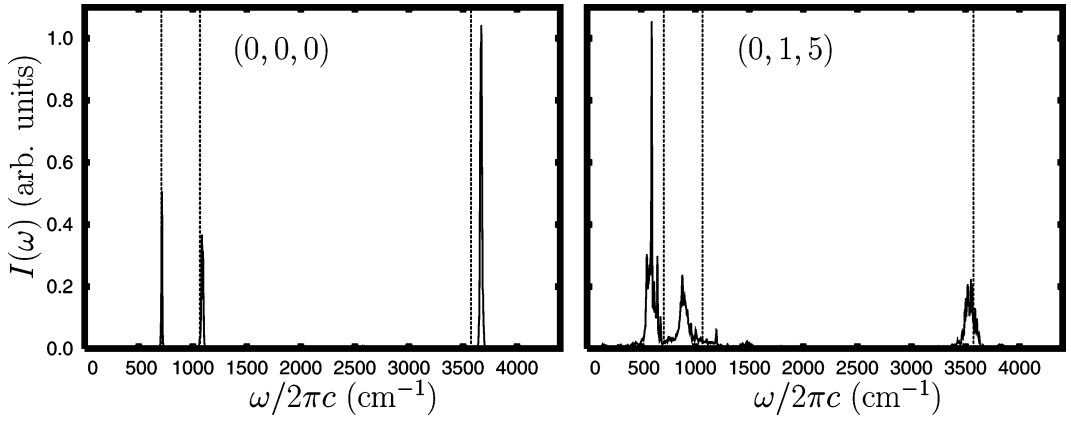


Figure 3. Power spectra $I(\omega)$ (cf. eq 14) for the ground state (0, 0, 0) and excited state (0, 1, 5). The power spectra were obtained by propagating ensembles of 100 trajectories and Fourier transformation of $Z(t) = \sum_i P_i(t)$ with Gaussian broadening $\exp\{-t^2/\tau^2\}$ and $\tau = 5$ ps. The propagation time was 14.5 ps. Fundamental transitions are indicated as dashed lines (cf. Table 1).

triatomic is fixed at the origin, it follows that

$$\mathbf{R} = -\frac{m_H}{M}\mathbf{s} \tag{12}$$

where m_H and M are the mass of a hydrogen and the total mass of the triatomic, respectively. The Lagrangian of the system can be derived straightforwardly as

$$\mathcal{L} = \frac{1}{2}m_H\left(1 - \frac{m_H}{M}\right)\dot{\mathbf{s}}^2 + \frac{1}{2}(2m_O)\dot{\mathbf{r}}^2 - V(\mathbf{s},\mathbf{r}) \tag{13}$$

where m_O is the oxygen mass and $V(\mathbf{s},\mathbf{r})$ is the potential energy surface. The configuration space $(\mathbf{s},\mathbf{r})$ of this Lagrangian is four-dimensional. This means that—because there are three internal coordinates—an individual trajectory is only unique up to an arbitrary rotation around the four-dimensional origin ($\mathbf{s} = 0$, $\mathbf{r} = 0$) corresponding to a rotation around the six-dimensional center of mass.

In the following, the vibrational eigenstates will be labeled as $\mathbf{n} = (n_1, n_2, n_3)$, where the quantum numbers correspond to the O−H stretch (n_1), H−O−O bend (n_2), and O−O stretch (n_3) vibrations. The initial conditions for the ensembles of trajectories corresponding to an eigenstate $\mathbf{n}$ were obtained by normal-mode sampling with energy rescaling,[25] where the energy was chosen to be the mean of the exact quantum mechanical energy for a split pair of levels. The normal-mode analysis with respect to a four-dimensional minimum, $(\mathbf{s}^{(\min)},\mathbf{r}^{(\min)})$, yields three vibrational normal modes plus an infinitesimal rotational vector. The normal modes are only guaranteed to be orthogonal to the infinitesimal rotational vector that corresponds to the minimum; for a displaced geometry, there is a slightly different rotational vector. Therefore, for each trajectory, the angular momentum was removed, and in order to keep the energy fixed, the momenta were rescaled.

To assess the quality of the initial state sampling, we have calculated power spectra of the sum of momenta $Z(t) = \sum_i P_i(t)$:

$$I(\omega) = \frac{1}{2\pi}\lim_{T\to\infty}\frac{1}{T}\left\langle\left|\int_0^T \mathrm{d}t\, Z(t)\exp(-i\omega t)\right|^2\right\rangle \tag{14}$$

where the brackets indicate an averaging over an ensemble of classical trajectories. Exemplary spectra for the ground state (0, 0, 0) at a total energy of 2676 cm^{-1} and state (0, 1, 5) at a total energy of 7000 cm^{-1}, respectively, are shown in Figure 3. The state (0, 1, 5) was chosen as being typical for a highly excited state. In both spectra, one may identify three main peaks that correspond to quasi-periodical motion with the three fundamental frequencies. For the ground-state spectrum, these peaks are narrow, indicating that phase space points on (or very close to) a certain invariant torus are sampled.[34] A comparison with the exact quantum mechanical fundamental transitions (dashed vertical lines) shows that this invariant torus almost coincides with the ground-state invariant torus. Conversely, in the spectrum of state (0, 1, 5), the main peaks are broadened and possess a substructure. The significant shifts of the peak maxima with respect to the quantum mechanical fundamentals are due to the anharmonicity of the potential. Thus, for the state (0, 1, 5), one expects the normal-mode sampling to be less efficient. Nevertheless, the appearance of three distinct peaks indicates that an invariant torus corresponding to the quantum state exists at least approximately.

For the MM method, the tunneling direction is defined such that the O−O distance vector $2\mathbf{r}$ is held fixed for the tunneling process; that is, at each point $(\mathbf{s},\mathbf{r})$, the tunneling direction could be calculated from

$$\mathbf{d}_{MM} = \begin{pmatrix}\mathbf{r}/|\mathbf{r}|\\ 0\end{pmatrix} \tag{15}$$

a choice which was also used for malonaldehyde in refs 24 and 26. If the hydrogen would be located in the left well (i.e., $\mathbf{s}\cdot\mathbf{r} < 0$), then a straight line along $\mathbf{d}_{MM}$ leads automatically to a C_{2v} configuration of the triatomic. This configuration is a point on the symmetry surface Σ that is defined by $\mathbf{s}\cdot\mathbf{r} = 0$ in the present case.

However, the molecule will rotate along a straight line defined by Cartesian directions as given by eq 15. This is because only the hydrogen moves along these straight lines, and the complementary counter rotation of the O−O axis is missing. Because of the mass ratio of about 1:16 between hydrogen and oxygen, the error introduced by the straight-

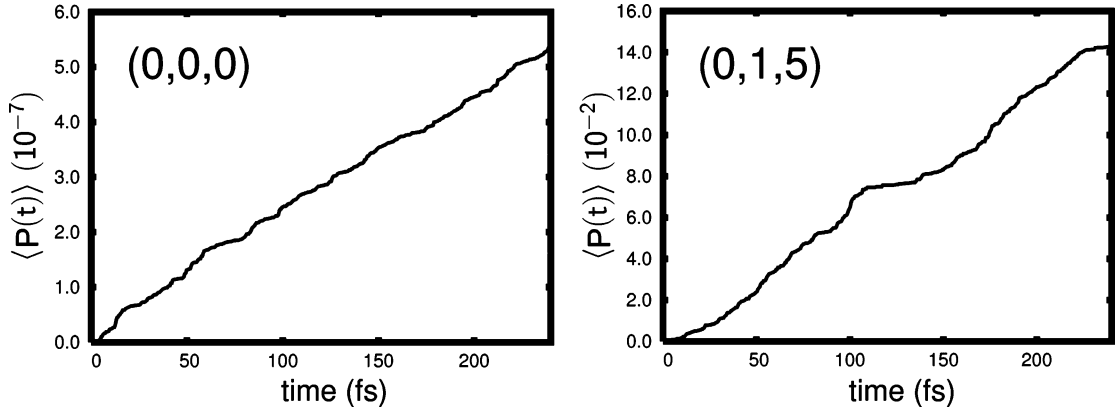


Figure 4. Cumulated EMM tunneling probabilities $\langle P(t)\rangle$ (ensemble of 1000 trajectories) for the ground state (0, 0, 0) and excited state (0, 1, 5).

line approximation may be small, although its magnitude is unknown a priori. However, a rotation-free tunneling direction and also a related rotation-free tunneling path can be constructed. To this end, let us consider a modified tunneling direction that is given by

$$\mathbf{d}_{\mathrm{MM}}^{\mathrm{R}} = \frac{\mathbf{d}_{\mathrm{MM}} - (\mathbf{d}_{\mathrm{MM}}\rho)\rho}{|\mathbf{d}_{\mathrm{MM}} - (\mathbf{d}_{\mathrm{MM}}\rho)\rho|} \tag{16}$$

where $\rho = \rho(\mathbf{s}_1,\mathbf{r}_1)$ is the infinitesimal rotational vector at the (turning) point $(\mathbf{s}_1,\mathbf{r}_1)$. With definition eq 16, the rotation in the vicinity of the turning point is removed. A complete rotation-free path may be obtained by solving the ordinary differential equation:

$$\frac{\mathrm{d}\mathbf{X}}{\mathrm{d}\xi} = \mathbf{d}_{\mathrm{MM}}^{\mathrm{R}}[\mathbf{X}(\xi)] \tag{17}$$

where $\mathbf{X}(\xi)$ is a 4D vector as a function of the coordinate ξ along the curved path. The tunneling direction eq 16 is considered to be **X**-dependent, such that ρ is evaluated at each point on the path while $\mathbf{d}_{\mathrm{MM}}$ is held fixed.

However, it turned out that the correction to the tunneling splittings due to the rotation-free tunneling path, eq 17, is negligible for the present molecule. Therefore, the MM tunneling integral, eq 1, was computed by following a straight-line tunneling path defined by the rotation-free direction eq 16, and turning points are given by the zeros of $\mathbf{P}\cdot\mathbf{d}_{\mathrm{MM}}^{\mathrm{R}}$, where **P** is the four-dimensional momentum.

In analogy to eq 15, the EMM turning point condition was modified such that, for each point (**s**,**r**), the set of vectors $\{\rho,\mathbf{d}_1,\mathbf{d}_2,\mathbf{d}_3\}$ is Schmidt-orthogonalized. Here, $\mathbf{d}_i$ corresponds to the normal-mode direction of the three fundamental vibrations at the minimum. The procedure guarantees that, for a turning point $\mathbf{P}\cdot\mathbf{d}_i' = 0$ with Schmidt-orthogonalized normal-mode vector $\mathbf{d}_i'$, the parity of motion is flipped to -1 for a direction that is orthogonal to the rotational direction at that turning point. The generalized trajectories are then generated for a reduced space that is spanned by the three Schmidt-orthogonalized normal-mode directions $\mathbf{d}_i'$. Notice that this is a generalization of the condition used for the MM tunneling directions and turning points. In particular, the straight line generated by direction eq 16 can be expressed as a linear combination of the $\mathbf{d}_i'$ values. The integration of a trajectory is stopped when the symmetry surface Σ has been reached, that is, when $\mathbf{s}\cdot\mathbf{r} \geq 0$.

For each vibrational level, **n** ensembles of 1000 classical trajectories were propagated using a fifth-order symplectic integrator[35] with a fixed step size of 0.48 fs. For the accurate determination of turning points, the step size was reduced to 0.048 fs. The propagation time was 240 fs, which corresponds to about 5 times the vibrational period of the weakest mode. The energy was bound within a typical maximum tolerance of 0.1 cm^{-1}.

In Figure 4, cumulated EMM tunneling probabilities $\langle P(t)\rangle$ for, respectively, the ground state (0, 0, 0) and state (0, 1, 5) are shown, which correspond to the power spectra in Figure 3. Compared to the ground state, there is an oscillation superimposed on the approximate linear increase of $\langle P(t)\rangle$ for state (0, 1, 5). This can be attributed to the sampling method; that is, the trajectories are only near the invariant torus.[25] Only those states are considered for which the linear regression of the data points yields a relative standard deviation of less than 1%. Notice that this excludes the states $(1, n_2, n_3)$ which show a larger deviation. The sampling error was estimated by performing five distinct runs for the (0, 0, 1) level; the relative standard deviation was 8% of the tunnel splitting.

3. Results

In Table 1, tunneling splittings obtained by diagonalizing the Hamiltonian (QM data taken from ref 32) are compared with semiclassical results obtained within, respectively, the MM and EMM approaches. When comparing the MM results for, for example, the fundamental of the O−O vibration (0, 0, 1) with those for the QM, one finds an underestimation of almost 1 order of magnitude. Conversely, the EMM result for that vibrational state agrees quite well with the QM. The same finding applies to overtones of the O−O vibration up to state (0, 0, 5). For the fundamental of the H−O−O bend vibration (0, 1, 0), the tunneling splittings are overestimated by the EMM. The same applies for states (0, 2, 0) and (0, 3, 0). Also, the MM yields too-large tunneling splittings for the two states (0, 2, 0) and (0, 3, 0). Finally, we note that, for the MM method, we have investigated the rotational error due to the use of a tunneling path which is not rotation-free for the lowest 10 states in Table 1. It turns out that, for the

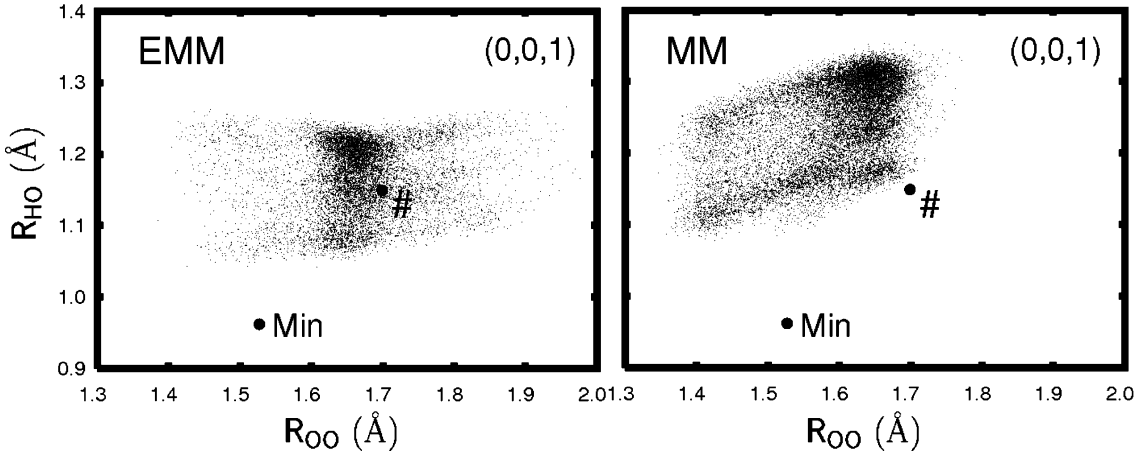


Figure 5. Point cloud plot of the tunneling probability, $\exp(-2W_n/\hbar)$, at the symmetry surface (C_{2v} configuration), where W_n is the absolute value of the imaginary part of the action obtained for a path from the nth turning point to the point (R_{HO}, R_{OO}) on the symmetry surface. The local density of points at (R_{HO}, R_{OO}) is chosen to be proportional to $\langle\sum_n (W_n)^{-4}\rangle$. The saddle point is indicated (#); the minimum is not a point on the symmetry surface, but the minimum bond length values are given for comparison and marked by "Min".

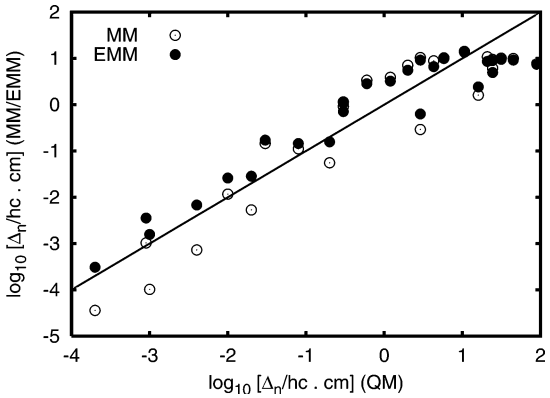


Figure 6. Mode-specific tunneling splittings Δ_n (logarithmic). Values obtained by MM and EMM vs the exact QM result. The solid line corresponds to perfect agreement.

reported number of digits, only for state (0,0,3) does the tunneling splitting change (to 0.0008 cm^{-1}).

In Figure 5 the magnitude of the tunneling action W_n at the symmetry surface Σ is visualized for the state (0, 0, 1) by point-cloud plots for EMM and MM. For visual clarity, the local density of points has been chosen to be proportional to $\sum_n W_n^{-4}(R_{HO}, R_{OO})$: a higher density corresponds to smaller action and, thus, a larger tunneling probability. For state (0, 0, 1), the maximum of the EMM and MM is at (R_{HO}, R_{OO}) $\approx$ (1.2, 1.65) Å and (1.3, 1.65) Å, respectively. Thus, the EMM prefers tunneling at shorter R_{HO} values that are closer to the transition state. This is clearly the reason for the larger tunneling probabilities, leading in turn to a larger tunneling splitting.

To visualize the performance of the EMM as compared to that of the MM more globally, the semiclassical results are plotted versus the QM results in Figure 6 on a double logarithmic scale. The EMM results do systematically overestimate the QM results for splittings smaller than about 2 cm^{-1}. The correlation between MM and QM seems to be more erratic than that for EMM and QM. However, the absolute error of the log values (concerning all values shown in Figure 6) is similar, namely, 0.55 and 0.45 for MM and EMM, respectively. To show that there is a significant improvement concerning the EMM results, the log values for all splittings $\leq$ 2 cm^{-1} were fitted by a linear function, $f(x) = A + Bx$, with $x = \Delta$(QM). For MM and EMM, it was found that (A, B) = (0.5 $\pm$ 0.2, 1.3 $\pm$ 0.1) and (0.4 $\pm$ 0.1, 1.04 $\pm$ 0.05), respectively. For an exact correlation, one would expect (0, 1). Thus, the difference between the EMM and QM results is mainly the constant factor $10^{0.4} \approx 2.5$. This suggests immediately that a *calibration* of the results would be reasonable if certain exact splittings are known. As far as the MM splittings are concerned, there is a significant deviation from slope 1, and the results cannot be corrected by simply introducing a constant factor. For states with splittings larger than about 2 cm^{-1}, there are also underestimations by EMM. These states are rather highly excited; that is, the normal-mode sampling is less efficient, which is most likely a major source of error. Moreover, the quantum spectrum is dominated by a 2:3 Fermi resonance between the OH bend and the OO vibration.[31] This leads to a perturbation of the spectra that cannot be accounted for by the present semiclassical methods.

Finally, the results are analyzed with respect to the number of parities that were flipped. Figure 7 shows EMM tunneling splittings for selected fundamentals and overtones (cf. Table 1) and tunneling splittings that were obtained by keeping only those W_n values in the sum eq 8 for which there were, respectively, one ("1"), two ("2"), or three ("3") negative parities when the generalized trajectory crossed the symmetry surface Σ; the sum ("sum") corresponds to the values given in Table 1. The case of three negative parities is irrelevant. Surprisingly, both cases with one and two parities contribute equal amounts (in a logarithmic sense). For the two-dimensional symmetric coupling potential, the MM tunneling direction $\mathbf{d} = (-1, 0)^t$ becomes equivalent to one of the normal modes for vanishing coupling.[29] For the hydroperoxyl anion, however, the normal modes do not coincide with the definition of the straight line direction, eq 16. Thus, the reason for the better performance of the EMM for the

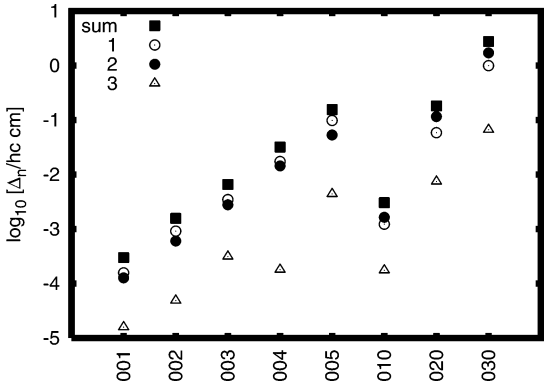


Figure 7. Mode-specific tunneling splittings Δ_n (logarithmic) obtained by the EMM: values for tunneling with one ("1"), two ("2"), and three ("3") parities and the sum ("sum") (cf. Table 1) are shown.

hydroperoxyl anion can be attributed to the more physical treatment of the connection condition of the semiclassical wave function across the caustic surfaces separating the regions of real, complex, and imaginary actions.

4. Conclusions

The recently proposed extension of the Makri−Miller model[29] has been applied to calculate tunneling splittings of vibrational modes for a three-dimensional ab initio potential energy surface describing the hydrogen-atom transfer in the hydroperoxyl anion.[32] For moderate excitation energies, it was found that the results of the extended Makri−Miller model systematically overestimate the quantum mechanical values, while the conventional Makri−Miller results show over- and underestimations in an erratic manner. It was pointed out that a calibration of the extended Makri−Miller method results by a constant factor is reasonable. It should be stressed that both methods rely on several ad hoc assumptions that can only be justified by numerical tests, as in this study. The present findings suggest that the extended Makri−Miller method gives a more realistic description of the connection of the semiclassical wave function across the different regions of configuration space. However, whether this feature is more general remains to be shown. Furthermore, it turned out that tunneling with one and two parity changes is almost equally important (in a logarithmic sense). This observation should be of high relevance for current attempts to use generalized trajectories for time-dependent wave-packet dynamics.[36]

Finally, the present method has to be brought into the context of ab initio molecular dynamics simulations with an on-the-fly calculation of the forces which has been used in connection with the Makri−Miller model.[26] Because the extended Makri−Miller method is trajectory-based as well, there seems to be no principal problem with an implementation of an ab initio molecular dynamics scheme. However, it should be emphasized that the benefits of this method come at the price of an increased numerical complexity. First, significantly more tunneling trajectories need to be calculated, and several of them are subsequently discarded. Second, the solution of Hamilton's equation of motion for the parity-based trajectories will typically require a smaller step size than the calculation of the potential energy along a straight line. Nevertheless, the present study shows that this new protocol may deserve further investigation.

Acknowledgment. This work was supported by the Deutsche Forschungsgemeinschaft (Project Ma 515/19-2). We would like to thank Dr. W.-T. Chan (Toronto) for providing the Fortran code for the potential energy surface, Prof. K. Takatsuka and Dr. H. Ushiyama (Tokyo) for helpful discussions, and Prof. J. Manz (Berlin) for his continued support of this work.

Note Added after ASAP Publication. This article was released ASAP on March 18, 2006, with the incorrect Received Date. The correct version was posted on April 12, 2006.

References

(1) Bell, R. P. *The Tunnel Effect in Chemistry*; Chapman and Hall: London, 1980.

(2) Benderskii, V. A.; Makarov, D. E.; Wright, C. A. *Adv. Chem. Phys.* **1994**, *88*, 1.

(3) Landau, L. D.; Lifshitz, E. M. *Quantum Mechanics*; Pergamon: Oxford, 1975.

(4) Delos, J. B. *Adv. Chem. Phys.* **1985**, *65*, 161.

(5) Huang, Z. H.; Feuchtwang, T. E.; Cutler, P. H.; Kazes, E. *Phys. Rev. A* **1990**, *41*, 32.

(6) Takada, S.; Nakamura, H. *J. Chem. Phys.* **1994**, *100*, 98.

(7) Takada, S.; Nakamura, H. *J. Chem. Phys.* **1995**, *102*, 3977.

(8) Tautermann, C. S.; Voegele, A. F.; Loerting, T.; Liedl, K. R. *J. Chem. Phys.* **2002**, *117*, 1967.

(9) Kleinert, H. *Path Integrals in Quantum Mechanics, Statistics, Polymer Physics, and Financial Markets*, 3rd ed.; World Scientific: New Jersey, 2004.

(10) Benderskii, V. A.; Grebenshchikov, S. Y.; Mil'nikov, G. V.; Vetoshkin, E. V. *Chem. Phys.* **1994**, *188*, 19.

(11) Benderskii, V.; Vetoshkin, E.; Irgibaeva, I.; Trommsdorff, H. *Chem. Phys.* **2000**, *262*, 393.

(12) Smedarchina, Z.; Fernandez-Ramos, A.; Siebrand, W. *J. Comput. Chem.* **2001**, *22*, 787.

(13) Smedarchina, Z.; Fernandez-Ramos, A.; Siebrand, W. *Chem. Phys. Lett.* **2004**, *395*, 339.

(14) Mil'nikov, G.; Nakamura, H. *J. Chem. Phys.* **2001**, *115*, 6881.

(15) Mil'nikov, G. V.; Yagi, K.; Taketsugu, T.; Nakamura, H.; Hirao, K. *J. Chem. Phys.* **2003**, *119*, 10.

(16) Mil'nikov, G. V.; Nakamura, H. *J. Chem. Phys.* **2005**, *122*, 124311.

(17) Mil'nikov, G. V.; Kühn, O.; Nakamura, H. *J. Chem. Phys.* **2005**, *123*, 074308.

(18) Miller, W. H. *J. Phys. Chem. A* **2001**, *105*, 2942.

(19) Mandelshtam, V. A.; Ovchinikov, M. *J. Chem. Phys.* **1998**, *108*, 9206.

(20) Shalashilin, D. V.; Child, M. S. *J. Chem. Phys.* **2001**, *114*, 9296.

(21) Giese, K.; Kühn, O. *J. Chem. Phys.* **2004**, *120*, 4207.

(22) Ben-Nun, M.; Martinez, T. J. *J. Chem. Phys.* **2000**, *112*, 6113.

(23) Makri, N.; Miller, W. H. *J. Chem. Phys.* **1989**, *91*, 4026.

(24) Sewell, T. D.; Guo, Y.; Thompson, D. L. *J. Chem. Phys.* **1995**, *103*, 8557.

(25) Guo, Y.; Li, S.; Thompson, D. L. *J. Chem. Phys.* **1997**, *107*, 2853.

(26) Ben-Nun, M.; Martinez, T. J. *J. Phys. Chem. A* **1999**, *103*, 6055.

(27) Guallar, V.; Gherman, B. F.; Miller, W. H.; Lippard, S. J.; Friesner, R. A. *J. Am. Chem. Soc.* **2002**, *124*, 3377.

(28) Guo, Y.; Thompson, D. L. *J. Phys. Chem. A* **2002**, *106*, 8374.

(29) Giese, K.; Ushiyama, H.; Kühn, O. *Chem. Phys. Lett.* **2003**, *371*, 681.

(30) Takatsuka, K.; Ushiyama, H.; Inoue-Ushiyama, A. *Phys. Rep.* **1999**, *322*, 347.

(31) Chan, W.-T.; Hamilton, I. P. *J. Chem. Phys.* **1996**, *105*, 5907.

(32) Chan, W.-T.; Hamilton, I. P. *Chem. Phys. Lett.* **1998**, *292*, 57.

(33) Maslov, V. P.; Fedoriuk, M. V. *Semiclassical Approximations in Quantum Mechanics*; D. Reidel: Dordrecht, The Netherlands, 1981.

(34) Gutzwiller, M. C. *Chaos in Classical and Quantum Mechanics*; Springer-Verlag: New York, 1990.

(35) Gray, S. K.; Noid, D. W.; Sumpter, B. G. *J. Chem. Phys.* **1994**, *101*, 4062.

(36) Ushiyama, H.; Takatsuka, K. *J. Chem. Phys.* **2004**, *120*, 4561.

CT0502918

JCTC Journal of Chemical Theory and Computation

Distributed Replica Sampling

Tomas Rodinger,[†,‡,§] P. Lynne Howell,[†,‡] and Régis Pomès*[,†,‡]

Structural Biology and Biochemistry, The Hospital for Sick Children, 555 University Ave., Toronto, Ontario, Canada M5G 1X8, Department of Biochemistry, University of Toronto, 1 King's College Circle, Toronto, Ontario, Canada M5S 1A8, and Institute of Biomaterials and Biomedical Engineering, University of Toronto, 4 Taddle Creek Road, Rm. 407, Rosebrugh Building, Toronto, Ontario, Canada M5S 3G9

Received December 1, 2005

Abstract: We present a simple and general scheme for efficient Boltzmann sampling of conformational space by computer simulation. Multiple replicas of the system differing in temperature T or reaction coordinate λ are simulated independently. In addition, occasional stochastic moves of individual replicas in T or λ space are considered one at a time on the basis of a generalized Hamiltonian containing an extra potential energy term or bias that depends on the distribution of all replicas. The algorithm is inherently suited for shared or heterogeneous computing platforms such as a distributed network.

Introduction

Despite rapid advances in computer technology, statistical sampling of systems governed by rugged potential energy surfaces remains a challenge, which underscores the need for efficient computational algorithms.[1−6] For example, the time scales of conformational transitions of biomolecular systems span many orders of magnitude. Many of these events are not directly accessible by atomistic computer simulations, precluding Boltzmann sampling of conformational space.

A number of statistical mechanical methods using generalized ensembles have been developed in recent years to enhance sampling efficiency. Some of these approaches are designed to induce a random walk either in temperature, potential energy, spatial coordinates, or transformation parameters in the Hamiltonian. In simulated tempering (ST),[7,8] a random walk in temperature is achieved by taking periodic steps from absolute inverse temperature β_i to β_j subject to acceptance of the following Metropolis Monte Carlo test:

$$p_{\text{accept}} = \min(1, \exp\{-[(\beta_j - \beta_i)E(\mathbf{q}) + (a_j - a_i)]\}) \quad (1)$$

where β_i and β_j are the inverse of the Boltzmann constant multiplied by absolute temperature, the potential energy $E(\mathbf{q})$ is a function of atomic coordinates $\mathbf{q}$, and a_i and a_j are weight factors that must be determined in advance using an iterative procedure. Similarly, the multicanonical algorithm (MUCA)[9,10] results in a random walk in potential energy space. Extending the Hamiltonian to include a transformation coordinate λ is another way to get around barriers because the higher dimensionality of phase space results in an increased number of alternate routes by which barriers can be avoided. A random walk along a transformation coordinate can be realized via adaptive umbrella sampling (AUS) methods,[11−15] where the Hamiltonian governing movement is defined as

$$H(\mathbf{q},\lambda) = E(\mathbf{q},\lambda) + U(\lambda) \quad (2)$$

In eq 2, the potential energy E is a function of atomic coordinates $\mathbf{q}$ and transformation coordinate λ, and $U(\lambda)$ is a biasing potential that must be adapted such that it removes any barriers that exist along the transformation coordinate. In ST, MUCA, or AUS, random walk behavior is achieved only once the system "adapts" or learns the correct weight factors for movement along the parameter of interest, a system-dependent procedure that introduces extra complexity as well as possible artifacts if not done carefully (see ref 16

* Corresponding author tel.: 416-813-5686; fax: 416-813-5022; e-mail: pomes@sickkids.ca.

† The Hospital for Sick Children.

‡ Department of Biochemistry, University of Toronto.

§ Institute of Biomaterials and Biomedical Engineering, University of Toronto.

10.1021/ct050302x CCC: $33.50
Published on Web 03/22/2006

for more details). Furthermore, although adaptive methods clearly improve sampling uniformity along the coordinate, in some cases, uniform sampling may only be achieved in very long simulations because of ruggedness in orthogonal degrees of freedom.[17]

In the past decade, replica exchange (RE),[18−20] also known as multiple Markov chain or parallel tempering, has emerged as a powerful alternative to adaptive methods. In RE, multiple noninteracting copies or replicas of a system, each governed by the same potential energy function but differing in temperature, are simulated at once. Periodically, the simulations are halted, and replicas i and j with neighboring temperatures T_i and T_j are swapped with a probability given by the following Metropolis Monte Carlo condition:

$$p_{\text{accept}} = \min(1, \exp\{-[\beta_i - \beta_j][E(\mathbf{q}_j) - E(\mathbf{q}_i)]\}) \quad (3)$$

Temperature swaps cause each replica to undergo a random walk in temperature, providing a means (for any replica) to escape from a potential energy trap. Like adaptive methods, RE may also be used to attain a random walk in a parameter of the Hamiltonian,[21−23] λ. Swaps between adjacent replicas i and j are accepted with a probability given by

$$p_{\text{accept}} = \min(1, \exp\{-\beta[E(\mathbf{q}_i,\lambda_j) - E(\mathbf{q}_i,\lambda_i) + E(\mathbf{q}_j,\lambda_i) - E(\mathbf{q}_j,\lambda_j)]\}) \quad (4)$$

When data from all replicas are taken together, perfect sampling uniformity along the temperature or reaction coordinate is attained. Furthermore, no adaptation of a biasing potential is required because the weight factors are known a priori. In the realm of biomolecular studies, large-scale applications of RE have pushed back limitations in the scope of problems and the size of systems studied.[19,23−31] Recent studies have applied the method to folding a 46 amino acid protein domain,[32] and the approach has been extended to two-dimensional random walks in pressure and temperature.[31] The principal drawback of RE is that it requires a homogeneous and often large computing cluster to implement efficiently. Because the number of replicas required follows the square root of the number of degrees of freedom in the system,[18] the simulation of complex systems eventually becomes impractical.

Here, we present distributed replica (DR) sampling, a simple and general scheme for efficient Boltzmann sampling of conformational space by computer simulations. As in RE, multiple replicas of the system, covering a preassigned range in temperature or reaction coordinate space, are simulated independently. The main difference with RE is that, instead of performing pairwise exchanges of replicas, stochastic moves of individual replicas in T or λ space are considered one at a time. The target distribution of replicas is enforced by a generalized Hamiltonian containing an extra potential energy bias that depends on the distribution of all replicas. Contrary to adaptive methods, DR does not require preliminary determination of the bias. We show with a simple but relevant example that the DR algorithm leads to random-walk movement along the parameter of interest with an efficiency comparable to that of RE. Because it avoids the need for all replicas to run synchronously, DR is intrinsically suited for implementation on a shared or inhomogeneous computing platform such as a large-scale distributed network.

Method

Let us consider N noninteracting copies (or "replicas") of a system of interest, all of which are governed by an identical potential energy function, $E(\mathbf{q}_i,\lambda_i)$, where $\mathbf{q}_i$ represents the atomic coordinates of the atoms in replica i and λ_i is the coupling parameter for the reaction coordinate of interest (the reaction in question may be either an alchemical or a spatial transformation). The DR method makes use of an additional potential energy term, $D(\lambda_1,\lambda_2,...,\lambda_N)$, henceforth referred to as the distributed replica potential energy (DRPE), which enforces the distribution of replicas across the range of the transformation coordinate (i.e., an energy penalty is associated with a nonideal distribution). The generalized Hamiltonian for state $X = \{\mathbf{q}_1,\lambda_1,\mathbf{q}_2,\lambda_2,...,\mathbf{q}_N,\lambda_N\}$, combining all replicas together with the DRPE, is given by

$$\mathscr{H}_\lambda(X) = \sum_{i=1}^{N} E(\mathbf{q}_i,\lambda_i) + D(\lambda_1,\lambda_2,...,\lambda_N) \quad (5)$$

The weight factor for state X is given by

$$W(X) = \exp[-\beta\, \mathscr{H}_\lambda(X)] \quad (6)$$

We consider one λ move at a time. Suppose that the λ value of replica m is to be changed from λ_m to $\lambda_m + \delta\lambda_m$, thus taking state X to state X':

$$X = \{\mathbf{q}_1,\lambda_1,...,\mathbf{q}_m,\lambda_m,...,\mathbf{q}_N,\lambda_N\} \rightarrow X' = \{\mathbf{q}_1,\lambda_1,...,\mathbf{q}_m,\lambda_m + \delta\lambda_m,...,\mathbf{q}_N,\lambda_N\} \quad (7)$$

In order for the exchange process to converge toward the equilibrium distribution, it is sufficient to impose the detailed balance condition on the transition probability $p(X\rightarrow X')$:

$$W(X)\, p(X\rightarrow X') = W(X')\, p(X'\rightarrow X) \quad (8)$$

From eqs 5, 6, and 8, we have

$$\frac{p(X\rightarrow X')}{p(X'\rightarrow X)} = \exp(-\beta\Delta) \quad (9)$$

where

$$\Delta = E(\mathbf{q}_m,\lambda_m + \delta\lambda_m) - E(\mathbf{q}_m,\lambda_m) + D(\lambda_1,\lambda_2,...,\lambda_m + \delta\lambda_m,...,\lambda_N) - D(\lambda_1,\lambda_2,...,\lambda_m,...,\lambda_N) \quad (10)$$

This can be satisfied using the Metropolis Monte Carlo criterion:

$$p_{\text{accept}} = \min[1, \exp(-\beta\Delta)] \quad (11)$$

Like RE, the DR method may also be performed in temperature space. To this end, we first define $X = \{\mathbf{q}_1,\beta_1,...,\mathbf{q}_N,\beta_N\}$. The corresponding dimensionless generalized Hamiltonian is

$$\mathscr{H}_\beta^*(X) = \sum_{i=1}^{N} \beta_i\, E(\mathbf{q}_i) + D^*(\beta_1,\beta_2,...,\beta_N) \quad (12)$$

Table 1. Example of a Distributed Replica Potential Energy Calculation

index	1	2	3	4	5	6
nominal (or ideal) λ positions of replicas	0.0	0.1	0.2	0.4	0.6	1.0
current λ_i positions of replicas (after several λ moves and one effective λ swap)	0.0	0.2	0.11	0.4	0.6	1.0
$\lambda_{i,\text{sorted}}$ values after step 1 (sorting)	0.0	0.11	0.2	0.4	0.6	1.0
$\lambda_{i,\text{unit}}$ values after step 2 (spacing transformation)	1.0	2.1	3.0	4.0	5.0	6.0
DRPE result after step 3 (assuming $c = 0.1$ kcal/mol)	0.1 kcal/mol					

where the asterisk indicates a dimensionless quantity. The appropriate Monte Carlo acceptance probability of a move involving a change $\delta\beta_m$ in the inverse temperature of replica m is

$$p_{\text{accept}} = \min[1, \exp(-\Delta^*)] \tag{13}$$

where

$$\Delta^* = \delta\beta_m E(\mathbf{q}_m) + D^*(\beta_1,\beta_2,...,\beta_m + \delta\beta_m,...,\beta_N) - D^*(\beta_1,\beta_2,...,\beta_m,...,\beta_N) \tag{14}$$

Many possibilities exist for calculating the DRPE, and several variations were tested as a part of the development of this method. One that was found to work very well, which was used in the analysis reported here, is as follows. The DRPE function is calculated from three algorithmic steps: (1) The λ (or β) values for all replicas are sorted in ascending order. The following holds true for the new order, $\lambda_{i,\text{sorted}}$:

$$\lambda_{i,\text{sorted}} > \lambda_{i-1,\text{sorted}} \text{ for } i = 2 \text{ to } N \tag{15}$$

(2) The spacing system is transformed to a uniform unit spacing arrangement to give $\lambda_{i,\text{unit}}$:

$$\lambda_{i,\text{unit}} = f^{-1}(\lambda_{i,\text{sorted}}) \tag{16}$$

where f^{-1} is the inverse of a function f, which maps the replica index to the nominal λ value of that replica (i.e., the λ position where the replica started). Note that because replica indices are integers, f is constructed by linearly interpolating between adjacent points.

(3) This step involves the following equation:

$$D = c\sum_{i=1}^{N}\sum_{j=1}^{N}[(\lambda_{i,\text{unit}} - \lambda_{j,\text{unit}}) - (i - j)]^2 \tag{17}$$

where c is a parameter that scales the DRPE function for the purpose of adjusting the move acceptance probability. Note that the DRPE must necessarily be a state function (i.e., it must conserve energy). See Table 1 for an example DRPE calculation.

Note that the DRPE defined in eqs 15−17 acts to reinforce spacing between the λ_i values of the replicas, and it has no effect on the absolute positions of the λ_i values. Therefore, to prevent a concerted drift of all replicas away from the region of interest, a few extra nonmoving, nonsimulated dummy replicas can be included, typically positioned just beyond the endpoints.

A DR simulation is realized as follows. Initially, each replica i is created at a different position, λ_i, spanning the transformation coordinate, and is optionally equilibrated. The spacing between adjacent λ_i values is chosen on the basis of the application at hand and may be uniform or nonuniform. The following two steps are then iterated: (1) Each replica is run as an independent molecular dynamics or Monte Carlo simulation for a set period of time or number of steps, typically on its own CPU in a computing cluster. The method does not require all of the simulations to run simultaneously nor does it require simulations to finish (or be halted) in a coordinated manner as is typical in standard replica exchange algorithms. λ_i values are fixed during the course of these simulations. (2) Periodically, one replica is considered for a λ move. The probability of acceptance is $p(X \rightarrow X')$ (see eq 11).

There are no restrictions on the intervals between replica move attempts, although some optimal interval will exist for a given application. Frequent move attempts allow a greater mobility of the λ_i values, but at the cost of increased overhead. The distance by which λ_i changes during a λ move is also not restricted and can be optimized for the application at hand.

Test Application

The performance of the DR method is illustrated with a two-dimensional model system designed such that a significant energy barrier exists in a degree of freedom x orthogonal to a reaction coordinate λ. As depicted in Figure 1, the reaction involves the conversion of a double-well potential to a single well. A total of 51 replicas spaced uniformly along the reaction coordinate, each initially at $x = -1$, were constructed. For each replica, the particle evolved on the 1D energy landscape governed by Monte Carlo moves at a temperature of 298 K with a fixed step size of 0.01. A total of 10^7 moves along x were simulated per replica. The calculation was performed with fully independent replicas, with DR, and with RE. After 10 steps elapsed for all DR replicas (20 steps for RE), a λ move (or a swap between neighbors) was attempted for one replica (or one replica pair) chosen at random, giving the particles a chance to move in the λ dimension. When either DR or RE was used, each replica experienced on average about 19 600 move attempts along λ (one swap event in RE is equivalent to two λ moves in DR).

Sequential moves of DR replicas in λ space are illustrated in Figure 2. An effective random walk is achieved by individual replicas over the entire simulation (Figure 2b). The occasional rapid replacement of one replica by another is tantamount to a pairwise exchange reminiscent of RE (Figure 2a). However, the fact that DR replicas move one at

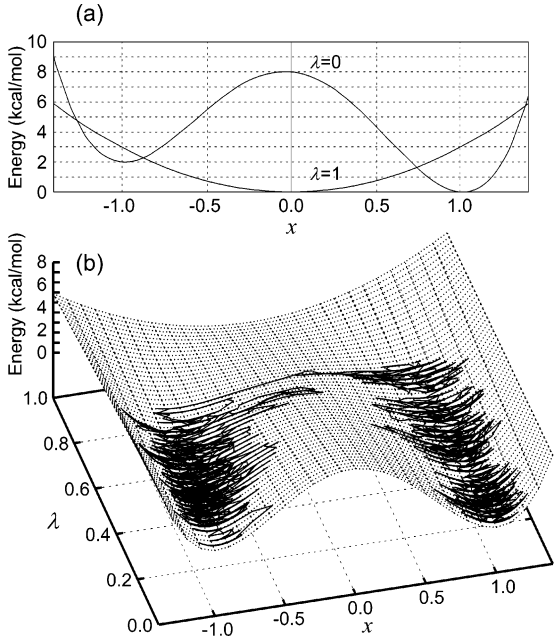


Figure 1. Model potential energy surface used to test DR sampling. (a) The energy profile gradually switches from a bistable well at $\lambda = 0$ (with minima at $x \approx \pm 1$) to a single well at $\lambda = 1$ and $x = 0$. (b) Two-dimensional representation; a DR trajectory illustrating barrier avoidance by diffusion in λ space is also shown.

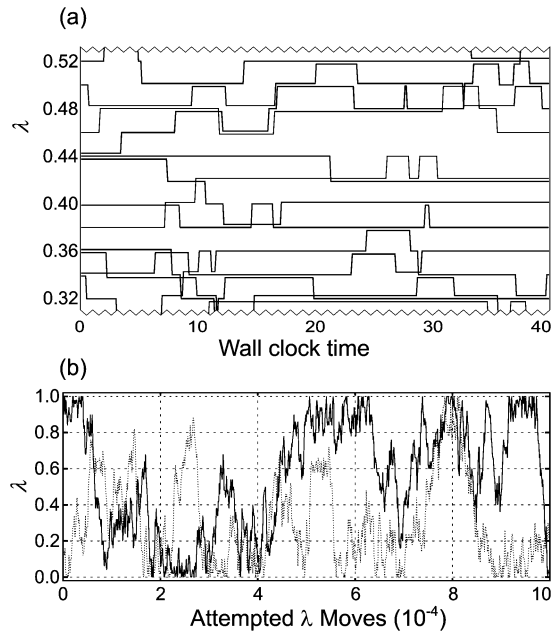


Figure 2. Diffusion of distributed replicas in λ space. (a) Representative trajectories of a few DR replicas over a small interval of wall-clock time corresponding to approximately 40 move attempts. Lines were slightly shifted vertically for clarity. (b) Two representative complete trajectories illustrate a random walk in λ.

a time also leads to overlaps of two or more replicas and temporary gaps in λ space. The DRPE function acts to enforce spreading of the replicas in λ space. On average, sampling is close to the target uniform distribution (Figure 3a). There is a nonlinear relationship between the root-mean-square deviation from uniformity and the acceptance ratio

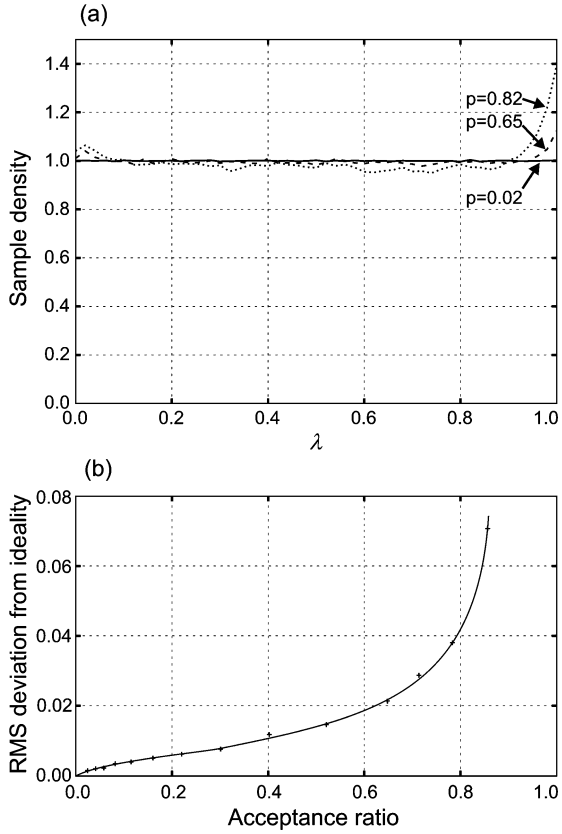


Figure 3. Relationship between sampling profiles and acceptance ratios in DR simulations. (a) Sampling profiles along the reaction coordinate λ obtained from three DR simulations, each with 51 replicas but differing in acceptance ratio, $p =$ 0.02, 0.65, and 0.82. (b) Root-mean-square deviation from sampling uniformity as a function of acceptance ratio. The lines are included to guide the eye.

(Figure 3b). Both of these quantities are controlled by the stiffness of the DRPE function (i.e., by the scaling constant c in eq 17). While sampling uniformity is achieved only for an infinitely forbidding DRPE ($c \rightarrow \infty$), deviations from the target distribution only become significant for relatively large acceptance ratios (loose DRPE). With intermediate acceptance ratios (i.e., 20−70%), the DR method nearly achieves the desired sampling profile.

In the limit of independent replicas, severe deviations from the theoretical sampling distribution along x are evident at small values of λ (Figure 4). For some replicas, the energy barrier prevents crossings altogether during the entire simulation, trapping the particle in the metastable well at $x = -1$. Considerably longer simulations would be necessary to achieve statistical convergence. The energy barrier can be bypassed if the system is allowed to evolve freely on pathways through λ space as well as x. The sampling distributions obtained using DR (with an acceptance ratio of 65%) and RE simulations are shown in Figure 4c,d. Both cases show dramatic improvements in sampling efficiency over independent replicas (Figure 4b) and good agreement with the theoretical distribution (Figure 4a). Accordingly, these two simulations yield mean force and free energy profiles within the statistical error from each other (Figure 5), indicating that DR is as efficient as RE in the given test

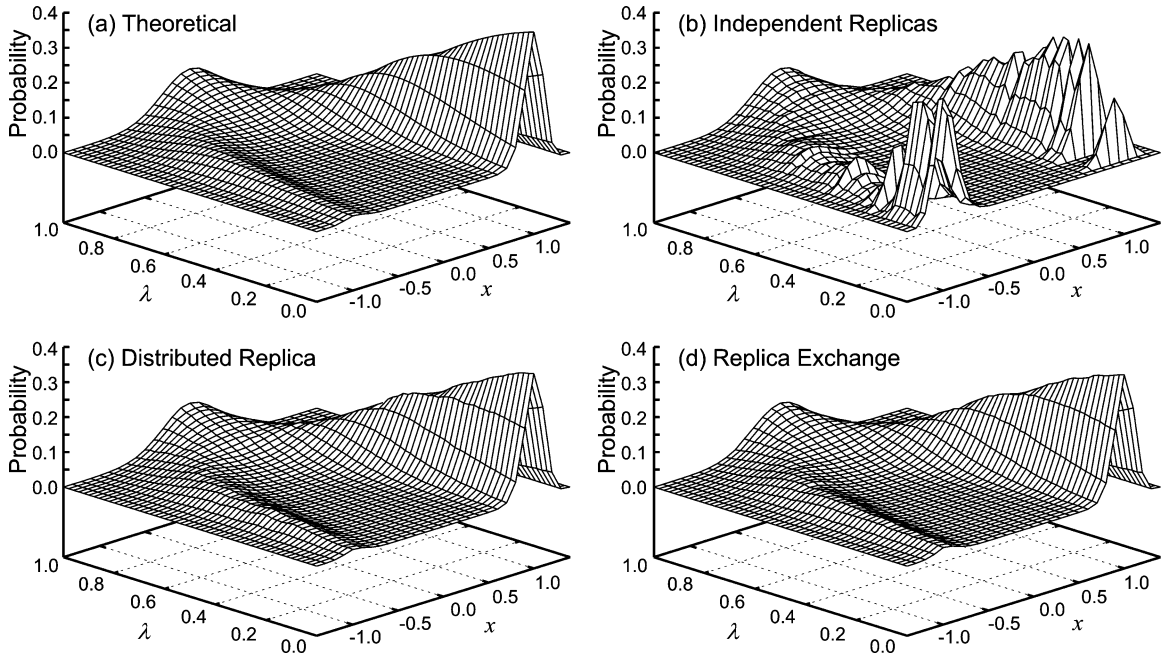


Figure 4. Sampling distribution for the model system depicted in Figure 1: exact (analytical) distribution (a) and distributions obtained with 51 replicas spanning the range $\lambda = 0-1$, successively from independent replicas (no λ move) (b), from DR with a 65% acceptance ratio (c), and from RE (d).

case. Importantly, the equally good agreement with the exact solution also shows that DR is as accurate as RE.

The effectiveness of the DR method is compromised at both low and high acceptance ratios. If the ratio is low, the systematic sampling error is large because of poor sampling along x (as is illustrated by the results obtained with independent replicas; see Figures 4b and 5). If the acceptance is high (soft DRPE), then the replicas accumulate in regions of lower energy (i.e., at the endpoints in the present case) and statistical sampling errors along λ grow due to the fact that high-lying regions do not get sampled effectively. Thus, it is desirable to strike a balance between the acceptance ratio and the deviation from the target distribution. In the present case, an acceptance ratio as high as 65% leads to acceptable deviations from sampling uniformity and to accurate results (see Figures 4 and 5). By comparison, it should be noted that 96% acceptance is achieved in the RE simulations on the same set of replicas. By construction, RE never departs from sampling uniformity; the desired acceptance ratio is modulated instead by trial simulations with different numbers and spacings of replicas. With DR, the step size taken by λ during a replica move can be optimized for the application at hand. This can be done either a priori or on the fly. Step sizes can be any distance and can be made as small as required in order to decrease the system energy penalty and increase the acceptance probability.

Although this is not apparent in the simple model system tested here, another advantage of DR over RE is expected in applications where a large energy penalty is associated with λ moves (such as in simulations involving atomic displacements in dense media). Because the exchange of two replicas is associated with two separate conformational energy penalties (one for each replica), the probability of accepting a move is the product of two small probabilities. This ultimately leads to low mobility in λ space and, thus,

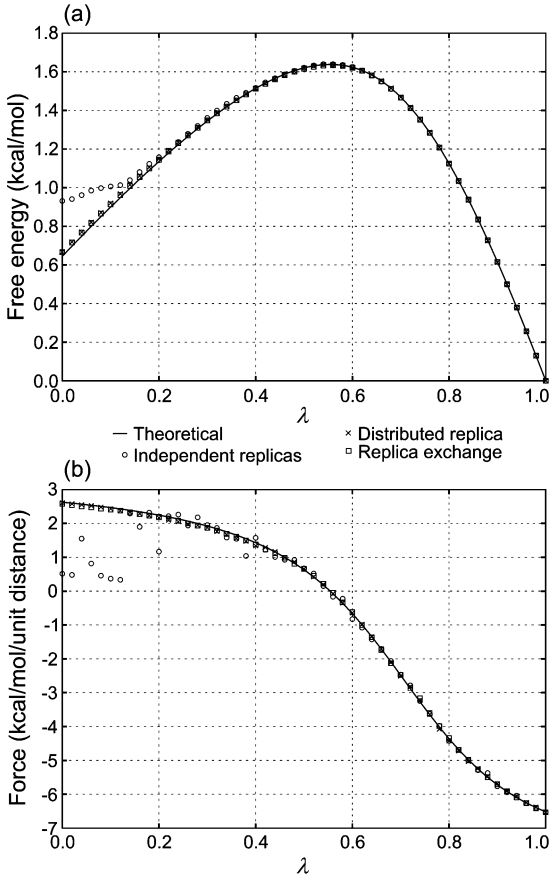


Figure 5. Potential of mean force (a) and mean force (b) acting on the particle along the reaction coordinate. The theoretical curves are shown with continuous lines; results obtained from independent replicas (circles), from distributed replica sampling (crosses), and from replica exchange (squares) are also shown. In the latter two cases, the size of the symbols is larger than the statistical sampling error.

poor sampling efficiency. The energy penalty for an exchange can be reduced by decreasing the spacing between nominal λ values of the replicas and, thus, increasing the number of replicas that span the reaction coordinate. However, smaller separation distances between replicas also hamper coordinate mobility. In contrast, replica moves in the DR method are associated with a single conformational energy penalty because, although the DRPE is analogous to a potential energy, it is not generally a function of system complexity. Furthermore, the severity of this penalty can be adjusted to some extent through the DRPE scaling constant c. For this reason, the sampling efficiency in DR is expected to be better than that in RE as the system complexity increases, although only in applications involving a spatial reaction coordinate and a dense medium. This advantage does not exist in temperature space.

Concluding Discussion

We report a simple yet powerful simulation technique that can be used in a range of applications utilizing both physical and alchemical transformations as well as in exhaustive sampling where a random walk in temperature is key. The method is of comparable efficiency to that of RE. Because it shares essential characteristics with RE, DR could be used with similar extensions. For example, it could be combined advantageously with adaptive-type methods such as ST and MUCA.[24−28]

The essential advantage of DR over RE lies in the ease of implementation on computer platforms. Traditional RE methods require dedicated CPUs in a homogeneous computing system. This is due to the need for replicas to finish simultaneously so that a swap event can occur. Thus, efficient implementation on shared or unreliable computing clusters or on large-scale distributed computing systems is very difficult to realize because CPUs often sit idle, waiting for other replicas to finish a simulation segment. For this reason, efforts to distribute replica exchange simulations have focused on minimizing this wait time through an optimized allocation of jobs on the basis of CPU speed.[33−35]

As the need for a larger number of replicas increases (in cases of complex systems), the feasibility advantage of DR over RE rises. Because DR undergoes replica jumps rather than replica exchanges, no CPU need ever wait for others to finish. A 100% utilization of all available CPU resources is thus realized. Furthermore, the method is not sensitive to changes in CPU speed or CPU availability, both of which cannot be considered fixed on most clusters. DR will continue running even if some replicas are suspended because of a drop in the number of available CPUs. DR sampling is, therefore, a different approach designed from the ground up to naturally suit shared or large-scale distributed computing platforms.

Acknowledgment. We thank Radford Neal, Jeremy Schofield, and Stuart G. Whittington for providing comments and gratefully acknowledge the Canadian Institutes of Health Research for support.

Note Added after ASAP Publication. This article was released ASAP on March 22, 2006, with the incorrect Received Date. The correct version was posted on April 12, 2006.

References

(1) Rodinger, T.; Pomès, R. *Curr. Opin. Struct. Biol.* **2005**, *15*, 164−170.

(2) Beveridge, D. L.; Dicapua, F. M. *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 431−492.

(3) Straatsma, T. P.; McCammon, J. A. *Annu. Rev. Phys. Chem.* **1992**, *43*, 407−435.

(4) Chipot, C.; Pearlman, D. A. *Mol. Simul.* **2002**, *28*, 1−12.

(5) van Gunsteren, W. F.; Daura, X.; Mark, A. E. *Helv. Chim. Acta* **2002**, *85*, 3113−3129.

(6) Simonson, T.; Archontis, G.; Karplus, M. *Acc. Chem. Res.* **2002**, *35*, 430−437.

(7) Lyubartsev, A. P.; Martsinovski, A. A.; Shevkunov, S. V.; Vorontsovvelyaminov, P. N. *J. Chem. Phys.* **1992**, *96*, 1776−1783.

(8) Marinari, E.; Parisi, G. *Europhys. Lett.* **1992**, *19*, 451−458.

(9) Berg, B. A.; Neuhaus, T. *Phys. Lett. B* **1991**, *267*, 249−253.

(10) Berg, B. A.; Neuhaus, T. *Phys. Rev. Lett.* **1992**, *68*, 9−12.

(11) Darve, E.; Pohorille, A. *J. Chem. Phys.* **2001**, *115*, 9169−9183.

(12) Darve, E.; Wilson, M. A.; Pohorille, A. *Mol. Simul.* **2002**, *28*, 113−144.

(13) Hénin, J.; Chipot, C. *J. Chem. Phys.* **2004**, *121*, 2904−2914.

(14) Mezei, M. *J. Comput. Phys.* **1987**, *68*, 237−248.

(15) Bartels, C.; Karplus, M. *J. Comput. Chem.* **1997**, *18*, 1450−1462.

(16) Hénin, J.; Chipot, C. *J. Chem. Phys.* **2004**, *121*, 2904−2914.

(17) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781−1802.

(18) Hukushima, K.; Nemoto, K. *J. Phys. Soc. Jpn.* **1996**, *65*, 1604−1608.

(19) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141−151.

(20) Tesi, M. C.; van Rensburg, E. J. J.; Orlandini, E.; Whittington, S. G. *J. Stat. Phys.* **1996**, *82*, 155−181.

(21) Woods, C. J.; Essex, J. W.; King, M. A. *J. Phys. Chem. B* **2003**, *107*, 13703−13710.

(22) Woods, C. J.; Essex, J. W.; King, M. A. *J. Phys. Chem. B* **2003**, *107*, 13711−13718.

(23) Sugita, Y.; Kitao, A.; Okamoto, Y. *J. Chem. Phys.* **2000**, *113*, 6042−6051.

(24) Mitsutake, A.; Okamoto, Y. *J. Chem. Phys.* **2004**, *121*, 2491−2504.

(25) Mitsutake, A.; Okamoto, Y. *Chem. Phys. Lett.* **2000**, *332*, 131−138.

(26) Mitsutake, A.; Sugita, Y.; Okamoto, Y. *J. Chem. Phys.* **2003**, *118*, 6676−6688.

(27) Mitsutake, A.; Sugita, Y.; Okamoto, Y. *J. Chem. Phys.* **2003**, *118*, 6664−6675.

(28) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **2000**, *329*, 261−270.

(29) García, A. E.; Onuchic, J. N. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13898−13903.

(30) Nymeyer, H.; García, A. E. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13934−13939.

(31) Paschek, D.; García, A. E. *Phys. Rev. Lett.* **2004**, *93*, 238105.

(32) Gnanakaran, S.; Nymeyer, H.; Portman, J.; Sanbonmatsu, K. Y.; García, A. E. *Curr. Opin. Struct. Biol.* **2003**, *13*, 168−174.

(33) Woods, C. J.; Ng, M. H.; Johnston, S.; Murdock, S. E.; Wu, B.; Tai, K.; Fangohr, H.; Jeffreys, P.; Cox, S.; Frey, J. G.; Sansom, M. S. P.; Essex, J. W. *Philos. Trans. R. Soc. London, Ser. A* **2005**, *363*, 2017−2035.

(34) Rhee, Y. M.; Pande, V. S. *Biophys. J.* **2003**, *84*, 775−786.

(35) Earl, D. J.; Deem, M. W. *J. Phys. Chem. B* **2004**, *108*, 6844−6849.

CT050302X

JCTC Journal of Chemical Theory and Computation

Efficient Generalized Born Models for Monte Carlo Simulations

Julien Michel,[†] Richard D. Taylor,[‡] and Jonathan W. Essex*,[†]

School of Chemistry, University of Southampton, Highfield, Southampton SO17 1BJ, U.K., and Astex Therapeutics Ltd., 436 Cambridge Science Park, Cambridge CB4 0QA, U.K.

Received February 17, 2006

Abstract: The Generalized Born Surface Area theory (GBSA) has become a popular method to model the solvation of biomolecules. While efficient in the context of molecular dynamics simulations, GBSA calculations do not integrate well with Monte Carlo simulations because of the nonlocal nature of the Generalized Born energy. We present a method by which Monte Carlo Generalized Born simulations can be made seven to eight times faster on a protein−ligand binding free energy calculation with little or no loss of accuracy. The method can be employed in any type of Monte Carlo or Hybrid Monte Carlo-molecular dynamics simulation and should prove useful in numerous applications.

Introduction

A proper representation of water is essential for the study of numerous important biomolecular processes. The traditional microscopic approach explicitly models thousands of water molecules to solvate a protein and then applies periodic boundary conditions to the whole system. While rigorous, this approach entails a heavy computational cost. Continuum electrostatics provides an elegant way to model the influence of the solvent on the solute structure and dynamics. In this approach the solvent is considered as a high dielectric continuum that is interacting with partial charges that are embedded in a solute molecule of lower dielectric. The solute response to the reaction field of the solvent dielectric can then be modeled by applying laws of classical electrostatics. The dramatic reduction in the number of degrees of freedom of the simulated system increases the rate of convergence of properties that are averaged over many snapshots. Among the many flavors of continuum electrostatics theories that have been developed in the last two decades, the Generalized Born Surface Area (GBSA) method has become popular in biomolecular simulations.[1] This is in no doubt due to the efficiency of the method and its reasonable accuracy. Furthermore, under certain approximations, the GBSA theory can be expressed in a pairwise form that can be implemented relatively easily within traditional biomolecular simulation packages.[2−4]

In a GBSA simulation, the force field is augmented with an extra term, the solvation energy ΔG_{solv}, taken as the sum of a polarization energy ΔG_{pol} and a nonpolar energy term ΔG_{nonpol}.

$$\Delta G_{solv} = \Delta G_{pol} + \Delta G_{nonpol} \tag{1}$$

In eq 1 ΔG_{nonpol} is often taken as proportional to the solvent accessible surface area (SASA) of the system. Alternative more elaborate treatments have recently been proposed in the literature.[3] The term ΔG_{pol} is obtained from eq 2.

$$\Delta G_{pol} = -\frac{1}{2}\left(\frac{1}{\epsilon_{vac}} - \frac{1}{\epsilon_{solv}}\right)\sum_i\sum_j \frac{q_i q_j}{\sqrt{r^2_{ij} + B_i B_j e^{(-r^2_{ij}/4B_iB_j)}}} \tag{2}$$

ϵ_{vac} and ϵ_{solv} are the dielectric constants of the vacuum and the solvent respectively, q_i is the atomic partial charge of atom i, r_{ij} is the distance between a pair of atoms ij, and B_i is the effective Born radius of atom i.

The effective Born radius B_i is in essence the spherically averaged distance of the solute atom to the solvent. Modifications to eq 2 have been occasionally proposed,[5] but much of the improvements in the accuracy of the Generalized Born

* Corresponding author e-mail: J.W.Essex@soton.ac.uk.
[†] University of Southampton.
[‡] Astex Therapeutics Ltd.

10.1021/ct600069r CCC: $33.50
Published on Web 04/08/2006

models has come from a better calculation of the Born radii.[6−8] One method to compute this quantity is the Pairwise Descreening Approximation (PDA) of Hawkins et al.[2] described by eq 3. While the PDA is not the most accurate way to calculate a Born radius, it is one of the fastest and has proven popular.

$$\frac{1}{B_i} = \frac{1}{\alpha_i} - \frac{1}{2}\sum_{j\neq i}\frac{1}{L_{ij}} - \frac{1}{U_{ij}} + \frac{r_{ij}}{4}\left(\frac{1}{U_{ij}^2} - \frac{1}{L_{ij}^2}\right) + \frac{1}{2r_{ij}}\ln\frac{L_{ij}}{U_{ij}} + \frac{S_j^2\alpha_j^2}{4r_{ij}}\left(\frac{1}{L_{ij}^2} - \frac{1}{U_{ij}^2}\right) \quad (3)$$

$$L_{ij} = 1 \text{ if } r_{ij} + S_j\alpha_j \leq \alpha_i$$

$$L_{ij} = \alpha_i \text{ if } r_{ij} - S_j\alpha_j \leq \alpha_i < r_{ij} + S_j\alpha_j$$

$$L_{ij} = r_{ij} - S_j\alpha_j \text{ if } \alpha_i \leq r_{ij} - S_j\alpha_j$$

$$U_{ij} = 1 \text{ if } r_{ij} + S_j\alpha_j \leq \alpha_j$$

$$U_{ij} = r_{ij} + S_j\alpha_j \text{ if } \alpha i < r_{ij} + S_j\alpha_j$$

In eq 3 r_{ij} is the distance between a pair of atoms ij and α_i is the intrinsic Born radius of atom i, that is, the Born radius that atom i would adopt if it were completely isolated. Finally S_j is a scaling factor which compensates for systematic errors introduced by this approximate Born radii calculation.

GBSA is widely used in the context of molecular dynamics simulations. For example, some interesting studies of GBSA molecular dynamics simulations of RNAs are discussed by Sorin et al.,[9,10] while Felts et al. used GBSA molecular dynamics to study the potential of mean force of small peptides.[11] To date, few Monte Carlo GBSA simulations have been reported in the literature. The flexible docking algorithm of Taylor et al. uses Monte Carlo moves and a GBSA model of water.[12] The Concerted Rotation with Angles (CRA) algorithm of Ulmschneider et al. uses novel Monte Carlo moves to fold peptides in a GBSA force field.[13,14] While the treatment of small peptides with a Monte Carlo GBSA method is still efficient compared to the explicit solvent alternative, or desirable in the case of Monte Carlo protein backbone moves, the method quickly loses its appeal as the system size increases. A Monte Carlo simulation of a biomolecule requires many more moves than molecular dynamics time steps because only portions of the system under study are updated at every move. Because most of the system does not change coordinates, and the force field terms are usually separable, it is generally sufficient to calculate only the change in energy of the part that has moved, which is very efficient. However, inspection of eq 3 shows that the Born radius of atom i depends on the position of every other atom j in the system. In turn, this means that the pairwise energies from eq 2 have the same dependency. As a result, the energy between atoms that did not move must be recomputed after every Monte Carlo move, and a full GB energy calculation must be performed after every Monte Carlo move. For even a mid-sized protein the computational cost can be very high. This is not a problem in a molecular dynamics simulation because the total energy of the system is calculated after every step in any case.

The aim of this article is to introduce methods that can overcome the limitations of a standard GBSA implementation within the framework of a Monte Carlo simulation. This work is motivated by the availability of powerful Monte Carlo methods such as concerted rotations[13] or configurational bias for sampling protein systems,[15] and the efficiency that can be attained by combining them with a GBSA model.

System Setup

In this work, the GBSA method was implemented in a modified version of the Monte Carlo package ProtoMS2.1.[16] Polarization energies were computed using eq 2, and the Born radii were calculated with eq 3. The Surface Area calculations were implemented using the method of Shrake and Rupley,[17] and a probe of 1.4 Å radius was used. The parameter set used for this GBSA model comes from a previous study ("PDAnum2" in ref 18).

To test the approximations introduced below we selected as a test case a set of protein−ligand relative binding free energy calculations which are shown in Figure 1. The perturbations are typical of the mutations performed in a protein ligand binding free energy calculation and cover apolar to apolar (**1to2**), polar to apolar (**1to3** and **4to5**), and polar to polar (**4to6**) perturbations. The two different proteins considered exhibit a very different binding site. Neuraminidase has a polar, solvent exposed, binding site, while cyclooxygenase-2 has a buried, fairly hydrophobic binding site. The test case should therefore represent a broad class of protein−ligand interactions that are studied by free energy perturbation methodologies. The binding mode of the inhibitors was inferred on the basis of a similar ligand complexed to a monomer of the N2 strain of influenza A (PDB code 1BJI)[19] or the PDB structure of murine cox2 complexed to SC-558 (PDB code 1CX2).[20] When necessary, hydrogens were added to the crystallographic structure using the program reduce.[21] Sugars, cofactors, crystallographic waters, and ions were removed. The protein was set up with the AMBER99 force field, inhibitors were set up with the GAFF force field, and the atomic partial charges were derived using the AM1/BCC method[22] as implemented in the package AMBER8.[23] The system was energy minimized using the Sander module of AMBER8 and a Generalized Born force field (the igb keyword was set to 1).[23] The backbone of the energy minimized protein was kept rigid for subsequent Monte Carlo simulations which were conducted with a modified version of the ProtoMS2.1 package.[16] To reduce the computational cost, only the protein residues that have one heavy atom within 15 Å of any heavy atom of the ligands were conserved. The bond angles and torsions of the protein side chains within 10 Å of any heavy atom of the ligand and all the bond angles and torsions of the ligand were sampled during the simulation, with the exception of rings. The bond lengths of the protein and ligand were kept rigid. The total charge of the system was brought to zero by neutralizing lysine residues lying in the outer (frozen) part of the scoop (residues number 511 and 532 for cox2, 432 and 273 for neuraminidase). The protonation state of the

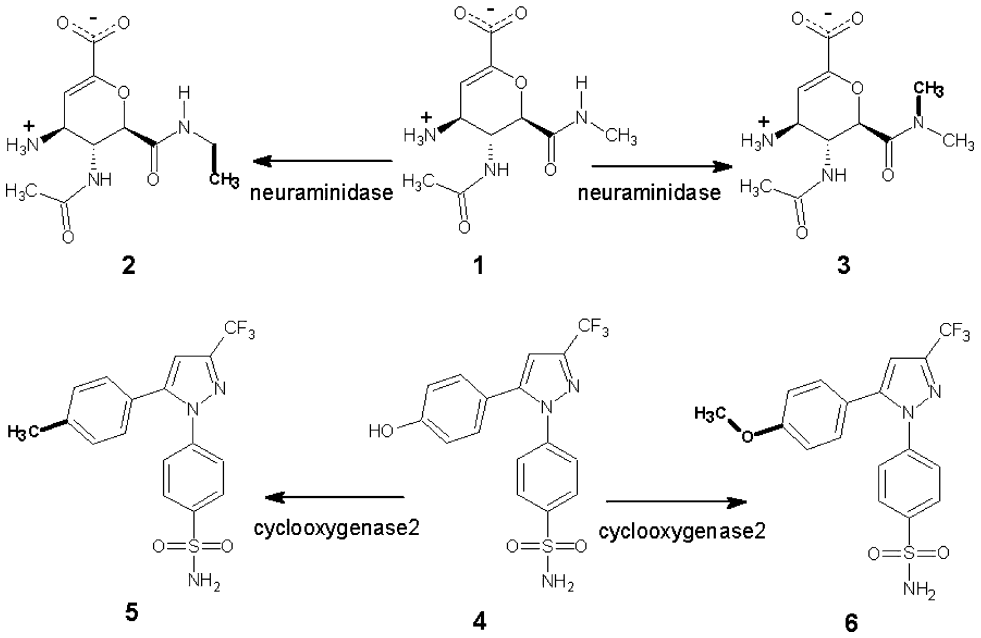


Figure 1. Representation of the ligands considered in this free energy study. For visual emphasis, the parts of the ligands being perturbed are highlighted by bold, straight lines.

histidine residues was decided by visual inspection of the crystallographic structures. The resulting model of cox2 had 155 residues and neuraminidase 145 residues. A 10 Å switched residue based cutoff was employed in all simulations. In the Generalized Born simulations, a cutoff of 20.0 Å for the calculation of the Born radii was applied.

Replica exchange thermodynamic integration[24,25] (RETI) was applied to these systems, and the necessary ensemble of states were formed using Metropolis Monte Carlo sampling[26] at a temperature of 25 °C. In the RETI protocol, standard finite difference thermodynamic integration (FD-TI)[27] is performed at each value of the coupling parameter λ ($\Delta\lambda$=0.001). Occasionally, moves that exchange system coordinates between replica i at $\lambda = A$ of energy $E_A(i)$ and replica j at $\lambda = B$ of energy $E_B(j)$ are attempted, subject to the following acceptance test.

$$\exp[\beta([E_B(j) - E_B(i)] - [E_A(j) - E_A(i)])] \geq \text{rand}(0, 1) \quad (4)$$

The occasional exchange of coordinates between the different simulations enhances configurational sampling and hence convergence of the calculated properties, while the acceptance test ensures that each replica converges the simulation to the correct distribution of states.[24,25]

Solute moves were attempted 10% of the time, with the remainder being protein side chain moves. In the unbound state, 2000 (2K) moves of equilibration were performed before 200K moves of data collection. In the bound state, the system was preequilibrated at one value of λ for 600K moves. The resulting configuration was distributed over 12 values of the coupling parameter λ (0.00, 0.10, ..., 0.90, 0.95, 1.00), and further equilibration was performed for 100K moves. Data were collected over the remaining 900K moves. Replica exchange moves were attempted every 5K configurations.

The error on the free energy gradients was calculated by taking the standard error of batch averages (size 1K). The standard error of these averages was then integrated over the λ coordinate to yield the maximum error.

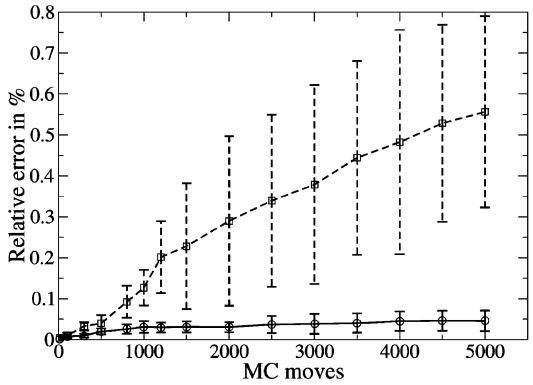
The speedup reported in the next sections are calculated as ratios of the time taken to complete 1000 MC moves on the test systems between two particular simulation protocols.

Approximated Generalized Born Potential

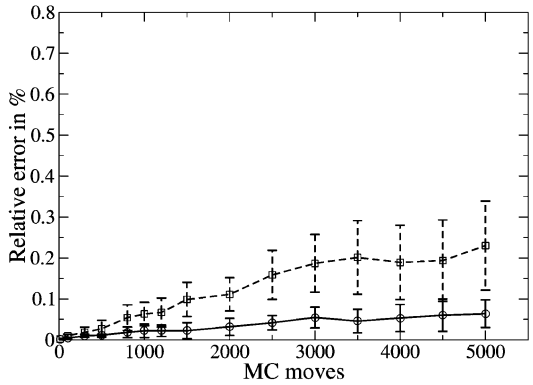
A rigorous GB calculation means that the GB energy between each pair of atoms must be recalculated after every move. However, the impact of a moving atom on the Born radius of a distant atom is small. We have therefore structured the implementation of the GB calculation such that the energy of a pair of atoms is recalculated only if the Born radius of either atom has changed by more than a specified threshold value after a MC move. A large number of pair interactions can be skipped in this fashion, resulting in a significant speedup. In this implementation, only the necessary old and new GB energy pair terms are recalculated to update the total GB energy. This keeps additional memory requirements low and makes the method easily applicable to larger system.

This approximation may have unwanted effects. For example, the total energy would not be completely conserved in a hypothetical Monte Carlo simulation in the NVE ensemble. However, the fact that useful results can be obtained from Molecular Dynamics simulations where fluctuations in the total energy are introduced because of errors in the integrator suggests that as long as the impact of the approximation is small, the resulting ensemble will closely mirror the correct one.

To assess the impact of this approximation, we have run a series of short free energy simulations with a GBSA force field at different values of the threshold parameter. The plots reported in Figure 2 are constructed by running a simulation for N steps with a specified value of the threshold parameter. The total energy of the last generated configuration is then recorded and compared to the value that is obtained by

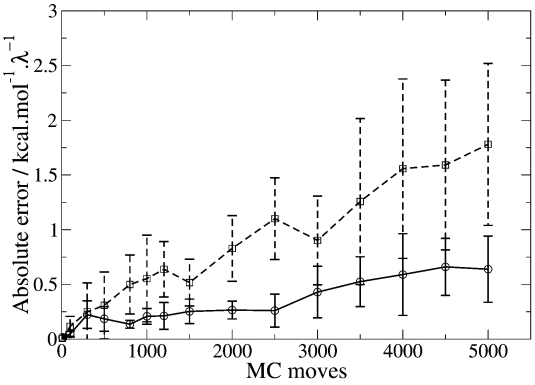


(a) Perturbation **1to3**, neuraminidase

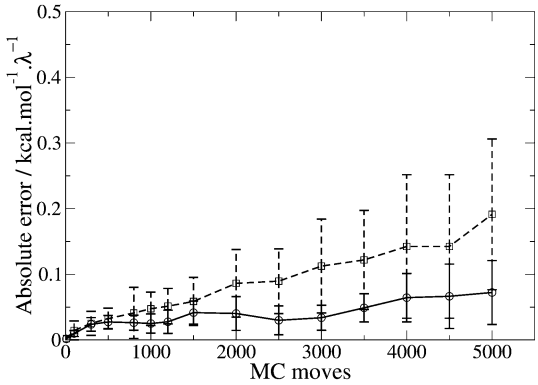


(b) Perturbation **4to5**, cox2

Figure 2. Relative error in percent of the total energy as a function of the number of Monte Carlo moves for a threshold of 0.005 Å (black line, circles) and 0.05 Å (dashed line, squares). Each point is the average of 10 different simulations, and the error bar represents the associated standard error. Similar plots for the perturbations **1to2** and **4to6** are observed (data not shown).



(a) Perturbation **1to3**, neuraminidase



(b) Perturbation **4to5**, cox2

Figure 3. Absolute error in the free energy gradients as a function of the number of Monte Carlo moves for a threshold of 0.001 Å (black line, circles) and 0.005 Å (dashed line, squares). Each point is the average of 10 different simulations, and the error bar represents the associated standard error. Similar plots for the perturbations **1to2** and **4to6** are observed (data not shown).

calculating the total energy with no approximation. This procedure is repeated 10 times for a number of values of *N*. An arguably acceptable error on the total energy would be about 0.1% since this is in the accepted range of MD integrator errors.[28] The systems run in neuraminidase are more sensitive to the approximation, and at a high threshold value, the deviations become quickly large. The systems run on cox2 appear much less sensitive. In both systems, at a threshold of 0.005 Å and up to 5000 MC moves the error is below 0.1%.

An added requirement for a free energy calculation is that the free energy gradients are not too sensitive to this approximation. In this application, we use a finite difference scheme, and the gradients are formed from the difference in total energy at a value of $\lambda - d\lambda$ and $\lambda + d\lambda$. In Figure 3 the protocol described previously is applied to report the free energy gradients accumulated at $\lambda = 0.50$. The gradients are formed from the difference of two large numbers and are therefore more sensitive to small errors in the total energies. The free energy gradients of the perturbations run on neuraminidase are seen to be much more sensitive to the threshold than those run on cox2. Because the binding site of neuraminidase is much more solvent exposed and comprises several polar amino acids, a rigorous treatment of the GB energy appears more important than for cox2, where the buried, hydrophobic binding site is less sensitive to solvent effects. After 1000 MC moves, at a threshold of 0.001, the average error on the free energy gradients is still 0.20 ± 0.07 kcal·mol^{-1}·λ^{-1} for **1to3**, while for **4to5** it is essentially negligible at high and low threshold values. From a computational perspective, the cost of a full GBSA calculation after 1000 approximate GBSA calculations is small. By updating completely the GB energy every 1K MC moves and with a suitably small threshold parameter, the errors on the total energy and the free energy gradients can be kept sufficiently low such that they have a small or negligible influence on the computed free energy.

To verify more rigorously the sensitivity of the systems to the threshold, a series of GBSA free energy simulations is run for each system with a varying threshold parameter. The impact of the threshold parameter on the calculated free energy is shown in Figure 4. For the perturbations in cox2, the calculated free energies are within the statistical error of the exact simulation over the range of thresholds studied. For the perturbations in neuraminidase, the free energy is more sensitive to the value of the threshold parameter, and a high value of the threshold yields results that deviate significantly from the rigorous calculations; this is more

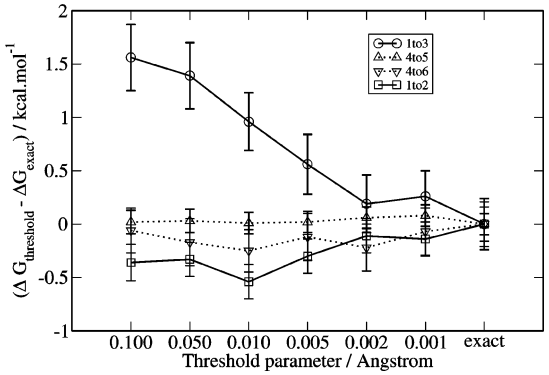


Figure 4. Influence of the threshold parameter on the calculated free energy for the selected perturbations in the bound state. The error bar represents the associated statistical error.

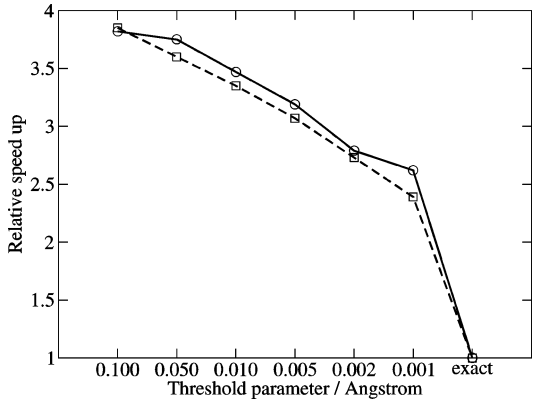


Figure 5. Relative speedup that can be achieved in the simulation of the perturbations **1to3** (circles, black line) and **4to6** (squares, dashed line) in the bound state with a varying threshold parameter. The speedups are based on the time taken to perform 1K Monte Carlo moves.

marked for **1to3** than **1to2**, with results agreeing to within a statistical sampling error from a threshold value of 0.002 Å or less. These results are consistent with the increased sensitivity of the free energy gradients to the threshold for the neuraminidase systems observed in Figure 3. Figure 5 shows the speedup relative to a full GBSA calculation. Because the computational expense is similar for the systems run on the same protein, speedups are shown for **1to3** and **4to5** only. Even with a threshold as low as 0.001 Å, a considerable speedup is achieved because the Born radii of several protein atoms are insensitive to the displacement of a distant residue. On these systems and over the range of thresholds studied, the simulations run 2.4−3.8 times faster.

On the range of systems studies here, the influence of the threshold parameter on the calculated binding free energies has been shown to be negligible (cox2) or minor (neuraminidase) and can be minimized by reducing sufficiently the threshold parameter, at the cost of additional simulation time. That a balance can be struck between speedup and accuracy can prove useful. In applications where accuracy is important, almost rigorous calculations can be made with a sufficiently low threshold. On the other hand, less accurate calculations that could be useful in the context of fast free energy calculations could be run with a higher threshold. The optimum value of the threshold varies according to the system but can be estimated rapidly by plotting the drifts in the free energy gradients for a series of short simulations. For the systems studied here, it appears that a good compromise would be achieved with a threshold of 0.005 Å.

Simplified Sampling Potential

Theory. A novel methodology to perform Monte Carlo simulations has recently been proposed by Gelb.[29] He shows that it is possible to perform a Monte Carlo simulation in which the potential energy is evaluated using an approximate, less expensive potential E_ζ than a more realistic potential E_π and still form an ensemble of states that are distributed according to the rigorous potential. The method is very powerful as in principle any kind of simplified potential/expensive potential combination can be devised.

The method can be briefly summarized as follows:

1. Start a simulation in state *i*.

2. Performs *N* steps of standard Metropolis sampling with a simple potential E_ζ of limiting distribution ζ until a state *j* is reached.

3. Set state i = state j with probability $\chi = (\pi_j\zeta_i)/(\pi_i\zeta_j)$. In this equation π_i and ζ_i are the probability of state *i* in the two different distributions π and ζ.

4. Accumulate any property of interest that is a function of the coordinates of state *i*.

5. Return to 1 or terminate after a number of iterations

In essence, a standard Monte Carlo simulation is conducted for *N* steps with a potential chosen for its convenience (usually computational efficiency). However, because the probabilities of state *i* and *j* in the two distributions π and ζ generally differ, it is necessary in step three to correct for any bias introduced by the potential E_ζ. This acceptance test makes sure that the ensemble formed during the simulation converges toward the distribution π instead of ζ. In the NVT ensemble, step three amounts to accepting state *j* according to

$$\exp[\beta([E_\pi(j) - E_\pi(i)] - [E_\zeta(j) - E_\zeta(i)])] \geq \text{rand}(0, 1) \quad (5)$$

The acceptance test is therefore based on the difference of differences of energies of state *i* and *j* between the two potentials E_π and E_ζ. With this method, no statistics for the target ensemble π can be collected during step two, and the number of data points accumulated is reduced compared to a traditional Monte Carlo simulation. This does not necessarily affect convergence because subsequent configurations in the Markov chain are typically highly correlated and do not contribute new information to the running average. That is to say, it is equally good to sample the distribution of interest less often if the samples are less correlated.

To date, applications of this methodology have been reported by Hetenyi et al.[30,31] (who seem to have developed a similar method independently). Hetenyi reported a 3.0−4.7 speedup in the simulation of a Lennard-Jones fluid using for E_ζ a potential similar to E_π but with a shorter cutoff. Gelb reported similar results on a similar system.[29] In a second application by Hetenyi, a MC Ewald sum simulation

of water was running 4.5−7.5 times faster using this methodology. This method has been employed by Iftimie et al. to perform ab initio simulations using a classical potential.[32]

If the change of energy in going from state *i* to *j* is similar with the two Hamiltonians, the probability of accepting the configuration *j* will be close to unity. On the other hand, if the two potentials differ too much, then the acceptance rate will drop, and the method will lose efficiency since all the steps performed with E_ξ have been wasted. Therefore a good approximate potential E_ξ must be faster than, and yet very similar to, E_π. This is of course difficult to achieve.

Application to a GBSA Model. The complete application of the GBSA theory requires the calculation of a Surface Area (SA) dependent term to yield a solvation free energy. The inclusion of this term can be expensive and becomes significant once the GB calculations have been accelerated with the use of a threshold. For example, the simulation of **4to5** with a GB threshold of 0.005 is about 1.8 times slower once the SA calculations are enabled.

The fluctuations in the SA term are known to be small compared to the other energy components of the force field. This observation has led other workers to devise schemes where the SA term is only periodically updated.[33,12,14] While reasonable, this approximation is not completely rigorous. Other workers have developed faster, approximate SASA calculations schemes, but these algorithms do not calculate reliably the small changes in SASA associated with the small conformational changes observed between MC moves.[34]

However, the effect of the SA term can be rigorously included in the GBSA simulation by adopting a particular simplified sampling potential methodology. The simple potential E_ξ corresponds to a GB simulation run without a SA term, while the correct potential E_π includes SA calculations.

In addition, we consider other means to further speed up the calculations by adopting a less rigorous solvation model for the simplified potential. Here two different simplified solvation models are investigated: a distance dependent dielectric (DDD) force field and a simplified GB force field (fastGB). In the DDD force field the GB equations are replaced by a $\epsilon(r) = 4r$ distance dependent dielectric. In the fastGB force field smaller cutoffs are applied: a residue based cutoff, Born radii cutoff, and thresholds of 6.0, 12, and 0.05 Å, respectively. Furthermore, since no statistics are collected with fast GB, it is not necessary to compute free energy gradients, which avoids the expensive GB energy calculations for the perturbed states. The rigorous potential is taken as a GB simulation with a threshold of 0.005 Å and a rigorous SA calculation.

Table 1 lists the average acceptance rate of the correction step for the two different potentials as a function of the number of moves performed. The speedup compared to the rigorous GBSA simulation is also reported. The parameter *M* is the number of moves performed with the quick potential before attempting to add the generated configuration to the ensemble. As this quantity increases, the acceptance rate diminishes. As has been pointed out, a tradeoff must be made between computational efficiency and sampling efficiency.[29]

Table 1. Acceptance Rate at the Correction Step and Relative Speedup for Different Combinations of Potentials and Number of Moves *M* with the Approximate Potential[a]

M	DDD rate[b] (%)	DDD speedup[c]	fastGB rate[b] (%)	fastGB speedup[c]
		1to3		
5	56.7	2.2	90.2	2.1
10	36.7	2.4	83.4	2.3
20	18.2	2.9	76.3	2.7
25	13.6	3.1	71.5	2.8
		4to5		
5	65.4	2.2	91.2	2.0
10	47.7	2.5	87.1	2.3
20	27.9	2.9	80.8	2.5
25	22.5	3.1	78.2	2.7

[a] The results for **1to2** and **4to6** are similar to **1to3** and **4to5**. [b] Average across all values of λ. [c] Relative to a GBSA simulation with a threshold of 0.005 Å.

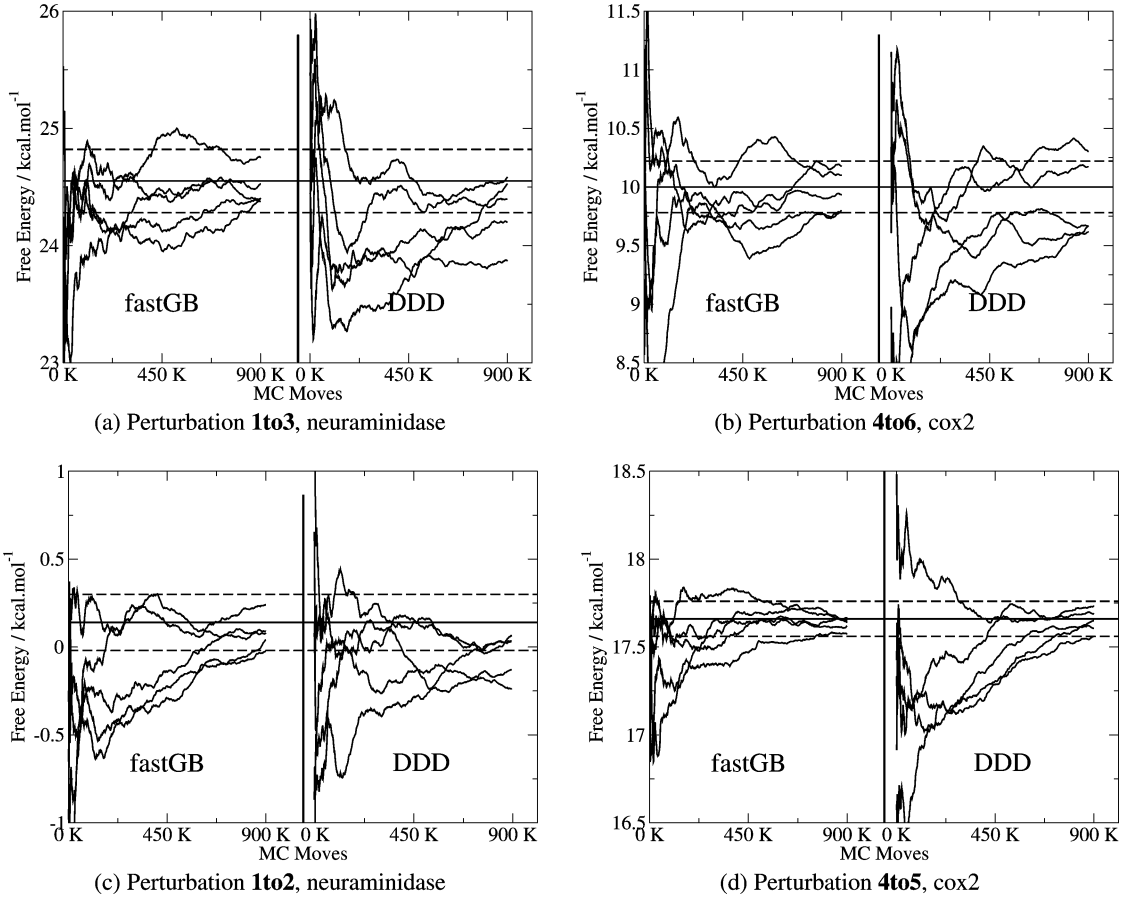
With the DDD model, the acceptance rate decreases faster than the speedup increases, and a short value of *M* is favored. Even after only 5 steps, the acceptance rate is only 55−65%. For the systems in neuraminidase, the acceptance rate of the correction step is actually similar to using vacuum conditions (data not shown). This illustrates that the configurations favored by a GBSA force field are rather different from those preferred by a DDD force field.

With the fastGB model the decreases in the acceptance rate are more or less counterbalanced by the increase in speedup, and no value of *M* is clearly favored. In addition, the acceptance rates are much higher and around 90% for *M* equal to 5.

If we make the assumption that simulations run with the DDD and fastGB force field explore the configurational space at the same rate, then these results suggest that the combination of the two potentials fastGB/GBSA yields more efficient sampling than the DDD/GBSA combination.

To demonstrate this more decisively, in Figure 6 we investigate the convergence of the calculated free energies in the bound state for 5 independent simulations performed with the different protocols and a value of *M* set to 10. After 900K moves of data collection, almost all the fastGB simulations have converged to within the error bounds of the results obtained with GBSA 0.005 Å. With the DDD protocol, the results are more spread out, and several simulations are outside the error bounds. It is apparent that the fastGB protocol converges better the free energies than the DDD protocol for the same number of iterations.

Taken together, the results in Table 1 and Figure 6 suggest that a low acceptance rate for the correction step hinders convergence. The DDD simulations are slightly faster than the fastGB simulations. However, since the simulation results are much better converged with the fastGB protocol, it should be preferred over a DDD model. By combining the fastGB potential with the value of *M* set to 10 and with a GBSA 0.005 Å potential described in the previous section, an approximately 2.3-fold speedup over a standard MC simulation run with GBSA 0.005 Å can be achieved. The present results demonstrate that the simplified sampling potential methodology, applied here to increase the efficiency of



(a) Perturbation **1to3**, neuraminidase (b) Perturbation **4to6**, cox2

(c) Perturbation **1to2**, neuraminidase (d) Perturbation **4to5**, cox2

Figure 6. The convergence of the calculated free energies in the bound state using different sampling potentials. The estimated free energy is plotted as a function of the number of Monte Carlo moves performed. In each figure, on the left-hand side, 5 independent simulations run with fastGB are shown. On the right-hand side, 5 independent simulations run with DDD are shown. The horizontal line is the estimate of the free energy obtained after 900K moves with the potential GBSA 0.005 Å. The dashed lines represents the statistical error associated with this number.

Generalized Born calculations for the first time, allows significant computational savings without additional approximations.

Conclusion

A novel methodology by which free energy calculations in a Generalized Born framework can be made more efficient within Monte Carlo simulations has been proposed. It can be summarized as follows:

1. An approximate Generalized Born potential in which the energy of the system is only partially updated after a MC move—the impact of this approximation on the calculated free energies can be made arbitrarily small at the expense of computational time by adjusting a single parameter.

2. Sampling driven by an inexpensive potential with a special Monte Carlo acceptance test that removes any bias in the distribution introduced by the cheap potential—this allows in addition the rigorous incorporation of surface area calculations at a minimum computational cost.

In Table 2 timings for various combinations of these approximations on two of the four systems are reported. Protocol 4, which combines the two approximations is seven to eight times faster compared to protocol 1 which would correspond to a standard implementation of a GBSA force field. Protocol 4 is also only about 4.1−4.3 times slower than a simulation run in a vacuum (protocol 5) which compares favorably with the typical efficiency of molecular dynamics GBSA simulations.[23] While the increased efficiency has been demonstrated on a free energy calculation, improvements in standard MC simulations could be sought with the same method.

Table 2. Time Required To Complete a Block of 1K Moves for Selected Approximations[a]

protocol	solvation model	simplified potential	time **1to3** (s)	time **4to5** (s)
1	GBSA exact	no	633.4	746.3
2	GBSA threshold 0.001 Å	no	241.8	312.0
3	GBSA threshold 0.005 Å	no	198.6	243.0
4	GBSA threshold 0.005 Å	fastGB $M = 10$	79.8	104.2
5	vacuum	no	19.5	24.1

[a] ProtoMS2.1 on a Pentium IV 2 GHz compiled with g77.

We may ask whether the partial rigidity of the system and the simplified treatment of solvation affects the accuracy of the calculated binding free energy. The simulation results can be compared with experimental figures by constructing a thermodynamic cycle, which requires the ligand perturbations in the unbound state to be performed. We stress that free energy calculations in the unbound state are extremely rapid (on the order of a few minutes), and there is no need to introduce the methods developed to speed up the simula-

Table 3. Calculated and Experimental Binding Free Energies of the Tested Systems with the ApproxGB+SA Protocol[a]

perturbation	$\Delta\Delta G_{exp}$	$\Delta\Delta G_{bind}$	ΔG_{prot}	ΔG_{wat}
1to2	−1.6	-0.3 ± 0.3	0.1 ± 0.2	0.4 ± 0.1
1to3	−2.7	-2.7 ± 0.3	24.5 ± 0.3	27.2 ± 0.2
4to5	<−4.6	-2.4 ± 0.1	17.7 ± 0.1	20.1 ± 0.1
4to6	<−5.6	-3.1 ± 0.2	10.0 ± 0.2	13.1 ± 0.1

[a] The threshold was set to 0.005 Å. The figures are in kcal·mol^{-1}. The experimental figures were taken from ref 35 for the neuraminidase inhibitors and from ref 36 for the cox2 inhibitors.

tions in the bound state. The calculated binding free energies are listed in Table 3. The implicit solvent protocol reproduces well the relative binding free energy of **1to3** but underestimates somewhat the binding free energy of the three other systems, although the trends are respected. The main emphasis of this work was to introduce a novel methodology to perform Generalized Born Monte Carlo free energy calculations efficiently. A thorough investigation of the influence of the solvation model on the relative binding free energies will require the comparison of explicit and implicit solvent simulations on a larger set of systems. Such studies will be made computationally tractable by the present technique.

Acknowledgment. We thank Astex Therapeutics and the University of Southampton for funding this work and EPSRC (GR/R06137/01) for providing computational resources. J. Michel is grateful to Dr. C. Woods for helpful discussions.

References

(1) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127−6129.

(2) Hawkins, C. J.; Cramer, D. *Chem. Phys. Lett.* **1995**, *246*, 122−129.

(3) Gallicchio, E.; Levy, R. M. *J. Comput. Chem.* **2004**, *25*, 479−499.

(4) Qiu, D.; Shenkin, P. S.; Hollinger, F. P.; Still, W. C. *J. Phys. Chem. A* **1997**, *101*, 3005−3014.

(5) Jayaram, B.; Liu, Y.; Beveridge, D. L. *J. Chem. Phys*. **1998**, *109*, 1465−1471.

(6) Ghosh, A.; Rapp, C. S.; Friesner, R. A. *J. Phys. Chem. B* **1998**, *102*, 10983−10990.

(7) Onufriev, A.; Case, D. A.; Bashford, D. *J. Comput. Chem.* **2002**, *23*, 1297−1304.

(8) Feig, M.; Onufriev, A.; Lee, M. S.; Im, W.; Case, D. A.; Brooks, L. C. *J. Comput. Chem.* **2004**, *25*, 265−284.

(9) Sorin, E. J.; Engelhardt, M. A.; Herschlag, D.; Pande, V. S. *J. Mol. Biol.* **2002**, *317*, 493−506.

(10) Sorin, E. J.; Rhee, Y. M.; Nakatani, B. J.; Pande, V. S. *Biophys. J.* **2003**, *85*, 790−803.

(11) Felts, A. K.; Harano, Y.; Gallicchio, E.; Levy, R. M. *Proteins: Struct., Funct., Bioinformatics* **2004**, *56*, 310−321.

(12) Taylor, R. D.; Jewsbury, P. J.; Essex, J. W. *J. Comput. Chem.* **2003**, *24*, 1637−1656.

(13) Ulmschneider, J. P.; Jorgensen, W. L. *J. Chem. Phys.* **2003**, *118*, 4261−4271.

(14) Ulmschneider, J. P.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2004**, *126*, 1849−1857.

(15) Siepmann, J. I.; Frenkel, D. *Mol. Phys.* **1992**, *75*, 59−70.

(16) Woods, C. J.; Michel, J. ProtoMS2.1; in house Monte Carlo Code, 2005.

(17) Shrake, A.; Rupley, J. A. *J. Mol. Biol.* **1973**, *79*, 351−371.

(18) Michel, J.; Taylor, R. D.; Essex, J. W. *J. Comput. Chem.* **2004**, *25*, 1760−1770.

(19) Taylor, N. R.; Cleasby, A.; Singh, O.; Skarzynski, T.; Wonacott, A. J.; Smith, P. W.; Sollis, S. L.; Howes, P. D.; Cherry, P. C.; Bethell, R.; Colman, P.; Varghese, J. *J. Med. Chem.* **1998**, *41*, 798−807.

(20) Kurumbail, R. G.; Stevens, A. M.; Gierse, J. K.; Mcdonald, J. J.; Stegeman, R. A.; Pak, J. Y.; Gildehaus, D.; Miyashiro, J. M.; Penning, T. D.; Seibert, K.; Isakson, P. C.; Stallings, W. C. *Nature* **1996**, *384*, 644−648.

(21) Word, J. M.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C. *J. Mol. Biol.* **1999**, *285*, 1735−1747.

(22) Jakalian, A.; Jack, D. B.; Bayly, C. I. *J. Comput. Chem.* **2002**, *23*, 1623−1641.

(23) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Wang, B.; Pearlman, D. A.; Crowley, M.; Brozell, S.; Tsui, V.; Gohlke, H.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Schafmeister, C.; Caldwell, J. W.; Ross, W. S.; Kollman, P. A. AMBER 8; University of California, San Francisco, 2004.

(24) Woods, C. J.; King, M. A.; Essex, J. W. *J. Phys. Chem. B* **2003**, *107*, 13711−13718.

(25) Woods, C. J.; King, M. A.; Essex, J. W. *J. Phys. Chem. B* **2003**, *107*, 13703−13710.

(26) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. *J. Chem. Phys.* **1953**, *21*, 1087−1092.

(27) Leach, A. R. *Molecular Modelling, Principles and Applications;* Longman: Harlow, U.K., 1996.

(28) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids;* Oxford University Press: 1989.

(29) Gelb, L. D. *J. Chem. Phys.* **2003**, *118*, 7747−7750.

(30) Hetenyi, B.; Bernacki, K.; Berne, B. J. *J. Chem. Phys.* **2002**, *117*, 8203−8207.

(31) Bernacki, K.; Hetenyi, B.; Berne, B. J. *J. Chem. Phys.* **2004**, *121*, 44−50.

(32) Iftimie, R.; Salahub, D.; Wei, D. Q.; Schofield, J. *J. Chem. Phys.* **2000**, *113*, 4852−4862.

(33) Zhu, J. A.; Shi, Y. Y.; Liu, H. Y. *J. Phys. Chem. B* **2002**, *106*, 4844−4853.

(34) Weiser, J.; Shenkin, P. S.; Still, W. C. *J. Comput. Chem.* **1999**, *20*, 217−230.

(35) Wall, I. D.; Leach, A. R.; Salt, D. W.; Ford, M. G.; Essex, J. W. *J. Med. Chem.* **1999**, *42*, 5142−5152.

(36) Price, M. L. P.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2000**, *122*, 9455−9466.

CT600069R

JCTC Journal of Chemical Theory and Computation

MPW1K Performs Much Better than B3LYP in DFT Calculations on Reactions that Proceed by Proton-Coupled Electron Transfer (PCET)

Mark Lingwood,[†] Jeff R. Hammond,[†] David A. Hrovat,[†,‡] James M. Mayer,[†] and Weston Thatcher Borden*[,†,‡]

Department of Chemistry, University of Washington, Box 351700, Seattle, Washington 98195-1700, and Department of Chemistry, University of North Texas, P. O. Box 305070, Denton, Texas 76203-5070

Received November 15, 2005

Abstract: Density functional theory calculations have been performed with the B3LYP and MPW1K functionals on the hydrogen atom abstraction reactions of ethenoxyl with ethenol and of phenoxyl with both phenol and α-naphthol. A comparison with the results of G3 calculations shows that B3LYP seriously underestimates the barrier heights for the reaction of ethenoxyl with ethenol by both proton-coupled electron transfer (PCET) and hydrogen-atom transfer (HAT) mechanisms. The MPW1K functional also underestimates the barrier heights, but by much less than B3LYP. Similarly, a comparison with the results of experiments on the reaction of a phenoxyl radical with α-naphthol indicates that the barrier height for the preferred PCET mechanism is calculated more accurately by MPW1K than by B3LYP. These findings indicate that the MPW1K functional is much better suited than B3LYP for calculations on hydrogen abstraction reactions by both HAT and PCET mechanisms.

Many hydrogen atom abstraction reactions proceed by a classical hydrogen-atom transfer (HAT) mechanism, involving three electrons distributed among three atomic orbitals (AOs).[1] As shown schematically in Figure 1a, the proton and one of the electrons in the X−H bond being broken are both transferred to the singly occupied orbital on radical Y•.

However, in recent years, both experimental and computational studies have found that, when the abstracting radical center carries at least one unshared pair of electrons and the hydrogen to be abstracted is bonded to an atom that also has an unshared pair of electrons, a proton-coupled electron transfer (PCET) mechanism may be preferred over a HAT mechanism.[2] As illustrated in Figure 1b, such a PCET mechanism involves a total of five atomic orbitals. The proton in the X−H bond is transferred to a lone pair of electrons on radical Y•, and simultaneously, an electron is transferred from a lone pair on X to the singly occupied orbital on Y. Thus, unlike the case in HAT, where the proton and the electron of the hydrogen atom are transferred from X to the same AO on Y, in PCET, the proton is transferred between one pair of AOs on X and Y and the electron is transferred between another pair of AOs on these two atoms.

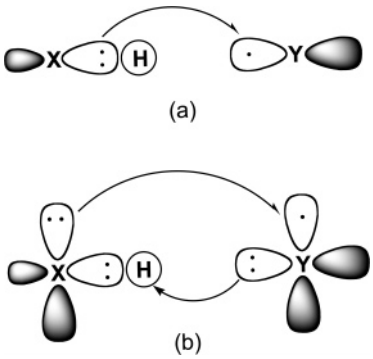


Figure 1. Schematic depiction of (a) HAT and (b) PCET mechanisms for the abstraction of a hydrogen atom from an X−H bond by radical Y•.

* Corresponding author phone: (940) 565-3658; fax: (940) 565-4460; e-mail: borden@unt.edu.
[†] University of Washington.
[‡] University of North Texas.

10.1021/ct050282z CCC: $33.50
Published on Web 03/08/2006

We have reported the results of unrestricted (U)B3LYP calculations on the preferred mechanism for the degenerate hydrogen abstraction reactions of a benzyl radical with toluene, a methoxyl radical with methanol, and a phenoxyl radical with phenol.[3] Our calculations found that, for the first two of these reactions, a HAT mechanism is favored. However, for the reaction of phenoxyl with phenol, our (U)B3LYP calculations found a PCET mechanism to be preferred,[4] and we have presented an analysis of why the mechanism of this reaction differs from that of the reaction of methoxyl with methanol.

Although Becke's three-parameter functional,[5] when combined with the correlation functional of Lee, Yang, and Parr,[6] usually gives good results for reactions of closed-shell molecules, the same is not true for reactions of radicals. Truhlar and co-workers have pointed out that (U)B3LYP calculations underestimate the barrier heights for a set of 40 HAT radical reactions, with a mean signed error of −4.8 kcal/mol.[7] In the same paper, Truhlar et al. showed that MPW1K—a modified version of the Perdew−Wang gradient-corrected exchange functional, with one parameter optimized to give the best fit to the kinetic data for these 40 radical reactions—reduced the mean signed error in the barrier heights for these reactions to −1.3 kcal/mol.

It is not known whether B3LYP makes similar or, perhaps, even larger errors for PCET reactions than for these 40 HAT reactions. It is also not known whether MPW1K is more or less accurate than B3LYP in computing the barrier heights for PCET reactions.

To assess how well these two functionals perform in calculations on a PCET reaction, we have carried out calculations on the degenerate abstraction of the hydroxyl hydrogen of ethenol (the enol of ethanal) by an ethenoxyl radical via a PCET mechanism.

$$H_2C{=}CH{-}OH + {\bullet}O{-}CH{=}CH_2 \rightarrow H_2C{=}CH{-}O{\bullet} + HO{-}CH{=}CH_2 \quad (1)$$

We chose this very simple reaction because we wanted to be able to compare the performance of both B3LYP and MPW1K against that of a high-quality ab initio method, which could be anticipated to give reliable results. For this purpose, we selected the G3 method.[8]

In this manuscript, we report a comparison of the results of (U)B3LYP and (U)MPW1K density functional theory (DFT) calculations with the results of G3 calculations for computing the HAT and PCET barrier heights for O−H hydrogen abstraction from ethenol by an ethenoxyl radical. We have also computed the (U)MPW1K barrier heights for both types of mechanisms in the reaction of phenol with a phenoxyl radical, which we studied previously with (U)B3LYP calculations,[3] and in addition, we have performed (U)MPW1K and (U)B3LYP calculations on the reaction of a phenoxyl radical with α-naphthol. For the last of these three reactions, we report a comparison of the (U)MPW1K and (U)B3LYP computational results with the experimental results obtained by Foti et al.[9]

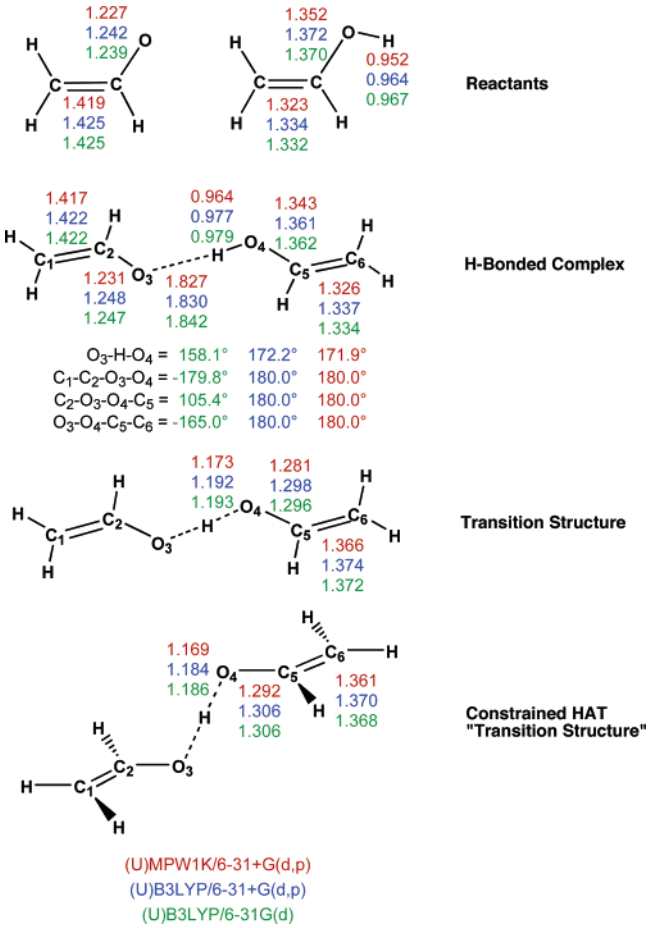


Figure 2. Important geometrical parameters of the stationary points, located by three different types of calculations, in the reaction between ethenoxyl and ethenol.

Computational Methodology

Truhlar and co-workers used the 6-31+G(d,p) basis set[10] in their comparisons of the B3LYP and MPW1K functionals,[7] so we elected to use the same basis set in this study for both types of DFT calculations. We also carried out single-point (U)B3LYP and (U)MPW1K calculations with the aug-cc-pVTZ basis set on the reaction of an ethenoxyl radical with ethenol.[11]

The geometries of the stationary points for the G3 calculations were obtained by performing (U)B3LYP calculations with the 6-31G(d) basis set.[12] Unrestricted wave functions were used for all of the calculations on radicals. Geometries were optimized, transition structures (TSs) located, and vibrational analyses performed using the Gaussian 03 package of electronic structure programs.[13] The unscaled vibrational frequencies were used to obtain the zero-point energies and heat capacities that were necessary in order to convert the differences in electronic energies into differences in enthalpies at 298 K.

Results and Discussion

Hydrogen Abstraction from Ethenol by an Ethenoxyl Radical. Figure 2 gives the most important bond lengths and bond angles in ethenol, ethenoxyl, the hydrogen-bonded complex formed from them, and the transition structure for abstraction of the hydroxyl hydrogen of the alcohol by the

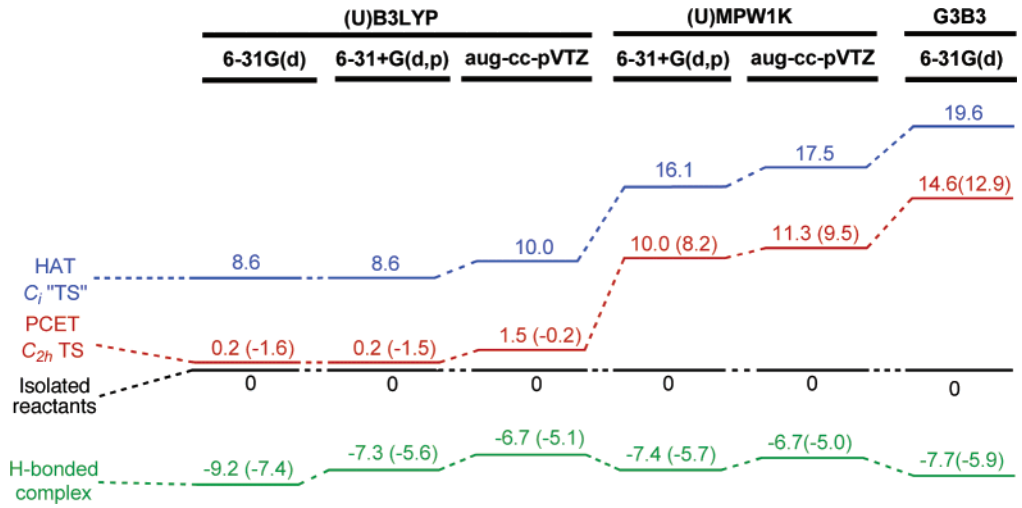


Figure 3. Relative energies (and enthalpies) of stationary points, computed by different methods with different basis sets, for the reaction between ethenoxyl and ethenol. The designation G3B3/6-31G(d) means that the geometries for the G3 calculations were optimized with the (U)B3LYP functional and the 6-31G(d) basis set. When the results of (U)B3LYP/6-31G(d) vibrational analyses were used, zero-point and heat capacity corrections were removed from the G3 enthalpies, so that the purely electronic energies of the PCET TSs and the constrained HAT C_i "TSs" could be compared, not only at the (U)B3LYP and (U)MPW1K levels of theory but also at the G3 level.

oxygen of the radical. (U)MPW1K/6-31+G(d,p), (U)B3LYP/6-31+G(d,p), and (U)B3LYP/6-31G(d) geometrical parameters are all provided.[14] Complete descriptions of all of the optimized geometries are available in the Supporting Information.

PCET, via a planar transition structure, was found to be the preferred mechanism for hydroxyl hydrogen abstraction from ethenol by ethenoxyl. In the C_{2h} transition structure, as the proton in the hydrogen-bonded complex is transferred from ethenol to a σ lone pair in ethenoxyl, a π electron is transferred from the alcohol to the radical.

Attempts to find the transition structure for hydrogen abstraction by a HAT mechanism were unsuccessful. To estimate what the energy of such a transition structure would be, we optimized a partially constrained geometry in C_i symmetry. In this "transition structure", the double bonds were constrained to planarity (i.e., the H−C−C−H and H−C−C−O dihedral angles were frozen at 0° or 180°) and the C═C−O−O dihedral angle was fixed at 90°. The most important bond lengths and bond angles in the (U)B3LYP and (U)MPW1K geometries, optimized with these constraints, are also given in Figure 2, and a full description is provided in the Supporting Information.

The relative energies (and enthalpies) at the (U)B3LYP, (U)MPW1K, and G3 levels of theory of the separated reactants, hydrogen-bonded complex, PCET TS, and the partially constrained C_i geometry (HAT C_i "TS") are shown graphically in Figure 3. The results given in this figure reveal that, with the 6-31+G(d,p) basis set, (U)B3LYP and (U)-MPW1K give nearly the same energy as the G3 calculations for the strength of the hydrogen bond formed between ethenol and ethenoxyl. However, with the larger aug-cc-pVTZ basis set, the (U)B3LYP and (U)MPW1K energies of the hydrogen-bonded complex, relative to the isolated reactants, are both about 1 kcal/mol higher than the G3 energy.

Although, relative to the isolated reactants, (U)B3LYP and (U)MPW1K give very similar energies for the hydrogen-bonded complex between ethenol and ethenoxyl, the PCET TS and the constrained HAT "TS" are calculated to be, respectively, 9.8 and 7.5 kcal/mol lower in energy by (U)B3LYP/6-31+G(d,p) than by (U)MPW1K/6-31+G(d,p). Calculations with the aug-cc-pVTZ basis set give essentially the same differences between the (U)B3LYP and (U)MPW1K results as the calculations with the 6-31+G(d,p) basis set. Obviously, hydroxyl hydrogen abstraction from ethenol by ethenoxyl, by either a PCET or HAT mechanism, is predicted to be much more facile by (U)B3LYP than by (U)MPW1K.

We assume that the G3 calculations give reliable results for the hypothetical model reaction of ethenoxyl with ethenol. If this assumption is valid, the G3 results, which are also shown graphically in Figure 3, indicate that (U)MPW1K gives a much more realistic estimate of the PCET and HAT barrier heights than does (U)B3LYP.

For the PCET mechanism, the G3 barrier height of $\Delta H^{\ddagger} = 12.9$ kcal/mol, relative to the isolated reactants, is 13.1 kcal/mol higher than the (U)B3LYP/aug-cc-pVTZ barrier height, but only 3.4 kcal/mol higher than the (U)MPW1K/aug-cc-pVTZ barrier height. Thus, it appears that (U)B3LYP and (U)MPW1K both underestimate the PCET barrier height, but (U)B3LYP underestimates it by much more than (U)-MPW1K.[15]

(U)B3LYP, (U)MPW1K, and G3 all find that the HAT "TS" is considerably higher in energy than the PCET TS. This qualitative agreement lends credence to the preference for a PCET over a HAT mechanism, which is predicted for the hypothetical reaction of ethenol with ethenoxyl by all three computational methods.

Hydrogen Abstraction from Phenol by a Phenoxyl Radical. The finding that (U)MPW1K apparently gives much more accurate results than (U)B3LYP for hydrogen abstrac-

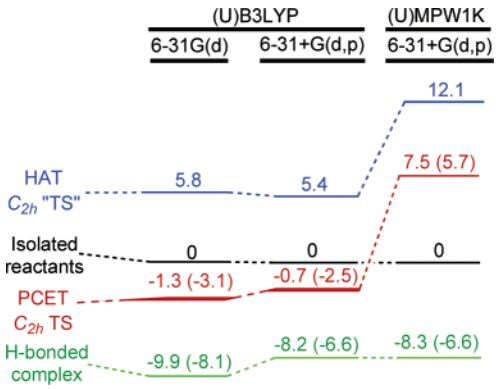


Figure 4. Relative energies (and enthalpies) of stationary points, computed by (U)B3LYP/6-31G(d), (U)B3LYP/6-31+G(d,p), and (U)MPW1K/6-31+G(d,p), for the reaction between phenoxyl and phenol.[3]

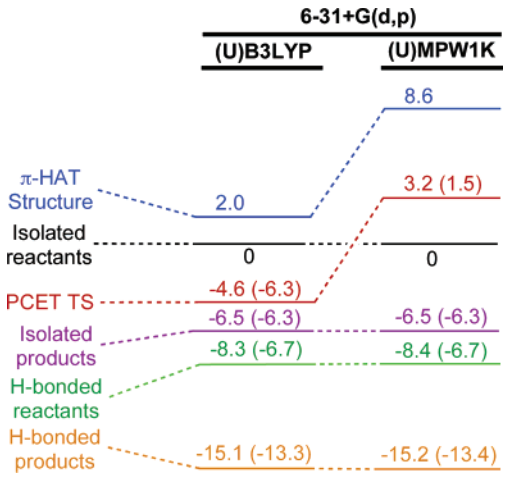


Figure 5. Relative energies (and enthalpies) of stationary points, computed by (U)B3LYP/6-31+G(d,p) and (U)MPW1K/6-31+G(d,p), for the reaction between phenoxyl and α-naphthol.

tion from ethenol by ethenoxyl led us to reinvestigate hydrogen abstraction from phenol by phenoxyl, using (U)MPW1K/6-31+G(d,p) calculations. The (U)MPW1K results are shown graphically in Figure 4. For comparison, the results that we previously obtained at the (U)B3LYP/6-31G(d) level of theory,[3] as well as (U)B3LYP/6-31+G(d,p) results, are also shown in Figure 4.

Both sets of DFT calculations find that the PCET transition structure is lower in energy than the C_{2h} constrained HAT "TS". The energy difference is computed to be 4.6 kcal/mol by (U)MPW1K/6-31+G(d,p), 6.1 kcal/mol by (U)B3LYP and the same basis set, and 7.1 kcal/mol by (U)B3LYP/6-31G(d). When the 6-31+G(d,p) basis set is used, the two functionals place the hydrogen-bonded complex between phenoxyl and phenol below the isolated reactants by almost exactly the same energy and enthalpy.

As shown in Figure 4, the major difference between the results obtained with the two functionals is in the enthalpy that is computed to be required in order to pass over the PCET TS. Starting from the isolated reactants, the activation enthalpy for the PCET reaction of phenoxyl with phenol is predicted to be 8.2 kcal/mol lower by (U)B3LYP/6-31+G(d,p) calculations than that by (U)MPW1K/6-31+G(d,p). This difference in predicted barrier heights is only 15% smaller than the difference of 9.7 kcal/mol between the (U)B3LYP/6-31+G(d,p) and (U)MPW1K/6-31+G(d,p) values of $\Delta H^{\ddagger}$ for the PCET reaction of ethenol with ethenoxyl.

Hydrogen Abstraction from α-Naphthol by a Phenoxyl Radical. The activation parameters for degenerate hydrogen exchange between phenol and phenoxyl have not been measured, so a comparison of the very different values of $\Delta H^{\ddagger}$, predicted by (U)B3LYP and (U)MPW1K, with an experimental value is not possible. However, the exothermic abstraction of the hydroxyl hydrogen atom from α-naphthol by a phenoxyl radical has been found to have $E_a = 2.2 \pm 0.3$ kcal/mol and $\log A = 8.9 \pm 0.3$ kcal/mol.[9] Therefore, we performed calculations on this reaction, to assess how well (U)B3LYP and (U)MPW1K do in calculating the activation enthalpy for it.

As shown in Figure 5, (U)B3LYP/6-31+G(d,p) predicts a PCET TS energy that is 7.8 kcal/mol lower than that computed by (U)MPW1K/6-31+G(d,p). (U)B3LYP unequivocally predicts a barrierless reaction between phenoxyl and α-naphthol. If this bimolecular reaction really were barrierless, it would be expected to be found experimentally to have a negative energy of activation[16] rather than the small but positive value measured by Foti et al.[9]

In contrast to (U)B3LYP, (U)MPW1K predicts an activation energy that is in very good agreement with the value of $E_a = 2.2 \pm 0.3$ kcal/mol, reported by Foti et al.,[9] because the calculated activation enthalpy should be based on the isolated reactants rather than on the hydrogen-bonded complex between them.[17,18] After converting the (U)MPW1K/6-31+G(d,p) value of $\Delta H^{\ddagger} = 1.5$ kcal/mol for the isolated reactants to $E_a = 2.7$ kcal/mol, the (U)MPW1K activation energy is within 0.5 kcal/mol of the experimental value.

Conclusions

A comparison of the results of (U)B3LYP and (U)MPW1K calculations on the reaction of ethenoxyl with ethenol with the results obtained by G3 calculations shows that all three methods find a PCET reaction mechanism to be favored over a HAT mechanism. However, the G3 results indicate that the enthalpies of activation, computed by DFT with these two functionals, are too low for both the PCET and HAT reaction mechanisms.

Nevertheless, the errors in barrier heights made by (U)MPW1K are much smaller than those made by (U)B3LYP. For example, for the favored PCET mechanism for the reaction of ethenoxyl with ethenol, the (U)MPW1K/aug-cc-vPTZ enthalpy of activation is 9.7 kcal/mol closer than the (U)B3LYP/aug-cc-pVTZ value to the G3 value of $\Delta H^{\ddagger} = 12.9$ kcal/mol. Similarly, the barriers to the PCET reactions of a phenoxyl radical with phenol and with α-naphthol are calculated by (U)MPW1K/6-31+G(d,p) to be higher than those predicted by (U)B3LYP/6-31+G(d,p) by, respectively, 8.2 and 7.8 kcal/mol. The (U)MPW1K activation energy of $E_a = 2.7$ kcal/mol for the latter reaction is in excellent agreement with the experimental value of $E_a = 2.2 \pm 0.3$ kcal/mol.

Because several recent studies of PCET reactions have been based on (U)B3LYP calculations,[3,4a−c,g] the barrier heights for these reactions were almost certainly underestimated by the calculations. For future DFT calculations on PCET reactions, the use of Truhlar's MPW1K functional is strongly recommended by the results reported here, because MPW1K is likely to provide much more reliable results than B3LYP.

Supporting Information Available: Geometries and energies of all the minima and TSs on the potential energy surfaces for the hydrogen transfer reactions of ethenol/ethenoxyl, phenol/phenoxyl, and α-naphthol/phenoxyl (40 pages). This information is available free of charge via the Internet at http://pubs.acs.org.

Acknowledgment. We thank the National Science Foundation for support of this research at both U. W. and U. N. T. The Robert A. Welch Foundation and NIH also provided partial support of the research done respectively at U. N. T. and U.W.

References

(1) See, for example, (a) *Free Radicals*; Kochi, J. K., Ed.; Wiley: New York, 1973. (b) Perkins, M. J. *Free Radical Chemistry*; Ellis Horwood: New York, 1994. (c) Olah, G. A.; Molnár, Á. *Hydrocarbon Chemistry*; Wiley: New York, 1995.

(2) (a) Cukier, R. I.; Nocera, D. G. *Annu. Rev. Phys. Chem.* **1998**, *49*, 337−369. (b) Stubbe, J.; van der Donk, W. A. *Chem. Rev.* **1998**, *98*, 705−762. (c) Hammes-Schiffer, S. *Acc. Chem. Res.* **2001**, *34*, 273−281. (d) Stubbe, J.; Nocera, D. G.; Yee, C. S.; Chang, M. C. Y. *Chem. Rev.* **2003**, *103*, 2167−2201. (e) Mayer, J. M. *Annu. Rev. Phys. Chem.* **2005**, *55*, 363−390.

(3) Mayer, J. M.; Hrovat, D. A.; Thomas, J. L.; Borden, W. T. *J. Am. Chem. Soc.* **2002**, *124*, 11142−11147.

(4) For some other recent computational studies of PCET reactions, see: (a) Siegbahn, P. E. M.; Blomberg, M. R. A.; Crabtree, R. H. *Theor. Chem. Acc.* **1997**, *97*, 289−300. (b) Siegbahn, P. E. M.; Eriksson, L.; Himo, F.; Pavlov, M. *J. Phys. Chem. B* **1998**, *102*, 10622−9. (c) O'Malley, P. J. *J. Am. Chem. Soc.* **1998**, *120*, 11732−11737. (d) Carra, C.; Iordanova, N.; Hammes-Schiffer, S. *J. Am. Chem. Soc.* **2003**, *125*, 10429−10436. (e) Olivella, S.; Anglada, J. M.; Solé, A.; Bofill, J. M. *Chem.—Eur. J.* **2004**, *10*, 3404−3410. (f) Anglada, J. M. *J. Am. Chem. Soc.* **2004**, *126*, 9809−9820. (g) DiLabio, G. A.; Ingold, K. U. *J. Am. Chem. Soc.* **2005**, *127*, 6693−6699.

(5) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.

(6) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.

(7) Lynch, B. J.; Fast, P. L.; Harris, M.; Truhlar, D. G. *J. Phys. Chem. A* **2000**, *104*, 4811−4815.

(8) Baboul, A. G.; Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. *J. Chem. Phys.* **1999**, *110*, 7650−7657.

(9) Foti, M.; Ingold, K. U.; Lusztyk, J. *J. Am. Chem. Soc.* **1994**, *116*, 9440−9447. Similar activation parameters for abstraction of the hydroxyl hydrogen atom from β-naphthol by a phenoxyl radical were also reported in this paper.

(10) Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, *80*, 3265−3269. Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. v. R. *J. Comput. Chem.* **1983**, *4*, 294−301. Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257−2261.

(11) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796−6806.

(12) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213−222.

(13) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.

(14) Selected geometries were also reoptimized with the 6-311+G(2df,2p) basis set. However, the relative (U)MPW1K and (U)B3LYP energies at these geometries were essentially the same as the single-point aug-cc-pVTZ energies, computed at the geometries optimized with the 6-31+G(d,p) basis set.

(15) It should be recalled that MPW1K also underestimated the barrier heights in the 40-reaction test set used by Truhlar and co-workers, but B3LYP underestimated the barrier heights by nearly 4 times as much as MPW1K.[7]

(16) See, for example: Houk, K. N.; Rondan, N. G.; Mareda, J. *Tetrahedron* **1985**, *41*, 1555−1563.

(17) Although MPW1K predicts that the enthalpy of the hydrogen-bonded complex is 6.7 kcal/mol lower than that of the isolated reactants, the complex is also calculated to be lower in entropy by 28.1 cal/mol K. Thus, at 298 K, $\Delta G = 1.7$ kcal/mol for complex formation, and the equilibrium constant, in terms of mole fractions of phenoxyl, α-naphthol, and the hydrogen-bonded complex formed from them is $K = 5 \times 10^{-2}$. In the solvent mixtures used by Fotie et al., the equilibrium constant for concentrations, expressed in mol/L, is about a factor of 10 smaller than the value for mole fractions. The highest concentrations of α-naphthol used were 0.3 M, so under these conditions, the ratio of hydrogen-bonded complex to free phenoxyl radical would have been on the order of 10^{-4}. There are obvious inaccuracies in using enthalpy and entropy, computed for the gas phase, to calculate the equilibrium constant for hydrogen-bonded complex formation in the solvent mixtures used in ref 9. Nevertheless, the very small equilibrium constant that we do obtain provides good reason to believe that the phenoxyl

does not react with α-naphthol in an irreversibly formed hydrogen-bonded complex between them. Instead, our calculations predict that complex formation is reversible, so that the overall enthalpy of activation should be based on the free reactants, rather than on the hydrogen-bonded complex between them.

(18) A bimolecular reaction between phenoxyl and α-naphthol, rather than a unimolecular reaction of an irreversibly formed hydrogen-bonded complex, is more consistent with the experimental value of log $A = 8.9 \pm 0.3$ for the reaction.[9]

CT050282Z

JCTC Journal of Chemical Theory and Computation

Cation-π Interactions in Serotonin: Conformational, Electronic Distribution, and Energy Decomposition Analysis

Jaturong Pratuangdejkul,[†,‡,#] Pascale Jaudon,[○] Claire Ducrocq,[§] Wichit Nosoongnoen,[†,‡,¥] Georges-Alexandre Guerin,[∥] Marc Conti,[⊥] Sylvain Loric,[£] Jean-Marie Launay,[†,‡] and Philippe Manivet*[,†,‡]

Service de Biochimie et de Biologie Moléculaire, IFR 139, Hôpital Lariboisière, 2, rue Ambroise Paré, 75475 Paris Cedex 10, France, E.A. 3621, Laboratoire de Biologie Cellulaire, UFR des Sciences Pharmaceutiques et Biologiques, 4 Avenue de l'Observatoire, 75270 Paris, Cedex 06, France, I.C.M.M.O. Laboratoire de Chimie Structurale Organique, Université Paris-Sud, Bat. 410, 91405 Orsay, Cedex, France, Institut de Chimie des Substances Naturelles, CNRS, F-91198 Gif-sur-Yvette, France, BioQuanta Corp., 2850 South Parker Road (S) 720, Aurora, Colorado 80014, Service de Biochimie 1, Hôpital Bicêtre, 78, rue du Général Leclerc, 94275 le Kremlin-Bicêtre Cedex, France, 7086, 1 rue Guy de la Brosse, 75005 Paris, France, and Laboratoire de Biochimie-Génétique, Hôpital Henri Mondor, 51, rue du Maréchal de Lattre de Tassigny, 94010 Créteil, France

Received January 25, 2006

Abstract: An adiabatic conformational analysis of serotonin (5-hydroxytryptamine, 5-HT) using quantum chemistry led to six stable conformers that can be either +gauche (**Gp**), -gauche **(Gm**), and anti (**At**) depending upon the value taken by ethylamine side chain and 5-hydroxyl group dihedral angles ϕ_1, ϕ_2, and ϕ_4, respectively. Further vibrational frequency analysis of the **GmGp**, **GmGm**, and **GmAt** conformers with the 5-hydroxyl group in the anti position revealed an additional red-shifted N−H stretch mode band in **GmGp** and **GmGm** that is absent in **GmAt**. This band corresponds to the 5-HT side-chain N−H bond involved in an intramolecular nonbonded interaction with the 5-hydroxy indole ring. The influence of this nonbonded interaction on the electronic distribution was assessed by analysis of the spin−spin coupling constants of **GmGp** and **GmGm** that show a marked increase for C_2-C_3 and C_8-C_9 bonds in **GmGm** and **GmGp**, respectively, with a decrease of their double bond character and an increase of their length. The Atoms in Molecules (AIM), Natural Bond Orbital (NBO), and fluorescence and CD spectra (TDDFT method) analyses confirmed the existence in **GmGp** and **GmGm** of a through-space charge-transfer between the HOMO and the HOMO-1 π-orbital of the indole ring and the LUMO σ^* N−H antibonding orbital of the ammonium group. The strength of the cation-π interaction was determined by calculating binding energies of the NH_4^+/5-hydroxyindole complexes extracted from stable conformers. The energy decomposition analysis indicated that cationic-π interactions in the **GmGp** and **GmGm** conformers are governed by the electrostatic term with significant contributions from polarization and charge transfer. The lower stability of the **GmGm** over the **GmGp** comes from a higher exchange repulsion and a weaker polarization contributions. Our results provide insight into the nature of intramolecular forces that influence the conformational properties of 5-HT.

1. Introduction

Serotonin (5-hydroxytryptamine, 5-HT) belongs to the group of monoamines neurotransmitters that includes catecholamines (dopamine, noradrenaline, adrenaline) and histamine. 5-HT is widely distributed in various organisms and controls a wide variety of physiological and behavioral processes[1]

* Corresponding author e-mail: crcb.manivet@lrb.ap-hop-paris.fr.
† Hôpital Lariboisière.
‡ UFR des Sciences Pharmaceutiques et Biologiques.
§ Institut de Chimie des Substances Naturelles.
∥ BioQuanta Corp.
⊥ Hôpital Bicêtre.
○ Université Paris-Sud.
£ Hôpital Henri Mondor.
On leave from the Department of Microbiology, Faculty of Pharmacy, Mahidol University, 447 Sri-Ayudhaya Rd., Rajtaevee Phayathai, Bangkok 10400, Thailand.
¥ On leave from the Department of Pharmacy, Chaiyaphum Hospital, Chaiyaphum province, Thailand 36000.

10.1021/ct0600316 CCC: $33.50
Published on Web 03/08/2006

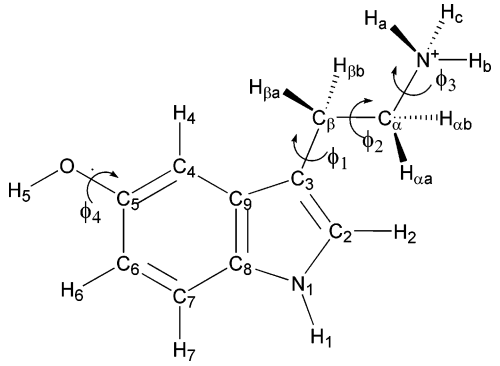


Figure 1. Atomic numbering scheme and definition of dihedral angles of the 5-HT molecule. The four dihedral angles are defined as ϕ_1 ($C_9-C_3-C_\beta-C_\alpha$), ϕ_2 ($C_3-C_\beta-C_\alpha-N$), ϕ_3 ($C_\beta-C_\alpha-N-H_a$), and ϕ_4 ($C_4-C_5-O-H_5$).

mainly through 15 different receptors types.[2] The intensity and duration of 5-HT signaling is determined by its reuptake into cells through the 12 transmembrane domain (TMDs) 5-HT transporter.[3] The important physiological roles of 5-HT and its implications in many pathological states has stimulated intensive research on the design of drugs targeting the serotonergic system. Since drug design needs a clear understanding of the physicochemical properties of the natural compound that is generally used as reference for in silico protocols, it is of prime importance to explore the conformational properties of 5-HT.

At physiological pH, 5-HT is protonated[4] and exists in several conformers due to the flexibility of its ethylamine side chain and the value taken by dihedral angles ϕ_1 and ϕ_2 (see Figure 1 for definitions of ethylamine side-chain dihedral angles). The two 5-HT dihedral angles ϕ_1 and ϕ_2 can be either in +gauche (**Gp**), −gauche (**Gm**), or anti (**At**) conformation, and the 5-HT conformers will be henceforth named in the text as for example **GpAt** when ϕ_1 is in the +gauche conformation and ϕ_2 in the anti conformation. Conformational analysis of 5-HT has been performed using various theoretical calculations and experimental methods. Dating back to the early 1970s, several theoretical calculations have shown the existence of low-energy conformers of 5-HT in the gas phase,[5] and empirical and semiempirical PCILO quantum mechanical calculations suggested that both the **GpGp** and **GpAt** conformers exist in aqueous solution,[6] a fact confirmed by NMR experiments.[7] In the crystalline state, 5-HT exhibits either a **GpGp**, **GpGm**, or **GpAt** conformation depending on the nature of the salt used.[8] Later, based on this information, a mixed model including explicit water solvent without periodic boundary conditions and an empirical force field containing a Coulombic term with a distance-dependent dielectric constant was used to perform molecular dynamics simulations on 5-HT.[9] These simulations showed the existence of the ϕ_1 and ϕ_2 in the **Gp**, **Gm**, and **At** conformations in water, whereas only the **Gp** and **Gm** conformations were observed in the gas phase. A protonated 5-HT conformational analysis has been performed at the B3LYP/6-31G(d) theory level.[10] The authors have reported a set of stable conformers including mirror images and corresponding transition states in the gas phase, with their relative conformational stability in various solvents modeled implicitly by using a polarized continuum model. More recently, gas-phase conformational analysis of protonated 5-HT has been performed at the B3LYP/6-31G(d) up to B3LYP/6-311++G(d,p) theory levels,[11] in aqueous solution using a continuum solvent model. Unfortunately, all the available theoretical studies on 5-HT lack a real description of the nature and amplitude of the intramolecular interactions that influence the conformational stability of 5-HT. The main reason comes from the small basis sets used in the calculations to describe weak nonbonded interactions such as charge transfer, one of the components of cation-π interactions, that need extended basis sets to be correctly described. Since 1971, it is widely admitted that cation-π interactions between indole π-electrons and the cationic head of 5-HT are at the origin of the higher stability of the **GmGp** and the **GmGm** conformers over the **GmAt** conformer.[5b,9] Hence, a non-negligible charge transfer could account mainly for the stabilizing forces in those cation-π interactions. Despite some theoretical calculations of energy decomposition that have been performed on ammonium/pyrole or benzene complexes,[12] until now, no quantum chemistry calculations have been carried out to assert the real existence of such interactions in 5-HT. The goal of the present study was thus to investigate the physicochemical nature of the intramolecular nonbonded forces driving the conformational behavior of 5-HT. For this purpose, we have first studied the whole molecule of 5-HT by performing an adiabatic conformational search in the gasphase in order to find local energy minima and transition-state conformations. The calculated NMR shifts and spin−spin coupling constants of three selected local minima, i.e. **GmGp**, **GmGm**, and **GmAt**, were in good agreement with available NMR experimental data. The calculated IR spectra as well as the topological and natural bond orbital analyses revealed the existence of a real through-space intramolecular charge transfer between the indole π-orbitals and the cationic head N−H antibonding orbital. To further appreciate the relative contribution of this charge transfer in the cation-π interaction with those of other dispersive and nondispersive forces, we performed an energy decomposition analysis on NH_4^+/5-hydroxyindole complexes extracted from the three whole 5-HT local minima optimized geometries.

Our investigations enabled us to provide additional and important insight on the cation-π interactions present in the 5-HT molecule. Furthermore, our results will be useful for the development of accurate force fields for 5-HT, 5-HT derivatives, and 5-HT/5-HT receptor or transporter modeling, that are important for the preparation of new drugs.

2. Computational Methods

An adiabatic conformational analysis of 5-HT was performed by setting the dihedral angle ϕ_1 to 90° and −90° and by rotating ϕ_2 by increments of 60°, while ϕ_3 was left free and ϕ_4 was assigned to 0° or 180° (for definition of dihedral angles, see Figure 1). At each point of the conformational search, the geometry of 5-HT was fully optimized using the DFT-B3LYP method[13] and the 6-31G(d) basis set in the gas phase. All the geometries and energies of stable conformers were further refined by performing a full optimization at the

B3LYP/6-31+G(d,p) theory level, followed by a vibrational frequency analysis in order to verify that each stationary point was a real minimum.

The transition state between two stable conformers was located using the QST3[14] method. The initial guess transition state was obtained from geometry selected along the molecular dynamics (MD) trajectory of 5-HT. Langevin MD simulations were performed in the microcanonical ensemble for 1000 ps. Bond lengths were constrained with the SHAKE algorithm, and the integration time step was 2 fs using the CHARMM molecular mechanics package (see the Supporting Information for details).[15] Finally, vibrational frequency analysis was performed on all optimized transition states in order to reassert their reality.

Infrared (IR) stretch spectra, i.e. wavenumbers (ν) in cm^{-1} and intensities (I) in km/mol, of 5-HT conformers were obtained using vibrational frequencies calculated at the B3LYP/6-31+G(d,p) level. A scaling factor of 0.9605 was applied in order to bring the fundamental stretch of indole N−H into agreement with previously published experimental data.[16]

The NMR isotropic shieldings (σ) were calculated using the Gauge-Including Atomic Orbitals (GIAO)[17] method at the B3LYP/6-311+G(2d,2p) theory level. Calculations were performed on the geometries optimized in the gas phase at the B3LYP/6-31+G(d,p) theory level. The NMR shifts (δ) for 5-HT were calculated using as references tetramethylsilane (TMS) for the C-atom and *tert*-butyl alcohol for the proton. The one-bond NMR spin−spin coupling constants $^1J(A,B)$[18] associated with atoms A and B were also determined.

A topological analysis was performed in order to calculate the charge density (ρ) and its second Laplacian derivative of charge density ($\nabla^2\rho$) for the bond critical points (BCP) using the Bader's Atoms in Molecules (AIM) theory.[19] To evaluate the direction and magnitude of the charge-transfer interactions, the Natural Bond Orbital (NBO)[20] analysis of 5-HT conformers was performed using the NBO 3.1 program.[21] Moreover, the effects of electron delocalization and the vertical excitation energies for the singlet excited states of 5-HT conformers were calculated in the gas phase using the time-dependent density functional theory (TDDFT) method.[22] All calculations were performed using the B3LYP/6-311+G(2d,2p) theory level based on the B3LYP/6-31+G(d,p) optimized geometries.

To get a better understanding on the nature and strength of cation-π interactions, NH_4^+/5-hydroxyindole complexes were generated using geometries of both the 5-hydroxyindole ring and the protonated amine group, which were extracted from whole optimized geometries of 5-HT local energy minima. A hydrogen was added to C_3(sp3) and N(sp3) atoms in order to fill the unsatisfied valence resulting from the broken C_β(sp3)−C_3(sp2) and C_α(sp3)−N(sp3) bonds. Binding energies (BE)[23] of the complexes were refined from single point energy calculations using the HF, B3LYP, and MP2 methods at the 6-311+G(d,p) basis set. The basis set superposition error (BSSE) for each complex was estimated using the full counterpoise method[24] at each level of calculation.

To emphasize the constraints imposed on internal coordinates in 5-HT by intramolecular nonbonded interactions, we fully relaxed and optimized the 5-hydroxyindole ring and the protonated amines group separately at the B3LYP/6-31+G(d,p) theory level. We therefore compared their internal energy within the complexes and after optimization by calculating the 5-HT deformation energy (ΔE_{def}).

Energy of the complexes was decomposed at the HF/6-311+G(d,p) theory level, into physical meaningful terms[25] of individual energetic components using the Kitaura and Morokuma (KM)[26] and reduced variational space self-consistent field (RVS)[27] decomposition analysis.

All quantum mechanics calculations were performed using the Gaussian 03W[28] program package, except for the KM and RVS decomposition analysis which were performed using Gordon and Chen[29] analysis in the GAMESS USA[30] package of programs. All 5-HT structures and molecular orbitals were visualized using the ChemCraft 1.4 beta software.[31] MD simulations were performed using CHARMM 31 running on Silicon Graphics O_2 R12000 workstation under the IRIX 6.5 operating system.

3. Results and Discussion

3.1. Conformational Analysis of 5-HT. 3.1.1. Gas-Phase Adiabatic Conformational Analysis of 5-HT. Potential energy surface (PES) scans are helpful for generating accurate empirical mechanics force field parameters for 5-HT. The adiabatic conformational search performed at the B3LYP/6-31G(d) theory level revealed the existence of six stable conformers of 5-HT and their six mirror images, in good agreement with previously published data.[10,11] Gas-phase energies in Hartrees, the values of 5-HT dihedral angles (ϕ_1, ϕ_2, and ϕ_4), and all the conformers and their relative energies are shown in Figure S1 and Tables S1 and S2 in the Supporting Information. The results show that conformers **1**−**6** and their corresponding mirror image conformers **7**−**12** are redundant since they have the same internal coordinates and relative energies. The results show that dihedral angles ϕ_4 of 5-hydroxyl group in conformers **1**−**3** are in the anti-conformation (OH-anti), while those in conformers **4**−**6** are in the syn-conformation (OH-syn). Depending upon the value of ϕ_2 dihedral angle, with ϕ_1 set in the **Gm** position, 5-HT conformers can be as follows: (i) +gauche (**GmGp**) when it folds on the phenol ring side (conformers **1**, **4**), (ii) −gauche (**GmGm**) when it folds on the pyrole ring side (conformer **2**, **8**), or (iii) anti (**GmAt**) when it extends from the indole ring (conformers **3**, **9**).

Energetically, OH-anti conformers are more stable than the corresponding OH-syn conformers. For OH-anti conformers, the **GmGp** conformer (**1**) is the global minima followed by the **GmGm** conformer (**2**). For OH-syn conformers, **GmGp** and **GmGm** are the next most stable. In the gas phase whatever the position (anti or syn) of the 5-hydroxy group, the lowest energy conformers correspond to the **GmGp** and **GmGm** conformers. However, their difference in energy is not significant. The origin of this stability is due to the stabilizing electrostatic interactions

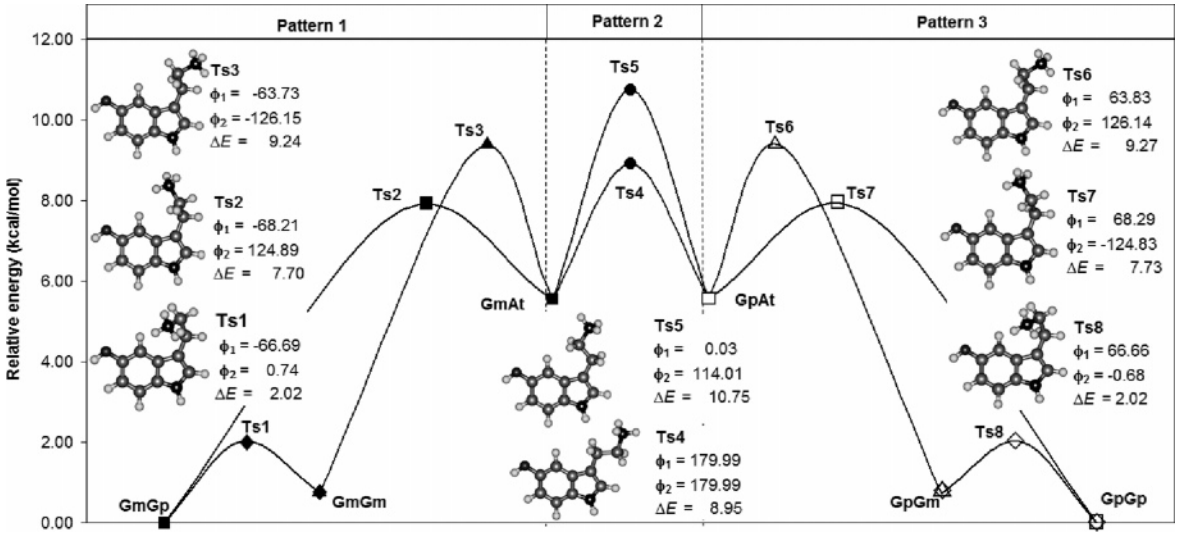


Figure 2. Relative energies (ΔE) in kcal/mol, dihedral angles (ϕ_1 and ϕ_2) in degree of local minima and Ts conformers, and structures of all Ts localized along the interconversion pathway between 5-HT OH-anti local minimum conformations. Calculations were performed using the QST3 Method at the B3LYP/6-31+G(d,p) theory level. Pattern 1 (black squares, triangles, and diamonds) and pattern 3 (white squares, triangles, and diamonds) refer to the two groups of mirror images defined by the value of the ethylamine side-chain dihedral angle ϕ_1. Pattern 2 (black and white circles) refers to the Ts structures between the two groups of mirror images.

occurring between the π-electron cloud of the indole ring and the cationic protonated amine group of 5-HT.

Reoptimization of all the conformers at the B3LYP/6-31+G(d,p) level did not modify significantly their internal coordinates value or their relative energies (see Tables S1 and S2 in the Supporting Information).

3.1.2. Location of Transition States on the 5-HT Potential Energy Surface (PES). To fully scan the PES of 5-HT, to understand the interconversion process of the ethylamine side chain between two neighboring stable conformers and to confirm the existence of the six local energy minima on the PES, we localized the transition-state (Ts) structures using the QST3 method at the B3LYP/6-31+G(d,p) theory level. To save calculation time for this study, only the most stable 5-HT OH-anti conformers and their corresponding mirror images were considered as representative of the entire population of 5-HT conformers. Using the above quantum chemistry PES study in the gas phase, we generated the first set of empirical mechanics parameters that were used to perform MD simulations of 5-HT. Guess conformations for the Ts study using QST3 calculations were extracted from Langevin MD simulation trajectories by a clustering method (additional MD data are provided in the Supporting Information, see Figure S2). The results of QST3 optimization are shown in Figure 2 including structures, dihedral angles, and energy barriers of Ts linking two stable conformers. The analysis of vibrational frequency spectra of Ts structures revealed only one imaginary frequency indicating that all the conformations were real transition states.

In a recent publication,[10] the transition states between two 5-HT minima have been located using the QST3 method. However, the conformers used as the reactant and the product for QST3 calculations were arbitrarily extracted from the 5-HT PES, without any respect for the reliability of the dynamics of the interconversion process of the ethylamine side chain. For instance, an **GmAt**-like conformer was found as a transition state between two conformers belonging to a different group of mirror images (conformers **2** and **8** in the present work). In our strategy we avoided bias in the choice of guess transition-state conformers for QST3 calculations. This was achieved by selecting adequately reactant and product conformers, based on the analysis of authorized rotation of the ethylamine side-chain dihedral angles along MD trajectories. This strategy requires a good set of empirical molecular mechanics force field parameters. We generated the first set of CHARMM parameters based on the adiabatic conformational analysis (Table S1 in the Supporting Information) performed at the B3LYP/6-31+G(d,p) level and used MD simulations to obtain guess transition-state conformers for the QST3 calculations. Finally, the transition states located using the QST3 method were used to refine the first set of CHARMM parameters and to confirm the reliability of the transition-state conformers. Only one imaginary frequency was found in vibrational frequency analysis for all optimized structures, confirming they are real transition states with a first-order saddle point. Visual inspection of the vectors of the atomic motions for the negative imaginary frequency along the reaction coordinates enabled us to check that the geometry of transition states corresponded to motions required for the mechanistic scheme.

Figure 2 shows that **pattern 1** and **pattern 3** are symmetric since they have the same energy barriers and symmetric geometries (dihedral angles in opposite signs), **Ts1**, **Ts2**, and **Ts3** being symmetric with **Ts8**, **Ts7**, and **Ts6**, respectively. In **pattern 1**, ethylamine side-chain interconversion between conformers **GmGp**, **GmGm**, and **GmAt** passes through transition states where the dihedral angle ϕ_2 is mostly driving the conformational changes. The dihedral angle ϕ_1 undergoes changes imposed by dihedral angle ϕ_2 in order to accommodate the conformation to the intramolecular nonbonded interactions. For **Ts1** and **Ts2**, dihedral angles ϕ_1 are similar.

Figure 3. Structures of the three most stable 5-HT conformers **1**, **2**, and **3** (**GmGp**, **GmGm**, and **GmAt**, respectively). 5-OH groups in anti position.

In **Ts3**, the attraction between the cationic head and the π-electrons of the five-membered ring constrains the dihedral angle ϕ_1 to accommodate to a value of $-63.73°$ to finally get the value of $-64.40°$ when **Ts3** evolves toward the **GmGm** conformer (conformer **2**). Those results, interpreted with the above observation of a higher stability of **GmGp** conformers compared to **GmGm**, suggest that the most important constraints imposed on 5-HT internal coordinates concern **GmGm** conformers. However, the relative energy between **GmGp** and **GmGm** conformers is very small. This reinforces the hypothesis of an existing balance in **GmGm** conformers between stabilizing intramolecular electrostatic interactions between the cationic head and π-electrons of the five-membered ring and strong destabilizing repulsive interactions.

The high relative energy of **Ts3** might be attributed to a progressive loss of stabilizing nonbonded interactions between the cationic head and π-electrons of the indole ring upon rotation of the ethylamine side chain from **GmGm** to **GmAt** conformers. These nonbonded interactions could be sufficient to attract the cationic head and maintain the stability of **Ts3** but too weak to balance the internal coordinate constraints imposed on ϕ_1. Surprisingly, the energy barrier between **GmGm** and **GmGp** is not high (2 kcal/mol) suggesting that interconversion between **GmGp** and **GmGm** conformers is relatively easy by rotation of the ethylamine side chain above the 5-HT ring plane. The largest constraints imposed by steric hindrance on internal coordinates in **Ts1** are bending angles θ_1 and θ_2 that are opened about 3° compared to **GmGp** and **GmGm** conformers (5° and 2° for the **GmAt** conformer θ_1 and θ_2 angles, respectively). The ethylamine side chain accommodates the repulsion upon rotation above the indole ring plane by opening θ_1 and θ_2 angles. During MD simulations, interconversion of the side chain between two groups of mirror images of **patterns 1** and **3** occurs only between **GmAt** and **GpAt**, through **Ts4** and **Ts5**. For **Ts4** and **Ts5** dihedral angle ϕ_1 the value is 0° or 180°, respectively, whereas the dihedral angle ϕ_2 is in anti orientation for both transition states.

3.2. Physicochemical Analysis of 5-HT. To better understand the nature of driving forces that influence the conformational properties of 5-HT, we further investigated the nature of intramolecular interactions and their relative weight in 5-HT. For this purpose, we have focused our attention on the three stable conformers **1**, **2**, and **3** (**GmGp**, **GmGm**, and **GmAt**, respectively) in which the ethylamine side chain is on the same side of the indole ring and 5-OH groups in the anti position (Figure 3 and Tables S2 and S3 in the Supporting Information).

3.2.1. Calculated Infrared (IR) Spectra of 5-HT. Harmonic vibrational frequencies (ν) and IR intensities (I) of three representative conformers (Figure 3) were calculated at the B3LYP/6-31+G(d,p) level. A scaling factor of 0.9605 was applied to frequencies of 5-HT conformers in order to bring the fundamental indole NH stretches into agreement with experimental results of the indole molecule (3525 cm^{-1}).[16a] This sharp band is characteristic of indole derivatives and commonly used for identification of tryptamine and tryptophan derivatives.[16b,32] Seventy-two normal vibrational modes were recorded for the three conformers. However, only frequencies above 2800 cm^{-1} corresponding to the fundamental vibration of stretch modes were analyzed in our study. The stretching modes can be divided into three groups: the heterocyclic system of the indole ring (aromatic CH and indole NH stretches), the ethylamine side chain (alkyl CH and amine NH stretches), and a hydroxyl group (OH stretch). All stretching frequencies and intensities of three 5-HT conformers are listed in Table S4 in the Supporting Information. IR-stretching bands are provided in the Supporting Information in Figure S3.

For the three conformers, the frequencies in the 2800−3050 cm^{-1} region belong to alkyl CH stretches of the ethylamine side chain. The lower frequencies in the 2881−2923 cm^{-1} and the 2985−2979 cm^{-1} regions were characterized as symmetric stretches of $CH_2(\beta)$ and $CH_2(\alpha)$groups, respectively. In addition, frequencies in the 2946−2971 cm^{-1} and the higher-frequency 3048−3035 cm^{-1} regions were identified as antisymmetric stretches of $CH_2(\beta)$ and $CH_2(\alpha)$ groups, respectively. The $CH_2(\beta)$ stretches of the **GmGp** and **GmGm** conformers are quite similar and have higher frequencies and lower intensities than those found for the **GmAt** conformer. The $CH_2(\alpha)$ stretches are similar for the **GmGm** and **GmAt** conformers and have higher frequencies and lower intensities than those of the **GmGp** conformer. The vibrational frequencies in the 3050−3130 cm^{-1} region are attributed to the aromatic CH stretches of the indole ring (3050−3080 cm^{-1} region for the phenyl ring and 3125−3131 cm^{-1} region for the pyrole ring) and are similar for the three conformers.

The alkyl CH stretch frequencies have been claimed to be sensitive to conformational changes in various indole derivatives including the neutral form of 5-HT,[32,33] containing a neutral ethylamine side chain through the interaction of the CH_2 group with the lone pair of amino nitrogen. In the

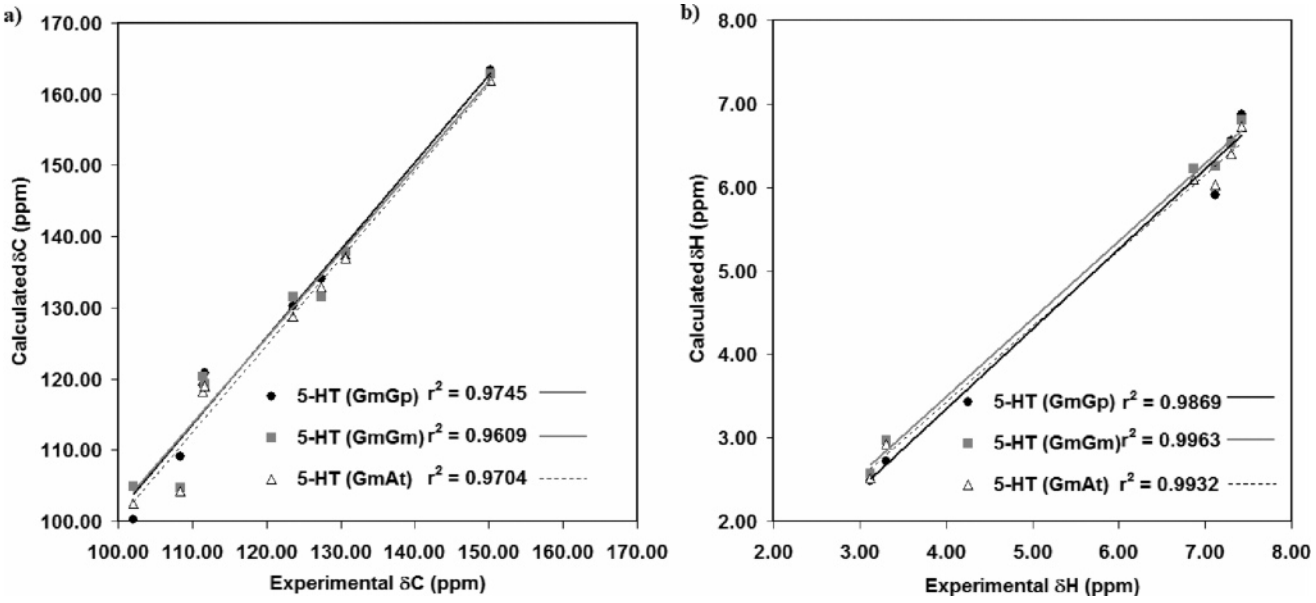


Figure 4. Correlation plots of (a) carbon chemical shift (δC) and (b) H chemical shift (δH) between available experimental and predicted values (Table 1) obtained from the GIAO calculation at the B3LYP/6-311+G(2d,2p) theory level for the three most stable conformers (OH anti, **GmGp**, **GmGm**, **GmAt**).

cationic form of 5-HT, the observed red-shift of $CH_2(\beta)$ stretch frequencies of the **GmGm** and **GmGp** conformers as compared to **GmAt** comes from a phenomenon involving electron transfer from indole π orbital to the CH antibonding orbital of the $CH_2(\beta)$ group. This is confirmed by NBO analysis (see below). The $CH_2(\alpha)$ group is not affected by π electron clouds since it is far away from the indole ring.

Particularities were observed for the three 5-HT conformers in the NH stretch region (3100−3530 cm^{-1}). All 5-HT conformers have four NH stretches: one indole NH stretch and three amine NH stretches. Amine NH stretch frequencies corresponding to the NH bond pointing toward the indole ring of **GmGp** and **GmGm** conformers are strongly red-shifted as compared to that of the **GmAt** conformer, by respectively 114 and 164 cm^{-1}. Hence, very high intensities of those bands were observed in **GmGp** and **GmGm** conformers. In the **GmAt** conformer, no N−H bond is pointing toward the indole ring, and the symmetric vibrational frequency of the three N−H bonds are similar. Two amine NH stretches were identified as antisymmetric stretches. The first antisymmetric NH stretch of **GmGp** and **GmGm** conformers is red-shifted at 3321 cm^{-1} with a higher intensity as compared to the **GmAt** conformer (3358 cm^{-1}). The last antisymmetric NH stretches of **GmGp** and **GmGm** were located in a similar region of frequencies (3377 and 3382 cm^{-1}, respectively), with a blue-shift as compared to the corresponding NH stretch of **GmAt** (3358 cm^{-1}) and a similar intensity. The indole NH and OH stretching modes of the three conformers are located in the 3523 ± 5 cm^{-1} and 3674 ± 3 cm^{-1} regions, respectively, with similar intensities. For amine NH stretches, the relevant red-shifts of harmonic vibrational frequencies and their increase in intensity for **GmGp** and **GmGm** conformers favor the existence of through-space NH-π/indole nonbonded interactions. In the **GmGp** and **GmGm** conformers, the amine $N-H_a$ bond, which points toward the indole ring, is involved in a charge transfer with the indole π-electrons cloud. This is supported by the existence of several physical criteria that are used to appreciate whether a hydrogen bond is established between a hydrogen bond donor and a lone pair of a hydrogen bond acceptor:[16b,34] (i) the elongation of the N−H bond due to the increase of the electron population of the N−H σ^* orbital, (ii) a red-shift of the N−H frequency with enhanced intensity, and (iii) the presence of a bond critical point in the vicinity of the hydrogen bond that we will detail below.

3.2.2. Calculated NMR Isotropic Shielding and Spin−Spin Coupling Constants of 5-HT Conformers. The correlation plots of NMR shifts (Figure 4) show a good agreement between the calculated and the experimental values of carbon chemical shifts δC,[35] for the three conformers. We also evaluated the root-mean-square (rms) deviation of predicted values from experimental ones for each atomic position (Table 1). The largest deviation of calculated δC for the three 5-HT conformers were observed for carbon C_5 (rms deviation of 12 ppm). Although the calculated δC were overestimated, their amplitudes at each position were quite similar to the experimental ones. The calculated H chemical shifts δH of H_2, H_4, H_6, H_7 and the two average values of the side chain H_α and H_β of the three conformers are also in good agreement with experimental data.[35] The three conformers show low rms deviation from predicted values (0.5−0.8 ppm) at each position, except at proton H4 for which rms deviation is 1.08 ppm.

The analysis of calculated NMR isotropic shieldings (σ) revealed that their values change depending upon the conformation adopted by 5-HT (Table S5 in Supporting Information). This is due to a change in the electron density of some atoms in the conformers (see Table S6 in the Supporting Information). This electron redistribution is promoted by the formation of intramolecular interactions between the $N-H_a$ bond of the ethylamine side chain and

Table 1. Experimental and Calculated B3LYP/6-311+G(2d,2p) GIAO δC and δH for the Three Most Stable 5-HT Conformers (OH Anti, **GmGp**, **GmGm**, **GmAt**)

position	exp.	GmGp	GmGm	GmAt	rms
		C-Chemical Shift (δC)			
C_2	123.61	130.29	131.58	128.79	6.71
C_3	108.38	109.06	104.76	104.24	3.20
C_4	101.97	100.24	104.95	102.24	2.00
C_5	150.26	163.36	162.86	161.39	12.31
C_6	111.39	119.17	120.29	118.32	7.91
C_7	111.72	120.92	119.30	118.84	8.02
C_8	130.73	136.82	137.96	137.62	6.75
C_9	127.39	134.09	131.52	132.97	5.57
		H-Chemical Shift (δH)			
H_2	7.30	6.56	6.52	6.35	0.83
H_4	7.12	5.91	6.25	5.99	1.08
H_6	6.87	6.21	6.22	6.13	0.68
H_7	7.42	6.87	6.80	6.65	0.65
H_β(average)	3.12	2.57	2.63	2.52	0.55
H_α(average)	3.31	2.72	2.97	2.94	0.45

the indole ring orbitals. The charge density (calculated with the help of the CHelpG Scheme[36] modifications) concerns mainly carbon atoms C_2, C_3, C_4, C_8, and C_9. In the **GmGm** and **GmGp** conformers, the N−H_a bond points toward C_8−C_9 and C_2−C_3 bonds, respectively, resulting in a withdrawal of electron density located at those bonds. The total potential derived charge of the C_8−C_9 bond (charge C_8 + charge C_9) decreases from 0.23, through 0.28 to 0.32 e^- when 5-HT adopts the **GmGm**, **GmAt**, and **GmGp** conformations, respectively. This effect is most marked on carbon C_9 (from 0.09, through 0.11 to 0.21 e^-). The decreasing electronegativity at carbon C_9 for the **GmGp** conformer and the redistribution of π orbital electrons upon intramolecular interactions led to the decrease of the calculated NMR shieldings $\sigma(C_9)$ that shifted by 1.2 and 2.8 ppm compared to the **GmAt** and **GmGm** conformers. The carbon C_8 charge increases from 0.18, through 0.14 to 0.11 e^- when 5-HT adopts the **GmAt**, **GmGm**, and **GmGp** conformations, respectively, while the **GmGp** calculated NMR shieldings $\sigma(C_8)$ shift by 1.3 and 0.8 ppm in the **GmGm** and **GmAt** conformers, respectively. The total bond charge of the C_2−C_3 bond decreases from −0.24, through −0.20, to −0.14 e^- when 5-HT adopts the **GmGp**, **GmAt**, and **GmGm** conformations, respectively. This effect is not particularly predominant at carbon C_2 or C_3 and the correlation between charge decreases and calculated NMR shieldings is more difficult to establish since electron delocalization and redistribution mitigates the influence of N−H_a. Carbon C_3 undergoes the influence of ethylamine side chain and the decrease of charge (from −0.16, through −0.12, to −0.04 e^- when 5-HT adopts the **GmGp**, **GmAt**, and **GmGm** conformation, respectively), is accompanied by a decrease of the **GmGm** conformer calculated NMR shieldings $\sigma(C_3)$ that shifted by 0.5 ppm compared to the **GmAt** conformer and an increase of $\sigma(C_3)$ that shifted by 5.3 ppm compared to the **GmGp** conformer. This lack of correlation between charges and NMR-carbon shifts $\delta(C_3)$ could be attributed to a greater deviation of the calculated shieldings from the experimental ones. Indeed, Table 1 shows that the greatest deviation between experimental and calculated shifts concerns $\delta(C_3)$.

An additional atomic center of interest is represented by C_4. Carbon C_4 undergoes a change in the natural charge from −0.38, through −0.39, to −0.44 e^- when 5-HT adopts the **GmAt**, **GmGm**, and **GmGp** conformations, respectively. A decrease of the **GmGp** conformer calculated NMR shielding $\sigma(C_4)$ shifts by 2.2 and 4.5 ppm in the **GmAt** and **GmGm** conformers, respectively.

For protons, the most significant modifications of NMR isotropic shieldings upon 5-HT conformational changes are observed for the amine proton H_a in gauche conformers where the N−H_a bond interacts with the indole ring π orbital (Table S5 in the Supporting Information). In those conformers, the proton H_a is localized within the diamagnetic deshielding zone of the six- and the five-membered rings for the **GmGp** and **GmGm** conformers, respectively. As a result, chemical shifts of H_a in the the **GmGm** and **GmGp** conformers are upfield (decrease of $\sigma(H_a)$) with respect to its shift in the **GmAt** conformer and to the shifts of the two other protons of the amine side chain H_b and H_c. To support this we calculated the shielding increment, $\Delta\delta$H, for proton H_a in the **GmGm** and **GmGp** conformers by subtracting H isotropic shielding calculated for optimized ammonium molecule alone to the corresponding H isotropic shielding in the whole 5-HT molecule at the same level of theory (Table S7 in the Supporting Information). An increase of the shielding increment, $\Delta\delta$H, for H_a confirmed the deshielding of the proton in the presence of the π orbitals in 5-HT. An additional deshielding effect on proton H_a is due to the electronegativity of the cationic head nitrogen atom that decreases the electronic density around the proton H_a. This phenomenon is due to the strong polarization of the N−H_a bond when it points either to the C_2−C_3 or to the C_8−C_9 bond in the **GmGm** or the **GmGp** conformer, respectively. We will demonstrate this polarization below, using the energy decomposition analysis.

The observation of calculated one-bond nuclear spin−spin coupling constants (^1J) of 5-HT shows a decrease of the ^1J upon conformational change from the **GmAt** to the **GmGp** and **GmGm** conformers. This is indicated in Table 2 that shows significant differences of calculated ^1J for C_2−C_3, C_3−C_9, C_4−C_9, and C_8−C_9 bonds and modification of their length. If the **At** conformer is used as a reference for internal coordinates, we can show for C_2−C_3, C_3−C_9, and C_4−C_9 bonds an elongation by 0.0018, 0.0024, 0.0041, and 0.0058 Å for the **GmGp** conformer and a bond length increase by 0.0059, 0.0055, and 0.0012 Å for the **GmGm** conformer. As compared to the **GmAt** conformer, calculated $^1J(C_2,C_3)$, $^1J(C_3,C_9)$, and $^1J(C_4,C_9)$ decreased by (i) 2.3, 3.5, and 2.4 Hz and (ii) 4.7, 2.7, and 0.8 Hz for the **GmGp** and **GmGm** conformers, respectively. The decrease of $^1J(C_2,C_3)$, $^1J(C_3,C_9)$, and $^1J(C_9,C_4)$ coupling constants in the **GmGp** and **GmGm** conformers is mainly due to a decrease of their FC contribution (diamagnetic spin−orbit (DSO), paramagnetic spin−orbit (PSO). Fermi contact (FC) and spin dipolar (SD) contributions of $^1J(C_2,C_3)$, $^1J(C_3,C_9)$, and $^1J(C_4,C_9)$ are provided in Table S8 in the Supporting Information). Since

Table 2. Spin−Spin Coupling Constants ^1J(A,B) in Hz of Atoms in 5-Hydroxyindole Ring Using GIAO Method at the B3LYP/6-311+G(2d,2p) Level and Bond Distance in Å (Differ from Bonds in At)

	^1J(A,B)			bond distance		
bond	**Gp**	**Gm**	**At**	**Gp**	**Gm**	**At**
N_1-C_2	10.3	10.4	10.6	−0.0001	−0.0014	1.3713
C_2-C_3	75.1	72.7	77.4	0.0018	0.0059	1.3805
C_3-C_9	53.8	54.6	57.3	0.0041	0.0055	1.4425
C_8-C_9	57.5	57.8	58.0	0.0024	0.0008	1.4192
N_1-C_8	11.0	10.6	10.8	−0.0027	0.0005	1.3870
C_4-C_9	65.8	67.4	68.2	0.0058	0.0012	1.4068
C_4-C_5	77.4	77.1	77.8	0.0016	0.0006	1.3909
C_5-C_6	66.4	66.2	67.0	0.0006	0.0011	1.4140
C_6-C_7	65.8	65.9	66.1	−0.0004	−0.0010	1.3900
C_7-C_8	70.5	70.4	71.0	0.0005	0.0005	1.3980
C_5-O	24.3	24.4	24.2	−0.0032	−0.0034	1.3696
$N-H_a$	50.47	49.43	50.55	0.0097	0.0130	1.0258
$N-H_b$	50.17	51.45	50.09	−0.0007	−0.0017	1.0254
$N-H_c$	50.39	50.17	50.46	−0.0017	−0.0012	1.0256

the FC operator[18] probes the s-electrons at the sites of coupling nuclei, the decrease of the FC contribution of these ^1J coupling constants infers a decreasing s-character of the σ orbital of C_2-C_3, C_3-C_9, and C_4-C_9 bonds. We have also noted a decrease of the SD contribution of the multiple bonds, indicating a decrease of the π-character leading to a weakening of the bonds.[37] However this decrease is more marked for some bonds, depending on the conformation adopted by 5-HT, **GmGp** or **GmGm**. For instance, concerning the **GmGm** conformer, the SD contribution decreased dramatically for the C_2-C_3 bond and to a lesser extent for the C_3-C_9 bond, since the $N-H_a$ bond points to C_2-C_3. For the **GmGp** conformer, the SD contribution decreased for the C_4-C_9 bond. This is consistent with our above observations on electron redistribution, isotropic NMR shielding, and FC contributions. All of these observations favor the existence of a plausible cation-π interactions between $N-H_a$ and the indole ring orbitals where a charge transfer from the π orbital of the 5-hydroxyindole ring to the protonated amine group could occur. In turn, this should affect the conjugated double bond in the indole ring, modifying their bond character. In the **GmGp** conformer we note a decrease of the s-character of the C_8-C_9 bond σ orbital that accounts for the decrease of the FC contribution to the $^1J(C_8,C_9)$ coupling. We also note a decrease of the SD contribution, indicating a weakening of the C_8-C_9 bond. Surprisingly, the value of $^1J(C_8,C_9)$ remains quite constant for the three conformers. The calculated SD and PSO contributions of $^1J(C_8,C_9)$ change (decrease of SD and increase of PSO) as compared to the **GmGm** and **GmAt** conformers, reflecting the change in bond multiplicity. Since the variation of the SD and PSO contribution of $^1J(C_8,C_9)$ compensate each other and the FC contribution change remains subtle, $^1J(C_8,C_9)$ is only weakly different among the three conformers.

3.2.3. Topological Analysis of the Electronic Density. To characterize the formation of intramolecular hydrogen bonds in 5-HT conformers, the topological analysis of the electron density using the "atoms in molecules" (AIM) theory was applied.[38] We used this method to detect bond paths inside the 5-HT molecule and critical points associated with them. Two main kinds of parameters are provided by AIM analysis: the charge densities (ρ) and the corresponding density laplacians ($\nabla^2\rho$) at the bond critical points (BCPs). The value of the electronic density at the BCP for a given bond, $\rho(r_c)$, can be correlated to the concept of bond order[38] with higher values of $\rho(r_c)$ corresponding to stronger bonds. The positive sign of the Laplacian of the charge density at the BCP, $\nabla^2\rho(r_c)$, corresponds to a closed shell interaction that is responsible for the bonding[38] where the electronic charge is concentrated around each nucleus. This is the typical case for ionic or hydrogen bonds. For 5-HT conformers, the existence of a hydrogen bond between the $N-H_a$ bond and the π orbital of the indole ring was confirmed by the presence of BCPs linking amine H_a and indole C_9 ($\rho(r_c) = 0.011564$, $\nabla^2\rho(r_c) = 0.05566$) or C_3 ($\rho(r_c) = 0.019181$, $\nabla^2\rho(r_c) = 0.06119$) atom respectively in conformers **GmGp** and **GmGm**. Our results are in reasonable agreement with the Popolier et al.[38b] criteria used to characterize a D−H···A type hydrogen bond, where D and A are hydrogen bond donor and acceptor atoms, respectively, even if the values of charge density at BCP $\rho(r_c)$ for the **GmGp** and **GmGm** conformers are relatively low compared to the values obtained with strong hydrogen bonded systems with canonical hydrogen bonds. The Popelier et al. criteria also include an increasing charge of the hydrogen atom involved in the hydrogen bond. In our results, the charge of H_a atom of gauche conformers increases as compared to the **GmAt** conformer. We note also for the **GmGm** conformers a higher electron density at the bond critical point as compared to **GmGp.** This is in good agreement with the above observations that suggest the presence of a stronger hydrogen bond in **GmGm**, shown by a more red-shifted and more intense band of N−Ha stretch vibration. Hence, the interatomic distances between H_a and C9 or C3 show that the hydrogen bond is weaker for the **GmGp** ($H_a\cdots C_9 = 2.36$ Å) conformer than for the **GmGm** ($H_a\cdots C_3 = 2.26$ Å) one. All of the above observations confirm the existence of an intramolecular hydrogen bond particularly in the **GmGm** conformer but also in the **GmGp** conformer. The geometrical constraints imposed on the internal coordinates of the ethylamine side chain in the **GmGm** and **GmGp** conformers prevent the formation of an ideal hydrogen bond that is known to have strong anisotropic characteristics. Indeed, it has been shown that the position of the ammonium cation in interaction with the benzene ring in the most stable cation-π complex is bidendate, with two hydrogen atoms pointing toward the benzene ring to form an ideal hydrogen bond.[39] In the **GmGp** and **GmGm** conformers, the position of the cationic head are largely shifted from the centroid of the six- and five-membered rings for the **GmGp** and **GmGm** conformers, respectively. However, in the **GmGm** conformer the cationic head seems to be in a better spatial configuration for the formation of a hydrogen bond than in the **GmGp** one, with a more important overlap of the π_{indole} and σ^*_{N-H} orbitals. We also observed the existence, for the **GmGm** conformer, of a ring critical point between the π_{indole} orbital

Table 3. NBO Analysis of NH-π Indole Remote Contacts in the 5-HT Conformers Calculated Using the B3LYP/6-311+G(2d,2p) Level of Theory

charge-transfer orbital		
donor$_{(i)}$/occupancy (e)	acceptor$_{(j)}$/occupancy (e)	$E^{(2)}$ (kcal/mol)
GmGp		
$\pi(C_8-C_9)$/1.60344	$\sigma^*(N-H_a)$/0.02012	0.56
GmGm		
$\pi(C_2-C_3)$/1.84178	$\sigma^*(N-H_a)$/0.02797	2.80
$\sigma(N-H_a)$/1.99267	$\pi^*(C_2-C_3)$/0.36524	0.56

and the $N-H_a$ bond, confirming a strong through-space cation-π interaction.

3.2.4. Analysis of Electron Distribution in the Ground and Excited States of 5-HT. Since we suspected the presence of a charge transfer between the $N-H_a$ bond and π orbitals of the indole ring, we further studied electron delocalization and their strength by performing the Natural Bond Orbital (NBO) analysis on the ground state of 5-HT. Electron delocalizations were assessed by the estimation of the occupation number and their magnitude determined from an analysis of the off-diagonal elements in the Fock matrix in the NBO basis, taking into account all of the possible donor−acceptor interactions and then calculating the strength of all of them according to second-order perturbation theory, ($E^{(2)}$) (Table 3). The schematic representation of energy levels and contour plots of important molecular orbitals (MOs) in the electronic structures are presented in Figure 5. The results of NBO analysis for the three most stable 5-HT conformers clearly show that the σ^*_{N-Ha} antibonding orbital occupation number in the amine group of the **GmGp** and **GmGm** conformers is strong, suggesting that this orbital is involved in the **GmGp** and **GmGm** conformers in a charge-transfer process, thereby leading to the elongation of the $N-H_a$ bond length, followed by a downshift of stretching vibrational frequency. The two other antibonding orbitals σ^*_{N-Hb} and σ^*_{N-Hc} of the **GmGp** and **GmGm** conformers and the three σ^*_{N-H} orbitals of the **GmAt** conformer have a lower occupation number indicating that they are not involved in any charge transfer. A very strong charge transfer from the π_{C2-C3} orbital to the σ^*_{N-Ha} antibonding orbital is found for the **GmGm** conformer with a stabilization energy $E^{(2)}$ of 2.80 kcal/mol. Hence, a backward direction charge transfer was also found from the donor σ_{N-Ha} orbital to the acceptor π^*_{C2-C3} with a stabilization energy $E^{(2)}$ of 0.56 kcal/mol. A less important, but non-negligible, charge transfer from the π_{C8-C9} orbital to the σ^*_{N-Ha} antibonding orbital was shown for the **GmGp** conformer with a stabilization energy $E^{(2)}$ of 0.56 kcal/mol. Second-order properties, such as the NMR spin−spin coupling constant, can be used together with HOMO-n and LUMO+n gap (HLG) energy to interpret simply the electronic redistribution. The inverse of the energy gap directly enters into the general expression for the coupling constant.[40] The smaller the HLG energy, the larger the spin−spin coupling constants are expected to be. If HLG is taken as a rough estimate of these trends, we can see in Table 2 that the trend for ^1J in the three conformers follow the inverse of HLG. The increase of HLG found for the **GmGp** and **GmGm** conformers as compared to the **GmAt** conformer is due to stabilization of the highest occupied orbitals labeled 46 and 47 (HOMO-1 and HOMO) which are mainly indole ring π orbitals. The increase of the average C−C distances of the indole ring weakens the antibonding character, increases the electronic density at the atomic centers, and thus stabilizes the orbitals. The value of ^1J(C4,C9) and ^1J(C2,C3) for the **GmGp** and **GmGm** conformers respectively are smaller than for the **GmAt** conformer and inversely correlated with HLG. For ^1J(C8,C9) the inverse correlation with HLG is less marked since, as mentioned above, this bond undergoes many electron rearrangements, where SD and PSO contribution compensate and where the density of HOMO and HOMO-1 orbital electrons is high, those orbitals mixing together under perturbations from the nuclear spins.

To further study the electronic structure of 5-HT conformers and rationalize spectroscopic properties of 5-HT, we analyzed the vertical excitation energies of the six low-lying singlet excited states calculated at the optimized geometry of the ground state by using the TDDFT.

The major composition of MOs, the ground states orbital HLG, vertical excitation energies (ΔE) are listed in Table 4. We also reported calculated wavelengths (λ) and oscillator strengths for each state of the three conformers in order to fit with the available experimental spectra. The experimental fluorescence spectra of 5-HT have shown a maximum emission at 336 nm in a wide pH range from 2 to 10 in aqueous solution.[41] This band shifts to 350 nm at pH 12.5 and to 380 nm with a very low quantum yield upon laser excitation at higher pH. The 336, 350, and 380 nm emissions have been assigned to protonated and singly and doubly deprotonated 5-HT, respectively. As shown in Table 4, the first singlet excited state (1^1A) of the **GmGp** conformer corresponds to a calculated wavelength of 336.77 nm, which fits almost perfectly to the emission spectra of 336 nm. Thus, our result strongly confirms this assignment to the cationic form of 5-HT. It is interesting to note that the maximum peak of 5-HT at neutral pH (336 nm) is blue-shifted as compared to that of tryptophan (350 nm) and tryptamine (356 nm), which could be attributed to the hydroxyl group in 5-HT. Those longer wavelength emissions (around 350 nm) have been assigned to the indole-NH_3^+ group interaction.[41] The red-shift of emission at high pH of these molecules has been previously proposed to be due to reduced electrostatic repulsion between the NH_3^+ group and the indole ring.[42] These experimental results provide a satisfactory explanation for the electronic excitation of the **GmGp** conformer. The 1^1A excited state of the **GmGm** conformer presents a calculated wavelength of 357.58 nm (Table 4 and Table S9 in the Supporting Information), which is close to that of tryptamine and to a lesser extent to that of tryptophan. This indicates that the indole-NH_3^+ group interaction is stronger in the **GmGm** conformer than that in the **GmGp** one. Our results can be used to confirm the presence of a charge-transfer between the cationic head of 5-HT and indole π orbitals, shown by the NBO analysis a stronger stabilization energy for **GmGm** than that for **GmGp**. No band could be observed for **GmAt** using the TDDFT calculation, indicating

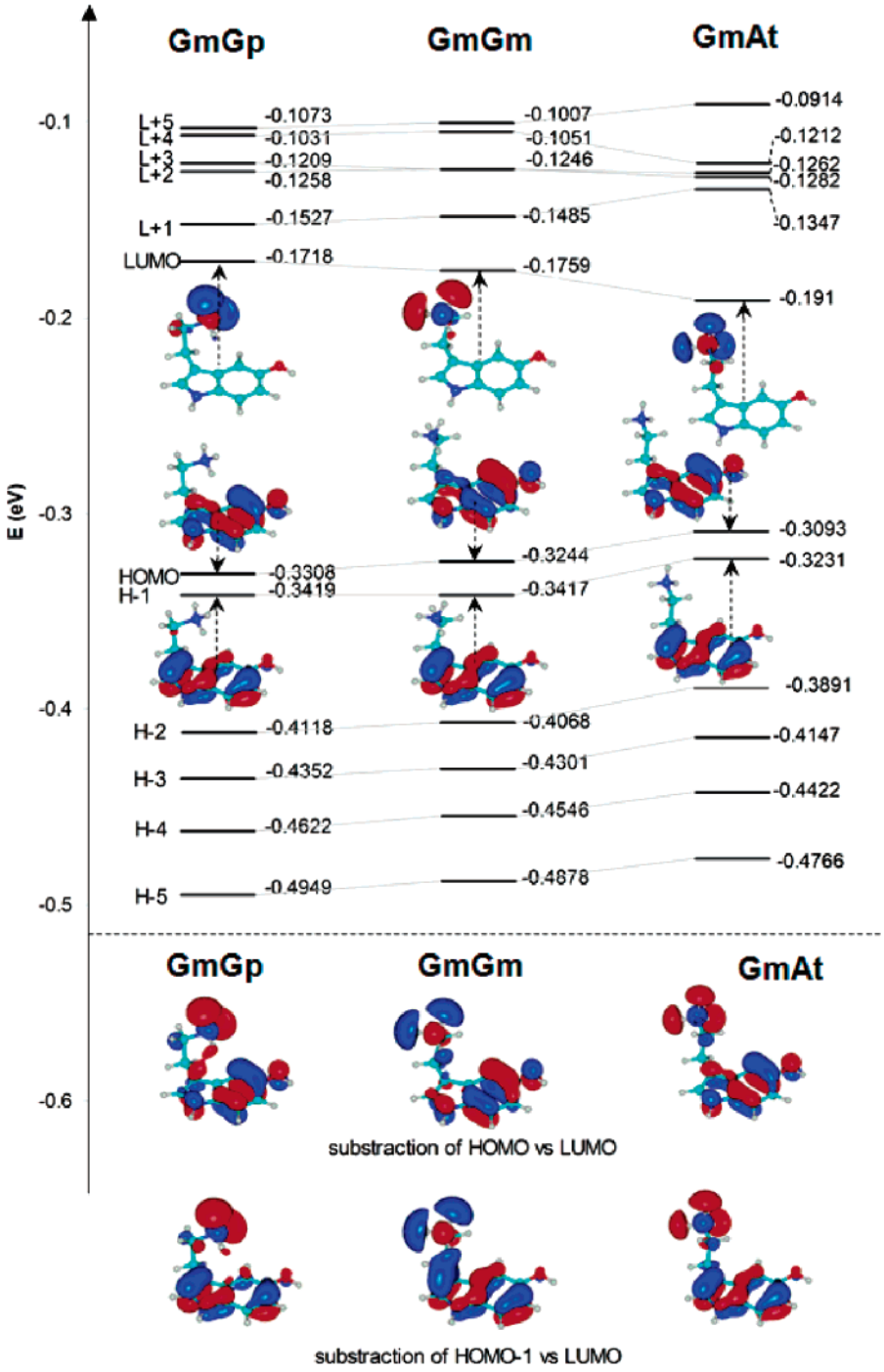


Figure 5. Energy levels (eV) of the 5-HT conformers (**GmGp**, **GmGm**, and **GmAt**) at ground-state calculated using population analysis in gas phase at the B3LYP/6-311+G(2d,2p) level of theory. The HOMO, HOMO-1, LUMO, and the density difference between HOMO vs LUMO and HOMO-1 vs LUMO molecular orbitals were plotted with contour values from 0.05 to −0.05. Red and blue indicate positive and negative contributions to molecular orbitals.

no excited indole-NH_3^+ interaction. Moreover, the oscillator strength (*f*) shows the highest value at the 3^1A state for the three conformers. This 3^1A state corresponds to a wavelength in the range of 288−293 nm, which is in a good agreement with the 296 nm band obtained from magnetic circular dichroism (MCD) of 5-HT stabilized by hydrogen oxalate in 0.1 N HCl.[43]

The major MOs contributions involved in donor/acceptor exchanges concern the 1^1A and the 2^1A excited states for the three conformers and correspond respectively to HOMO/LUMO and HOMO-1/LUMO transfers. Figure 5 shows the electron cloud of these MOs distributed in each conformer and the different electron densities between two corresponding MOs. The electrons of HOMO and HOMO-1 orbitals are mainly localized within π indole ring orbitals. For LUMO, electrons are distributed to the NH_3^+ group, within a σ^* antibonding orbital. The analysis of MOs after subtraction of electron density between HOMO and LUMO, and HOMO-1 and LUMO, clearly shows that the excitation from the ground state allowed the 1^1A and 2^1A states respectively for both the **GmGp** and **GmGm** conformers. This confirms the existence of an indole/NH_3^+ through-space charge

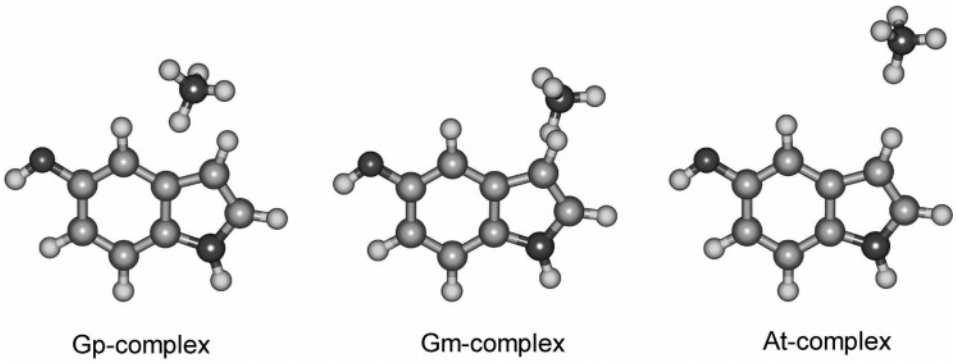


Figure 6. Structures of three NH_4^+/5-hydroxyindole complexes.

Table 4. Major Orbital Contribution, Calculated Wavelength (λ in nm), Oscillator Strength (f), and Electronic Excitation Energies (ΔE in eV) of Six Singlet-Excited States of the **GmGp**, **GmGm**, and **GmAt** Conformers of 5-HT Calculated Using the TDDFT Method at the B3LYP/6-311+G(2d,2p) Level of Theory, Together with Corresponding Ground-State Gap Energies between HOMO-n and LUMO+n (HLG in eV)

	major orbital contribution					
excited state	character	coefficients	λ (nm)	f	ΔE (eV)	HLG (eV)
			GmGp			
1^1A	HOMO→LUMO	0.69821	336.77	0.0002	3.6816	0.1590
2^1A	HOMO-1→LUMO	0.69706	309.89	0.0025	4.0009	0.1701
3^1A	HOMO→LUMO+1	0.63991	287.66	0.0725	4.3101	0.1781
			GmGm			
1^1A	HOMO→LUMO	0.70296	357.58	0.0015	3.4674	0.1485
2^1A	HOMO-1→LUMO	0.70081	318.36	0.0179	3.8945	0.1658
3^1A	HOMO→LUMO+1	0.64599	290.68	0.0874	4.2654	0.1759
			GmAt			
1^1A	HOMO→LUMO	0.70189	454.78	0.0004	2.7262	0.1183
2^1A	HOMO-1→LUMO	0.70167	402.00	0.0006	3.0842	0.1321
3^1A	HOMO→LUMO+1	0.66497	292.91	0.0777	4.2328	0.1746

transfer with a major $\pi \rightarrow$s* contribution. This is also observed by the partial electron density loss of the indole ring and their delocalization to one of the N−H bonds in the cationic head. This result again confirms the previous assignment of maximum fluorescence spectra for 5-HT at the 1^1A state. Oscillator strength of the 3^1A state of the three conformers is higher than that of the 1^1A and 2^1A states, which indicates that the low-energy absorption band in MCD of 5-HT is mainly derived from HOMO to LUMO+1 with $\pi \rightarrow$p* transfer (the plot is not shown), while this excited state is absent in the fluorescence spectra of 5-HT.

3.3. Cation-π Interactions in NH_4^+/5-Hydroxyindole Complexes. To characterize the nature and strength of the nonbonded intramolecular interactions in the 5-HT conformers, the cationic head and the 5-hydroxy indole ring of the three stable conformers **GmGm**, **GmGp**, and **GmAt** were extracted from the whole 5-HT molecule optimized geometries, and a hydrogen atom was added to the cationic head and the 5-hydroxy indole ring to fill the unsatisfied valency resulting in the suppression of the $C_{sp3}-C_{sp3}$ ethyl side chain. The resulting three complexes of NH_4^+/5-hydroxyindole named **Gp**-complex, **Gm**-complex, and **At**-complex are shown in Figure 6. For each complex we evaluated the total binding energy and its different energy term contributions. Since the results of this analysis were obtained from single point calculations, they are useful for understanding, by extrapolation, the amplitude and the nature of the intramolecular nonbonded interactions acting in the whole 5-HT molecule.

3.3.1. Intermolecular Interaction Potentials of NH_4^+/5-Hydroxyindole Complexes. Intermolecular interaction energies were calculated for the three NH_4^+/5-hydroxyindole complexes using HF and MP2 methods and the 6-311+G(d,p) basis set. The total interaction energy and BSSE correction for the three complexes at all levels of theory are listed in Table 5. The actual interaction energies were not corrected for either zero point energy or for the thermal contribution in order to directly appreciate the correlation effect acting on the electronic energy of the nonoptimized complexes. As we can see, the binding energy is important for the three complexes and is the greatest for the **Gp**-complex, regardless of the theory level, with or without correction by the BSSE (by 3.3 and 9.6 kcal/mol with BSSE correction as compared with **Gm-** and **At-**complexes respectively). The large difference between HF and MP2 calculations for the two gauche conformers can be explained because MP2 calculations take into account the electron transitions between the occupied and the unoccupied orbitals that affect the stability of the whole 5-HT molecule. The fact that the electronic correlation effect is major for the **Gm-** and **Gp**-complexes emphasizes the important

Table 5. Absolute Binding Energies (BE) without Counterpoise Correction, Basis Set Superposition Error Correction (BSSE), and Deformation Energy (E_{def}) in kcal/mol of Three Complexes Calculated Using the HF, B3LYP, and MP2 Methods at the 6-311+(d,p) Basis Set

HF/6-311+G(d,p)			B3LYP/6-311+G(d,p)			MP2/6-311+G(d,p)		
BE	BSSE	E_{def}	BE	BSSE	E_{def}	BE	BSSE	E_{def}
				Gp				
−11.34	0.34	0.95	−14.23	0.25	0.90	−16.53	1.62	0.75
				Gm				
−8.19	0.32	0.75	−11.41	0.23	0.70	−12.98	1.38	0.58
				At				
−5.17	0.13	0.28	−6.42	0.09	0.49	−6.80	0.51	0.44

electronic redistribution occurring in the gauche conformer through a transfer of charge.

However, the relative binding energies remain quite similar when we use either the HF or the MP2 method, with or without counterpoise correction. We also observe the same tendency when we compare the relative binding energies obtained with MP2 and DFT methods. The difference between the MP2 and B3LYP results may be due to the missing long-range (1/r6) dispersion in B3LYP calculations. For each conformer, the effect of the dispersion energy is smaller than the electron correlation effects. Hence, electron correlation and dispersion energy contributions to the HF and B3LYP binding energy respectively are greater for the **Gp**- (30% and 10%) and **Gm**-complexes (40% and 10%) than for the **At** one (30% and negligible). This suggests that the position of the ammonium group of 5-HT, relative to the 5-hydroxyindole ring, is not driven by dispersion energy but rather by other energy terms. This observation was confirmed by the energy decomposition analysis.

The geometry deformation energy (E_{def}) represents the constraints imposed on the internal coordinates of the two fragments of the complexes and is calculated by the summation over the two monomers of a complex, of the energy difference between a monomer taken in the conformation it adopts inside 5-HT and after full optimization in the gas phase. As we can see in Table 5, E_{def} is negligible for the three complexes, suggesting that the deformation of the monomer geometries is modest and does not alter the relative stability of the three conformers of 5-HT brought about by nonbonded interactions. This observation suggests that, besides repulsive and attractive interactions, a strong constraint applied on the internal coordinates of the ethyl side chain of 5-HT (dihedral angles ϕ_1 and ϕ_2, opening of angles θ_1 and θ_2) could also participate in the decrease of the relative energy gap to around 1 kcal/mol between the two gauche 5-HT conformers.

Several recent studies have pointed out that the dispersion, quantified in terms of dynamical electronic correlation, provides a crucial contribution to the interaction energy between organic cations and aromatic groups and, in particular, between ammonium ions and a benzene ring.[39,44] A direct quantitative comparison of binding energies obtained from three NH_4^+/5-hydroxyindole complexes to those available for the NH_4^+/benzene complex is not possible since the indole ring of 5-HT and benzene molecule differ by their numbers of atoms and the nature of their bonds. However, we are able to compare qualitatively the geometries of the two kinds of dimers which may allow us to appreciate the importance of their dispersion energy. Lee et al. have shown that the most stable conformer for the NH_4^+/benzene complex is the one with two hydrogen atoms directed toward the benzene ring.[39b] This conformation allows an optimal through-space interaction between the benzene π−HOMO orbital and the ammonium σ^*_{N-H} LUMO antibonding orbital. If we compare the **Gp**- and **Gm**-complexes to the NH_4^+/benzene complex, only one N−H bond interacts with the five-membered ring π orbitals. Hence, the σ^*_{N-Ha} LUMO antibonding orbital and the π HOMO orbital are also tilted and cannot lead to an optimal interaction through a maximum overlap.

3.3.2. Energy Decomposition Analysis of NH_4^+/5-Hydroxyindole Complexes. To better understand the nature and weight of forces involved in the intramolecular interactions of the 5-HT conformers we analyzed the importance of individual energy components (electrostatic, polarization, charge transfer, repulsion) of the total interaction energy of NH4+/5-hydroxyindole complexes corresponding to the **GmGp**, **GmGm**, and **GmAt** conformers. The total interaction energies along with the energetic components through the KM and RVS methods using the HF theory and the 6-31G(d) and 6-311+G(d,p) basis sets are listed in Table S10 in the Supporting Information. Whatever the method used for energy decomposition, the total interaction energies (ΔE_{TOT}) and the corresponding energetic components of the **Gp**- and **Gm**-complexes are higher than those obtained for the **At**-complex at the HF/6-31G(d) level, indicating a major contribution of intermolecular interaction between the NH_4^+ and 5-hydroxy indole in gauche complexes. The largest ΔE_{TOT} is found for the **Gp**-complex. It is greater than those found for the **Gm**- and **At**-complexes by 3 and 6 kcal/mol, respectively. In the **At**-complex, the most important contribution for ΔE_{TOT} comes from the polarization effects, since ΔE_{POL} accounts for 70−80% of the total interaction energy. Clearly, there are no cation-π interactions in the **At**-complex, due to the very low charge transfer term, especially using the RVS scheme. This observation confirms the results obtained with NBO and the topological analysis of electron density. Hence, the electrostatic (ΔE_{ES}) and the exchange repulsion (ΔE_{EX}) terms, combined in the ΔE_{CEX}, have a minor contribution of around 0.5 kcal/mol.

The ΔE_{ES} and ΔE_{EX} were found to be predominant for the **Gp**- and **Gm**-complexes. The electrostatic attractive

energy in the **Gp**- and **Gm**-complexes is similar (about 0.5 kcal/mol) while ΔE_{EX}, the exchange repulsion energy, is mostly marked for the **Gm**-complex, 1.6 kcal/mol higher than that of the **Gp**-complex. The balance of these two components can be simply evaluated using the ΔE_{CEX} term, which is the summation of the ΔE_{ES} and ΔE_{EX} terms, indicating the presence of a stronger repulsion between NH_4^+ and 5-hydroxyindole in the **Gm**-complex. The polarization term (ΔE_{POL}) provides an important contribution to the total energy of the three complexes. It is important to note that the ΔE_{POL} obtained from the KM scheme appears to be overestimated, particularly for the **Gp**- (by −43.3 kcal/mol) and **Gm**-complex (no convergence found for ΔE_{POL} calculation) when the increase of basis set sets up to 6-311+G(d,p). Thus, the analysis of ΔE_{POL} is strictly considered by using the more physical RVS method. The energy decomposition analysis also reinforces the importance of charge-transfer effects (ΔE_{CT}). The stabilization due to charge transfer between the 5-hydroxyindole ring and the NH_4^+ group in the **Gp**- and **Gm**-complexes contributes for 30% and 40% (the KM method) or 20% and 30% (the RVS method) of the total interaction energy, respectively.

The above results are mainly based on the HF level with a very moderate 6-31G(d) basis set. However, the basis set dependence of the results was investigated by performing KM and RVS decomposition analysis with the 6-311+G-(d,p) basis set. As expected, the basis set superposition error (BSSE) is reduced with an enlarged basis set, in particular with the inclusion of polarization and diffusion functions. The ΔE_{ES} and ΔE_{EX} terms stay relatively constant for the **At**-complex upon extension of the basis set, while those terms change respectively by −0.35 and 1.83 kcal/mol for the **Gp**-complex and by −1.27 and 2.72 kcal/mol for the **Gm**-complex with respect to the 6-31G(d) basis set. For the RVS method, the ΔE_{POL} term increases by −1.26 and −1.16 kcal/mol for the **Gp**- and **Gm**-complexes, respectively, while a slight change (−0.43 kcal/mol) occurs for the **At**-complex. The charge-transfer energy (ΔE_{CT}) of the **Gp**-, **Gm**-, and **At**-complexes respectively decreases by 0.71, 0.65, and 0.35 kcal/mol. The total interaction energy (ΔE_{TOT}), which is the summation of all energetic terms in the RVS method, decreases by about 1 kcal/mol for gauche complexes but does not change significantly for the **At**-complex when the basis set increases from 6-31G(d) to 6-311+G(d,p).

Our results confirm that the electrostatic component (ΔE_{ES}) is significant in stabilizing cation-π interactions within the NH_4^+/5-hydroxyindole complexes. The ΔE_{ES} value is quantitatively dominant among the other components and is also the only term undergoing the highest shift when the ethylamine side chain is changed from the gauche to the anti conformation. The largest contribution to the electrostatic term is thought to be due to the interaction between the positive charge of the ammonium nitrogen and the permanent electric multipoles of the 5-hydroxyindole ring. A large number of theoretical studies, performed on several different complexes, have also shown that the cationic-π interactions are strongly governed by electrostatic interactions between a cation and the permanent quadrupole moment of an aromatic ring.[39a,45] Furthermore, the presence of an ion−dipole interaction has been proposed for a series of tetramethylammonium-aromatic complexes.[44a] Recently, the energy decomposition analysis of ammonium-π complexes with benzene, furan, pyrrole, and thiophene have shown that electrostatic interactions drive the stability of a complex formed between a positive charge centered on the ammonium nitrogen and the permanent electric multipoles of aromatic rings.[12] The authors have reported both the ion−dipole and the ion−quadrupole to significantly contribute to the electrostatic term using the KM approach.

The exchange repulsion term (ΔE_{EX}) appears to be more important in the **Gm**- than in the **Gp**-complex, both being much larger than in the **At**-complex. The increase of short-range repulsion in gauche conformers comes from the frontier orbitals overlap, which is smaller for the **At**-complex since the two monomers are far from each other. The ΔE_{EX} term also increases when we add diffusion functions on hydrogen atoms in the basis set. The polarization (ΔE_{PL}) and charge-transfer (ΔE_{CT}) terms, which contribute to cation-π interactions of the NH_4^+/5-hydroxyindole complexes are evaluated using the more accurate RVS method, which has been claimed to allow a better control over the antisymmetrization of the wave functions.[12] After ΔE_{ES}, ΔE_{PL} provides the next important contribution to the NH_4^+-5-hydroxyindole total interaction energies, which confirms the significance of the polarization effect on cation−π interactions, as noted by others.[46] The ΔE_{PL} term is assumed to mainly account from an ion-induced multipole interaction, its strength depending on the amplitude of induced multipoles in the 5-hydroxyindole ring. The charge transfer (ΔE_{CT}) between the π system and the cation fragment contributes significantly to the total interaction energy and stabilizes the NH_4^+/5-hydroxyindole complexes. This confirms the postulate of the existence of a stabilizing charge transfer in 5-HT proposed in the 1970s by Kang and Cho.[5b] This charge transfer has been precisely analyzed using the frontier orbitals theory for some ammonium−benzene complexes and indicated the existence of a through-space $\pi-\sigma^*$ charge-transfer.[39b,44g]

4. Conclusion

The present work provides a clear view on the intramolecular energy components that are driving the conformational properties of 5-HT. Our calculations show a very good agreement with experimental IR and CD spectra and also NMR data. We confirm the existence of three local minima conformers of 5-HT with OH in the anti position (**GmGp**, **GmGm**, and **GmAt**) in the gas phase, the **GmGp** and **GmGm** conformers being 5 kcal/mol more stable than the **GmAt** conformer. This higher stability for the **GmGp** and **GmGm** conformers arises from strong nonbonded interactions established between the cationic head of 5-HT and the indole ring. The previously suspected charge transfer occurring in 5-HT is confirmed by the analysis of the IR spectra. Indeed, analysis of stretch frequencies and intensities have confirmed the presence of the red-shift of the amine N−H_a stretch frequencies and elongation of the N−H_a in gauche conformers. The analysis of calculated carbon and proton shifts along with experimental data shows modifications of the strength and length of indole ring bonds upon

the influence of the cationic head. For the **GmGp** and **GmGm** conformers C_8-C_9 and C_2-C_3 bonds are elongated respectively and show a decrease in their double bond character due to a loss of electronic density in the vicinity of the bond promoted by the charge transfer. The topological analysis of charge density and natural bond orbital analyses confirms the existence of such charge transfer by revealing a bond critical point in the gauche conformers between H_a and C_3 and H_a and C_9 in the **GmGm** and **GmGp** conformers, respectively. The direction and magnitude of those charge-transfers derived from the Kohn−Sham analogue of the Fock matrix within the NBO basis shows that the charge transfers occur between the π HOMO-1 and HOMO orbitals of indole ring and the σ^*_{N-Ha} LUMO antibonding orbital. The analysis of the binding energies and their energy components in the NH_4^+/5-hydroxyindole complexes, which were generated from geometries of whole 5-HT conformers, shows that the contribution of the charge-transfer energy, although not negligible, is minor compared to other terms such as electrostatic and polarization energies.

To summarize, the two gauche conformers are stabilized by strong intramolecular nonbonded interactions. The **GmGp** and **GmGm** conformers are stabilized mainly by electrostatic interactions with a similar charge-transfer contribution. However, the higher stability of the **GmGp** over the **GmGm** conformer comes from stronger polarization and weaker exchange repulsion energy contributions. Internal coordinates energy constraints imposed on ethylamine side chain bending and dihedral angles appear to penalize the stability of the **GmGm** conformer.

To date, our work is the first to provide in detail the physical nature of the different energy contributions that influence the conformational properties of 5-HT and gives the means to interpret 5-HT experimental IR, CD, and NMR spectra. Hence, this work can be used as a solid basis for further studies such as, for example, the influence of implicit or explicit water solvent on 5-HT conformation or the chemical route leading to some physiologically active metabolites.

Acknowledgment. This study was partially supported by BioQuanta *Inside cell* Corp, CO, and the Conseil Régional de l'Ile-de-France (CRIF), grant no. 03.DAI.61F. We gratefully thank Dr. Mike Briley from NeuroBiz Consulting & Communications for the English corrections.

Supporting Information Available: Structure of 12 local minima of 5-HT (Figure S1) and their energies (in Hartrees) (Table S1) including energies of eight transition states; the values of 5-HT dihedral angles (ϕ_1, ϕ_2, and ϕ_4), all the conformers, and their relative energies (Table S2); molecular dynamics method and time evolution plots of the dihedral angles ϕ_1 and ϕ_2 during trajectory analysis (Figure S2) indicating guess transition state (Ts) of 5-HT for the QST3 calculations; geometric parameters of the three stable conformers of 5-HT: OH anti, **GmGp**, **GmGm**, **GmAt** (Table S3); predicted stretching IR spectra of the three stable conformers at the B3LYP/6-31+G(d,p) level of theory (Figure S3) and their harmonic stretching frequencies (ν) and intensities (I) (Table S4); absolute isotropic NMR shielding tensors (σ in ppm) of C and H atoms calculated using the GIAO method at the B3LYP/6-311+G(2d,2p) level of theory (Table S5); electrostatic potential derived charges of atoms in the 5-HT conformers calculated at the B3LYP/6-311+G(2d,2p) using ChelpG Scheme (Table S6); the shielding increment of amine hydrogen ($\Delta\delta$H) calculated using the GIAO method at the B3LYP/6-311+G(2d,2p) level of theory (Table S7); Fermi contact (FC), spin dipolar (SD), (paramagnetic spin−orbit PSO) contributions using the GIAO method at the B3LYP/6-311+G(2d,2p) level of theory (Table S8); electronic information of six singlet-excited states of three 5-HT conformers calculated using the TDDFT method at the B3LYP/6-311+G(2d,2p) level of theory (Table S9); and energy components of the three NH_4^+/5-hydroxyindole complexes calculated using KM and RVS energy decomposition analyses at the HF/6-31G(d) and the HF/6-311+G(d,p) levels of theory (Table S10). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Hoyer, D.; Hannon, J. P.; Martin, G. R. *Pharmacol. Biochem. Behav.* **2002**, *71*, 533−554.

(2) Hoyer, D.; Clarke, D. E.; Fozard, J. R.; Hartig, P. R.; Martin, G. R.; Mylecharane, E. J.; Saxena P. R.; Humphrey, P. P. *Pharmacol. Rev.* **1994**, *46*, 157−203.

(3) (a) Blakely, R. D.; De Felice, L. J.; Hartzell, H. C. *J. Exp. Biol.* **1994**, *196*, 263−281. (b) Uhl, G. R.; Johnson, P. S. *J. Exp. Biol.* **1994**, *196*, 229−236. (c) Nelson, N. *J. Neurochem.* **1998**, *71*, 1785−1803.

(4) Chattopadhyay, A.; Rukmini, R.; Mukherjee, S. *Biophys J.* **1996**, *71*, 1952−1960.

(5) (a) Courrière, P.; Coubeils, J. L.; Pullman, B. *C. R. Acad. Sci. Paris D* **1971**, *272*, 1697−1700. (b) Kang, S.; Cho, M. H. *Theor. Chim. Acta* **1971**, *22*, 176−183. (c) Kang, S.; Johnson, C. L.; Green, J. P. *Mol. Pharmacol.* **1973**, *9*, 640−648. (d) Port, G. N. J.; Pullman, B. *Theor. Chim. Acta* **1974**, *33*, 275−278.

(6) (a) Kumbar, M.; Sankar S. D. V. *J. Am. Chem. Soc.* **1975**, *97*, 7411−7416. (b) Pullman, B.; Courriere, P.; Berthod, H. *J. Med. Chem.* **1974**, *17*, 439−447.

(7) Ison, R. R.; Partington, P.; Roberts, G. C. K. *J. Pharm. Pharmacol.* **1972**, *24*, 82−85.

(8) (a) Karle, I. L.; Dragonette, K. S.; Brenner, S. A. *Acta Crystallogr.* **1965**, *19*, 713−716. (b) Thewalt, U.; Bugg, C. E. *Acta Crystallogr.* **1972**, *B28*, 82−92. (c) Amit, A.; Mester, L.; Klewe, B.; Furberg, S. *Acta Chem. Scand.* **1978**, *A32*, 267−270.

(9) Edvardsen O.; Dahl S. G. *Brain Res. Mol. Brain Res.* **1991**, *9*, 31−37.

(10) Pisterzi, L. F.; Almeida, D. R. P.; Chass, G. A.; Torday, L. L.; Papp, J. G.; Varro, A.; Csizmadia, I. G. *Chem. Phys. Lett.* **2002**, *365*, 542−551.

(11) Alagona, G.; Ghio, C.; Nagy, P. I. *J. Chem. Theory Comput.* **2005**, *1*, 801−816

(12) Aschi, M.; Mazza, F.; Nola, A. D. *J. Mol. Struct. (THEOCHEM)* **2002**, *587*, 177−188.

(13) (a) Beck, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652. (b) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789. (c) Stevens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623−11627.

(14) (a) Peng, C.; Schlegel, H. B. *Isr. J. Chem*. **1994**, *33*, 449−454. (b) Peng, C.; Ayala, P. Y.; Schlegel, H. B.; Frisch, M. J. *J. Comput. Chem.* **1996**, *17*, 49−56.

(15) (a) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem*. **1983**, *4*, 187−212. (b) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. C. *J. Comput. Phys*. **1977**, 23, 327−341.

(16) (a) Carney, J. R.; Zwier, T. S. *J. Phys. Chem. A* **1999**, *103*, 9943−9957. (b) Dian, B. C.; Longarte, A.; Mercier, S. *J. Chem. Phys.* **2002**, *117*, 10688−10702.

(17) (a) Ditchfield, R. *Mol. Phys.* **1974**, *27*, 789−807. (b) Dodds, J. L.; McWeeny, R.; Sadlej, A. J. *Mol. Phys.* **1980**, *41*, 1419−1430. (c) Wolinski, K.; Hilton, J. F.; Pulay, P. *J. Am. Chem. Soc.* **1990**, *112*, 8251−8260.

(18) Ramsey, N. F. *Phys. Rev*. **1953**, *91*, 303−307.

(19) (a) Bader, R. F. W. *Chem. Rev*. **1991**, *91*, 893−928. (b) Cioslowski, J.; Nanayakkara, A.; Challacombe, M. *Chem. Phys. Lett.* **1993**, *203*, 137−142.

(20) (a) Reed, A. E.; Curtiss, L. A.; Weinhold, F. *Chem. Rev.* **1988**, *88*, 899−926. (b) Weinhold, F. A. *J. Mol. Struct. (THEOCHEM)* **1997**, *398*, 181−197.

(21) Gledening, E. D.; Reed, A. E.; Carpenter, J. A.; Weinhold, F. NBO Version 3.1.

(22) Runge, E.; Gross, E. K. U. *Phys. Rev. Lett.* **1984**, *52*, 997−1000.

(23) Xantheas, S. S. *J. Chem. Phys.* **1996**, *104*, 8821−8824.

(24) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553−566.

(25) Chalasinski, G.; Szczesniak, M. M. *Chem. Rev*. **2000**, *100*, 4227−4252.

(26) Kitaura, K.; Morokuma K. *Int. J. Quantum Chem*. **1976**, *10*, 325−340.

(27) Stevens, W. J.; Fink, W. H. *Chem. Phys. Lett.* **1987**, *139*, 15−22.

(28) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03, Revision A.1*; Gaussian, Inc.: Pittsburgh, PA, 2003.

(29) Chen, W.; and Gordon, M. S. *J. Phys. Chem.* **1996**, *100*, 14316−14328.

(30) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. J.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem*. **1993**, *13*, 1347−1363.

(31) Zhurko, G. A.; Zhurko D. A.; Romanov A. Chemcrafts version 1.5 b248, a tool for visualization of quantum chemistry data (http://www.chemcraftprog.com).

(32) Bayari, S.; Die, S. *Spectrochim. Acta Part A* **2003**, *59*, 1255−1263.

(33) van Mourik T.; Emson Laura E. V. *Phys. Chem. Chem. Phys.* **2002**, *4*, 5863−5871.

(34) Barnes, A. J. *J. Mol. Struct.* **2004**, *704*, 3−9.

(35) (a) Sapper, H.; Lohmann, W. *Mol. Pharmacol.* **1976**, *12*, 605−611. (b) Kang, S. *Mol. Pharmacol.* **1979**, *16*, 1031−1039.

(36) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361−373.

(37) Sychrovsky, V.; Grafenstein, J.; Cremer, D. *J. Chem. Phys.* **2000**, *113*, 3530−3547.

(38) (a) Bader, R. F. W. *Atoms In Molecules. A Quantum Theory*; Oxford University Press: New York, 1995. (b) Popelier, P. *Atoms In Molecules, An Introduction*; Prentice Hall: Englewood Cliffs, NJ, 2000.

(39) (a) Kim, K. S.; Tarakeshwar, P.; Lee, J. Y. *Chem. Rev.* **2000**, *100*, 4145−1486 and references therein. (b) Lee, J. Y.; Lee, S. J.; Choi, H. S.; Cho S. J.; Kim, K. S.; Ha, T.-K. *Chem. Phys. Lett.* **1995**, *232*, 67−71.

(40) Pople, J. A.; Santry, D. P. *Mol. Phys.* **1964**, *8*, 1−18.

(41) Wang, G.; Geng, L. *Anal. Chem.* **2005**, *77*, 20−29.

(42) Eftink, M. R.; Jia, Y.; Hu, D.; Ghiron, C. A. *J. Phys. Chem.* **1995**, *99*, 5713−5723.

(43) Sprinkel, F. M.; Shillady D. D.; Strickland, R. W. *J. Am. Chem. Soc*. **1975**, *97*, 6653−6657.

(44) (a) Pullman, A.; Berthier, G.; Savinelli, R. *J. Comput. Chem.* **1997**, *18*, 2012−2022. (b) Tan, X. J.; Jiang, H. L.; Zhu, W. L.; Chen K. X.; Ji, R. Y. *J. Chem. Soc., Perkin Trans. 2* **1999**, *1*, 107−112. (c) Pullman, A.; Berthier, G.; Savinelli, R. *J. Am. Chem. Soc*. **1998**, *120*, 8553−8554. (d) Minoux, H.; Chipot, C. *J. Am. Chem. Soc*. **1999**, *121*, 10366−10372. (e) Mecozzi, S.; West Jr, A. P.; Dougherty D. A. *J. Am. Chem. Soc.* **1996**, *118*, 2307−2308. (f) Berthomieu, D.; Brenner, V.; Ohanessian, G.; Denhez, J.; Millie, P.; Audier, H. E. *J. Phys. Chem.* **1995**, *99*, 712−720. (g) Kim, S. K.; Lee, J. Y.; Ha, T. K.; Kim, D. H. *J. Am. Chem. Soc*. **1994**, *116*, 7399−7400.

(45) (a) Dougherty D. A. *Science* **1996**, *271*, 163−168. (b) Luhmer, M.; Bartik, K.; Dejaegere A.; Bovy, P.; Reisse, J. *Bull. Soc. Chim. Fr*. **1994**, *131*, 603−606. (c) Williams, J. H. *Acc. Chem. Res.* **1993**, *26*, 593−598. (d) Hobza, P.; Selzle, H. L.; Schlag, E. W. *J. Am. Chem. Soc*. **1994**, *116*, 3500−3506.

(46) (a) Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc*. **1995**, *117*, 4177−4178. (b) Eriksson, M. A. L.; Morgantini, P.-Y.; Kollman, P. A. *J. Phys. Chem. B* **1999**, *103*, 4474−4480. (c) Cubero, E.; Luque, F. J.; Orozco, M. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 5976−5980. (d) Cubero, E.; Orozco, M.; Luque, F. J. *J. Phys. Chem. A* **1999**, *103*, 315−321.

CT0600316

JCTC Journal of Chemical Theory and Computation

On the Aromaticity of the Planar Hydrogen-Bonded $(HF)_3$ Trimer

Ying-Chan Lin and Dage Sundholm*

Department of Chemistry, University of Helsinki, P. O. Box 55 (A. I. Virtanens plats 1), FIN-00014, Helsinki, Finland

Jonas Jusélius

Department of Chemistry, University of Tromsø, N-9037 Tromsø, Norway

Received November 30, 2005

Abstract: The nuclear magnetic shieldings and magnetically induced ring currents have been calculated for the planar ring-shaped hydrogen fluoride trimer $(HF)_3$ at correlated ab initio and density functional theory levels. Calculations of the magnetically induced current densities using the gauge-including magnetically induced current (GIMIC) method show that, contrary to a recent suggestion, $(HF)_3$ has, at the MP2/TZVPP level, a very small ring-current susceptibility of 0.37 nA/T. Thus, only a weak net current is passing across the H···F hydrogen bond. An external magnetic field perpendicular to the ring plane induces strong edge currents circling around each HF molecule giving rise to a nonvanishing magnetic shielding at the center of the ring. The GIMIC results are supported by calculations of the long-range magnetic shielding function; the long-range magnetic shielding is very small, indicating that the magnetically induced ring-current is very weak. The surprisingly large nucleus-independent chemical shift (NICS) value for $(HF)_3$ was recently taken as an indication of "H-bonded aromaticity". The NICS value calculated at the CCSD/QZ2P level is 2.77 ppm. The present GIMIC and aromatic ring-current shielding study shows that some care has to be taken when using NICS values as aromaticity indices.

I. Introduction

Aromatic molecules were originally defined as planar organic molecules fulfilling the $4n + 2$ rule for the electron count, with benzene as the archetypal example.[1] Aromatic molecules were also found to sustain diamagnetic currents when exposed to external magnetic fields, whereas antiaromaticity was coined to describe molecules sustaining paramagnetic ring currents; that is, the current circles in the opposite direction as compared to the aromatic ones.[2−4] The aromaticity concept has lately been extended to comprise many kinds of molecules with electron delocalization.[5] Nowadays, it is used not only in organic chemistry but also to describe inorganic systems.[6,7] Metal clusters are found to be both aromatic and antiaromatic.[8−10] Molecules sustaining ring currents in the σ orbitals are denoted σ aromatic. Homoaromatic molecules are bent organic species that sustain strong ring currents even though they have aliphatic bonds in the ring, which at first sight seem to interrupt the electron delocalization of the ring.[11−13] Spherical aromatic molecules consist of approximately spherical (e.g., icosahedral) molecules fulfilling a $2(N + 1)^2$ rule for the number of valence electrons.[14−18] These are the magic numbers of electrons for closed-shell model systems with an indefinitely thin spherically symmetric shell potential. Simpler three-dimensional cage-shaped molecules are found to possess 3D aromaticity.[19]

Chemists are very fond of the aromaticity concept, even though, or because, it is not very well defined. Computational chemists have invented new computational tools to predict and define molecular aromaticity and the degree of aromaticity. As a result, regularly, new aromaticity concepts are born, some more successful than others. Recently, Datta et al.[20,21] proposed the term *H-bonded aromaticity* as a new

* Corresponding author e-mail: sundholm@chem.helsinki.fi.

10.1021/ct050297x CCC: $33.50
Published on Web 03/02/2006

kind of aromaticity. Computationally, they found a significant diamagnetic shielding at the center of cyclic HX (X = F, Cl, and Br) trimers and interpreted it as a result of ring currents passing the hydrogen bonds between the monomers. They have also reported similar results for water clusters $(H_2O)_n$.[21] In their studies, nucleus-independent chemical shifts (NICSs) were used to determine the degree of aromaticity.[22] The NICS value is an often-engaged tool to assess molecular aromaticity.[23]

The discovery of "H-bonded aromaticity" involving strong ring currents across hydrogen bonds would indeed be very intriguing. In principle, the NICS index is supposed to be linearly related to the strength of the induced ring current. In practice, however, the magnetic shielding at the center of the molecule can be strongly affected by the local electronic structure. Thus, care must be taken in the interpretation of calculated NICS values.[13,24−27] In this work, we use the aromatic ring-current shielding (ARCS) index[28] and the newly developed gauge-including magnetically induced current (GIMIC) method[29] to study the nature of the induced currents and to assess the degree of aromaticity of the $(HF)_3$ ring.

II. Computational Methods

The molecular structure was optimized at the second-order Møller−Plesset perturbation level (MP2) using the TURBOMOLE program package.[30] In the structure optimization, we employed the Karlsruhe standard triple-ζ valence basis sets augmented with double polarization functions (TZVPP) as well as Dunning's correlation consistent quadruple-ζ (cc-pVQZ) and quintuple-ζ (cc-pV5Z) basis sets.[31] The molecular structure was also optimized at the density functional theory level (DFT) employing Becke's three-parameter functional in combination with the Lee−Yang−Parr correlation functional (B3LYP)[32,33] using the TZVPP basis sets.

The nuclear magnetic shieldings were calculated at the coupled-cluster singles and doubles (CCSD) level using Ahlrichs' quadruple-ζ (QZ) quality (11s7p/6s4p) basis sets augmented with 2d1f polarization functions for F. For H, Ahlrichs' QZ basis set (7s/4s) was augmented with 2p1d polarization functions.[30] We denote these basis sets QZ2P. Magnetic shielding calculations were also performed at the MP2 and B3LYP levels. The magnetic shielding calculations at the MP2 and CCSD levels were performed with the ACES II program,[34−37] whereas the B3LYP and some MP2 calculations of the magnetic shieldings were done with TURBOMOLE.[38,39] The magnetically induced current densities were deduced from the one-particle density matrix and the magnetically perturbed density matrices calculated at the MP2 level using the GIMIC approach.[29] The ring-current susceptibility was obtained by numerical integration of the current density passing cut planes perpendicular to the molecular ring. The ring-current susceptibility was also estimated from the long-range magnetic shielding along the symmetry axis perpendicular to the molecular plane using the ARCS approach. The ring-current susceptibility can be obtained by fitting the ARCS function to Biot-Savart's expression for a circular and infinitely thin conducting wire.

Table 1. Molecular Structure of $(HF)_3$ Optimized at the MP2 and B3LYP Levels Using Different Sizes of Basis Sets[a]

level	*R*(H−F)	*R*(H···F)
MP2/TZVPP	93.7	173.9
MP2/cc-pVQZ	93.5	174.2
MP2/cc-pV5Z	93.5	175.5
B3LYP/TZVPP	94.4	173.4
B3LYP/cc-pVQZ	94.3	173.0
B3LYP/6-311G++(d,p)[b]	94	183

[a] The HF hydrogen bond is denoted with H···F. The bond lengths are given in pm. [b] Ref 20.

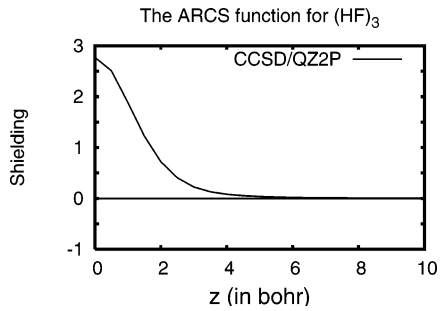


Figure 1. Nuclear magnetic shielding function (in ppm) calculated at the CCSD/QZ2P level. The shielding function is calculated from the ring center along the symmetry axis perpendicular to the ring.

NICS values were obtained as the negative magnetic shielding in the center of the molecular ring.[22,23]

III. Molecular Structures

The molecular structure optimized at the MP2 and B3LYP levels are given in Table 1. At the MP2 level, the length of the hydrogen bond increases slightly with the size of the basis set, because the basis-set superposition error causes some overbinding, which decreases with increasing basis-set size. The length of the hydrogen bond is found to be slightly longer at the MP2 level than at the B3LYP level. In their recent study, Rehaman et al.[20] obtained 10-pm-longer hydrogen bonds at the B3LYP/6-311G++(p,d) level.

IV. Magnetically Induced Currents and Magnetic Shieldings

The calculation of the magnetic shielding function along the symmetry axis (the ARCS function) shows that the long-range magnetic shielding vanishes outside the electron charge density, indicating that the molecular ring does not sustain any strong ring current. The ARCS function calculated at the CCSD/QZ2P level is shown in Figure 1. The ring-current susceptibility deduced from the ARCS function is 0.9 nA/T as compared to the benzene value of 8 ppm.[28] The ARCS fit is not very accurate because the long-range magnetic shielding is small as a result of the weak ring current. In the ARCS fit, the angular coefficient is 1.88 and not 1.5 as it should be. For comparison, a ring-current susceptibility of 0.34 nA/T was obtained by performing a numerical integration of the current density passing a cut plane of the molecular ring through the center of the HF bond. The current density was calculated at the MP2/TZVPP level using

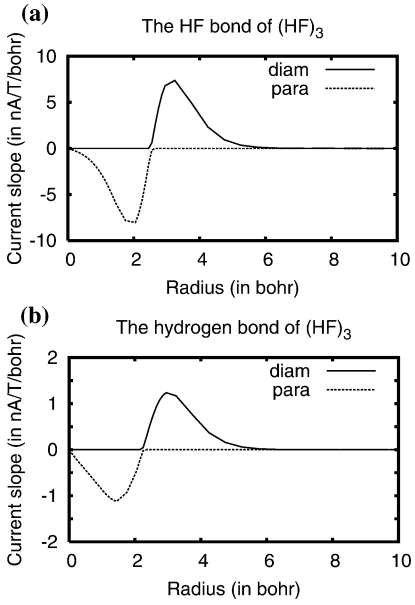


Figure 2. Diamagnetic and paramagnetic currents passing a cut plane (a) through the HF bond and (b) through the hydrogen bond of the $(HF)_3$ trimer. Note that the maximum for the hydrogen bond is about a factor of 6 smaller than that for the HF molecule. The calculation is performed at the MP2/TZVPP level.

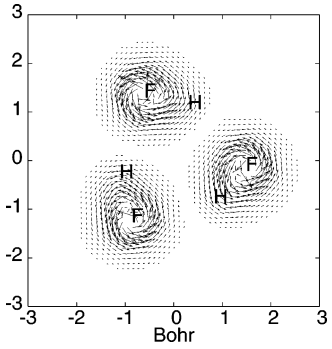


Figure 3. Magnetically induced current density in the molecular plane calculated at the MP2/TZVPP level.

the GIMIC approach. At the B3LYP/TZVPP level, we obtained, in the similar GIMIC calculation, a ring-current susceptibility of 0.44 nA/T. The ring-current susceptibility function calculated at the MP2/TZVPP level is shown in Figure 2. The ring-current susceptibility for benzene calculated at the CCSD/TZP level is 11.8 nA/T.[29] For comparison, the ring-current susceptibility for the edge current circling around the HF molecules calculated at the MP2/TZVPP and B3LYP/TZVPP levels are 9.1 and 8.8 nA/T, respectively. The magnetically induced current density in the molecular plane is shown in Figure 3. The current is diamagnetic on the outside of the molecule, whereas the return current on the inside of the HF monomer becomes paramagnetic (negative). At the MP2/TZVPP level, the integration of the current passing the cut plane across the hydrogen bond yields a ring-current susceptibility of 0.37 nA/T (0.46 nA/T at the B3LYP/TZVPP level), which is very close to the sum of the diamagnetic and paramagnetic currents circling the HF molecules, as it should be. The ARCS and GIMIC calculations of the ring-current susceptibility show that $(HF)_3$ should not be considered aromatic according to the ring-current criterion.

The NICS values obtained at different levels of theory are 2.89 ppm (MP2/QZ2P), 2.77 ppm (CCSD/QZ2P), 2.97 ppm (B3LYP/cc-pVQZ), and 3.24 ppm (B3LYP/TZVPP). Thus, according to the NICS values, the ring current circling the ring should be about one-third of the benzene value. Thus, the NICS calculation is not capable of providing the correct degree of aromaticity for this system, showing that NICS values as aromaticity indices cannot be used as a black-box tool but require careful consideration. Lazzeretti has also recently pointed out that the NICS value is not a very reliable aromaticity index.[40] This notion is supported by our study.

Conclusion

The magnetically induced current density and magnetic shieldings of the planar $(HF)_3$ trimer have been studied at ab initio and DFT levels using large basis sets. The calculation of the long-range magnetic shieldings and the explicit integration of the magnetically induced ring-current susceptibility yield ring-current susceptibilities of 0.37 nA/T (MP2/TZVPP), showing that $(HF)_3$ cannot be considered to be aromatic, as recently proposed by Rehaman et al.[20] The obtained NICS values are 2.77−3.24 ppm, depending on the computational level. The surprisingly large NICS value is not due to a magnetically induced ring-current and is, therefore, not an indication of molecular aromaticity but a result of the strong edge currents circling around each HF monomer in the ring. The ARCS and GIMIC calculations clearly show that the $(HF)_3$ ring is nonaromatic and that NICS values as aromaticity indices should be used with care.

Acknowledgment. We acknowledge financial support from the European research training network on "Understanding Nanomaterials from a Quantum Perspective" (NANOQUANT), Contract No. MRTN-CT-2003-506842, from the Nordisk Forskerakademi network for research and research training (NorFA Grant No. 030262) on "Quantum Modeling of Molecular Materials" (QMMM), from The Academy of Finland (FA Projects 53915, 200903, and 206102), and from Magnus Ehrnrooth's Foundation. We also thank Prof. Reinhart Ahlrichs (Karlsruhe) and Prof. Jürgen Gauss (Mainz) for up-to-date versions of the TURBOMOLE and ACES II program packages.

References

(1) Balaban, A.; von Ragué Schleyer, P.; Rzepa, H. S. *Chem. Rev.* **2005**, *105*, 3436−3447.

(2) Pople, J. A.; Untch, K. G. *J. Am. Chem. Soc.* **1966**, *88*, 4811−4815.

(3) Lazzeretti, P. *Prog. Nucl. Magn. Res. Spectrosc.* **2000**, *36*, 1−88.

(4) Jusélius, J.; Sundholm, D. *Phys. Chem. Chem. Phys.* **2001**, *3*, 2433−2437.

(5) Chen, Z.; Heine, T.; Sundholm, D.; von Ragué Schleyer, P. Aromaticity Indices from Magnetic Shieldings. In *Quantum Chemical Calculation of Magnetic Resonance Properties*; Kaupp, M., Bühl, M., Malkin, V., Eds.; Wiley-VCH: Weinheim, Germany, 2004.

(6) Bleeke, J. R. *Chem. Rev.* **2001**, *101*, 1205−1228.

(7) Nyulászi, L. *Chem. Rev.* **2001**, *101*, 1229−1246.

(8) Li, X.; Kuznetsov, A. E.; Zhang, H. F.; Boldyrev, A. I.; Wang, L. S. *Science* **2001**, *291*, 859−861.

(9) Kuznetsov, A. E.; Birch, K. A.; Boldyrev, A. I.; Li, X.; Zhai, H. J.; Wang, L. S. *Science* **2003**, *300*, 622−625.

(10) Boldyrev, A. I.; Wang, L. S. *Chem. Rev.* **2005**, *105*, 3716−3757.

(11) Childs, R. F. *Acc. Chem. Res.* **1984**, *17*, 347−352.

(12) Williams, R. V. *Chem. Rev.* **2001**, *101*, 1185−1204.

(13) Jusélius, J.; Patzschke, M.; Sundholm, D. *THEOCHEM* **2003**, *633*, 123−136.

(14) Hirsch, A.; Chen, Z.; Jiao, H. *Angew. Chem., Int. Ed.* **2000**, *39*, 3915−3917.

(15) Bühl, M.; Hirsch, A. *Chem. Rev.* **2001**, *101*, 1153−1183.

(16) Reiher, M.; Hirsch, A. *Chem.*−*Eur. J.* **2003**, *9*, 5442−5452.

(17) Johansson, M. P.; Sundholm, D.; Vaara, J. *Angew. Chem., Int. Ed.* **2004**, *43*, 2678−2681.

(18) Johansson, M. P.; Jusélius, J.; Sundholm, D. *Angew. Chem., Int. Ed.* **2005**, *44*, 1843−1846.

(19) King, R. B. *Chem. Rev.* **2001**, *101*, 1119−1152.

(20) Rehaman, A.; Datta, A.; Mallajosyula, S. S.; Pati, S. K. *J. Chem. Theory Comput.* **2006**, *2*, 30−36.

(21) Datta, A.; Pati, S. K. *Int. J. Quantum Chem.* **2005**, in press.

(22) von Ragué Schleyer, P.; Maerker, C.; Dransfeld, A.; Jiao, H.; van Eikema Hommes, N. J. R. *J. Am. Chem. Soc.* **1996**, *118*, 6317−6318.

(23) Chen, Z.; Wannere, C. S.; Corminboeuf, C.; Puchta, R.; von Ragué Schleyer, P. *Chem. Rev.* **2005**, *105*, 3842−3888.

(24) Aihara, J. *Chem. Phys. Lett.* **2002**, *365*, 34−39.

(25) Lin, Y. C.; Jusélius, J.; Sundholm, D.; Gauss, J. *J. Chem. Phys.* **2005**, *122*, 214308.

(26) Faglioni, F.; Ligabue, A.; Pelloni, S.; Soncini, A.; Viglione, R. G.; Ferraro, M. B.; Zanasi, R.; Lazzeretti, P. *Org. Lett.* **2005**, *7*, 3457−3460.

(27) Lin, Y. C.; Jusélius, J.; Sundholm, D.; Cui, L. F.; Li, X.; Zhai, H. J.; Wang, L. S. *J. Phys. Chem. A* **2006**, in press.

(28) Jusélius, J.; Sundholm, D. *Phys. Chem. Chem. Phys.* **1999**, *1*, 3429−3435.

(29) Jusélius, J.; Sundholm, D.; Gauss, J. *J. Chem. Phys.* **2004**, *121*, 3952−3963.

(30) Ahlrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165−169. Current version: see http://www.turbomole.de.

(31) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007−1023.

(32) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.

(33) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.

(34) Gauss, J. *Chem. Phys. Lett.* **1992**, *191*, 614−620.

(35) Gauss, J. *J. Chem. Phys.* **1993**, *99*, 3629−3643.

(36) Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **1995**, *102*, 251−253.

(37) Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **1995**, *103*, 3561−3577.

(38) Kollwitz, M.; Gauss, J. *Chem. Phys. Lett.* **1996**, *260*, 639−646.

(39) Kollwitz, M.; Häser, M.; Gauss, J. *J. Chem. Phys.* **1998**, *108*, 8295−8301.

(40) Lazzeretti, P. *Phys. Chem. Chem. Phys.* **2004**, *6*, 217−223.

CT050297X

JCTC Journal of Chemical Theory and Computation

Constrained Density Functional Theory and Its Application in Long-Range Electron Transfer

Qin Wu* and Troy Van Voorhis

Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received December 9, 2005

Abstract: Recently, we have proposed an efficient method in the Kohn−Sham density functional theory (DFT) to study systems with a constraint on their density (*Phys. Rev. A* **2005**, *72*, 24502). In our approach, the constrained state is calculated directly by running a fast optimization of the constraining potential at each iteration of the usual self-consistent-field procedure. Here, we show that the same constrained DFT approach applies to systems with multiple constraints on the density. To illustrate the utility of this approach, we focus on the study of long-range charge-transfer (CT) states. We show that constrained DFT is size-consistent: one obtains the correct long-range CT energy when the donor−acceptor separation distance goes to infinity. For large finite distances, constrained DFT also correctly describes the 1/*R* dependence of the CT energy on the donor−acceptor separation. We also study a model donor−(amidinium−carboxylate)−acceptor complex, where experiments suggest a proton-coupled electron-transfer process. Constrained DFT is used to explicitly calculate the potential-energy curves of both the donor state and the acceptor state. With an appropriate model, we obtain qualitative agreement with experiments and estimate the reaction barrier height to be 7 kcal/mol.

1. Introduction

Density functional theory (DFT) as in the framework of Hohenberg, Kohn, and Sham[1,2] is a ground-state theory because there is a one-to-one mapping between the ground-state density and the external potential. Time-dependent DFT (TDDFT) is based on the analogous mapping between the time-dependent density and external potential[3] and is an extremely useful tool for calculating excitation energies through the linear response of the ground-state density.[4,5] TDDFT, in principle, exactly produces the whole electronic spectrum. In practice, even with the seemingly crude adiabatic local density approximation (ALDA) where the ground-state time-independent exchange correlation functional is used in place of the time-dependent exchange correlation action, TDDFT calculations give excitation energies of low-lying states that agree well with experiments.[6,7] However, there are also a few well-known failures of ALDA and the approximate ground-state functional employed in TDDFT. For example, Rydberg states can only be correctly described by TDDFT after an asymptotic correction to the ground-state Kohn−Sham (KS) potential.[8−10] Here, the exchange correlation potential is forced to decay as $-1/r$ asymptotically with r being the distance from the nuclei. Recent studies also indicate the importance of the intermediate part of the exchange correlation potential.[11,12] Another failure of ALDA is that excitation energies of long-range charge-transfer (CT) states are seriously in error because of the local approximation in the exchange correlation kernel.[13−15] Long-range charge transfer is a key energy-transfer step in biological photosynthetic complexes,[16] dye-sensitized solar cells,[17] and organic light emitting diodes.[18] The failure of TDDFT for CT states puts its application to these systems in jeopardy. Therefore, much effort has been spent to explore the origin of the failure, and corrections have been proposed.[19−24] The origin of the failure has been linked to the derivative discontinuity of the exchange correlation potential[25] as well as self-interaction error. The energy of a long-range CT excitation can be approximated as the energy required to remove an electron from the electron donor, that is, the ionization potential (IP) of the

* Corresponding author fax: 617-253-7030; e-mail: qinwu@mit.edu.

10.1021/ct0503163 CCC: $33.50
Published on Web 03/15/2006

donor, minus the energy gain of putting the electron to the acceptor, that is, the electron affinity (EA) of the acceptor, plus a correction of the Coulomb interaction (cf. section 3.1 below). However, in DFT, though the highest occupied molecular orbital (HOMO) energy is the IP, the energy of the lowest unoccupied molecular orbital (LUMO) cannot be associated with EA because of the derivative discontinuity.[25] ALDA, on the other hand, gives a long-range CT excitation energy that is simply the energy difference between the HOMO of the donor and the LUMO of the acceptor because the kernel is local and provides no correction for long-range Coulomb interactions between the particle and hole. Thus, the ALDA excitation energies of long-range CT states are not reliable.

Recently, we have shown[26] that long-range CT states can be accurately captured through a constrained DFT method.[27] Constrained DFT is based on time-independent ground-state DFT. It focuses on a system whose ground-state density is required to satisfy some specific constraint. With the Lagrange multiplier approach, a constraining potential is introduced to enforce the desired density constraint. The key observation we make is that any excited state that has a qualitatively different density from the ground state can be obtained using constrained DFT. For instance, in a system consisting of an electron donor (D) and an electron acceptor (A), the ground state might have both the donor and acceptor being neutral (DA). In many applications, one is interested in the particular CT excited state that has a positive charge on the donor and a negative charge on the acceptor (D^+A^-). One can obtain this low-energy CT state by minimizing the total energy under an explicit constraint that the electron density has to correspond to D^+A^-. Because constrained DFT calculates the CT state directly rather than as a time-dependent response of the ground state, it avoids the problems associated with the exchange correlation kernel in ALDA. Therefore, constrained DFT gives good results for long-range CT state energies.[26]

Constrained optimization is, of course, not new to DFT. There is a basic constraint that has to be satisfied by all DFT methods when they minimize the energy: the electron density integrates to give the total number of electrons. The important quantity associated with this constraint is chemical potential (for a detailed discussion of chemical potential, see ref 28). In the Kohn−Sham formalism of DFT, where the lowest-energy orbitals of a noninteracting system are used to construct the electron density, the chemical potential can be chosen as any value between the energy levels of HOMO and LUMO. However, the constraining potentials associated with other constraints such as those for the CT states are not so obvious to determine. Therefore, most applications of constrained DFT rely on scanning over various potentials to find the one that meets the target criterion.[27,29−31] This is equivalent to applying various perturbations to the ground state and so is also referred to as the "perturbed ground state" method.[32] Recently, we have developed a direct optimization method to calculate the constraining potential.[26] Our method starts from the usual Lagrange multiplier approach for constrained optimization problems. However, instead of treating the constraining potential and orbitals as independent variables, we make use of the stationary equations for the orbitals and the technique of optimized effective potential[33,34] to transform the energy functional into a concave function of the constraining potential only.[35] The maximum then exactly corresponds to the target state. Because both the first and second derivatives are easily calculated, the optimization is very efficient. Also, our method yields the target state with one self-consistent field (SCF) calculation, which is computationally much less intensive compared to the scanning method. We have shown that constrained DFT gives very good results for long-range CT states:[26] the excitation energy is accurate and displays a negative inverse dependence on the donor−acceptor distance as it should. Recently, there have been suggestions of doing constrained calculations with wave-function-based methods.[36,37] We look forward to more exciting results in that aspect also.

In this report, we present the direct optimization method to handle systems with multiple constraints. Details of our implementation are described, such as the direct inverse of iterative subspace (DIIS) approach used in assisting SCF convergence and a comparison of different electron population schemes. We then test the method with calculations of infinitely separated CT states as well as a model system for proton-coupled electron transfer. Finally, we make a connection between constrained DFT and the Marcus theory of electron transfer[38,39] and briefly discuss possible applications.

2. Method

2.1. Constrained DFT with Multiple Constraints. Suppose there are m constraints that one wants to impose on the electron density, and the kth constraint can be written, in general, as

$$\sum_{\sigma} \int w_k^{\sigma}(\mathbf{r})\, \rho_{\sigma}(\mathbf{r})\, \mathrm{d}\mathbf{r} - N_k = 0 \tag{1}$$

where σ stands for either α or β spin, w_k is the weight function that defines the constrained property, and N_k is the constraint value. ρ_σ is calculated as $\sum_i^{N_\sigma} |\phi_{i\sigma}(\mathbf{r})|^2$ in the KS scheme, with $\phi_{i\sigma}$ being the ith lowest energy orbital and N_σ the number of σ electrons. The sum of ρ_α and ρ_β gives the total electron density ρ. Our goal is to minimize the electronic energy, a functional of ρ, under all constraints in eq 1. Creating a Lagrange multiplier, V_k, for each constraint and adding every product of the corresponding multiplier and constraint to the energy functional forms a new functional:

$$W[\rho,\{V_k\}] = E[\rho] + \sum_k^m V_k\Big[\sum_{\sigma} \int w_k^{\sigma}(\mathbf{r})\, \rho_{\sigma}(\mathbf{r})\, \mathrm{d}\mathbf{r} - N_k\Big] \tag{2}$$

$E[\rho]$ in the KS scheme is calculated by

$$E[\rho] = \sum_{\sigma}\sum_{i}^{N_\sigma} \left\langle \phi_{i\sigma} \middle| -\frac{1}{2}\nabla^2 \middle| \phi_{i\sigma} \right\rangle + \int \mathrm{d}\mathbf{r}\, v_n(\mathbf{r})\, \rho(\mathbf{r}) + J[\rho] + E_{\mathrm{xc}}[\rho^{\alpha},\rho^{\beta}] \tag{3}$$

where J is the classical Coulomb energy, E_{xc} is the exchange correlation energy, and v_n is the external potential. The

stationary equations of W with respect to the orbitals, which are required to be normalized, then are

$$\left[-\frac{1}{2}\nabla^2 + v_n(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,d\mathbf{r}' + v_{xc\sigma}(\mathbf{r}) + \sum_k^m V_k\, w_k^\sigma(\mathbf{r})\right]\phi_{i\sigma} = \epsilon_{i\sigma}\phi_{i\sigma} \quad (4)$$

with a similar one for $\phi_{i\sigma}^*$. These are the standard KS equations except for the addition of constraining potentials, $\sum_k^m V_k\, w_k^\sigma(\mathbf{r})$, in the effective Hamiltonian. Equation 4 is solved together with eq 1 to yield V_k and ϕ_i. In the following, when there is no confusion, we simply use V_c to represent the set of $\{V_k\}$, where $k = 1, ..., m$, and the same applies to N_c and w_c.

Because, for any given V_c, eq 4 uniquely determines a set of orbitals, if only the orbitals from eq 4 are to be explored, W becomes a function of V_c only. Furthermore, $W(V_c)$ is a concave function of V_c as we now show. Following ref 35, the first derivatives of $W(V_c)$ are

$$\frac{\partial W}{\partial V_k} = \sum_\sigma \sum_i^{N_\sigma} \left(\frac{\delta W}{\delta \phi_{i\sigma}^*}\frac{\partial \phi_{i\sigma}^*}{\partial V_k} + \text{cc}\right) + \frac{\partial W}{\partial V_k} = \sum_\sigma \int w_k^\sigma(\mathbf{r})\,\rho_\sigma(\mathbf{r})\,d\mathbf{r} - N_k \quad (5)$$

Here, the fact that $\delta W/\delta\phi_{i\sigma}^* = 0$, that is, eq 4, has been used. The stationary point of $W(V_c)$, which means $\partial W/\partial V_k = 0$, then restores the constraints of eq 1 automatically. It remains to show that the second derivative matrix (Hessian) of $W(V_c)$ is nonpositive definite to prove that $W(V_c)$ is concave. Starting from eq 5,

$$\frac{\partial^2 W}{\partial V_k \partial V_l} = \sum_\sigma \sum_i^{N_\sigma} \int w_k^\sigma(\mathbf{r})\,\phi_{i\sigma}^*(\mathbf{r}) \frac{\delta \phi_{i\sigma}(\mathbf{r})}{\delta[V_l\, w_l^\sigma(\mathbf{r}')]} w_l^\sigma(\mathbf{r}')\,d\mathbf{r}\,d\mathbf{r}' + \text{cc}$$

$$= \sum_\sigma \sum_i^{N_\sigma} \int w_k^\sigma(\mathbf{r})\,\phi_{i\sigma}^*(\mathbf{r}) \sum_{a\neq i} \frac{\phi_{a\sigma}^*(\mathbf{r}')\,\phi_{i\sigma}(\mathbf{r}'')}{\epsilon_{i\sigma}-\epsilon_{a\sigma}} \phi_{a\sigma}(\mathbf{r})\, w_l^\sigma(\mathbf{r}')\,d\mathbf{r}\,d\mathbf{r}' + \text{cc}$$

$$= 2\sum_\sigma \sum_i^{N_\sigma} \sum_{a>N_\sigma} \frac{\langle\phi_{i\sigma}|w_k^\sigma|\phi_{a\sigma}\rangle\langle\phi_{i\sigma}|w_l^\sigma|\phi_{a\sigma}\rangle}{\epsilon_{i\sigma}-\epsilon_{a\sigma}} \quad (6)$$

Here, the first-order perturbation theory is used to evaluate $\delta\phi_{i\sigma}(\mathbf{r})/\delta[V_l\, w_l^\sigma(\mathbf{r}')]$. In the final expression, the index i goes over occupied orbitals, while a only has to go over the unoccupied orbitals because the summand is antisymmetric with the exchange of i and a. The Hessian is nonpositive definite because, for any V_c,

$$\sum_{k,l}^m V_k \frac{\partial^2 W}{\partial V_k \partial V_l} V_l = 2\sum_\sigma \sum_i^{N_\sigma} \sum_{a>N_\sigma} \frac{\langle\phi_{i\sigma}|\sum_k^m V_k\, w_k^\sigma|\phi_{a\sigma}\rangle^2}{\epsilon_{i\sigma}-\epsilon_{a\sigma}} \le 0 \quad (7)$$

This is true because in the KS scheme the occupied orbitals are chosen as the lowest eigenstates; hence, $\epsilon_{i\sigma} - \epsilon_{a\sigma} < 0$, while the numerator is non-negative. This implies that there is only one stationary point of $W(V_c)$ and that it is a maximum. Thus, by optimizing W through varying V_c, one can find the constraining potential that exactly yields the ground state of the constrained system. Because both the first and second derivatives are easily calculated, the optimization can be done efficiently.

To study charge-transfer states, which are the focus of this work, one can put two separate constraints simultaneously on the donor and acceptor charges. Alternatively, one can impose just one constraint on the charge difference between the donor and the acceptor; that is, $N_c = (N_D - N_A)/2$, where N_D and N_A are the net charges on D and A.[26] This can be done by defining the weight function in eq 1 to be positive on the donor and negative on the acceptor, which effectively constrains both the donor and the acceptor, and should give the same result for the charge-separated state as that using two constraints. Therefore, in what follows, we choose to use one constraint for simplicity.

2.2. Convergence. We have implemented the direct optimization approach in NWChem.[40] Our method is formulated as an outer loop and an inner loop. First, to solve the KS equations, a SCF procedure is employed, wherein orbitals are iteratively improved. This constitutes the outer loop. Second, at each SCF iteration, an optimization is carried out to find the optimal V_c, which is the inner loop. For the optimization, the availability of both the first and second derivatives recommends the use of Newton's method (see, for example, ref 41), which can reduce the number of inner iterations. This is important because every trial V_c demands a matrix diagonalization to obtain the corresponding orbitals. However, because only V_c is updated and the Fock matrix is not rebuilt at each inner iteration, the inner loop is still relatively cheap, at least for atom-centered basis sets where the number of basis functions per atom is small.

DIIS is a powerful tool to assist the convergence of SCF procedures. With DIIS, the current Fock matrix at the nth iteration is replaced by a linear combination of previous Fock matrices. A detailed description of how to determine the linear coefficients is provided by Pulay.[42,43] Now that there is an extra inner loop, constrained DFT can still make use of DIIS to accelerate the outer-loop convergence. But because of changes to the KS equations in eq 4, the Fock matrix needs to be handled carefully. The way we use DIIS in our calculations can be summarized as follows: (1) Construct the current Fock matrix $\mathbf{F}$ from the current density matrix $\mathbf{P}$. (2) Use the optimal V_c from the last iteration to build the constrained Fock matrix $\mathbf{F}_c = \mathbf{F} + V_c\mathbf{w}_c$. (3) Determine the DIIS linear coefficients d^i, and replace the current Fock matrix with $\mathbf{F}^* = \sum_i^n d^i \mathbf{F}_c^i$. (4) Fix $\mathbf{F}^*$, and optimize V_c again until the constraints are satisfied. (5) Obtain the new density matrix from $\mathbf{F}^*$ and the optimized V_c. The new $\mathbf{P}$ and the optimized V_c are then fed into the next iteration, and the above steps are repeated until convergence. We find that the modified SCF procedure is as robust as that in normal DFT calculations.

2.3. Electron Population in a Molecule. For CT, the density constraints are generally expressed as constraints on the number of electrons associated with an atom (or a group

of atoms) in the molecule, that is, the electron population. There is no unambiguous way of defining an atom in a molecule; nevertheless, this concept must be defined for chemists.[44] When we calculate the electron population, we first consider the easiness of its implementation in existing DFT codes. We will further explore the optimal definition in our future work.

It is preferable to have the electron number of the constrained area calculated by $N_c = \mathrm{Tr}(\mathbf{P}\mathbf{w}_c)$, where Tr stands for the matrix trace, $\mathbf{P}$ is the density matrix, and $\mathbf{w}_c$ is a weight matrix in the basis set space defining both the constrained area, that is, the donor and acceptor groups, and the constrained property, that is, the charge difference between the donor and the acceptor. $\mathbf{w}_c$ should be formed at the beginning of a calculation and remain unchanged throughout the calculation. Below, we show how the Mulliken population, the Löwdin population, and Becke's multicenter integration scheme, all of which are readily available in most DFT codes, can be easily turned into a weight matrix.

For Mulliken populations,[45] the electron population of a group, for example, C, of atoms in a molecule is calculated as

$$N_c = \sum_{\mu \in C} (\mathbf{PS})_{\mu\mu} = \mathrm{Tr}(\mathbf{P}\mathbf{w}_c^{\mathrm{M}})$$

where $\mathbf{S}$ is the overlap matrix and the Mulliken weight matrix $\mathbf{w}_c^{\mathrm{M}}$ is

$$w_{c\mu\nu}^{\mathrm{M}} = \begin{cases} S_{\mu\nu} & \text{if } \mu \in C \text{ and } \nu \in C \\ \frac{1}{2}S_{\mu\nu} & \text{if } \mu \in C \text{ or } \nu \in C \\ 0 & \text{if } \mu \ni C \text{ and } \nu \ni C \end{cases}$$

For Löwdin populations,[46] the population of the same group of atoms is given by

$$\begin{aligned} N_c &= \sum_{\mu \in C} (\mathbf{S}^{1/2}\mathbf{P}\mathbf{S}^{1/2})_{\mu\mu} \\ &= \sum_{\mu \in C}\sum_{\nu\lambda} S_{\mu\nu}^{1/2} P_{\nu\lambda} S_{\lambda\mu}^{1/2} \\ &= \sum_{\nu\lambda} P_{\nu\lambda} \sum_{\mu \in C} S_{\lambda\mu}^{1/2} S_{\mu\nu}^{1/2} \\ &= \mathrm{Tr}(\mathbf{P}\mathbf{w}_c^{\mathrm{L}}) \end{aligned}$$

where $w_{c\lambda\nu}^{\mathrm{L}} = \sum_{\mu \in C} S_{\lambda\mu}^{1/2} S_{\mu\nu}^{1/2}$, defining the Löwdin weight matrix.

Becke's multicenter integration scheme[47] is widely used in DFT calculations. Here, the integration of the whole molecule is decomposed into single-center integrations by assigning a relative weight function $w_n(\mathbf{r})$ to each nucleus n. $w_n(\mathbf{r})$ is equal to 1 around its own nucleus and drops to 0 rapidly but continuously when in a region closer to another nucleus. Hence, these weight functions partition the molecule into fuzzy cells, and each cell can be viewed as an atom. Projecting the weight functions $w_n(\mathbf{r})$ of the atoms in C from real space to the basis set space then forms the Becke weight matrix $\mathbf{w}_c^{\mathrm{B}}$ with $w_{c\lambda\nu}^{\mathrm{B}} = \sum_{n \in C} \int \mathrm{d}\mathbf{r}\, \chi_\lambda(\mathbf{r})\, w_n(\mathbf{r})\, \chi_\nu(\mathbf{r})$.

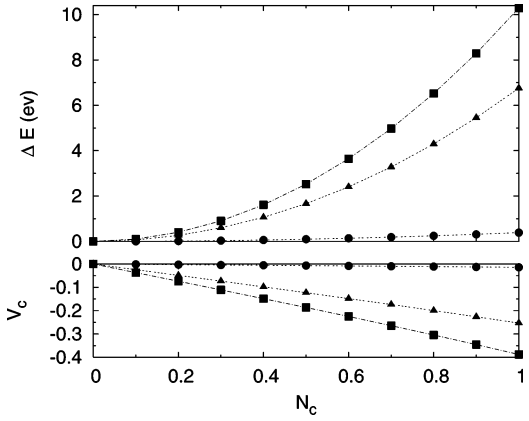


Figure 1. Energy and constraining potential as against the charge separation in N_2 with different population schemes. Squares: Becke weights population. Triangles: Löwdin population. Dots: Mulliken population. Calculations are done using Becke's three-parameter hybrid functional[48] with Lee−Yang−Parr correlation[49] and a 6-31G* basis set.

First, we compare the results of the above three population schemes. In Figure 1, the constraining potential and the increase of total energy are plotted against the amount of charge separation between the two N atoms in the N_2 molecule with a bond length of 1.12 Å. The Mulliken population apparently gives very different results from the Löwdin population and the Becke weights population. The unusually small changes in the potential and energy indicate that the results of the Mulliken population are qualitatively wrong. The reason is that the Mulliken population is not defined by a projection operator[50]—for example, a Mulliken population can be negative—therefore, Mulliken populations do not give well-defined constraints. The difference between the Löwdin population and the Becke weights population, though not small for N_2 where the two atoms are very close to each other, becomes insignificant when the donor and the acceptor are far apart. This point is well-illustrated in Figure 2, where the constraining potentials and energies of the CT states for the zincbacteriochlorin−bacteriochlorin (ZnBC−BC) complex[21] are shown. Here, $N_c = 1$ is the $\mathrm{ZnBC}^+-\mathrm{BC}^-$ state and $N_c = -1$ is the $\mathrm{ZnBC}^- - \mathrm{BC}^+$ state. Only results of the Löwdin population (black triangles) and the Becke weights population (lines) are presented, and they are nearly indistinguishable. Because the Löwdin population is easier to compute and more commonly used, it will be our first choice in all calculations. We have also tested a variant of the Löwdin population, namely, the atomic-orthogonalized Löwdin population,[51] and found the same qualitative results as those for the Löwdin population.

3. Calculations and Results

3.1. Charge-Separated States at an Infinite Distance. A good test of our method is to calculate the energy of a charge-separated (CS) state at an infinite distance, $E_{\mathrm{CS}}(\infty)$. Consider a system of an electron donor and an acceptor that are both neutral initially. The CS state is formed by transferring one electron from the donor to the acceptor. When the donor and the acceptor are far apart and there is nothing between them, one can make the valid approximation to treat both parts as point charges and obtain the relationship between

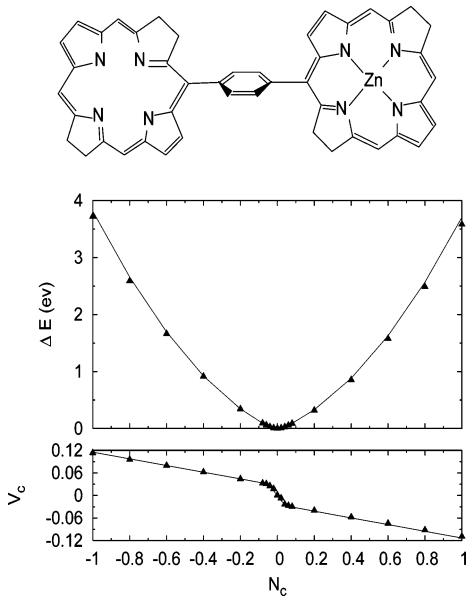


Figure 2. Energy and constraining potential as against the charge separation in the ZnBC−BC complex (shown above). Results of the Becke weights population form the line, and the black triangles are data of the Löwdin population. Calculations are done with Becke's exchange functional[52] and LYP correlation functional[49] using a 6-31G* basis set.

Table 1. Comparison of Charge-Separated State Energies with the Electron Donor (D) and Acceptor (A) Infinitely Apart and the Sum of Ionic Energies of D and A[a]

D	A	$E_D^+ + E_A^-$	$E_{CS}(\infty)$
N_2	N_2	−218.360 411	−218.361 386
H_2O	F_2	−275.391 972	−275.392 850
C_2F_4	C_2H_4	−553.595 853	−553.595 591

[a] All numbers are in hartrees.

the energy of the CS state and the separation distance (R) as In our previous report,[26] we have shown the excellent linear

$$E_{CS}(R) \approx -\frac{1}{R} + E_{CS}(\infty) \tag{8}$$

dependence of $E_{CS}(R)$ on $1/R$ given by constrained DFT. Here, we study the value $E_{CS}(\infty)$. It is evident that $E_{CS}(\infty)$ should equal the sum of individually calculated energies of D^+ and A^-. Therefore, by fitting $E_{CS}(R)$ and $1/R$ data to a straight line, the intersection parameter, which is $E_{CS}(\infty)$, should have the same value as $E_D^+ + E_A^-$.

We have tested CS state energies of three donor and acceptor molecular pairs as listed in Table 1. All calculations are done with the B3LYP functional[48,49,52] and 6-31G* basis set. Molecules are placed in two parallel planes that are perpendicular to the axis connecting their centers. The distance between the planes is defined as R, and five different R values are used, starting from $R = 8$ Å and increasing by 0.5 Å at each point. The differences between $E_{CS}(\infty)$ and $E_D^+ + E_A^-$ in all three cases are less than 1 millihartree, which are attributable to the point charge approximation and fitting errors. Thus, constrained DFT has the correct limit for a CS state. Together with the linear relationship shown in our previous work,[26] these tests demonstrate that constrained DFT is very appropriate to study long-range charge-separated states.

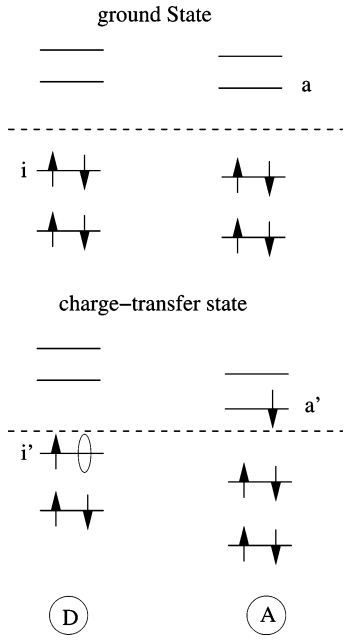


Figure 3. Illustrative picture of the frontier orbitals of the ground and CT states. The donor and acceptor molecules are far apart, so an orbital from D does not overlap with that from A. Note that orbitals *i* and *a* are different from *i′* and *a′*.

In the CS state calculations, it is important to use the unrestricted Kohn−Sham (UKS) method to break the spin symmetry because the ions are both spin-polarized and calculated with UKS. Though restricted KS (RKS) energies display a linear dependence on $1/R$ too, RKS approaches a different asymptotic value. For instance, the RKS $E_{CS}(\infty)$ for the H_2O-F_2 pair is −275.273 965 hartree, significantly higher than the unrestricted value −275.392 850 hartree.

Before we present more results, we want to explore the reason constrained DFT works for long-range CT states while TDDFT fails. Here, we follow the argument of Dreuw and Head-Gordon.[21] Figure 3 is a sketch of the frontier orbitals of the donor and acceptor molecules for the ground and CT states. Zero overlap between the orbitals of D and A is assumed. Because TDDFT uses the ground state as the reference to calculate the CT state (through linear response), it requires exact exchange to correctly account for the electron−hole Coulomb attraction, which is missing in existing pure density functionals. It also suffers from the fact that the LUMO energy in DFT cannot be connected with the electron affinity when it uses orbital a of the ground state for such a purpose. Constrained DFT builds the CT state directly, so the reference is not the ground state anymore, but the real CT state. Therefore, the electron−hole Coulomb interaction is calculated classically. Moreover, orbital a' of A^- is now the HOMO, which can be used as an approximation to EA of A.[53] The excitation energy of the CT state then is approximately

$$\begin{aligned}\omega_{CT} &\approx -(i'i'|a'a') - \epsilon_i + \epsilon_{a'} \\ &\approx -1/R + IP^D - EA^A \end{aligned} \tag{9}$$

which is the same formula as that used in ref 22.

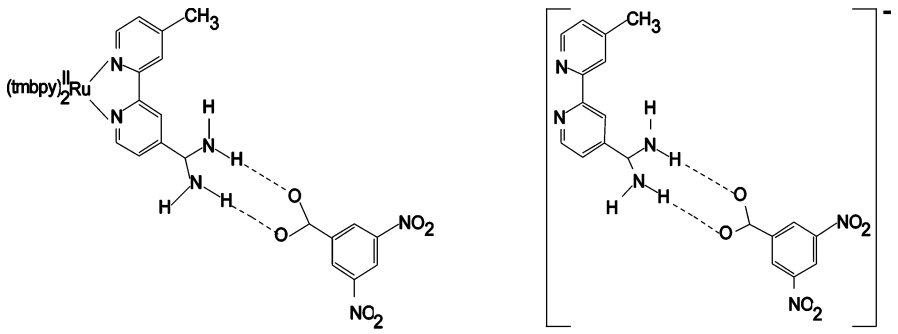


Figure 4. Real (**1**) and model systems (**2**) in studying electron transfer through hydrogen-bonded interfaces.

3.2. Electron Transfer through Hydrogen-Bonded Interfaces. Next, we study the electron transfer in a donor−(amidinium−carboxylate)−acceptor complex. Amidinium and carboxylate form a hydrogen-bonded interface that models the aspartate−arginine salt bridges in many important biological structures[54−59] and is a well-defined model for studying proton-coupled electron-transfer (PCET) reactions. Extensive studies on the kinetics of this complex have been reported experimentally and theoretically.[60−63] To predict the kinetics theoretically, it is very important to obtain the potential-energy surface for proton transfer, which can be calculated with constrained DFT.

Here, we pick the system **1**, as shown in Figure 4, for our study,[61] where tmbpy = 3,3′,4,4′-tetramethyl-2,2′-bipyridine. In this system, the electron donor is $(tmbpy)_2Ru^{II}$(Mebpy), where Mebpy = 4-methyl-2,2′-bipyridine, and the acceptor is 3,5-dinitrobenzene. As a first attempt, we optimized the geometry of **1** with the B3LYP functional and 6-31G* basis set, which is done with the Gaussian 03 program.[64] Unfortunately, one of the bridge protons moves away from the nitrogen and forms a bond with the oxygen. Thus, the structure of **1** in Figure 4 is not the most stable configuration in the gas phase, probably because the positive charge on Ru^{II} repels the bridge protons so that one of them is pushed away. What makes **1** stable in solution is solvent molecules, which are difficult to include in our calculations. Therefore, we adopt the model system **2** shown in Figure 4. The model system has an overall negative charge because an electron transfers from ruthenium to Mebpy upon excitation. However, during the excitation process, the nuclear framework should not change significantly (i.e., we make the Franck−Condon approximation). Thus, we optimize the geometry of **2** at zero charge, that is, before the electron is injected. The optimized structure agrees qualitatively with the picture in Figure 4, with both bridge protons on the nitrogen side. The N−H bond length and the N−O distance are 1.0817 and 2.6181 Å, respectively, for the upper bridge and 1.0788 and 2.6261 Å, respectively, for the lower bridge. In both bridges, the N−O distance equals the sum of N−H and O−H distances, indicating that N, H, and O can be connected by a straight line. We then fix all nuclear positions except for the upper bridge proton, which we take as the one coupled with the electron transfer. This proton is moved manually from N to O at several points along the bridge, and the energy at each point is calculated. According to B3LYP calculations, the charge is almost evenly distributed over the donor and the acceptor at all proton positions. This is because of DFT methods' artificial stabilization of systems with fractional numbers of electrons.[65,66] Hence, normal unconstrained DFT calculations cannot predict the correct electronic configuration of this molecule. Instead, energies of the states that have the charge constrained on the donor (Mebpy) or the acceptor (dinitrobenzene) are calculated at each point. With the potential-energy curves of these two states, one can then tell the preferred electronic state as the proton transfers.

We use D^-NOA and DNOA^- to denote the donor state and the acceptor state, respectively. Experimental evidence suggests that a bridge proton moves from the nitrogen side to the oxygen side as an electron transfers from the donor to the acceptor.[61] Thus, one expects the energetic picture to be that D^-NOA has a lower energy when the proton stays with the nitrogen and DNOA^- has a lower energy after the proton transfers to bond with the oxygen. Figure 5a shows the constrained B3LYP results using the 6-31G* basis set and Löwdin population. In this graph as well as in Figures 5b and 6, the (0,0) coordinate represents the state with the charge constrained on the acceptor and the proton at the equilibrium position for the neutral complex. As the position coordinate increases, the proton moves closer to the oxygen. It is clear from this graph that DNOA^- is a more stable state at all proton positions. This is contradictory to the experimental results. The calculation results do not change qualitatively when the Becke weights population is used (lines in Figure 5b). We also tested a different functional: B97-2,[67] which is reported to be better in predicting reaction barriers,[68] but the same picture is obtained (pluses in Figure 5b). Finally, we used a much larger basis set of 6-311++G**, but the DNOA^- is still much lower in energy than D^-NOA (squares in Figure 5b).

To test the reliability of DFT methods in this case, we calculated the EA of Mebpy and dinitrobenzene. We took the geometry of Mebpy and dinitrobenzene as they are in the model system, used a hydrogen to terminate each group, and then calculated the vertical EA as E(neutral) − E(anion) at the same geometry. With B3LYP and 6-31G*, the vertical EA of Mebpy is −7.9 kcal/mol, which means its anion is not stable. This result is consistent with the fact that there have been no experimental observations of the EA for this molecule. The vertical EA of *m*-dinitrobenzene is 30.2 kcal/mol, in good agreement with the experimental value 38.05 kcal/mol.[69] This result concurs with other findings that DFT generally gives a good approximation to EA.[70−72] With such

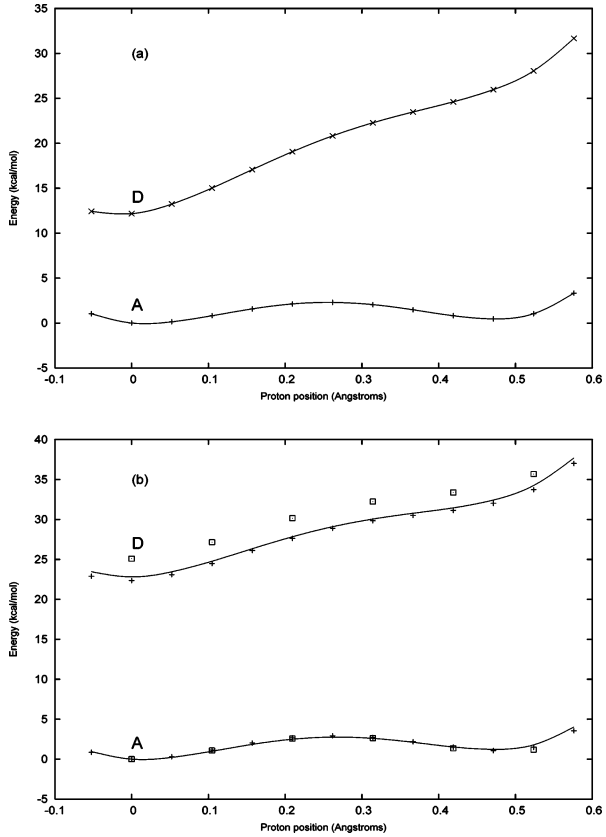


Figure 5. Constrained state energies of D^-NOA (D) and $DNOA^-$ (A) as a function of the proton position. (a) B3LYP/6-31G* and the Löwdin population. (b) Lines: B3LYP/6-31G* and Becke weights population. Crosses: B97-2/6-31G* and Becke weights population. Squares: B97-2/6-311++G** and Becke weights population.

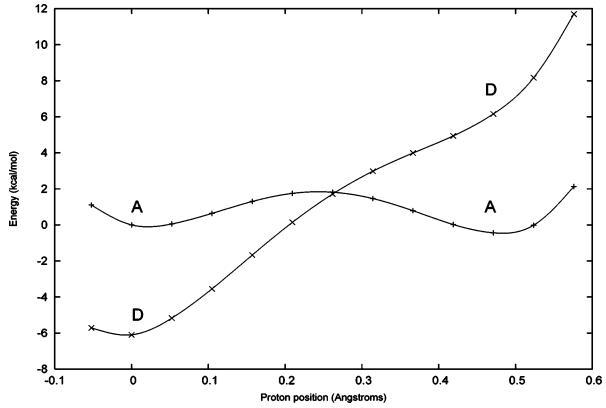


Figure 6. Constrained state energies of D^-NOA (D) and $DNOA^-$ (A) as a function of the proton position. A point charge of +0.3 is in place of Ru in the real system, and B3LYP/6-31G* and the Löwdin population are used.

a big difference in EA, it is then not strange that the electron prefers to be with dinitrobenzene—the electron acceptor in the model complex—which means $DNOA^-$ should be lower in energy than D^-NOA as found by constrained DFT.

Realizing that the discrepancy between our calculation results and the experimental results is probably not a weakness of DFT, we go back to check the model system that we chose. When we use the model system **2** to replace the real system **1** and put a negative charge on the model system, we neglect the positive hole on $Ru^{II}(tmbpy)_2$. While the actual charge distribution is hard to determine without a full scale calculation, its effect on the relative energy of the two locally charged states of the model system can be analyzed: with the positive fragment much closer to the donor than to the acceptor, it will bring down the energy of D^-NOA more than that of $DNOA^-$. To illustrate the effect, we simulate the positive fragment simply by putting a point charge at the location of Ru in the real system. The point charge is in the same plane as Mebpy and has a distance of 2.10 Å from both nitrogens in Mebpy. This choice is determined from the optimized geometry of $(tmbpy)_2Ru^{II}$-(Mebpy)-amH^+. Placing a whole charge of +1 would exaggerate the effect of the positive fragment and make D^-NOA lower in energy than $DNOA^-$ at all proton positions. A partial charge of +0.3 is found to give results in qualitative agreement with experiments (Figure 6). The point charge approach is no doubt a very crude simulation. Using a partial charge further makes the charge an ad hoc parameter. However, our purpose here is merely to demonstrate that, with an appropriate model system, constrained DFT generates qualitatively correct results.

From the results in Figure 6, the potential-energy difference (ΔG) between the D^-NOA state before proton transfer and the $DNOA^-$ state after proton transfer is about 5 kcal/mol, and the reaction barrier height is about 7 kcal/mol. Comparing Figure 6 with Figure 5a, one can see that the effect of the point charge is to vertically shift the two curves. Hence, the maximum possible barrier height and ΔG is about 13 kcal/mol (the minimum after proton transfer of the $DNOA^-$ curve has to be below the D^-NOA curve to make the reaction happen). It is known that DFT generally underestimates barrier heights[68,73−77] because DFT methods overstabilize the transition state because of greater self-interaction errors for systems of fractional charges than for systems of integer charges[66] (or, from an alternative point of view, because of the spurious buildup of nondynamical correlation[78]). In our calculations, however, we have largely reduced this effect by constraining the donor and the acceptor to have integer charges. Though some error still remains because the bridge charges are not constrained, we expect it to be small. Further, even without constraints, the B3LYP and B97-2 functionals have typical errors of about 5 and 3 kcal/mol, respectively, for barrier heights.[68] ΔG is calculated at equilibrium structures and has no known systematic errors for this type of system. Thus, we feel that the results in Figure 6 should be fairly accurate. Previous work[63] used multiconfiguration SCF (MCSCF) to study the same model system and found a much larger barrier height of 25 kcal/mol. However, the MCSCF calculation has only one electron in the active space, so it is effectively a Hartree−Fock (HF) calculation and does not include correlations. Hence, the lowering of the barrier height is probably due to the correlation effects that DFT methods account for. It is known that HF tends to overestimate barrier heights and needs correlations to lower them.[68] In this particular study, we note that this correlation-induced barrier lowering is *significant* and could potentially change our understanding of how PCET happens in this system.

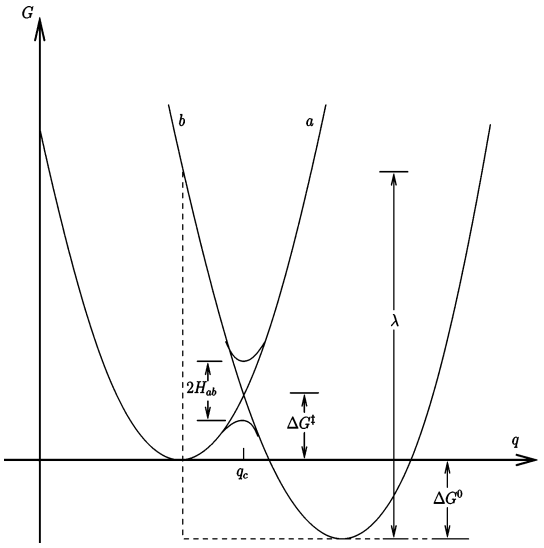


Figure 7. Picture of Marcus theory for electron-transfer reactions.

Conclusions

It has been shown that the direct optimization approach to constrained DFT can be generalized to include multiple constraints, where the constraint potentials are obtained efficiently within an expanded SCF calculation. When the proper Fock matrix is chosen, DIIS can be used to accelerate SCF convergence in a manner similar to that of usual DFT calculations, making constrained DFT truly efficient. Comparing different choices in defining the electron population of a molecule, we have found that the Löwdin population and the Becke weights population behave similarly and give nearly identical results for long-range charge-separated states. Generally, when the electron donor and acceptor have little density overlap, the constrained DFT results are well-defined. Hence, constrained DFT is naturally suited to long-range ET reactions. At short distances, the population scheme will be critical to the reliability of constrained DFT, and we will address this issue in future work.

To demonstrate the application of constrained DFT, we have done calculations on donor and acceptor molecular pairs that are separated far from each other and have shown that the energy of an infinitely separated CT state is equal to the energy sum of the donor and acceptor ionic states, as it should be. We have also studied a model system for the proton-coupled electron transfer in the donor−(amidinium−carboxylate)−acceptor complex. The potential-energy curves of the donor and acceptor states are generated using constrained DFT. With a partial point charge to simulate the positive fragment left out from the model system, we have obtained qualitative agreement with experimental results: the donor state is more stable before proton transfer, and the acceptor state is more stable after proton transfer. The barrier height is estimated to be 7 kcal/mol, which is significantly lower than the 25 kcal/mol barrier obtained with an uncorrelated MCSCF approach.

Before we close this report, we make the observation that the constrained potential-energy curves in the PCET case are actually *diabatic* curves, which implies that constrained DFT can fit well into Marcus theory[38,39] to describe electron-transfer reactions (Figure 7). Diabatic states *a* and *b*, which correspond to states before and after ET, are just two different constrained states. To make use of Marcus theory, however, requires locating the minimum energy point on each diabatic curve first. This means that one needs to be able to calculate the forces in constrained DFT and, therefore, perform geometry optimizations. Because of the variational nature of our method, it is not difficult to calculate the forces analytically in constrained DFT. We shall present an efficient computational scheme and its applications to electron-transfer systems in a future work.

As a final note, constrained DFT is not limited to electron-transfer calculations. By constraining the spin density instead of the total density, one can calculate exchange coupling constants in magnetic molecules.[79] There are other studies using constrained DFT on charge[29] and magnetization[30] fluctuations in solids, spin-dependent sticking of molecules on surfaces,[80] and parametrization of model Hamiltonians based on DFT calculations.[81] We expect our method to be useful for future applications in all of these areas.

Acknowledgment. We are grateful to Drs. Dreuw and Head-Gordon for providing us the ZnBC−BC complex structure. This work was made possible by a start-up grant from M. I. T.

Note Added after ASAP Publication. This article was released ASAP on March 15, 2006, with the incorrect Received Date. The correct version was posted on April 17, 2006.

References

(1) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864.

(2) Kohn, W.; Sham, L. *Phys. Rev.* **1965**, *140*, A1133.

(3) Runge, E.; Gross, E. K. U. *Phys. Rev. Lett.* **1984**, *52*, 997−1000.

(4) Casida, M. E. In *Recent Advances in Density Functional Methods, Part I*; Chong, D. P., Ed.; World Scientific: Singapore, 1995; pp 155−192.

(5) Petersilka, M.; Gossmann, U. J.; Gross, E. K. U. *Phys. Rev. Lett.* **1996**, *76*, 1212−1215.

(6) Bauernschmitt, R.; Ahlrichs, R. *Chem. Phys. Lett.* **1996**, *256*, 454−464.

(7) Hirata, S.; Head-Gordon, M. *Chem. Phys. Lett.* **1999**, *302*, 375−382.

(8) van Leeuwen, R.; Baerends, E. J. *Phys. Rev. A* **1994**, *49* (4), 2421−2431.

(9) Casida, M. E.; Jamorski, C.; Casida, K. C.; Salahub, D. R. *J. Chem. Phys.* **1998**, *108* (11), 4439−4449.

(10) Tozer, D. J. *Phys. Rev. A* **1998**, *58* (5), 3524−3527.

(11) Wu, Q.; Cohen, A. J.; Yang, W. *Mol. Phys.* **2005**, *103*, 711−717.

(12) Ayers, P.; Morrison, R.; Parr, R. *Mol. Phys.* **2005**, *103*, 2061−2072.

(13) Tozer, D. J.; Amos, R. D.; Handy, N. C.; Roos, B. O.; Serrano-Andres, L. *Mol. Phys.* **1999**, *97* (7), 859−868.

(14) Sobolewski, A. L.; Domcke, W. *Chem. Phys.* **2003**, *294* (1), 73−83.

(15) Fabian, J. *Theor. Chem. Acc.* **2001**, *106* (3), 199−217.

(16) Polıvka, T.; Sundström, V. *Chem. Rev.* **2004**, *104*, 2021−2071.

(17) Grätzel, M. *Nature* **2001**, *414*, 338−344.

(18) Mitschke, U.; Bäuerle, P. *J. Mater. Chem.* **2000**, *10*, 1471−1507.

(19) Tozer, D. J. *J. Chem. Phys.* **2003**, *119*, 12697.

(20) Dreuw, A.; Weisman, J. L.; Head-Gordon, M. *J. Chem. Phys.* **2003**, *119*, 2943−2946.

(21) Dreuw, A.; Head-Gordon, M. *J. Am. Chem. Soc.* **2004**, *126*, 4007−4016.

(22) Gritsenko, O. V.; Baerends, E. J. *J. Chem. Phys.* **2004**, *121* (2), 655−660.

(23) Tawada, Y.; Tsuneda, T.; Yanagisawa, S.; Yanai, T.; Hirao, K. *J. Chem. Phys.* **2004**, *120* (18), 8425−8433.

(24) Maitra, N. T. *J. Chem. Phys.* **2005**, *122* (23), 234104.

(25) Perdew, J. P.; Parr, R. G.; Levy, M.; Balduz, J. L., Jr. *Phys. Rev. Lett.* **1982**, *49*, 1691−1694.

(26) Wu, Q.; Van Voorhis, T. *Phys. Rev. A* **2005**, *72*, 024502.

(27) Dederichs, P. H.; Blügel, S.; Zeller, R.; Akai, H. *Phys. Rev. Lett.* **1984**, *53* (26), 2512−2515.

(28) Parr, R. G.; Yang, W. *Density-Functional Theory of Atoms and Molecules*; Oxford University Press: New York, 1989.

(29) Akai, H.; Blügel, S.; Zeller, R.; Dederichs, P. *Phys. Rev. Lett.* **1986**, *56* (22), 2407−2410.

(30) Zhang, Z.; Satpathy, S. *Phys. Rev. B* **1991**, *44* (24), 13319−13331.

(31) Cioslowski, J.; Stefanov, B. B. *J. Chem. Phys.* **1993**, *99* (7), 5151−5162.

(32) Prezhdo, O. V.; Kindt, J. T.; Tully, J. C. *J. Chem. Phys.* **1999**, *111* (17), 7818−7827.

(33) Talman, J. D.; Shadwick, W. F. *Phys. Rev. A* **1976**, *14*, 36−40.

(34) Yang, W.; Wu, Q. *Phys. Rev. Lett.* **2002**, *89*, 143002.

(35) Wu, Q.; Yang, W. *J. Chem. Phys.* **2003**, *118*, 2498−2509.

(36) Pan, X. Y.; Sahni, V.; Massa, L. *Phys. Rev. Lett.* **2004**, *93*, 130401.

(37) Pan, X. Y.; Sahni, V.; Massa, L. *Phys. Rev. A.* **2005**, *72*, 32505.

(38) Marcus, R. A. *Rev. Mod. Phys.* **1993**, *65*, 599−610.

(39) Barbara, P. F.; Meyer, T. J.; Ratner, M. A. *J. Phys. Chem.* **1996**, *100*, 13148−13168.

(40) High Performance Computational Chemistry Group. *NWChem, A Computational Chemistry Package for Parallel Computers*, version 4.6; Pacific Northwest National Laboratory: Richland, WA 99352, 2004.

(41) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes*; Cambridge University Press: Cambridge, U. K., 1992.

(42) Pulay, P. *Chem. Phys. Lett.* **1980**, *73* (2), 393−398.

(43) Pulay, P. *J. Comput. Chem.* **1982**, *3* (4), 556−560.

(44) Parr, R. G.; Ayers, P. W.; Nalewajski, R. F. *J. Phys. Chem. A* **2005**, *109*, 3957−3959.

(45) Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 1833−1840.

(46) Löwdin, P.-O. *J. Chem. Phys.* **1950**, *18* (3), 365−375.

(47) Becke, A. D. *J. Chem. Phys.* **1988**, *88* (4), 2547−2553.

(48) Becke, A. D. *J. Chem. Phys.* **1993**, *98* (7), 5648−5652.

(49) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(50) Davidson, E. R. *J. Chem. Phys.* **1967**, *46* (9), 3320−3324.

(51) Clark, A. E.; Davidson, E. R. *Int. J. Quantum Chem.* **2003**, *93*, 384−394.

(52) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098−3100.

(53) Wu, Q.; Ayers, P. W.; Yang, W. *J. Chem. Phys.* **2003**, *119*, 2978−2990.

(54) Ramirez, B. E.; Malmstrom, B. G.; Winkler, J. R.; Gray, H. B. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 11949.

(55) Brzezinski, P. *Biochemistry* **1996**, *35*, 5611.

(56) Berg, J. M. *Acc. Chem. Res.* **1995**, *28*, 14.

(57) Howell, E. E.; Villafranca, J. E.; Warren, M. S.; Oatley, S. J.; Kraut, J. *Science* **1986**, *231*, 1125.

(58) Crane, B. R.; Siegel, L. M.; Getzoff, E. D. *Science* **1995**, *270*, 59.

(59) Puglisi, J. D.; Chen, L.; Frankel, A. D.; Williamson, J. R. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 3680.

(60) Roberts, J. A.; Kirby, J. P.; Nocera, D. G. *J. Am. Chem. Soc.* **1995**, *117* (30), 8051−8052.

(61) Kirby, J. P.; Roberts, J. A.; Nocera, D. G. *J. Am. Chem. Soc.* **1997**, *119*, 9230−9236.

(62) Cukier, R. I.; Nocera, D. G. *Annu. Rev. Phys. Chem.* **1998**, *49*, 337−369.

(63) Soudackov, A.; Hammes-Schiffer, S. *J. Am. Chem. Soc.* **1999**, *121*, 10598−10607.

(64) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision B.05; Gaussian, Inc.: Pittsburgh, PA, 2003.

(65) Perdew, J. P.; Levy, M. *Phys. Rev. B* **1997**, *56*, 16021−16028.

(66) Zhang, Y.; Yang, W. *J. Chem. Phys.* **1998**, *109*, 2604−2608.

(67) Wilson, P. J.; Bradley, T. J.; Tozer, D. J. *J. Chem. Phys.* **2001**, *115*, 9233.

(68) Zhao, Y.; Pu, J.; Lynch, B. J.; Truhlar, D. G. *Phys. Chem. Chem. Phys.* **2004**, *6* (4), 673−676.

(69) *CRC Handbook of Chemistry and Physics 2001−2002*; CRC Press LLC: Boca Raton, FL, 2001.

(70) Galbraith, J. M.; Schaefer, H. F., III. *J. Chem. Phys.* **1996**, *105* (2), 862−864.

(71) Schumper, G. S.; Schaefer, H. F., III. *J. Chem. Phys.* **1997**, *107* (7), 2529−2541.

(72) de Oliveira, G.; Martin, J. M. L.; de Proft, F.; Geerlings, P. *Phys. Rev. A* **1999**, *60* (2), 1034−1045.

(73) Johnson, B. G.; Gonzales, C. A.; Gill, P. M.; Pople, J. A. *Chem. Phys. Lett.* **1994**, *221*, 100−108.

(74) Deng, L.; Branchadell, V.; Ziegler, T. *J. Am. Chem. Soc.* **1994**, *116*, 10645−10656.

(75) Baker, J.; Andzelm, J.; Muir, M.; Taylor, P. R. *Chem. Phys. Lett.* **1995**, *237*, 53.

(76) Hrouda, V.; Roeselova, M.; Bally, T. *J. Phys. Chem. A* **1997**, *101*, 3925−3935.

(77) Pu, J.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109* (5), 773−778.

(78) Gritsenko, O. V.; Ensing, B.; Schipper, P. R. T.; Baerends, E. J. *J. Phys. Chem. A* **2000**, *104*, 8558−8565.

(79) Rudra, I.; Wu, Q.; Van Voorhis, T. *J. Chem. Phys.* **2006**, *124*, 024103.

(80) Behler, J.; Delley, B.; Lorenz, S.; Reuter, K. M. S. *Phys. Rev. Lett.* **2005**, *94* (3), 036104.

(81) Hybertsen, M. S.; Schluter, M.; Christensen, N. *Phys. Rev. B* **1989**, *39* (13), 9028−9041.

CT0503163

JCTC Journal of Chemical Theory and Computation

The Carbon Allotrope Hexagonite and Its Potential Synthesis from Cold Compression of Carbon Nanotubes

M. J. Bucknum and E. A. Castro*

INIFTA, Theoretical Chemistry Division, Suc. 4, C.C. 16, Universidad de La Plata, 1900 La Plata, Buenos Aires, Argentina

Received January 3, 2006

Abstract: In a previous report, the approximate crystalline structure and electronic structure of a novel, hypothetical hexagonal carbon allotrope has been disclosed. Employing the approximate extended Hückel method, this C structure was determined to be a semiconducting structure. In contrast, a state-of-the-art density functional theory (DFT) optimization reveals the hexagonal structure to be metallic in band profile. It is built upon a bicyclo[2.2.2]-2,5,7-octatriene (barrelene) generating fragment molecule and is a Catalan network, with the Wells point symbol $(6^6)_2(6^3)_3$ and the corresponding Schläfli symbol (6, 3.4). As the network is entirely composed of hexagons and, in addition, possesses hexagonal symmetry, lying in space group *P*6/*mmm* (space group #191), it has been given the name hexagonite. The present report describes a density functional theory (DFT) optimization of the lattice parameters of the parent hexagonite structure, with the result giving the optimized lattice parameters of $a = 0.477$ nm and $c = 0.412$ nm. A calculation is then reported of a simple diffraction pattern of hexagonite from these optimized lattice parameters, with Bragg spacings enumerated for the lattice out to fourth order. Results of a synchrotron diffraction study of carbon nanotubes which underwent cold compression in a diamond anvil cell (DAC) to 100 GPa, in which the carbon nanotubes have evidently collapsed into a hitherto unknown hexagonal C polymorph, are then compared to the calculated diffraction pattern for the DFT optimized hexagonite structure. It is seen that a close fit is obtained to the experimental data, with a standard deviation over the 5 matched reflections being given by $\sigma_x = 0.003107$ nm/reflection.

1. Introduction

As a potential allotropic structure of C, the crystalline and electronic structure of the so-called, 3-dimensional (3D) hexagonite lattice[1] and some of its expanded 3D derivatives were first reported by Karfunkel et al. in 1992.[2] The description of the parent structure of hexagonite in their report[2] was substantially refined and clarified later on by Bucknum et al. in a preprint published in 2001, where an identification of the space group symmetry (*P*6/*mmm*, space group #191) and a complete set of crystallographic coordinates for the hexagonite unit cell were given.[3] Such a 3-dimensional (3D) hexagonite structure can be expanded into an indefinitely large number of derivative 3D structures, by the insertion of 1,4-dimethylene-2,5-cyclohexadieneoid organic spacers into the parent hexagonite structure.[2−3] Expanded hexagonites include 3D crystalline materials with arbitrarily large pores directed along the crystallographic *c*-axis; they occur in infinite families possessing orthorhombic (*Pmmm*), trigonal (*P*3*m*1), and hexagonal (*P*6/*mmm*) space group symmetries.[2−3]

It was reported in this paper,[3] that hexagonite could be realized from the elaboration of a bicyclo[2.2.2]-2,5,7-octatriene (barrelene) generating fragment molecule[4−6] in 3D, as is shown in Figure 1.

Thus the full elaboration of the 3D hexagonite network, from the barrelene generating fragment, can be seen in Figure 2 from a perspective normal to the *ab*-plane of the lattice.

* Corresponding author e-mail: castro@quimica.unlp.edu.ar.

10.1021/ct060003n CCC: $33.50
Published on Web 03/18/2006

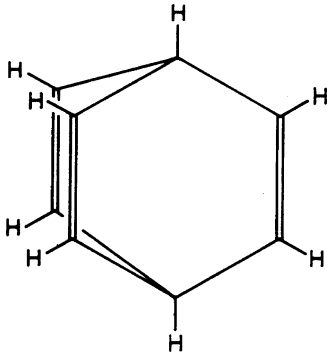


Figure 1. Structure of bicyclo[2.2.2]-2,5,7-octatriene.

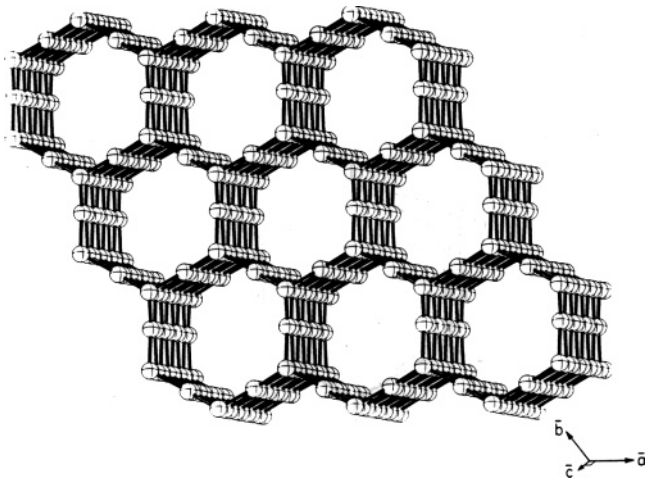


Figure 2. Extended drawing of the hexagonite lattice, viewed approximately normal to the *ab*-plane of the lattice.

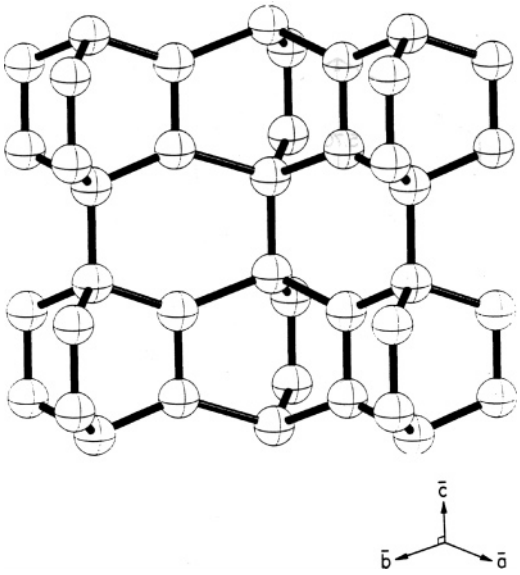


Figure 3. View of the hexagonite lattice from the perspective of the crystallographic *ab*-plane.

Yet another perspective of this hexagonite lattice is shown in Figure 3, where there is a view of it parallel to the *ab*-plane.[3]

One can see in these various views of the hexagonite lattice, given in Figures 2 and 3, the omnipresence of 6-ness in the structure. The organic tunnels apparent in Figure 2 are indeed hexagonal macrocyclic tunnels which are further built upon component hexagons. Thus in Figure 3, which is in the crystallographic *ab*-plane, we see illustrated the hexagon nature of these rings that are components of the larger rings directed along the *c*-axis and apparent in the view of Figure 2.

Table 1. Fractional Hexagonal Crystallographic Coordinates of Hexagonite from the Original Report[3]

atom no.	*x/a*	*y/b*	*z/c*	*a* (Å)	*c* (Å)
1	2/3	1/3	0.1935	4.89	3.88
2	2/3	1/3	0.8065	4.89	3.88
3	1/3	2/3	0.1935	4.89	3.88
4	1/3	2/3	0.8065	4.89	3.88
5	1/2	1/2	0.3265	4.89	3.88
6	1/2	1/2	0.6735	4.89	3.88
7	1/12	2/3	0.3265	4.89	3.88
8	1/12	2/3	0.6735	4.89	3.88
9	5/12	5/6	0.3265	4.89	3.88
10	5/12	5/6	0.6735	4.89	3.88

In the 2001 report by Bucknum et al. on hexagonite's structure,[3] the C−C single bonds were assumed to be 0.1500 nm, and the C═C double bonds were assumed to be 0.1350 nm, and all bond angles were assumed to be tetrahedral at 109.5°, except the trigonal C−C−C angles lying along the *c*-axis, which were constrained to 141°. This resulted in a crudely defined unit cell, with the lattice parameters given by $a = b = 0.4890$ nm and $c = 0.3880$ nm, and the set of fractional hexagonal coordinates, as listed in Table 1 .

Some of the topological character of the hexagonite lattice has been described previously.[3] An introduction to the topological characterization of crystalline networks is given by us elsewhere.[7] From a perspective entirely normal to the *ab*-plane, as shown in Figure 2, the lattice reveals itself in this aspect to be reminiscent of the familiar hexagonal tiling of the plane, represented by the Schläfli symbol (6, 3) (or the Wells point symbol 6^3) called the honeycomb tessellation. Remarkably, a view perpendicular to the *c*-axis, inclined by 30° from the *a*-axis of the unit cell, reveals yet a second perspective from which a perfect honeycomb tessellation emerges from the pattern of bonds within the hexagonite lattice. There are thus 2 views of this hexagonite pattern that reveal its high hexagonal symmetry, in space group *P*6/*mmm*, as manifested in 2 independent honeycomb motifs that are patterned in directions entirely perpendicular to each other from the perspective of the unit cell.

As hexagonite is a 3-,4-connected network it contains an admixture of 3-connected and 4-connected vertices in the unit cell. The overall connectivity of the lattice,[7] a weighted average of the 3- and 4-connected points taken from the stoichiometry of the network, is given by $p = 3^{2/5}$. While the other key topological parameter, called the polygonality,[7] is indeed simply $n = 6$, as inspection of Figure 2 and 3 will reveal. One can thus represent the topology of hexagonite by the Wells point symbol $(6^6)_2(6^3)_3$, and this, then, has the corresponding Schläfli symbol $(n, p) = (6, 3.4)$.[7] It is a Catalan C-network that can be expanded infinitely by insertion of 1,4-dimethylene-2,5-cyclohexadieneoid organic spacers between, and within, the barrelene moieties that make

up the parent hexagonite lattice. This has been described already by Karfunkel et al. in their 1992 paper.[2−3]

It is interesting here, in this regard, to see that hexagonite and the expanded hexagonites are represented by the collective Schläfli symbol given by $(n, p) = (6, 3^{x/x+y})$, where "x" represents the number of 4-connected points in the unit of pattern (which will always be 4) and "$x + y$" represents the sum of the numbers of 3- and 4-connected points in the unit of pattern (which will increase in increments, as the 1,4-dimethylene-2,5-cyclohexadieneoid organic spacers are added to the unit cell in the expanded hexagonites). Hexagonite and its expanded realizations, therefore, represent a related family of Catalan 3D C-based networks that provide an interesting contrast to the Archimedean family of C-based fullerenes.[8] In contrast to the Catalan hexagonites, the fullerenes collectively have the Schläfli index $(n, p) = (5^{x/x+y}, 3)$, where "x" is the number of hexagons in the polyhedron and "$x + y$" is the sum of the number of pentagons and hexagons in the polyhedron. It should be noted in this context that a Schläfli relation exists for the polyhedra, shown as eq 1 below, that is entirely rigorous for the innumerable fullereneic structures which collectively possess the Schläfli index $(n, p) = (5^{x/x+y}, 3)$. In eq 1, the parameter E is the number of edges in the fullereneic polyhedron (or polyhedron), n is the weighted average polygon size over the polygons in the polyhedron (for fullereneic structures it will always be an admixture of pentagons and hexagons) and p is the weighted average connectivity over the vertices in the polyhedron (for fullereneic structures, this will always be 3). The number of edges E is related to the number of vertices, V, and the number of faces, F, by the Euler identity,[7] given as $V - E + F = 2$.

$$\frac{1}{n} - \frac{1}{2} + \frac{1}{p} = \frac{1}{E} \tag{1}$$

In section 2 that follows, we report on the electronic structural characteristics of the C-based hexagonite structure from the point of view of the extended Hückel molecular orbital method (EHMO), which is an approximate solid-state electronic structure algorithm based upon the tight binding methodology.[9−12] Next, in section 3, we report on the details of a density functional theory (DFT) geometry optimization of the parent hexagonite structure and provide a calculation of a simple diffraction pattern of hexagonite,[13] out to fourth order in Bragg spacings of the crystalline structure. We then compare this theoretically calculated data to experimental data, reported in 2004, for a hexagonal polymorph of carbon produced by cold compression of C nanotubes to 100 GPa pressure in a diamond anvil cell (DAC) by Wang et al.[14]

2. Electronic Structure of Hexagonite

Using the approximate electronic structure algorithm called the extended Hückel method (EHMO),[9−12] a calculation of the approximate band structure and density of states (DOS) of hexagonite, as an allotrope of C, were carried out.[3] Figure 4 shows a representation of the approximate band structure of the C-based hexagonite. Figure 5 shows the corresponding density of states (DOS); this is derived from the band structure shown in Figure 4.

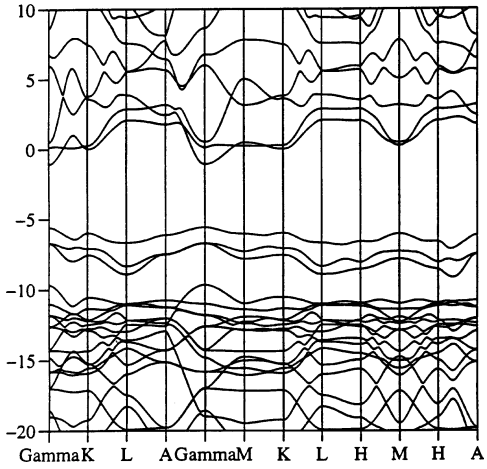


Figure 4. Electronic band structure of the hexagonite crystal structure.

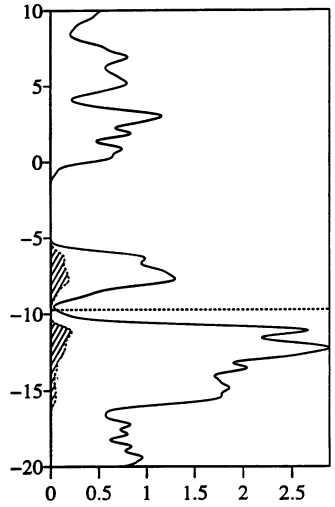


Figure 5. Density of states (DOS) of the hexagonite crystal structure.

Thus Figures 4 and 5 indicate that the hexagonite structure should be a C-based semiconductor,[3] in this approximation, where 3 unoccupied π^* bands are relatively low lying and separated from the σ^* manifold by several eV of energy.[3,15] One could therefore envision the doping of the pores of hexagonite with univalent alkali metals, divalent alkaline earth metals, or trivalent rare earth metals to form conducting or indeed, with the corresponding stoichiometry, insulating metal-organic composite structures, in which the metals readily ionize in the parent C-lattice and donate their valence shell electrons to the higher-lying, unoccupied π^* bands of the host hexagonite lattice.[3] The electronic features of such doped hexagonites and their potential applications in high technology remain an unexplored vista of both theory and, potentially some day, experiment.[16−18]

3. DFT Optimization of Hexagonite Structure

The C-based hexagonite structure, as described in previously in section 1, has been optimized with the DFT algorithm named CASTEP, and the details of the optimization calculations are reported herein.[19] The CASTEP code has been adequately described elsewhere.[20] The results indicate a much different optimized geometry for hexagonite than was

Table 2. CASTEP Optimized Fractional Hexagonal Crystallographic Coordinates and Lattice Parameters of Hexagonite

atom no.	*x/a*	*y/b*	*z/c*	*a* (Å)	*c* (Å)
1	2/3	1/3	0.1840	4.772	4.129
2	2/3	1/3	0.8160	4.772	4.129
3	1/3	2/3	0.1840	4.772	4.129
4	1/3	2/3	0.8160	4.772	4.129
5	1/2	1/2	0.3400	4.772	4.129
6	1/2	1/2	0.6600	4.772	4.129
7	0	1/2	0.3400	4.772	4.129
8	0	1/2	0.6600	4.772	4.129
9	1/2	0	0.3400	4.772	4.129
10	1/2	0	0.6600	4.772	4.129

assumed to pertain in the initial report by Bucknum et al. of the structure.[3] The C−C single bonds in the structure are found to be fairly closely related to each other with a uniform length of about 0.1521 nm. The C═C double bonds in hexagonite are 0.1326 nm in length. The C═C−C trigonal angles are about 115°, while the C−C−C trigonal angles are 130°, and the tetrahedral angles within the cage of the barrelene substructures are 103°, while outside this cage the tetrahedral angles are 115°. Finally, and most importantly, the lattice parameters optimized for the hexagonite structure by CASTEP are given as $a = 0.4772$ nm and $c = 0.4129$ nm. Therefore, the optimized coordinates of the 10 C atoms in the hexagonite unit cell are given in Table 2. The density[21] of hexagonite is calculated to be 2.449 g/cm^3; it lies between the density of the 3-connected carbon form, graphite, at 2.27 g/cm^3, and the density of the 4-connected carbon form, diamond, at 3.56 g/cm^3, but is closer to graphite's density than to diamond's density.

From the CASTEP optimized lattice parameters we can calculate a simple diffraction pattern for the C-based hexagonite structure,[13] and such a powder pattern is shown with Bragg spacings enumerated out to fourth order for the lattice in Table 3. For comparison in Table 3, we report a set of 5 Bragg reflections recorded from a sample of C nanotubes which have been cold compressed in a diamond anvil cell and have transformed into what is believed to be a novel hexagonal polymorph of C, which nonetheless possesses an unknown structure.[14] This latter C polymorph is able to be quenched at room pressure, as well. The Bragg spacings in this cold-compressed sample of C nanotubes were monitored using the specialized technique of energy-dispersive-X-ray-diffraction (EDXRD) on a specially constructed high-pressure DAC synchrotron beamline at the Cornell High Energy Synchrotron Source (CHESS) in Ithaca, NY.[14]

One can see in Table 3 that the 5 reflections in the EDXRD data set of the C nanotube sample, as cold-compressed in the DAC to over 100 GPa, readily match the set of Bragg spacings calculated theoretically from the optimized set of lattice parameters provided for the hexagonite lattice from the CASTEP algorithm.[19,20] The average deviation over the 5 reflections in the experimental data set matched to the calculated Bragg spacings of the optimized hexagonite structure is given by $\underline{\Delta x} = 0.003040$ nm/reflection, while the standard deviation over the 5 matched reflections in the experimental data set is given by $\sigma_x = 0.003107$ nm/reflection. Clearly all the data, over the 5 matched reflections in the experimental data set, have deviations that fit to within less than $3\sigma_x = 0.009321$ nm of the standard deviation of the data set, as is reflected in Table 3. In a parallel comparison of the experimental data set given here of the hexagonal C polymorph, with the commonly observed Bragg reflections from cubic diamond, it is important to point out that only the cubic diamond reflections (111), with a Bragg spacing of 0.2060 nm, and (220), with a Bragg spacing of 0.1261 nm, matched to the experimental set of reflections for the hexagonal C polymorph.[22]

A caution should be made here with regard to the size of the data set, with only 5 reflections to fit; clearly the model proposed here of the hexagonite lattice as an explanation of the experimentally derived EDXRD synchrotron data set for

Table 3. Observed Diffraction Data of Proposed Hexagonal C Polymorph Compared to Calculated Hexagonite Pattern from CASTEP Optimization

calculated hexagonite reflections $a = 0.4772$ nm, $c = 0.4129$ nm		C polymorph reflections[a]	absolute deviation per reflection[b]
(*hkl*)	*d*-spacing, nm	*d* spacing, nm	Δ*d*-spacing, nm
100	0.4133		
001	0.4129		
110	0.2386		
101	0.2921		
111	0.2066	0.2155	0.0089
200	0.2066	0.2155	0.0089
002	0.2065	0.2053	0.0012
102	0.1847		
120	0.1562		
201	0.1848		
211	0.1461	0.1495	0.0034
221	0.1146	0.1161	0.0015
212	0.1246	0.1248	0.0002
222	0.1033		
300	0.1378		
003	0.1376		
103	0.1306		
130	0.1146	0.1161	0.0015
301	0.1307		
311	0.1104		
331	0.07810		
313	0.08808		
333	0.06886		
203	0.1146	0.1161	0.0015
302	0.1146	0.1161	0.0015
320	0.09481		
223	0.09015		
232	0.08616		
332	0.07422		
323	0.07808		
321	0.09241		
312	0.1002		
213	0.1033		
104	0.1002		
401	0.1002		

[a] Wang, Z.; Zhao, Y.; Tait, K.; Liao, X.; Schiferl, D.; Zha, C.; Downs, R. T.; Qian, J.; Zhu, Y.; Shen, T. *Proc. Natl. Acad. Sci. (PNAS)* **2004**, *101* **(**38**)**, 13699−13702. [b] Average deviation over the 5 reflections compared is 0.003040 nm/reflection.

a proposed hexagonal polymorph of C is not constrained as much by observation as one would like it to be. With only 5 observations to work with, such a fit between experimental and theoretical diffraction data is little more than an educated guess.

Still, the authors of the report[14] state that their fit to the data can be interpreted (from the reflections that they have indexed to the data set of 5 Bragg spacings from a hexagonal model with lattice parameters $a = 0.249$ nm and $c = 0.412$ nm) from rules of systematic absences among the hexagonal space groups to a structure that lies in the hexagonal space group $P6m2$ (#190).[14] We offer here in response to this point that the hexagonite lattice lies in the very closely related hexagonal symmetry space group $P6/mmm$ (#191). It would therefore appear that the 2 sets of data, while only consisting of 5 matches, are closely connected together by considerations of the potential symmetry of the unknown C phase. These considerations thus include such a proposed model as the hexagonite lattice.

4. Discussion of Densities of Carbon Phases

One area of disagreement in the current comparison of theoretical and experimental data is in the density reported by Wang et al.[14,21] in their compression study of the C nanotubes and that reported for the candidate hexagonite structure from first principles theory. Experimentally, the density[14] of the hexagonal C polymorph of an unknown structure is given as 3.6 ± 0.2 g/cm^3; while the theoretically calculated value for the density of hexagonite[21] is some 32% lower than this experimental value, at 2.449 g/cm^3, it is thus important here to note that the density of the starting nanotube powder material in this compression synthesis can be estimated to be lower than the known density of fullerite at 1.75 g/cm^3.

Wang et al. proposed that the C material of unknown structure should have a density in excess of the density of diamond, at 3.56 g/cm^3, based upon thermodynamic arguments. They reasoned that a polymorph of C would have collapsed to a denser structure than diamond given the high pressures involved in the study to well over 100 GPa.[14] It is believed by the authors of the present study, however, that the inhomogeneous nature of the C material produced in the synchrotron study,[14] in which amorphous C was present in the quenched DAC sample, along with the novel crystalline phase, suggests that possibly kinetically stabilized products, such as the low-density C material hexagonite,[21,23] may have been formed. The current study does not rule out such a kinetically favored crystalline C product from forming at such high pressures exceeding 100 GPa.[23] Indeed it is physically reasonable to expect that the starting nanotube powder material, at a density of less than that of fullerite at 1.75 g/cm^3, could quite possibly have collapsed to the hexagonite lattice at a density of 2.449 g/cm^3, in their study, if such a lattice was thermodynamically a stable phase up to 100 GPa, otherwise.

It was reported in their study[14] that the hexagonal C polymorph produced under pressure partially indented the diamond anvils in the high-pressure cell used. It was conjectured by Wang et al. that the hexagonal, crystalline C phase, which was quenched from high pressure, was a superhard phase of C. They estimated a bulk modulus of at least 447 GPa for the C form in the study. In this instance, it is worth pointing out that a semiempirical estimate[24,25] of the bulk modulus of hexagonite, despite its low density of 2.449 g/cm^3, puts it at a value of $B_0 = 445$ GPa. This comes from a semiempirical formula for bulk modulus in materials, developed by Cohen et al.,[24,25] and is

$$B_0 = \frac{1972 - 220I}{\langle d \rangle^{3.5}} \frac{\langle N_c \rangle}{4} \qquad (2)$$

Here in formula 2, the parameter I represents the degree of ionicity of the bonding in the unit cell of a given material. For the C form of hexagonite this parameter is just 0, because the hexagonite structure is assumed to be a C polymorph in this instance. The parameter N_c is the averaged coordination number in the unit cell; this is just the connectivity in the lattice, p, which has a value of 3.4 for hexagonite, as was discussed above. Finally, the parameter d is the weighted average bond distance in the unit cell. The C−C single bonds have a length of 0.1521 nm, while the C═C double bonds are 0.1326 nm in length, in the CASTEP hexagonite optimization. They average out to a distance of about 0.1460 nm over the unit cell. These unit cell parameters thus lead to a semiempirical estimation of B_0 for hexagonite that is listed above. One can see that the bulk moduli from experiment and from theory are in good agreement with each other.

5. Conclusions

In this communication, a geometry optimization of the novel, hypothetical form of C called hexagonite[2,3] has been performed using a state-of-the-art DFT-based program called CASTEP.[20] The lattice parameters for the hexagonal unit cell are given as $a = 0.4772$ nm and $c = 0.4129$ nm. The density of the C polymorph has been optimized to be 2.449 g/cm^3. The coordinates for the 10 C atoms in the hexagonite unit cell are thus listed in Tables 1 and 2.[3]

From the optimized lattice parameters of hexagonite, a simple diffraction pattern has been calculated[13] of the Bragg spacings in the crystalline material out to fourth order. From this optimized diffraction pattern, a comparison has been carried out over a set of 5 Bragg reflections obtained from an experimental study in which C nanotubes have been cold compressed in a diamond anvil cell (DAC), on a specially designed high-pressure beam line.[14] Comparison of the 2 data sets yields a close fit over the 5 experimentally determined X-ray reflections, with an average deviation given by $\underline{\Delta x} = 0.003040$ nm/reflection and a standard deviation of $\sigma_x = 0.003107$ nm/reflection. It therefore appears possible that the cold compressed nanotubes in the high pressure study have collapsed under compressive forces into the hexagonite structure, which was first proposed as a likely structure for C to adopt by Karfunkel et al.[2] in 1992 and Bucknum et al.[3] in 2001.

It should be emphasized here that the parent structure, called hexagonite, that represents the first member of a possibly infinite family of related Catalan C polymorphs has

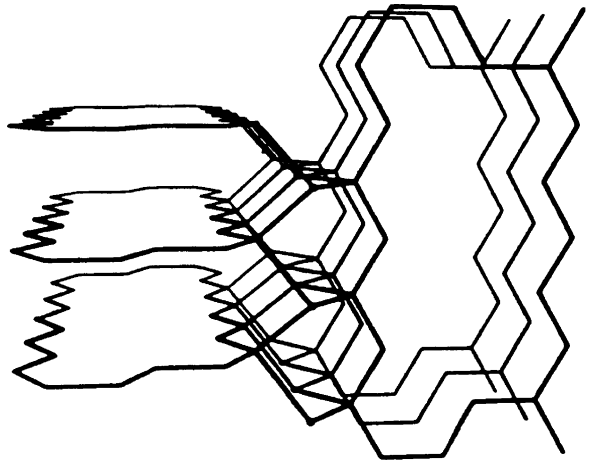


Figure 6. Structure of a variety of the "ortho-" graphite-diamond hybrids.

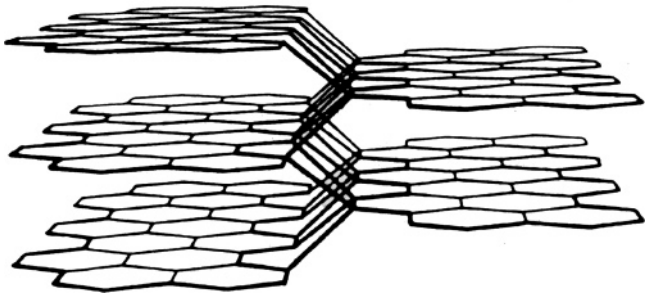


Figure 7. Structure of a variety of the "para-" graphite-diamond hybrids.

possibly been synthesized in the experimental study cited here by Wang et al.[14] This infinite family of 3D network C structures, that are collectively described as Catalan networks, represent an interesting contrast to the infinity of Archimedean fullerene structures first identified in 1985 by Kroto et al.[26] Collectively, the hexagonites and the fullerenes, being semiregular structures, can be contrasted, on one hand, with the graphite and diamond polytypes, which are regular (Platonic) structures of C, given by the Wells point symbols 6^3 and 6^6, respectively.[7] Speculatively, on the other hand, these semiregular structures can be contrasted, as well, with the Wellsean (or topologically irregular) tetragonal glitter network, which has been previously optimized and shown to have a stability close to that of hexagonite[22,23] and with the Wells point symbol given by[27,28] $(6^28^4)(6^28)_2$ and other such irregular networks enumerated by Merz et al.,[29] A. F. Wells,[30,31] and others[27−29] and the references therein. And finally, of course, the hexagonites can be speculatively compared directly to the infinite family of graphite-diamond hybrids proposed in 1993 by Balaban et al.[32] which share the same Catalan Wells point symbol given by $(n, p) = (6, 3^{x/x+y})$, as described above. A couple of varieties of the graphite-diamond hybrids are shown in Figures 6 and 7.[32] It is important to note that the graphite-diamond hybrids collectively circumscribe several infinities of potential C structures, all of which possess orthorhombic symmetry, and which, to date, have not been structurally optimized in any of their various forms.

Acknowledgment. M.J.B. thanks his wife Hsi-cheng Shen for much love and patience in his work on C allotropy and the subtle structural issues of C. The authors wish to thank Norman Goldberg, Ph.D. for producing the structural drawings of hexagonite while a postdoctoral associate in Professor Roald Hoffmann's theoretical chemistry group at Cornell University. The authors wish to thank Chris J. Pickard, Ph.D. of the Theoretical Condensed Matter (TCM) Group at Cambridge University, for his great help in carrying out the DFT-CASTEP optimization calculations of the hexagonite structure. The authors wish to thank Roald Hoffmann, Ph.D. for his suggestions in writing this manuscript. Finally, the authors wish to thank D.M.E. (Marian) Szebenyi, Ph.D. at Cornell High Energy Synchrotron Source (CHESS) for helpful discussions of the symmetry aspects of hexagonite.

Note Added after ASAP Publication. This article was released ASAP on March 18, 2006, with the incorrect Received Date. The correct version was posted on April 12, 2006.

References

(1) The C structure described in this communication, and elsewhere, with the name hexagonite is not to be confused with the inorganic mineral structure of the same name. The authors felt it appropriate to name the C structure, described herein, as hexagonite because of the special circumstance of its hexagonal symmetry space group (*P*6/*mmm*, #191), combined with its further 6-ness, as distinguished by its topological polygonality, given by $n = 6$, in which all of the smallest circuits in the corresponding network are hexagons.

(2) Karfunkel, H. R.; Dressler, T. *J. Am. Chem. Soc.* **1992**, *114* (7), 2285−2288.

(3) Bucknum, M. J. *Chem. Preprint Arch.* **2001**, *2001* (1), 75−110.

(4) Cotton, F. A. *Chemical Applications of Group Theory*, 3rd ed.; John Wiley & Sons Inc.: New York, 1990; pp 166−172.

(5) Zimmerman, H. E.; Paufler, R. M. *J. Am. Chem. Soc.* **1960**, *82*, 1514−1515.

(6) Wilcox, C. F., Jr.; Winstein, S.; McMillan, W. G. *J. Am. Chem. Soc.* **1960**, *82*, 5450−5454.

(7) Bucknum, M. J.; Castro, E. A. *MATCH Commun. Math. Comput. Chem.* **2005**, *54*, 89−119.

(8) Baggott, J. *Perfect Symmetry: The Accidental Discovery of Buckminsterfullerene*, 1st ed.; Oxford University Press: Oxford, United Kingdom, 1996; pp 1−150.

(9) Hoffmann, R. *J. Chem. Phys.* **1963**, *39*, 1397−1412.

(10) Hoffmann, R.; Lipscomb, W. N. *J. Chem. Phys.* **1962**, *37*, 2872−2883.

(11) Whangbo, M. H.; Hoffmann, R. *J. Am. Chem. Soc.* **1978**, *100*, 6093−6098.

(12) Whangbo, M. H.; Hoffmann, R.; Woodward, R. B. *Proc. R. Soc. A* **1979**, *366*, 23−46.

(13) Warren, B. E. *X-ray Diffraction*, 1st ed.; Dover Publications: Mineola, NY, 1990; pp 21−22. The formula used to calculate Bragg spacings in the hexagonal crystal system of hexagonite is given in the book by B. E. Warren in the following format: $1/d_{hkl}^2 = (4/3)((h^2 + hk + k^2)/a^2) + l^2/c^2$.

(14) Wang, Z.; Zhao, Y.; Tait, K.; Liao, X.; Schiferl, D.; Zha, C.; Downs, R. T.; Qian, J.; Zhu, Y.; Shen, T. *Proc. Natl. Acad. Sci. (PNAS)* **2004**, *101* (38), 13699−13702.

(15) The CASTEP-DFT method calculates the band structure of hexagonite to be metallic, in contrast to the result from the approximate EHMO method, which calculates the hexagonite structure to have a semiconducting band profile. It is believed by the authors of the present study that the EHMO calculations of semiconducting hexagonite are potentially closer to a true reflection of the electronic structure of this material than that provided by the DFT results, based upon the fact that hexagonite can be viewed as a layering of delocalized π bonding (sp^2) sandwiched between insulating layers of C σ bonding (sp^3) and thus cannot realistically be represented as a 3D metallic structure.

(16) Kaner, R. B.; MacDiarmid, A. G. *Sci. Am.* **1988**, *258*, 60−65.

(17) Burroughes, J. H.; Jones, C. A.; Friend, R. H. *Nature (London)* **1988**, *335*, 137−141.

(18) Epstein, A. J.; Miller, J. S. *Sci. Am.* **1979**, *241*, 48−57.

(19) CASTEP (Cambridge Serial Total Energy Package) is a plane wave pseudopotential code, based upon density functional theory (DFT), that was used to optimize the hexagonite structure in the present report. Therefore, for the present implementation of CASTEP, used to optimize the structural parameters of hexagonite, the local density approximation (LDA) was used, ultrasoft pseudopotentials were employed, the basis set had an energy cutoff of 400 eV. and k-point sampling was done with a $10 \times 10 \times 4$ mesh. The ultrasoft pseudopotentials used in the calculation are due to Vanderbilt (Vanderbilt, D. Soft Self-Consistent Pseudopotentials in a Generalized Eigenvalue Formalism. *Phys. Rev. B* **1990**, *41* (Rapid Communications), 7892−7895.) The Brillouin zone was sampled at a density of 0.004 nm^{-1}.

(20) Segall, M. D.; Lindan, P. J. D.; Probert, M. J.; Pickard, C. J.; Hasnip, P. J.; Clark, S. J.; Payne, M. C. *J. Phys.: Condens. Matter* **2002**, *14* (11), 2717−2743.

(21) The volume of the hexagonite unit cell was calculated with the formula $V = ((3)^{1/2}/2)(a^2c)$, and the corresponding density of hexagonite, with 10 C atoms in the unit cell, was found to be 2.449 g/cm^3. The density of the hexagonal C polymorph reported by Wang et al. in ref 14 is 32% greater than this calculated value, at 3.6 g/cm^3.

(22) Bucknum, M. J.; Stamatin, I.; Castro, E. A. *Mol. Phys.* **2005**, *103* (20), 2707−2715.

(23) On a stability scale at which diamond is at the 0 of energy, the glitter allotrope of C, described in ref 22, has an energy of formation of 0.5116 eV/C atom above that of diamond, by the CASTEP optimization, under the local density approximation (LDA). In contrast, the hexagonite allotrope of C, described herein, has an energy of formation of 0.5343 eV/C atom above that of diamond by the same computational method.

(24) Cohen, M. L. *Solid State Commun.* **1994**, *92* (1−2), 45−52.

(25) Liu, A. Y.; Cohen, M. L. *Science* **1989**, *245*, 841−842.

(26) Kroto, H. W.; Heath, J. R.; O'Brien, S. C.; Curl, R. F.; Smalley, R. E. *Nature (London)* **1985**, *318*, 162−163.

(27) Bucknum, M. J.; Castro, E. A. *J. Math. Chem.* **2004**, *36* (4), 381−408.

(28) Bucknum, M. J.; Hoffmann, R. *J. Am. Chem. Soc.* **1994**, *116*, 11456−11464.

(29) Merz, K. M., Jr.; Hoffmann, R.; Balaban, A. T. *J. Am. Chem. Soc.* **1987**, *109*, 6742−6751.

(30) Wells, A. F. *Three-Dimensional Nets and Polyhedra*, 1st ed.; John Wiley & Sons: NY, 1977; pp 1−150.

(31) Wells, A. F. *Further Studies of Three-dimensional Nets*; ACA Monograph #8, ACA Press: Pittsburgh, PA, 1979; pp 1−75.

(32) Balaban, A. T.; Klein, D. J.; Folden, C. A. *Chem. Phys. Lett.* **1994**, *217*, 266−270.

CT060003N

JCTC Journal of Chemical Theory and Computation

H_2, Ne, and N_2 Energies of Encapsulation into C_{60} Evaluated with the MPWB1K Functional

Zdeněk Slanina,*,†,‡ Peter Pulay,§ and Shigeru Nagase†

Department of Theoretical Molecular Science, Institute for Molecular Science, Myodaiji, Okazaki 444-8585, Aichi, Japan, Institute of Chemistry, Academia Sinica, Nankang, Taipei 11529, Taiwan−ROC, and Department of Chemistry and Biochemistry, University of Arkansas, Fayetteville, Arkansas 72701

Received December 27, 2005

Abstract: The recently suggested MPWB1K functional is tested on $H_2@C_{60}$, $Ne@C_{60}$, and $N_2@C_{60}$ as a tool for evaluations of stabilization energies upon encapsulation of nonmetallic species into fullerenes. It is found that the MPWB1K (modified Perdew−Wang and Becke functionals) values can be within a few kilocalories per mole from the MP2 or SCS-MP2 (spin-component scaled MP2) values so that further applications of the functional are clearly encouraged. The best estimates of the encapsulation-energy gains found for $H_2@C_{60}$, $Ne@C_{60}$, and $N_2@C_{60}$ are at least 4 kcal/mol, slightly less than 4 kcal/mol, and about 9 kcal/mol, respectively.

1. Introduction

The suitability of density-functional theory (DFT) for evaluations of weak molecular interactions and especially of dispersion energies has been an important computational issue.[1−10] Very recently, Zhao et al.[11−15] performed a series of test DFT calculations with a conclusion[15] that the MPWB1K functional (the modified Perdew and Wang exchange functional MPW[16] and Becke's meta correlation functional[17] optimized against a kinetics database) is the best combination for evaluations of nonbonded interactions with a relative averaged mean unsigned error of only 11%.

Fullerene encapsulations of nonmetal atoms and small molecules represent an interesting case[18−22] for this type of computation because there are also some related observations. A well-known example is represented by complexes of fullerenes with encapsulated rare gas atoms[23−27] prepared by heating under high pressure. Ion bombardment can produce $N@C_{60}$[28,29] and $P@C_{60}$[30] (and also $Li@C_{60}$[31]), though in very low yields. $N@C_{60}$ and its derivatives have been studied vigorously,[32−36] inter alia as a candidate for the implementation of qubits in an electron-spin-based quantum computer.[37] $N_2@C_{60}$ was prepared by pressure heating[38] and was also reported[39,40] in the chromatographic separation after the nitrogen-ion implantation into C_{60}. Recently, molecular hydrogen[41−43] (and also water molecule[21]) was placed inside an open-cage fullerene, and the cage with H_2 was subsequently closed.[44]

Previous DFT calculations on nonmetallic fullerene encapsulates used mostly[45−52] the B3LYP functional. The present paper reports calculations of $H_2@C_{60}$, $Ne@C_{60}$, and $N_2@C_{60}$ with the newly suggested MPWB1K functional in order to judge its performance for such encapsulation energetics.

2. Calculations

Three systems known from observations were selected for the present calculations: $H_2@C_{60}$, $Ne@C_{60}$, and $N_2@C_{60}$. The computations are carried out at fixed experimental geometries. The C_{60} geometry is taken from the gas-phase electron diffraction[53] where the 5/6 (i.e., between a pentagon and a hexagon) and 6/6 bonds (between two hexagons) were determined as 1.458 and 1.401 Å, respectively. The bond lengths of H_2 and N_2 are also from observed data.[54,55] The diatomic molecules in the cage are oriented toward a pair of its parallel pentagons so that the complexes exhibit D_{5d} symmetry. This arrangement is known from previous calculations[52] as the lowest energy structure, though energy

* Corresponding author e-mail: zdenek@ims.ac.jp.
† Institute for Molecular Science.
‡ Institute of Chemistry, Academia Sinica.
§ University of Arkansas.

10.1021/ct0503320 CCC: $33.50
Published on Web 04/08/2006

Table 1. Encapsulation Energies Computed in the Observed Geometries (kcal/mol)

level	$H_2@C_{60}$	$Ne@C_{60}$	$N_2@C_{60}$
B3LYP/6-31G**	1.56	−2.52	7.70
B3LYP/6-311G(2d, 2p)	1.34	−2.95	7.94
MPWB1K/6-31G**	−7.93[a]	−3.08	−14.93
MPWB1K/6-31G** & BSSE	−7.29	−0.22	−11.34
MP2/6-31G**	−3.99	−5.05	−17.05
MP2/6-31G** & BSSE	−2.36	−1.85	−8.90
MPWB1K/6-311G(2d, 2p)	−8.09	−7.20	−13.18
MPWB1K/6-311G(2d, 2p) & BSSE	−7.17	−3.83	−9.54
MP2/6-311G(2d, 2p)	−8.63	−5.39	−23.44
MP2/6-311G(2d, 2p) & BSSE	−6.07		−15.54
SCS-MP2/6-311G(2d, 2p)	−6.71	−4.35	−16.35
SCS-MP2/6-311G(2d, 2p) & BSSE	−4.20		−8.57
MPWB1K/(d, p)−6-311G**	−9.44		
MPWB1K/(d, p)−6-311G** & BSSE	−7.81		
MP2/(d, p)−6-311G**	−10.62		
MP2/(d, p)−6-311G** & BSSE	−6.94		

[a] The value after the full geometry optimization is −8.37 kcal/mol.

differences relative to other orientations are less than 0.1 kcal/mol. In the case of $Ne@C_{60}$, the Ne atom is placed in the cage center so that the system still exhibits icosahedral symmetry.

Single-point energy calculations were carried out with the MPWB1K functional in three well-known basis sets: 6-31G* (6-31G** for H atoms), 6-311G(2d,2p), and the augmented (d,p)-6-311G** basis set.[56] For comparison, some computations were also performed with Becke's[57] three-parameter exchange functional combined with the nonlocal Lee−Yang−Parr[58] correlation functional (B3LYP). The ultrafine grid in numerical integrations of the DFT functional and the tight self-consistent field convergency criterion were used. However, the comparisons are primarily related to the second-order Møller−Plesset (MP2) perturbation treatment[59,60] with the frozen core option carried out in the above three basis sets. Moreover, the conventional MP2 method is also complemented with the spin-component-scaled (SCS) MP2 calculations.[61] Finally, the basis set superposition error (BSSE) was estimated by the Boys−Bernardi counterpoise method.[62] The selected basis sets represent a practical computational limit for calculations of systems comparable to C_{60} at DFT and especially MP2 levels.

The Gaussian[63] and PQS[64] program packages were used for all calculations.

3. Results and Discussion

Table 1 reports the zero-point-exclusive encapsulation energies (also called the energies of interaction or stabilization) corresponding to the change in potential energy for the process:

$$X(g) + C_{60}(g) = X@C_{60}(g) \qquad (1)$$

where X is H_2, Ne, or N_2. To compare this to an experimental term, one would have to add the changes in zero-point energy and heat content function. Let us note that the quantity computed here is a thermodynamic one, although the kinetics of reaction 1 are also of interest.[65−67]

As already known from previous computations,[21,52] the B3LYP functional does not produce reliable values for nonbonding fullerene encapsulations. In the cases of $H_2@C_{60}$ and $N_2@C_{60}$, their stabilization energies become positive at the B3LYP/6-31G** and B3LYP/6-311G(2d,2p) levels, which does not agree well with the very fact that the endohedrals can be observed. Only for $Ne@C_{60}$ is the B3LYP/6-31G* encapsulation energy negative (binding), though the stabilization energy is too low compared to other calculations in Table 1 and also to previous evaluations.[20,68]

All of the MPWB1K and MP2 calculations in Table 1 show for $H_2@C_{60}$ a substantial stabilization energy. Obviously, the BSSE-corrected energies should be preferred, though it is not necessarily assured[69] that the true energy is closer to the BSSE-corrected term rather than to the uncorrected term. The MPWB1K and MP2 treatments with the 6-31G** basis set differ by some 4 kcal/mol before the BSSE correction. With respect to the remaining computations of $H_2@C_{60}$ in Table 1, it seems however likely that the MP2/6-31G** term underestimates the stabilization for the species. Upon moving to the 6-311G(2d,2p) and (d,p)-6-311G** basis sets, the MPWB1K and MP2 values differ only by about 1 kcal/mol. Application of the SCS approach tested for the 6-311G(2d,2p) basis set suggests, however, a reduction of the stabilization by about 2 kcal/mol compared to the conventional MP2 method. Overall, it can be stated that the encapsulation of H_2 into C_{60} brings a potential energy gain of at least 4 kcal/mol.

The effect of geometry optimizations was also checked with $H_2@C_{60}$. It turns out that, if the full geometry optimization is carried out at the MPWB1K/6-31G** level, the encapsulation energy is changed only by some 0.44 kcal/mol. The perturbation of the cage upon encapsulation changes the C−C bonds only on the fourth digit. The 5/6 C−C bonds in the cage optimized at the MPWB1K/6-31G** level vary between 1.4377 and 1.4388 Å, while for the 6/6 bonds, it is from 1.3792 to 1.3798 Å. It should also be noted that rotating the H_2 molecule toward two hexagons (instead of toward two pentagons) changes the energy by less than 0.1 kcal/mol so that the molecule can indeed nearly freely rotate inside the cage.[70]

Let us move now to the $Ne@C_{60}$ system. With the 6-31G** basis set, the MPWB1K and MP2 terms differ by some 2 kcal/mol both before and after the BSSE correction, $Ne@C_{60}$ showing up as more stable in the MP2/6-31G* treatment. However, with the 6-311G(2d,2p) basis set, the MPWB1K functional produces greater stabilization compared to the conventional MP2 value, which is further reduced by about 1 kcal/mol upon application of the SCS approach. Although the BSSE computations could not be finished in this case owing to computational difficulties, $Ne@C_{60}$ is likely stabilized by slightly less than 4 kcal/mol.

The differences between the conventional MP2 and SCS-MP2 treatments are more significant in the $N_2@C_{60}$ system, amounting to some 7 kcal/mol. However, if the MPWB1K/6-311G(2d,2p) and SCS-MP2/6-311G(2d,2p) values are compared, they are still within 3 kcal/mol before the BSSE correction and within 1 kcal/mol after correcting for BSSE. The best estimate gives about a 9 kcal/mol gain in the

potential energy upon $N_2@C_{60}$ formation. Interestingly enough, the MP2/6-31G* encapsulation energy without the BSSE correction for $N@C_{60}$ was computed[50] to be −7.3 kcal/mol, that is, roughly one-half of the value found here at the same level for $N_2@C_{60}$.

While the performance of the MPWB1K exchange-correlation potential is reasonably good for the inclusion compounds considered here, it does not perform well for π stacking. For instance, the binding energy predicted for the prototype benzene dimer, in the parallel displaced configuration, is only 0.11 kcal/mol after counterpoise correction, using the 6-311+G(2df,2pd) basis set. The best current estimate is[71,72] 2.78 kcal/mol. For the sandwich conformer, the counterpoise corrected MPWB1K potential is repulsive, although the best estimate of this quantity is 1.81 kcal/mol. Zhao and Truhlar have successfully used[13a] an empirical combination of density-functional and wave-function-based correlation methods to calculate this quantity accurately at a reasonable cost, and they also suggested[13b] an improved version of MPWB1K called PWB6K, with a good performance for stacking interactions. Although the benzene dimer issue is not directly related to the fullerene encapsulation energies, it indicates the fundamental limitations of local or semi-local DFT to describe the dispersion interaction.

It can be concluded that the values for nonmetallic encapsulations into C_{60} evaluated with the MPWB1K functional closely follow the results from the MP2 and SCS-MP2 treatments with the same basis set. This finding suggests the MPWB1K functional as a practical tool for the evaluation of encapsulations of small molecules into fullerenes and model nanotubes, though larger basis sets[73,74] and also other functionals[74,75] are to be still tested, especially for specific cases[76,77] like interactions with metals.

Acknowledgment. This research has been supported by a grant-in-aid for the NAREGI Nanoscience Project and for Scientific Research on Priority Area (A) from the Ministry of Education, Culture, Sports, Science and Technology of Japan. P.P. would like to thank the Institute for Molecular Science for a visiting professorship. The constructive comments from the reviewers are highly appreciated, too.

References

(1) Lackes, D. J.; Gordon, R. G. *Phys. Rev. A* **1993**, *47*, 4681.

(2) Kristyan, S.; Pulay, P. *Chem. Phys. Lett.* **1994**, *229*, 175.

(3) Perez-Jorda, J. M.; Becke, A. D. *Chem. Phys. Lett.* **1995**, *233*, 134.

(4) Meijer, E. J.; Sprik, M. *J. Chem. Phys.* **1996**, *105*, 8684.

(5) Tsuzuki, S.; Uchimaru, T.; Tanabe, K. *Chem. Phys. Lett.* **1998**, *287*, 202.

(6) Zhang, Y. K.; Pan, W.; Yang, W. T. *J. Chem. Phys.* **1997**, *107*, 7921.

(7) Wesolowski, T. A.; Parisel, O.; Ellinger, Y.; Weber, J. *J. Phys. Chem. A* **1997**, *101*, 7818.

(8) Patton, D. C.; Pederson, M. R. *Phys. Rev. A* **1997**, *56*, R2495.

(9) Perez-Jorda, J. M.; San-Fabian, E.; Perez-Jimenez, A. J. *J. Chem. Phys.* **1999**, *110*, 1916.

(10) Tsuzuki, S.; Lüthi, H. P. *J. Chem. Phys.* **2001**, *114*, 3949.

(11) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2004**, *108*, 6908.

(12) Zhao, Y.; Lynch, B. J.; Truhlar, D. G. *ChemPhysChem* **2005**, *7*, 43.

(13) (a) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 4209. (b) Zhao, Y.; Truhlar, D. G. *Phys. Chem. Chem. Phys.* **2005**, *7*, 2701.

(14) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 5656.

(15) Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2005**, *1*, 415.

(16) Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244.

(17) Becke, A. D. *J. Chem. Phys.* **1996**, *104*, 1040.

(18) Jiménez-Vázquez, H. A.; Cross, R. J. *J. Chem. Phys.* **1996**, *104*, 5589.

(19) Patchkovskii, S.; Thiel, W. *J. Chem. Phys.* **1997**, *106*, 1796.

(20) Bühl, M.; Patchkovskii, S.; Thiel, W. *Chem. Phys. Lett.* **1997**, *275*, 14.

(21) Iwamatsu, S.-I.; Uozaki, T.; Kobayashi, K.; Re, S.; Nagase, S.; Murata, S. *J. Am. Chem. Soc.* **2004**, *126*, 2668.

(22) Hu, Y. H.; Ruckenstein, E. *J. Am. Chem. Soc.* **2005**, *127*, 11277.

(23) Saunders, M.; Jiménez-Vázquez, H. A.; Cross, R. J.; Poreda, R. J. *Science* **1993**, *259*, 1428.

(24) Saunders, M.; Jiménez-Vázquez, H. A.; Cross, R. J.; Mroczkowski, S.; Freedberg, D. I.; Anet, F. A. L. *Nature* **1994**, *367*, 256.

(25) Saunders, M.; Cross, R. J.; Jiménez-Vázquez, H. A.; Shimshi, R.; Khong, A. *Science* **1996**, *271*, 1693.

(26) Cross, R. J.; Saunders, M.; Prinzbach, H. *Org. Lett.* **1999**, *1*, 1479.

(27) Rubin, Y.; Jarrosson, T.; Wang, G.-W.; Bartberger, M. D.; Houk, K. N.; Schick, G.; Saunders, M.; Cross, R. J. *Angew. Chem., Int. Ed.* **2001**, *40*, 1543.

(28) Murphy, T. A.; Pawlik, T.; Weidinger, A.; Höhne, M.; Alcala, R.; Spaeth, J.-M. *Phys. Rev. Lett.* **1996**, *77*, 1075.

(29) Knapp, C.; Dinse, K.-P.; Pietzak, B.; Waiblinger, M.; Weidinger, A. *Chem. Phys. Lett.* **1997**, *272*, 433.

(30) Knapp, C.; Weiden, N.; Kass, K.; Dinse, K. P.; Pietzak, B.; Waiblinger, M.; Weidinger, A. *Mol. Phys.* **1998**, *95*, 999.

(31) Gromov, A.; Krawez, N.; Lassesson, A.; Ostrovskii, D. I.; Campbell, E. E. B. *Curr. Appl. Phys.* **2002**, *2*, 51.

(32) Pietzak, B.; Waiblinger, M.; Murphy, T. A.; Weidinger, A.; Höhne, M.; Dietel, E.; Hirsch, A. *Chem. Phys. Lett.* **1997**, *279*, 259.

(33) Dietel, E.; Hirsch, A.; Pietzak, B.; Waiblinger, M.; Lips, K.; Weidinger, A.; Gruss, A.; Dinse, K.-P. *J. Am. Chem. Soc.* **1999**, *121*, 2432.

(34) Cao, B. P.; Peres, T.; Cross, R. J.; Saunders, M.; Lifshitz, C. *J. Phys. Chem. A* **2001**, *105*, 2142.

(35) Kobayashi, K.; Nagase, S.; Dinse, K.-P. *Chem. Phys. Lett.* **2003**, *377*, 93.

(36) Wakahara, T.; Matsunaga, Y.; Katayama, A.; Maeda, Y.; Kako, M.; Akasaka, T.; Okamura, M.; Kato, T.; Choe, Y. K.; Kobayashi, K.; Nagase, S.; Huang, H. J.; Atae, M. *Chem. Commun.* **2003**, 2940.

(37) Meyer, C.; Harneit, W.; Naydenov, B.; Lips, K.; Weidinger, A. *Appl. Magn. Reson.* **2004**, *27*, 123.

(38) Peres, T.; Cao, B. P.; Cui, W. D.; Khong, A.; Cross, R. J.; Saunders, M.; Lifshitz, C. *Int. J. Mass Spectrom.* **2001**, *210*, 241.

(39) Suetsuna, T.; Dragoe, N.; Harneit, W.; Weidinger, A.; Shimotami, H.; Ito, S.; Takagi, H.; Kitazawa, K. *Chem.—Eur. J.* **2002**, *8*, 5079.

(40) Suetsuna, T.; Dragoe, N.; Harneit, W.; Weidinger, A.; Shimotani, H.; Ito, S.; Takagi, H.; Kitazawa, K. *Chem.—Eur. J.* **2003**, *9*, 598.

(41) Carravetta, M.; Murata, Y.; Murata, M.; Heinmaa, I.; Stern, R.; Tontcheva, A.; Samoson, A.; Rubin, Y.; Komatsu, K.; Levitt, M. H. *J. Am. Chem. Soc.* **2004**, *126*, 4092.

(42) Iwamatsu, S.-I.; Murata, S.; Andoh, Y.; Minoura, M.; Kobayashi, K.; Mizorogi, N.; Nagase, S. *J. Org. Chem.* **2005**, *70*, 4820.

(43) Komatsu, K.; Murata, Y. *Chem. Lett.* **2005**, *34*, 886.

(44) Komatsu, K.; Murata, M.; Murata, Y. *Science* **2005**, *307*, 238.

(45) Williams, C. I.; Whitehead, M. A.; Pang, L. *J. Phys. Chem.* **1993**, *97*, 11652.

(46) Cioslowski, J. *Electronic Structure Calculations on Fullerenes and Their Derivatives*; Oxford University Press: Oxford, U. K., 1995.

(47) Scuseria, G. E. In *Modern Electronic Structure Theory, Part I*; Yarkony, D. R., Ed.; World Scientific: Singapore, 1995; p 279.

(48) Greer, J. C. *Chem. Phys. Lett.* **2000**, *326*, 567.

(49) Larsson, J. A.; Greer, J. C.; Harneit, W.; Weidinger, A. *J. Chem. Phys.* **2002**, *116*, 7849.

(50) Park, J. M.; Tarakeshwar, P.; Kim, K. S.; Clark, T. *J. Chem. Phys.* **2002**, *116*, 10684.

(51) Pietzak, B.; Weidinger, A.; Dinse, K.-P.; Hirsch, A. In *Endofullerenes: A New Family of Carbon Clusters*; Akasaka, T., Nagase, S., Eds.; Kluwer: Dordrecht, The Netherlands, 2002; p 13.

(52) Slanina, Z.; Kobayashi, K.; Nagase, S. *Mol. Phys.*

(53) Hedberg, K.; Hedberg, L.; Bethune, D. S.; Brown, C. A.; Dorn, H. C.; Johnson, R. D.; De Vries, M. *Science* **1991**, *254*, 410.

(54) Chase, M. W., Jr.; Davies, C. A.; Downery, J. R., Jr.; Frurip, D. J.; McDonald, R. A.; Syverud, A. N. *J. Phys. Chem. Ref. Data* **1985**, *14* (Suppl. 1); *JANAF Thermochemical Tables*, 3rd ed.; American Institute of Physics: College Park, Maryland, 1986; Vols. 1 and 2.

(55) Huber, K. P.; Herzberg, G. *Molecular Spectra and Molecular Structure, IV. Constants of Diatomic Molecules*; Van Nostrand Reinhold Company: New York, 1979; p 420.

(56) Tsuzuki, S.; Uchimaru, T.; Mikami, M.; Tanabe, K. *J. Phys. Chem. A* **1998**, *102*, 2091.

(57) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(58) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(59) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.

(60) Krishnan, R.; Pople, J. A. *Int. J. Quantum Chem.* **1978**, *14*, 91.

(61) Grimme, S. *J. Chem. Phys.* **2003**, *118*, 9095.

(62) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.

(63) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision C.01; Gaussian, Inc.: Wallingford, CT, 2004.

(64) Pulay, P.; Baker, J.; Wolinski, K. *Parallel Quantum Solutions*; PQS: Fayetteville, AR, 2005.

(65) Patchkovskii, S.; Thiel, W. *J. Am. Chem. Soc.* **1996**, *118*, 7164.

(66) Shimshi, R.; Khong, A.; Jiménez-Vázquez, H. A.; Cross, R. J.; Saunders, M. *Tetrahedron* **1996**, *52*, 5143.

(67) Patchkovskii, S.; Thiel, W. *Helv. Chim. Acta* **1997**, *80*, 495.

(68) Son, M.-S.; Sung, Y. K. *Chem. Phys. Lett.* **1995**, *245*, 113.

(69) (a) Schwenke, D. W.; Truhlar, D. G. *J. Chem. Phys.* **1985**, *82*, 2418; **1987**, *86*, 3760. (b) van Duijneveldt, F. B.; van Duijneveldt-van de Rijdt, J. G. C. M.; van Lenthe, J. H. *Chem. Rev.* **1994**, *94*, 1873.

(70) Cross, R. J. *J. Phys. Chem.* **2001**, *105*, 6943.

(71) Sinnokrot, M. O.; Valeev, E. F.; Sherrill, C. D. *J. Am. Chem. Soc.* **2002**, *124*, 10887.

(72) Sinnokrot, M. O.; Sherrill, C. D. *J. Phys. Chem. A* **2004**, *108*, 10200.

(73) Staroverov, V. N.; Scuseria, G. E.; Tao, J.; Perdew, J. P. *Phys. Rev. B* **2004**, *69*, 075102.

(74) Staroverov, V. N.; Scuseria, G. E.; Tao, J.; Perdew, J. P. *J. Chem. Phys.* **2003**, *119*, 12129; **2004**, *121*, 11507.

(75) Tao, J.; Perdew, J. P. *J. Chem. Phys.* **2005**, *122*, 114102.

(76) Strassner, T.; Taige, M. A. *J. Chem. Theory Comput.* **2005**, *1*, 848.

(77) Lee, E. P. F.; Dyke, J. M.; Chow, W.-K.; Chau, F.-T.; Mok, D. K. W. *Chem. Phys. Lett.* **2006**, *417*, 256.

CT0503320

JCTC Journal of Chemical Theory and Computation

Analytical Method for Predicting Ferromagnetic Properties of Benzyl-Radical Polymers Based on NBMO Theory

Yuuichi Orimoto[†] and Yuriko Aoki*[,†,‡]

Department of Material Sciences, Faculty of Engineering Sciences, Kyushu University, 6-1 Kasuga-Park, Fukuoka 816-8580, Japan, and Group, PRESTO, Japan Science and Technology Agency (JST), Kawaguchi Center Building, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan

Received January 5, 2006

Abstract: It has been demonstrated that the high-spin stability of benzyl radical oligomers can be predicted without any quantum chemical calculations. This method is composed of three steps: (1) predicting the shapes of nonbonding molecular orbitals (NBMOs), (2) counting up the mixings of NBMOs, and (3) formulating the mixings for the number of radical centers (*N*). This treatment enables us to predict the high-spin stability evaluated by ab initio MO calculations involving the post-Hartree−Fock method or the density functional theory (DFT) method.

1. Introduction

Recently, experimental interests in organic magnets have been roughly classified into two types: the synthesis of crystalline solids composed of small radical molecules or charge-transfer salts[1,2] and that of organic π-conjugated systems.[3−8] In particular, high-spin π-conjugated systems have received growing interest as new organic magnets. This is because high transition temperature is theoretically predicted for the π-conjugated systems due to a strong exchange interaction between radicals through π-conjugated networks.[3,9] In 2001, Rajca et al. actually succeeded in the synthesis of ultrahigh-spin π-conjugated polymers with spin quantum numbers (S) > 5000.[6] Numerous attempts have been made to explain the ferromagnetism in π-conjugated systems both experimentally[3−8,10−13] and theoretically.[14−17] Borden et al. conducted a molecular orbital (MO) approach to clarify the relationship between exchange interactions and ferromagnetic properties in organic conjugated systems.[18−20]

As another MO approach to predict ferromagnetism, one of the authors, Aoki et al., proposed a simple method to estimate the ferromagnetic properties of π-conjugated systems based on the $L_{ij}^{\min}$ value, which corresponds to the mixing between nonbonding molecular orbitals (NBMOs).[21]

In the NBMOs of alternant hydrocarbon systems, the carbon atoms with MO coefficients can be defined as active atoms (denoted by "∗"), and the other carbon atoms with no coefficients can be defined as inactive atoms (denoted by "0"). When an allyl radical molecule is considered as a NBMO unit, there are two types of linkages between two NBMO units (see Figure 1 in ref 21). One is a (0−0) linkage that connects two inactive carbons. The other is a (0-∗) linkage that connects an inactive carbon to an active carbon. In both types of linkages, two degenerated levels of NBMOs are obtained. In the (0−0) linkage, the NBMO level of each unit is preserved in the entire system because there is no interaction between the inactive carbon atoms. Thus, the (0−0) linkage corresponds to a "disjoint" type linkage. In contrast, in the (0-∗) linkage, the energy levels of the NBMOs are preserved because of the interaction between the inactive and active carbon atoms (see Figure 3 in ref 21). The (0-∗) linkage corresponds to a "nondisjoint" type linkage. That is to say, the two types of interactions that correspond to the stabilization and destabilization of the orbitals are canceled out by each other according to the pairing theorem[22] within the framework of the simple Hückel method. Our previous paper[21] used algebraic equations to prove that these two types of linkages, (0−0) and (0-∗), do not change the eigenvalues of the NBMOs.

* Corresponding author e-mail: aoki@cube.kyushu-u.ac.jp.
† Kyushu University.
‡ Group, PRESTO, Japan Science and Technology Agency (JST).

10.1021/ct060008k CCC: $33.50
Published on Web 03/30/2006

A new value

$$L_{ij} = \sum_r (C_{ir}C_{jr})^2 \quad (1)$$

was proposed by Aoki et al. to estimate the mixing between NBMOs for the prediction of the high-spin stability in "nondisjoint" type systems,[21] where C_{ir} is the coefficient of AO χ_r in the ith NBMO. In general, we can change the MO coefficients of degenerated NBMOs by using a unitary transformation (see Figure 2 in ref 21). Therefore, the stability of the high-spin ground state can be predicted by the smallest L_{ij} value, i.e., the $L_{ij}^{\min}$ value, after the unitary transformation. When L_{ij} has the smallest value, NBMO coefficients are selected to minimize the mixing between ith and jth NBMOs.

The total energy difference between excited singlet and triplet states at the Hartree−Fock MO level is expressed as

$$2K_{ij} = E(\mathrm{S}) - E(\mathrm{T}) \quad (2)$$

where E(S) and E(T) denote total energies in excited singlet and triplet states, respectively. The K_{ij} is the exchange integral between ith and jth NBMOs, that is

$$K_{ij} = \sum_r \sum_s \sum_t \sum_u C_{ir}C_{js}C_{it}C_{ju}(rs|tu) \quad (3)$$

where r, s, t, and u indicate atomic orbitals (AOs). C_{ir} is the coefficient of AO χ_r in the ith NBMO in the linear combination of AO (LCAO) approximation. The two-electron integral $(rs|tu)$ in eq 3 is expanded by

$$(rs|tu) = \int\int \chi_r(1)\chi_s(1)\frac{1}{r_{12}}\chi_t(2)\chi_u(2)d\tau_1 d\tau_2 \quad (4)$$

When the 2-electron integrals part of K_{ij} is approximated by a one-center two-electron integral, e.g., $(rr|rr)$, $\sum_r (C_{ir}C_{jr})^2(rr|rr)$ can be obtained from eq 3. Therefore, the L_{ij} value in eq 1 corresponds to the MO coefficients part of the one-center two-electron terms in eq 3 and corresponds to the exchange integral (K_{ij}) as well as the high-spin stability expressed by eq 2. Because the L_{ij} is composed of only MO coefficients, we can efficiently examine the stability of high-spin ground state by calculating the $L_{ij}^{\min}$ value using various methods from a simple Hückel method to an ab initio MO method after unitary transformation.

Although the $L_{ij}^{\min}$ value is useful for evaluating the ferromagnetic properties of NBMO systems, there are some problems with huge systems in which we cannot obtain coefficients of the whole MO. The purpose of this article is to predict the $L_{ij}^{\min}$ value analytically without performing any quantum chemical calculations and unitary transformations. Our new treatment enables us to predict the shapes of the NBMOs corresponding to the $L_{ij}^{\min}$ value by using a "zero sum rule" in the NBMO theory. In the present article, we focused on benzyl radical type systems as a first step of the analysis, because the benzyl radical unit is involved in typical organic ferromagnetic materials. It was found that we were able to predict the $L_{ij}^{\min}$ values of various types of benzyl radical type oligomers without using direct calculations.

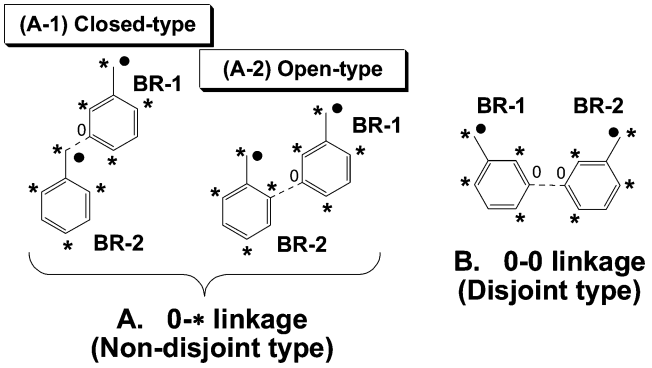


Figure 1. 0-∗ and 0−0 linkages between benzyl radical (BR) units in the benzyl radical dimer model. The broken line denotes the C−C bond connecting the BR units.

2. Method

2.1. Benzyl Radical Dimer Model. The benzyl radical dimer model in Figure 1 is composed of two benzyl radical (BR) units, BR-1 and BR-2. It should be noted that the structure of the benzyl radical oligomer in the present article is assumed to have a planar structure. According to the classification mentioned above, we can divide benzyl radical dimer models into two types: (A) nondisjoint type with a (0-∗) linkage and (B) disjoint type with a (0−0) linkage. Furthermore, the former type is divided into "(A-1) closed-type" and "(A-2) open-type." The nondisjoint closed-type model, (A-1), is defined as a model in which two BR units are connected at the sites of the inactive carbon atom (0) in the meta position of BR-1 and the radical center of BR-2. Here, the radical center is defined as the active carbon atom (∗) having the largest NBMO coefficient in the BR unit marked by a dot. In contrast, the nondisjoint open-type model, (A-2), is defined as the model that includes the connection between the inactive atom (0) in the meta position of BR-1 and the active atom (∗) in the ortho position of BR-2. In the present work, only the nondisjoint closed-type model, (A-1), is focused on, because we have not yet succeeded in predicting the ferromagnetic properties for the open-type model.

Figure 2 shows a method for predicting the shapes of NBMOs that provide the $L_{ij}^{\min}$ value for a benzyl radical dimer model. This procedure is the first step of the analytical prediction (AP) method to evaluate the high-spin stability of nondisjoint closed-type benzyl radical oligomers. Because the benzyl radical dimer model in Figure 2 has a (0-∗) linkage, the system keeps two degenerated NBMOs unchanged as a result of the interaction between the inactive and active carbon atoms. The procedures are summarized as follows:

(i) In the first step (see step 0 in Figure 2), the whole system is divided into two radical units, BR-1 and BR-2.

(ii) Next, we focus on the upper BR unit (BR-1) and consider assigning its NBMO coefficients using the letter "a" according to the conventional NBMO rule[23,24] (step 1a). Furthermore, another unit (BR-2) should be assigned based on the NBMO coefficients of BR-1 (step 2a). According to the "zero sum rule" in the NBMO theory, the NBMO

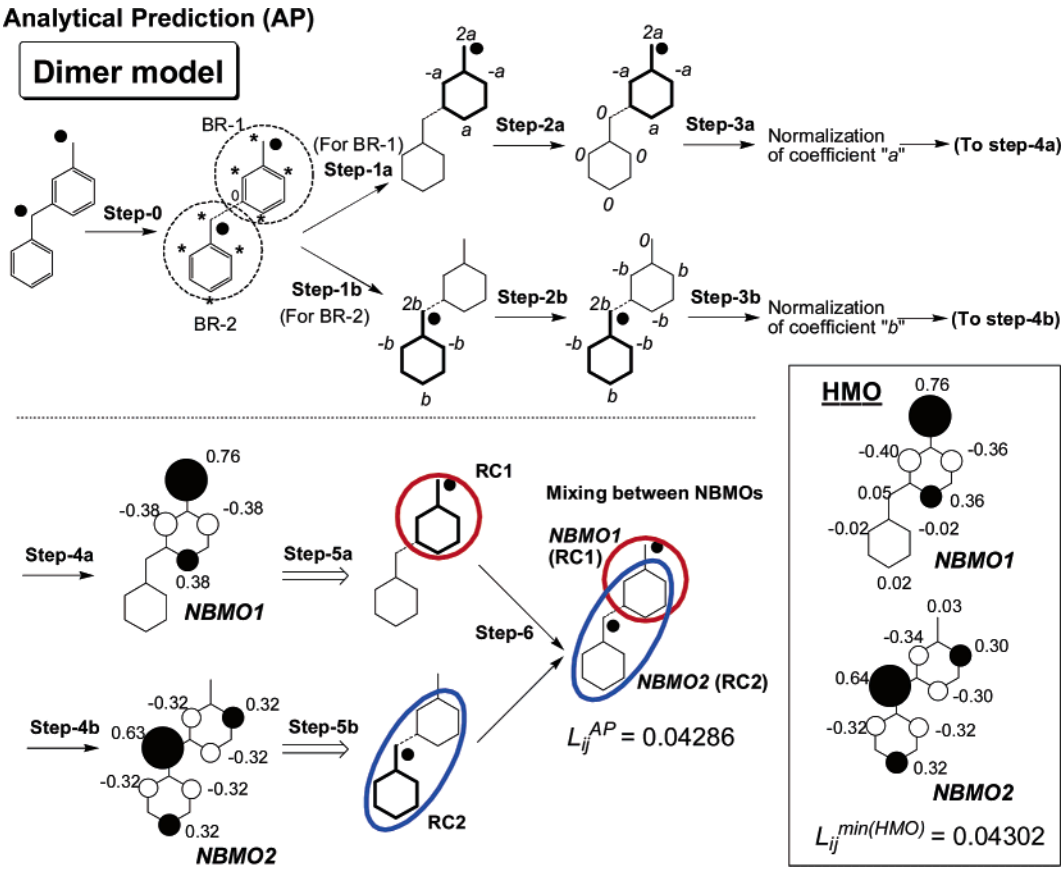


Figure 2. Procedures for predicting the shapes of NBMOs and the L_{ij} values by the analytical prediction (AP) method. The benzyl radical units under consideration are indicated by bold lines. The results by the simple Hückel method (HMO) are shown in the rectangular box.

coefficients of all the carbon atoms in BR-2 become zero. This implies that the NBMO coefficients of BR-1 cannot be delocalized over the region of BR-2.

(iii) Conversely, we focus on the lower BR unit (BR-2) and consider assigning the NBMO coefficients by using the letter "*b*" (step 1b), followed by assigning the NBMO coefficients of the BR-1 region (step 2b). In this case, one can assign the BR-1 region with nonzero NBMO coefficients except for the site of the radical center in BR-1. Thus, the NBMO coefficients of BR-2 can be delocalized over the region of BR-1 through a C−C bond connecting two BR units (broken line), and the NBMO spreads over two benzene rings.

(iv) The "*a*" and "*b*" are normalized for each NBMO by $\sum_r C_{ir}^2 = 1$

$$\text{Step 3a: } (2a)^2 + 2(-a)^2 + a^2 = 7a^2 = 1; \quad a = \frac{1}{\sqrt{7}} \approx 0.37796 \quad (5a)$$

$$\text{Step 3b: } (2b)^2 + 4(-b)^2 + 2b^2 = 10b^2 = 1; \quad b = \frac{1}{\sqrt{10}} \approx 0.31623 \quad (5b)$$

(v) Finally, we obtained NBMO1 and NBMO2 (see steps 4a and 4b). Here, we assigned the types of radical centers (RCs) according to the NBMO shape. In NBMO1, MO coefficients spread over one benzene ring (step 5a). In NBMO2, MO coefficients spread over two benzene rings (step 5b). The radical centers that belong to NBMO1 and NBMO2 are called RC1 and RC2, respectively.

(vi) The stability of the high-spin ground state is predicted by calculating the mixing between the two NBMOs (see step 6), i.e., $L_{ij}^{AP} = \sum_r (C_{ir}C_{jr})^2$, by

$$L_{ij}^{AP} = L_{ij}^{AP[NBMO1(RC1)-NBMO2(RC2)]} = (ab)^2 + (ab)^2 + (ab)^2 = 3a^2b^2 = 3\left(\frac{1}{\sqrt{7}}\right)^2\left(\frac{1}{\sqrt{10}}\right)^2 = \frac{3}{70} \approx 0.04286 \quad (6)$$

where $L_{ij}^{AP[NBMO1(RC1)-NBMO2(RC2)]}$ represents the mixing between NBMO1(RC1) and NBMO2(RC2).

Figure 2 also shows the calculation of NBMO coefficients by the simple Hückel method (HMO) and the L_{ij}^{min} value after unitary transformation ($L_{ij}^{min(HMO)} = 0.04302$). It was found that the difference in the NBMO coefficients between the AP and HMO methods is very small. Therefore, the L_{ij}^{AP} value agreed well with the $L_{ij}^{min(HMO)}$ value and provides an even better minimum value in L_{ij}. In the present case, it was found that the $L_{ij}^{min(HMO)}$ value is slightly larger than the L_{ij}^{AP} value. This is because the NBMO coefficients calculated by the HMO method are not suitable for the benzyl radical dimer model as the initial coefficients for unitary rotation, and thus NBMO coefficients do not lead to the minimum L_{ij} value

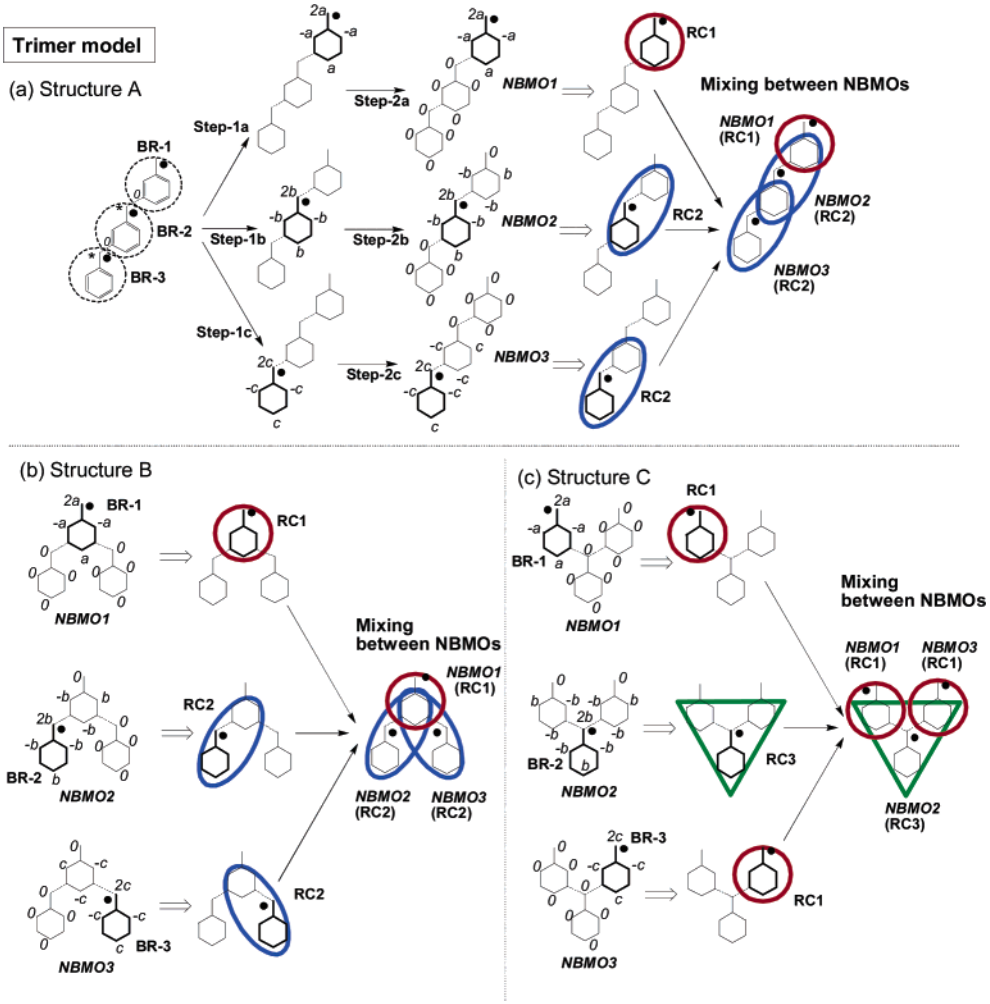


Figure 3. Analytical prediction of the L_{ij} value for three different structures, A−C in the benzyl radical trimer model. The benzyl radical units under consideration are indicated by bold lines.

after unitary transformation. This means that the L_{ij}^{AP} value estimated by the AP method provides us with the minimum value of L_{ij} regardless of the systems.

2.2. Benzyl Radical Trimer Model. The benzyl radical trimer model in Figure 3 is composed of three BR units (BR-1, BR-2, and BR-3). For the "nondisjoint closed-type" trimer model, we consider three types of structures A, B, and C with three NBMOs. We can predict the shapes of the NBMOs in the same way that we can predict the benzyl radical dimer model mentioned above.

Structure A is considered first (see Figure 3(a)). NBMO coefficients are assigned for each BR unit under consideration (steps 1a−c). Next, we consider the coefficients of each NBMO for the whole system (steps 2a−c), followed by normalization of the NBMO coefficients. The letters "*a*", "*b*", and "*c*" are used for the MO coefficients of NBMO1, NBMO2, and NBMO3, respectively. According to subsection 2.1, one can assign NBMO1, NBMO2, and NBMO3 as RC1, RC2, and RC2, respectively. In a manner similar to the procedure (iv) in subsection 2.1, these NBMO coefficients are calculated with $a = 1/\sqrt{7}$ (for RC1) and $b = c = 1/\sqrt{10}$ (for RC2). Therefore, the L_{ij}^{AP} value for structure A can be evaluated by

$$\text{Structure A: } L_{ij}^{AP} = L_{ij}^{AP[NBMO1(RC1)-NBMO2(RC2)]} + L_{ij}^{AP[NBMO2(RC2)-NBMO3(RC2)]} = 3 \times (ab)^2 + 3 \times (bc)^2 = \frac{3}{70} + \frac{3}{100} \approx 0.07286 \quad (7)$$

There is no mixing between NBMO1(RC1) and NBMO3(RC2) in structure A.

In structure B, NBMO1, NBMO2, and NBMO3 belong to RC1, RC2, and RC2, respectively (see Figure 3(b)). Thus, we obtained $a = 1/\sqrt{7}$ (for RC1) and $b = c = 1/\sqrt{10}$ (for RC2). The L_{ij}^{AP} value for B structure is calculated by

$$\text{Structure B: } L_{ij}^{AP} = L_{ij}^{AP[NBMO1(RC1)-NBMO2(RC2)]} + L_{ij}^{AP[NBMO1(RC1)-NBMO3(RC2)]} + L_{ij}^{AP[NBMO2(RC2)-NBMO3(RC2)]} = 3 \times (ab)^2 + 3 \times (ac)^2 + 3 \times (bc)^2 = \frac{3}{70} + \frac{3}{70} + \frac{3}{100} \approx 0.11571 \quad (8)$$

Finally, we consider structure C (see Figure 3(c)). Both NBMO1 and NBMO3 are assigned as RC1, and we obtained $a = c = 1/\sqrt{7}$ (for RC1). The NBMO2 in which the coefficients spread over three benzene rings is assigned as

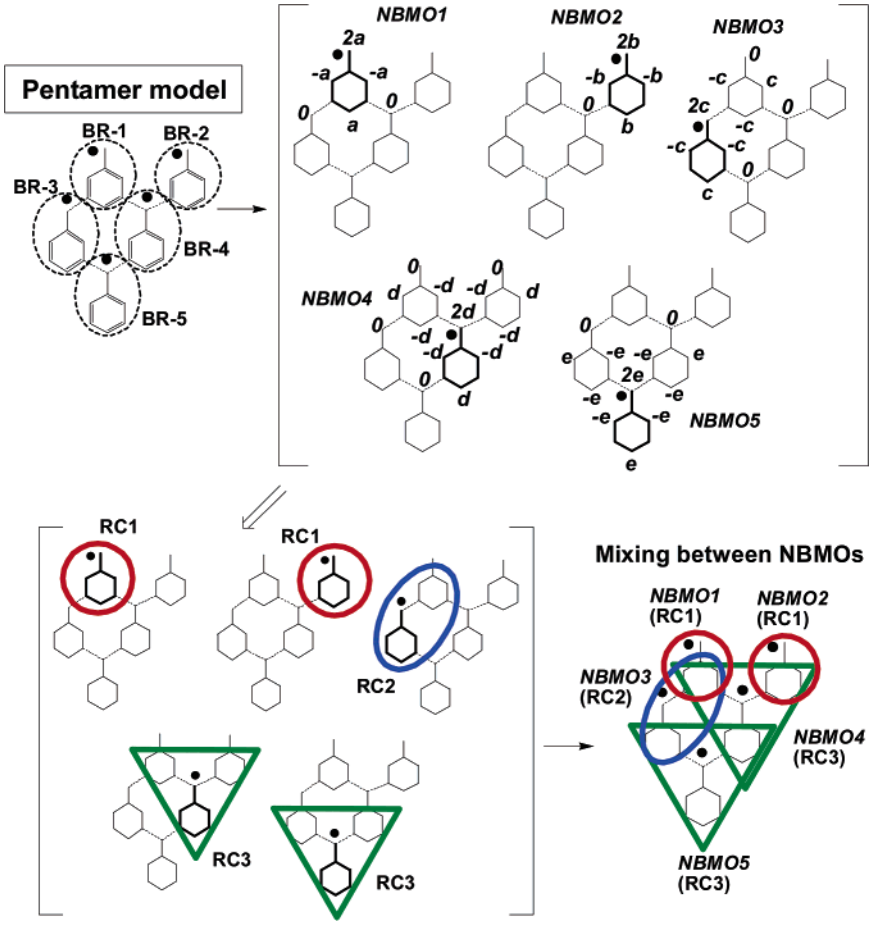


Figure 4. Analytical prediction of the L_{ij} value for a benzyl radical pentamer model. The benzyl radical units under consideration are indicated by bold lines.

RC3. The NBMO coefficients corresponding to RC3 are calculated by the normalization of value "*b*" in Figure 3(c), i.e.

$$(2b)^2 + 6(-b)^2 + 3b^2 = 13b^2 = 1;\ b = \frac{1}{\sqrt{13}} \approx 0.27735 \quad \text{(for RC3)} \quad (9)$$

Therefore, the L_{ij}^{AP} value for structure C is calculated by

$$\text{Structure C: } L_{ij}^{AP} = L_{ij}^{AP[NBMO1(RC1)-NBMO2(RC3)]} + L_{ij}^{AP[NBMO2(RC3)-NBMO3(RC1)]} = 3 \times (ab)^2 + 3 \times (bc)^2 = \frac{3}{91} + \frac{3}{91} \approx 0.06593 \quad (10)$$

There is no mixing between NBMO1(RC1) and NBMO3(RC1) in structure C.

2.3. Benzyl Radical Pentamer Model. Figure 4 shows a benzyl radical pentamer model with five benzyl radical units, BR-1−BR-5. This model includes five NBMOs, NBMO1−NBMO5. NBMO1 and NBMO2 belong to RC1. NBMO3 corresponds to RC2, and both NBMO4 and NBMO5 correspond to RC3. Therefore, the NBMO coefficients are calculated with $a = b = 1/\sqrt{7}$ (for RC1), $c = 1/\sqrt{10}$ (for RC2), and $d = e = 1/\sqrt{13}$ (for RC3). The L_{ij}^{AP} value can be estimated by

$$L_{ij}^{AP} = L_{ij}^{AP[NBMO1(RC1)-NBMO3(RC2)]} + L_{ij}^{AP[NBMO1(RC1)-NBMO4(RC3)]} + L_{ij}^{AP[NBMO2(RC1)-NBMO4(RC3)]} + L_{ij}^{AP[NBMO3(RC2)-NBMO4(RC3)]} + L_{ij}^{AP[NBMO3(RC2)-NBMO5(RC3)]} + L_{ij}^{AP[NBMO4(RC3)-NBMO5(RC3)]} = 3 \times (ac)^2 + 3 \times (ad)^2 + 3 \times (bd)^2 + 3 \times (cd)^2 + 3 \times (ce)^2 + 3 \times (de)^2 = \frac{3}{70} + \frac{3}{91} + \frac{3}{91} + \frac{3}{130} + \frac{3}{130} + \frac{3}{169} \approx 0.17270 \quad (11)$$

All the other combinations between the NBMOs in eq 11, for example, between NBMO1 and NBMO2, have no mixing.

2.4. Treatment of Methylene or Methylidyne Radical Units. Figure 5 shows a nondisjoint closed-type system including methylene (:CH_2) or methylidyne (:CH•) radical (MR) units, MR-1 and MR-2, in addition to benzyl radical units, BR-1 and BR-2. We can predict the shapes of the NBMOs and the L_{ij}^{AP} value for such a system in the following manner.

NBMO coefficients were assigned for each radical unit under consideration (see step 1). Note that the NBMO coefficient in the MR unit is initially assigned by "*2a*" for MR-1 and "*2c*" for MR-2. Next, we considered the coefficients of each NBMO in other radical units (see step 2) and obtained four NBMOs 1−4. NBMO1, 2, 3, and 4 can be assigned as RC1, RC1, RC2, and RC1, respectively (see step 3). After normalization of the NBMO coefficients, $a =$

$b = d = 1/\sqrt{7}$ (for RC1) and $c = 1/\sqrt{10}$ (for RC2) can be calculated. The L_{ij}^{AP} value is estimated (see step 4) by

$$L_{ij}^{AP} = L_{ij}^{AP[NBMO1(RC1)-NBMO2(RC1)]} + L_{ij}^{AP[NBMO1(RC1)-NBMO3(RC2)]} + L_{ij}^{AP[NBMO2(RC1)-NBMO3(RC2)]} + L_{ij}^{AP[NBMO3(RC2)-NBMO4(RC1)]} = 3 \times (ab)^2 + 3 \times (ac)^2 + 3 \times (bc)^2 + 3 \times (cd)^2 = \frac{3}{49} + \frac{3}{70} + \frac{3}{70} + \frac{3}{70} \approx 0.18980 \quad (12)$$

Therefore, we can treat methylene or methylidyne radical units in a manner similar to the benzyl radical units except for the initial assignment for the NBMO coefficients as shown in step 1 of Figure 5.

2.5. Generalization of the Estimation of the L_{ij}^{AP} Value. In the above subsections, we defined three types of radical centers according to the shape of their NBMOs. That is to say, RC1, RC2, and RC3 represent the NBMOs spreading over one benzene ring, two benzene rings, and three benzene rings, respectively. It should be noted that for a huge benzyl radical oligomer, only the three types of radical centers, RC1, RC2, and RC3, can be considered for a nondisjoint closed type system. Thus, only six types of mixings between the different shapes of *i*th and *j*th NBMOs are considered as

$$\text{RC1-RC1: } L_{ij}^{AP(RC1-RC1)} = 3 \times (\alpha_i^{RC1}\alpha_j^{RC1})^2 = 3 \times \left(\frac{1}{\sqrt{7}} \times \frac{1}{\sqrt{7}}\right)^2 = \frac{3}{49} \approx 0.06122 \quad (13a)$$

$$\text{RC1-RC2: } L_{ij}^{AP(RC1-RC2)} = 3 \times (\alpha_i^{RC1}\alpha_j^{RC2})^2 = 3 \times \left(\frac{1}{\sqrt{7}} \times \frac{1}{\sqrt{10}}\right)^2 = \frac{3}{70} \approx 0.04286 \quad (13b)$$

$$\text{RC1-RC3: } L_{ij}^{AP(RC1-RC3)} = 3 \times (\alpha_i^{RC1}\alpha_j^{RC3})^2 = 3 \times \left(\frac{1}{\sqrt{7}} \times \frac{1}{\sqrt{13}}\right)^2 = \frac{3}{91} \approx 0.03297 \quad (13c)$$

$$\text{RC2-RC2: } L_{ij}^{AP(RC2-RC2)} = 3 \times (\alpha_i^{RC2}\alpha_j^{RC2})^2 = 3 \times \left(\frac{1}{\sqrt{10}} \times \frac{1}{\sqrt{10}}\right)^2 = \frac{3}{100} = 0.03 \quad (13d)$$

$$\text{RC2-RC3: } L_{ij}^{AP(RC2-RC3)} = 3 \times (\alpha_i^{RC2}\alpha_j^{RC3})^2 = 3 \times \left(\frac{1}{\sqrt{10}} \times \frac{1}{\sqrt{13}}\right)^2 = \frac{3}{130} \approx 0.02308 \quad (13e)$$

$$\text{RC3-RC3: } L_{ij}^{AP(RC3-RC3)} = 3 \times (\alpha_i^{RC3}\alpha_j^{RC3})^2 = 3 \times \left(\frac{1}{\sqrt{13}} \times \frac{1}{\sqrt{13}}\right)^2 = \frac{3}{169} \approx 0.01775 \quad (13f)$$

where α_i^{RC1} indicates the NBMO coefficient of *i*th NBMO with an RC1-type radical center. A series of NBMO coefficients is provided in subsections 2.1 and 2.2 as $\alpha_i^{RC1} = \alpha_j^{RC1} = 1/\sqrt{7}$, $\alpha_i^{RC2} = \alpha_j^{RC2} = 1/\sqrt{10}$, and $\alpha_i^{RC3} = \alpha_j^{RC3} = 1/\sqrt{13}$. Therefore, the L_{ij}^{AP} value for the whole system

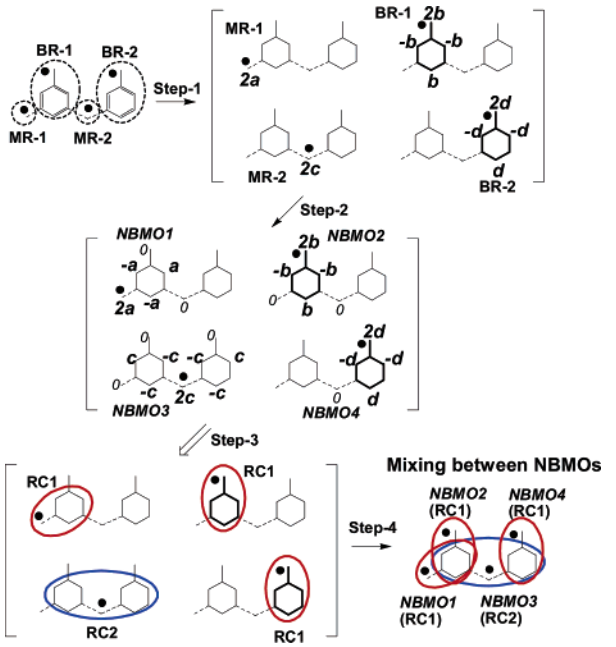


Figure 5. Analytical prediction of the L_{ij} value for a system including methylene (:CH_2) or methylidyne (:CH•) radical units. The benzyl radical units under consideration are indicated by bold lines.

can be calculated by counting all the NBMO mixings (eq 13a−f) over the whole system as

$$L_{ij}^{AP} = N^{AP(RC1-RC1)}L_{ij}^{AP(RC1-RC1)} + N^{AP(RC1-RC2)}L_{ij}^{AP(RC1-RC2)} + N^{AP(RC1-RC3)}L_{ij}^{AP(RC1-RC3)} + N^{AP(RC2-RC2)}L_{ij}^{AP(RC2-RC2)} + N^{AP(RC2-RC3)}L_{ij}^{AP(RC2-RC3)} + N^{AP(RC3-RC3)}L_{ij}^{AP(RC3-RC3)} \quad (14)$$

where $N^{AP(RC1-RC1)}$, $N^{AP(RC1-RC2)}$, etc. indicate the number of mixings, $L_{ij}^{AP(RC1-RC1)}$, $L_{ij}^{AP(RC1-RC2)}$, etc., respectively.

The process for counting the NBMO mixings can be conducted by judging whether the two NBMOs under consideration make mixings or not. For instance, for the benzyl radical pentamer model (see Figure 4) we can count the components of NBMO mixings using eqs 13 and 14 as

$$L_{ij}^{AP} = 0 \times L_{ij}^{AP(RC1-RC1)} + 1 \times L_{ij}^{AP(RC1-RC2)} + 2 \times L_{ij}^{AP(RC1-RC3)} + 0 \times L_{ij}^{AP(RC2-RC2)} + 2 \times L_{ij}^{AP(RC2-RC3)} + 1 \times L_{ij}^{AP(RC3-RC3)} = 1 \times \left(\frac{3}{70}\right) + 2 \times \left(\frac{3}{91}\right) + 2 \times \left(\frac{3}{130}\right) + 1 \times \left(\frac{3}{169}\right) \approx 0.17270 \quad (15)$$

Therefore, finally, we obtained the same results as eq 11. Similarly, one can count the NBMO mixings and calculate the L_{ij}^{AP} value for the system including methylene or methylidyne radical units.

3. Results and Discussion

Figure 6 shows model systems for a quasi-one-dimensional benzyl radical oligomer that are combined by the "nondisjoint closed-type", as shown in Figure 1(A-1). Models 1−3 in

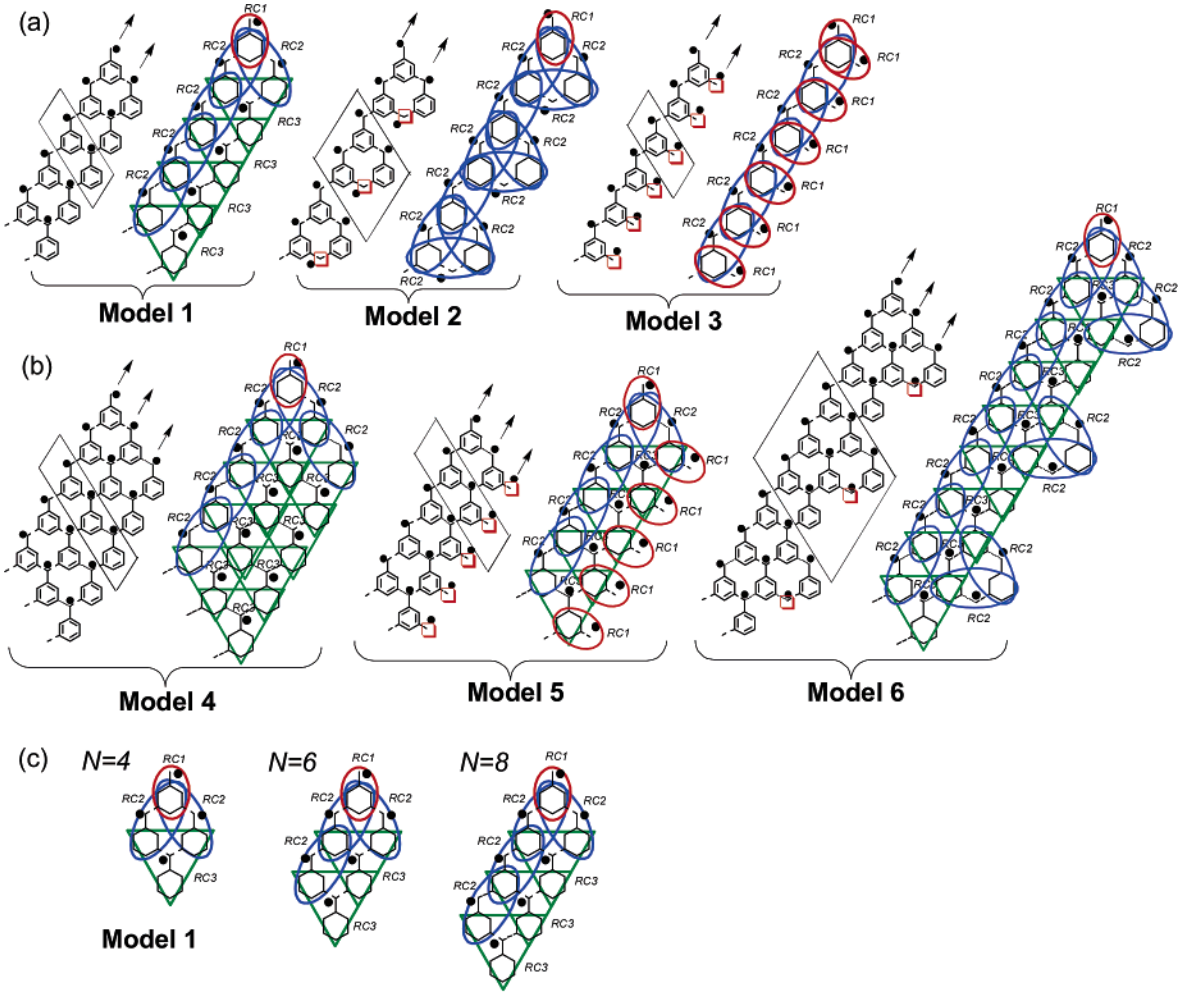


Figure 6. Nondisjoint closed-type models for quasi-one-dimensional benzyl radical oligomers with (a) two- and (b) three-line alignment of radical centers. The red open square indicates a methylene (:CH_2) or methylidyne (:CH•) radical unit. The periodic structure for each system is shown in a parallelogram. The arrows indicate the directions of the alignment of the radical centers. The mixings between NBMOs are illustrated in the right-hand figure of each model. (c) Model 1 with $N = 4$, 6, and 8 for the formulation of the L_{ij}^{AP} value in Table 1.

Figure 6(a) correspond to "two-line systems" in which radical centers are arranged in two lines as shown by the arrows indicating the directions of alignment. In contrast, models 4−6 in Figure 6(b) correspond to "three-line systems" in which radical centers are arranged in three lines. Note that models 2, 3, 5, and 6 include methylene (:CH_2) or methylidyne (:CH•) radical units as indicated by the red open squares in the figure.

3.1. Analytical Predictions for Two-Line Systems. In this subsection, we formulate the L_{ij}^{AP} value for the nondisjoint closed-type models 1−3 in the two-line system. When the number of radical centers is defined as "N", we can count the number of mixings between NBMOs according to the right figure of each model (see Figure 6(a)). Table 1 (a) and Figure 6(c) show the formulation of the L_{ij}^{AP} value for model 1. The number of L_{ij}^{AP}'s components, that is, $L_{ij}^{AP(RC1-RC2)}$, $L_{ij}^{AP(RC2-RC2)}$, etc., are counted for some sizes of periodic structures step by step ($N = 4$, 6, 8...). Eventually, we can estimate the number of each component for an arbitrary size of oligomer, that is, $N = n$. Therefore, for model 1 the L_{ij}^{AP} value is formulated as

$$\textbf{Model 1:}\ L_{ij}^{AP} = 2L_{ij}^{AP(RC1-RC2)} + \left(\frac{N}{2} - 1\right)L_{ij}^{AP(RC2-RC2)} + (N-2)L_{ij}^{AP(RC2-RC3)} + \left(\frac{N}{2} - 2\right)L_{ij}^{AP(RC3-RC3)} = \frac{6}{70} + \left(\frac{N}{2} - 1\right)\frac{3}{100} + (N-2)\frac{3}{130} + \left(\frac{N}{2} - 2\right)\frac{3}{169} \quad (16a)$$

where small "n" is replaced with capital "N". The increase of the L_{ij}^{AP} value per radical center is calculated by considering the N limits ($N \rightarrow \infty$) of L_{ij}^{AP}/N as

$$\lim_{N\to\infty}\frac{L_{ij}^{AP}}{N} = \lim_{N\to\infty}\left\{\frac{6}{70N} + \left(\frac{1}{2} - \frac{1}{N}\right)\frac{3}{100} + \left(1 - \frac{2}{N}\right)\frac{3}{130} + \left(\frac{1}{2} - \frac{2}{N}\right)\frac{3}{169}\right\} = \left(\frac{1}{2}\right)\frac{3}{100} + \frac{3}{130} + \left(\frac{1}{2}\right)\frac{3}{169} \approx 0.04695 \quad (16b)$$

In a similar way, we can estimate the L_{ij}^{AP} values (see also Table 1(b) and (c)) and the increases per radical center for models 2 and 3 as follows:

$$\textbf{Model 2:}\ L_{ij}^{AP} = 2L_{ij}^{AP(RC1-RC2)} + \left(\frac{7}{4}N - 4\right)L_{ij}^{AP(RC2-RC2)} = \frac{6}{70} + \left(\frac{7}{4}N - 4\right)\frac{3}{100} \quad (17a)$$

$$\lim_{N\to\infty}\frac{L_{ij}^{AP}}{N}=\lim_{N\to\infty}\left\{\frac{6}{70N}+\left(\frac{7}{4}-\frac{4}{N}\right)\frac{3}{100}\right\}=\left(\frac{7}{4}\right)\frac{3}{100}=0.0525 \quad (17b)$$

Model 3: $L_{ij}^{AP}=L_{ij}^{AP(RC1-RC1)}+(N-1)L_{ij}^{AP(RC1-RC2)}+\left(\frac{N}{2}-2\right)L_{ij}^{AP(RC2-RC2)}=\frac{3}{49}+(N-1)\frac{3}{70}+\left(\frac{N}{2}-2\right)\frac{3}{100}$ (18a)

$$\lim_{N\to\infty}\frac{L_{ij}^{AP}}{N}=\lim_{N\to\infty}\left\{\frac{3}{49N}+\left(1-\frac{1}{N}\right)\frac{3}{70}+\left(\frac{1}{2}-\frac{2}{N}\right)\frac{3}{100}\right\}=\frac{3}{70}+\left(\frac{1}{2}\right)\frac{3}{100}\approx 0.05786 \quad (18b)$$

Figure 7(a) shows the relationship between the L_{ij}^{AP} value and the number of radical centers (N) for models 1−3. Our AP method predicts the order of high-spin stability as "*model 3 > model 2 > model 1*" regardless of N. This order directly reflects the order of the increase of the L_{ij}^{AP} value per radical center, that is, $\lim_{N\to\infty}(L_{ij}^{AP}/N)$.

To confirm the validity of the L_{ij}^{AP} value, we compared the L_{ij}^{AP} value in this work with the $L_{ij}^{min(HMO)}$ value obtained by the HMO method in Figure 7(b) as the number of radical centers, $N = 8$ and 16, are selected here. It was found that the absolute value of the L_{ij}^{AP} value is underestimated compared with the $L_{ij}^{min(HMO)}$ value. However, the order of the L_{ij} value for the three models exhibits a very similar tendency toward each other in both $N = 8$ and $N = 16$. The only difference is the order of models 2 and 3 in $N = 8$. Such differences in the L_{ij} value between the AP and HMO methods result from the fact that in these models, after unitary rotation, the $L_{ij}^{min(HMO)}$ values do not correspond to the minimum L_{ij} value correctly, as mentioned in subsection 2.1.

Table 1. Formulation of L_{ij}^{AP} Value for Models 1−3 (Two-Line System)[a]

(a) Model 1[b]

	components of L_{ij}^{AP} value			
N	$L_{ij}^{AP(RC1-RC2)}$	$L_{ij}^{AP(RC2-RC2)}$	$L_{ij}^{AP(RC2-RC3)}$	$L_{ij}^{AP(RC3-RC3)}$
4	2	1	2	0
6	2	2	4	1
8	2	3	6	2
10	2	4	8	3
12	2	5	10	4
n	2	$(n/2)-1$	$n-2$	$(n/2)-2$

(b) Model 2

	components of L_{ij}^{AP} value	
N	$L_{ij}^{AP(RC1-RC2)}$	$L_{ij}^{AP(RC2-RC2)}$
4	2	3
8	2	10
12	2	17
n	2	$(7/4)n-4$

(c) Model 3

	components of L_{ij}^{AP} value		
N	$L_{ij}^{AP(RC1-RC1)}$	$L_{ij}^{AP(RC1-RC2)}$	$L_{ij}^{AP(RC2-RC2)}$
2	1	0	
4	1	3	0
6	1	5	1
8	1	7	2
10	1	9	3
n	1	$n-1$	$(n/2)-2$

[a] N represents the number of radical centers in the system. [b] Systems with $N = 4$, 6, and 8 in model 1 are shown in Figure 6(c).

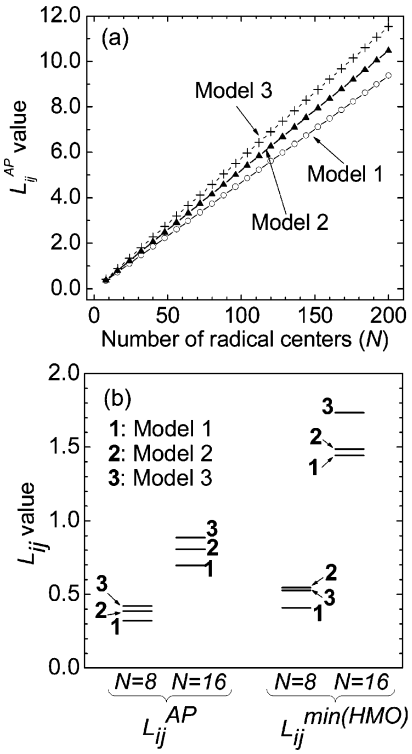


Figure 7. (a) Relationship between the L_{ij}^{AP} value and the number of radical centers (N) in two-line systems (models 1−3). (b) Comparison of the L_{ij} value between the analytical prediction (AP) method (L_{ij}^{AP}) and the HMO method ($L_{ij}^{min(HMO)}$). $N = 8$ and 16 are selected as the number of radical centers.

Next, we examined the relationship between the L_{ij}^{AP} values and the energetic stability of the high-spin ground states (ΔE(L−H)) estimated by single-point calculations using ab initio MO methods by the Gaussian03 program package.[25] High-spin stability ΔE(L−H) is defined as the difference in total energy between the lowest spin state (E(L)) and the highest spin state (E(H)), that is, ΔE(L−H) = E(L) − E(H). All the single-point calculations adopt the standard geometric parameters, as follows: [benzene ring] C−C = 1.395 Å, C−H = 1.100 Å; [CH or CH_2 group] C−C = 1.54 Å, C−H = 1.070 Å. All bond angles are fixed at 120°, and dihedral angles are fixed to keep a planar structure.

We calculated the ΔE(L−H) values for models 1−3 at N = 8 using various methods and basis sets within ab initio MO calculations (see Figure 8). The methods selected were the Hartree−Fock (HF) method, second-order Mϕller−Plesset perturbation theory (MP2) including frozen core (FC) approximation, and the density functional theory (DFT) method using the Becke-3−Lee−Yang−Parr (B3LYP) functional.[26−28] For calculating open-shell systems, we adopted the restricted open-shell Hartree−Fock (ROHF) method for each theory. The basis sets are selected as 3-21G

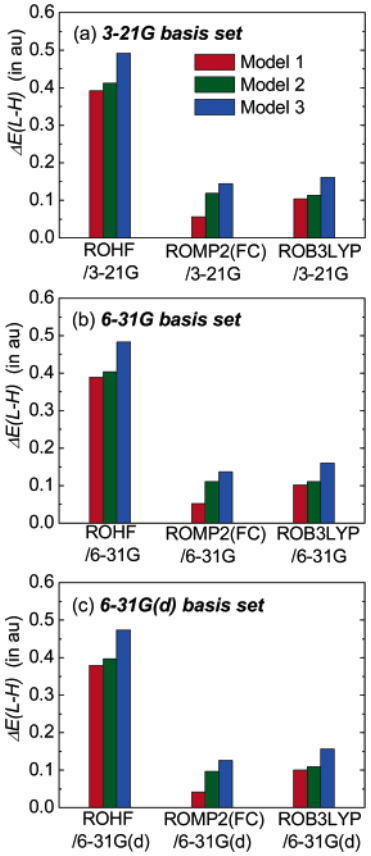


Figure 8. Ab initio MO calculations of high-spin stability ΔE(L−H) for two-line systems, models 1−3 ((a) 3-21G, (b) 6-31G, and (c) 6-31G(d) basis sets). ΔE(L−H) indicates the difference in total energy between the lowest and highest spin states, that is, ΔE(L−H) = E(L) − E(H). $N = 8$ is selected as the number of radical centers.

(in panel (a)), 6-31G (in panel (b)), and 6-31G(d) (in panel (c)). Figure 8(a)−(c) shows the order of high-spin stability as "*model 3* > *model 2* > *model 1*" regardless of the method and the size of the basis set. This order of high-spin stability, ΔE(L−H), agrees well with the order of L_{ij}^{AP} that is predicted by our AP method. It was found from Figure 8 that the ROMP2(FC) and ROB3LYP results exhibit smaller values of ΔE(L−H) than do the ROHF results. This is because the effects of electron correlations primarily stabilize the low-spin state. In other words, the HF calculation overestimates the stability of the high-spin state.

Furthermore, we examined the dependence of the order of ΔE(L−H) on the number of radical centers (N) (see Figure 9). We calculated the ΔE(L−H) value of models 1−3 for $N = 8$, 16, 24, and 32 at the levels of ROB3LYP/3-21G. The order of high-spin stability of ΔE(L−H) is "*model 3* > *model 2* > *model 1*" regardless of N. It was found that both the L_{ij}^{AP} value (by the AP method) and ΔE(L−H) (by ROB3LYP/3-21G calculation) show a similar tendency in the high-spin stability of systems with even larger values of N.

3.2. Analytical Predictions for Three-Line Systems. In this subsection, we formulate the L_{ij}^{AP} value for the "non-disjoint closed-type" models 4−6 in three-line systems (see Figure 6(b) and Table 2). We can count the number of mixings between NBMOs in a way similar to that of the

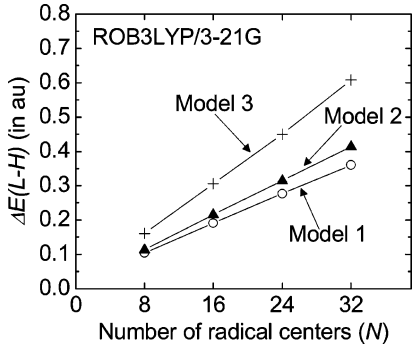


Figure 9. Dependence of high-spin stabilityΔE(L−H) on the number of radical centers (N) for two-line systems (models 1−3) at the ROB3LYP/3-21G level. ΔE(L−H) indicates the difference in total energy between the lowest and highest spin states.

two-line system. The L_{ij}^{AP} values and the increase per radical center ($\lim_{N\to\infty}(L_{ij}^{AP}/N)$) for models 4−6 were estimated as

$$\textbf{Model 4:}\quad L_{ij}^{AP} = 2L_{ij}^{AP(RC1-RC2)} + \left(\frac{N}{3}\right)L_{ij}^{AP(RC2-RC2)} + \left(\frac{2}{3}N\right)L_{ij}^{AP(RC2-RC3)} + \left(\frac{4}{3}N - 7\right)L_{ij}^{AP(RC3-RC3)} = \frac{6}{70} + \frac{N}{100} + \frac{2N}{130} + \left(\frac{4}{3}N - 7\right)\frac{3}{169} \quad (19a)$$

$$\lim_{N\to\infty}\frac{L_{ij}^{AP}}{N} = \lim_{N\to\infty}\left\{\frac{6}{70N} + \frac{1}{100} + \frac{2}{130} + \left(\frac{4}{3} - \frac{7}{N}\right)\frac{3}{169}\right\} = \frac{1}{100} + \frac{2}{130} + \frac{4}{169} \approx 0.04905 \quad (19b)$$

$$\textbf{Model 5:}\quad L_{ij}^{AP} = 3L_{ij}^{AP(RC1-RC2)} + \left(\frac{N}{3} - 1\right)(2L_{ij}^{AP(RC1-RC3)} + L_{ij}^{AP(RC2-RC2)} + 2L_{ij}^{AP(RC2-RC3)}) + \left(\frac{N}{3} - 2\right)L_{ij}^{AP(RC3-RC3)} = \frac{9}{70} + \left(\frac{N}{3} - 1\right)\left(\frac{6}{91} + \frac{3}{100} + \frac{6}{130}\right) + \left(\frac{N}{3} - 2\right)\frac{3}{169} \quad (20a)$$

$$\lim_{N\to\infty}\frac{L_{ij}^{AP}}{N} = \lim_{N\to\infty}\left\{\frac{9}{70N} + \left(\frac{1}{3} - \frac{1}{N}\right)\left(\frac{6}{91} + \frac{3}{100} + \frac{6}{130}\right) + \left(\frac{1}{3} - \frac{2}{N}\right)\frac{3}{169}\right\} = \frac{2}{91} + \frac{1}{100} + \frac{2}{130} + \frac{1}{169} \approx 0.05328 \quad (20b)$$

$$\textbf{Model 6:}\quad L_{ij}^{AP} = 2L_{ij}^{AP(RC1-RC2)} + \left(\frac{N}{2}\right)L_{ij}^{AP(RC2-RC2)} + \left(\frac{5}{4}N - 3\right)L_{ij}^{AP(RC2-RC3)} + \left(\frac{3}{8}N - 2\right)L_{ij}^{AP(RC3-RC3)} = \frac{6}{70} + \left(\frac{N}{2}\right)\frac{3}{100} + \left(\frac{5}{4}N - 3\right)\frac{3}{130} + \left(\frac{3}{8}N - 2\right)\frac{3}{169} \quad (21a)$$

$$\lim_{N\to\infty}\frac{L_{ij}^{AP}}{N} = \lim_{N\to\infty}\left\{\frac{6}{70N} + \left(\frac{1}{2}\right)\frac{3}{100} + \left(\frac{5}{4} - \frac{3}{N}\right)\frac{3}{130} + \left(\frac{3}{8} - \frac{2}{N}\right)\frac{3}{169}\right\} = \left(\frac{1}{2}\right)\frac{3}{100} + \left(\frac{5}{4}\right)\frac{3}{130} + \left(\frac{3}{8}\right)\frac{3}{169} \approx 0.05050 \quad (21b)$$

Figure 10 (a) shows the relationship between the L_{ij}^{AP} value and the number of radical centers (N) for models 4−6.

Table 2. Formulation of L_{ij}^{AP} Value for Models 4−6 (Three-Line System)[a]

(a) Model 4

	components of L_{ij}^{AP} value			
N	$L_{ij}^{AP(RC1-RC2)}$	$L_{ij}^{AP(RC2-RC2)}$	$L_{ij}^{AP(RC2-RC3)}$	$L_{ij}^{AP(RC3-RC3)}$
6	2	2	4	1
9	2	3	6	5
12	2	4	8	9
15	2	5	10	13
n	2	$n/3$	$(2/3)n$	$(4/3)n-7$

(b) Model 5

	components of L_{ij}^{AP} value				
N	$L_{ij}^{AP(RC1-RC2)}$	$L_{ij}^{AP(RC1-RC3)}$	$L_{ij}^{AP(RC2-RC2)}$	$L_{ij}^{AP(RC2-RC3)}$	$L_{ij}^{AP(RC3-RC3)}$
6	3	2	1	2	0
9	3	4	2	4	1
12	3	6	3	6	2
15	3	8	4	8	3
n	3	$2(n/3-1)$	$n/3-1$	$2(n/3-1)$	$n/3-2$

(c) Model 6

	components of L_{ij}^{AP} value			
N	$L_{ij}^{AP(RC1-RC2)}$	$L_{ij}^{AP(RC2-RC2)}$	$L_{ij}^{AP(RC2-RC3)}$	$L_{ij}^{AP(RC3-RC3)}$
8	2	4	7	1
16	2	8	17	4
24	2	12	27	7
32	2	16	37	10
n	2	$n/2$	$(5/4)n-3$	$(3/8)n-2$

[a] N represents the number of radical centers in the system.

From the results, our AP method predicts the order of high-spin stability as "*model 5 > model 6 > model 4*" regardless of N.

We calculated the ΔE(L−H) values for models 4−6 at $N = 24$ at the level of ROHF/3-21G and ROB3LYP/3-21G (see Figure 10(b)). Both the ROHF and ROB3LYP results predict the order of high-spin stability as "*model 5 > model 6 > model 4*", although the ROB3LYP results show a smaller value for ΔE(L−H) than do the ROHF results due to the stabilization of the low-spin state by the effects of electron correlation. Even for the three-line system, it was found that the L_{ij}^{AP} value in Figure 10(a) agrees well with the order of ΔE(L−H) by ab initio MO calculations.

As we have seen above, it is clear that the mixing between NBMOs controls the order of the L_{ij}^{AP} value and high-spin stability. Therefore, ferromagnetic systems with large L_{ij}^{AP} values can be designed by considering the following factors.

(1) *A system with strong ferromagnetic properties should involve large mixings between NBMOs such as* $L_{ij}^{AP(RC1-RC1)}$, $L_{ij}^{AP(RC1-RC2)}$, $L_{ij}^{AP(RC1-RC3)}$, *and so on*. According to subsection 2.5, the order of magnitudes of the NBMO mixings is

$$L_{ij}^{AP(RC1-RC1)} > L_{ij}^{AP(RC1-RC2)} > L_{ij}^{AP(RC1-RC3)} > L_{ij}^{AP(RC2-RC2)} > L_{ij}^{AP(RC2-RC3)} > L_{ij}^{AP(RC3-RC3)} \quad (22)$$

(2) *In a system with strong ferromagnetic properties, each NBMO should mix with many other NBMOs.* For example, in structure A of the trimer model (see Figure 3(a)), NBMO1 mixes only with NBMO2, whereas in structure B of the trimer model (Figure 3(b)), NBMO1 mixes with both NBMO2 and NBMO3. The latter case has an advantage in the L_{ij}^{AP} value compared with the former case.

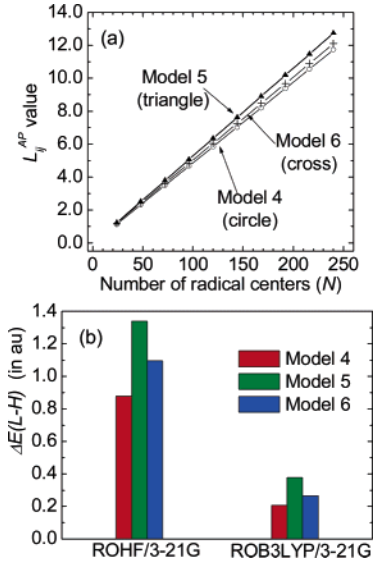


Figure 10. (a) Relationship between the L_{ij}^{AP} value and the number of radical centers (N) in three-line systems (models 4−6). (b) Ab initio MO calculations of high-spin stability ΔE(L−H) for three-line systems. $N = 24$ is selected as the number of radical centers. ΔE(L−H) indicates the difference in total energy between the lowest and highest spin states.

In particular, factor (1) plays a dominant role for determining the L_{ij}^{AP} value as long as we compare a similar size of oligomers. Therefore, model 3 with many $L_{ij}^{AP(RC1-RC2)}$ and model 5 with many $L_{ij}^{AP(RC1-RC3)}$ show the largest L_{ij}^{AP} value within the two- and three-line systems. Generally speaking, these results provide us with a few guidelines on the structure of radical units to design ferromagnetic systems with a large L_{ij}^{AP} value. First, each benzene ring should be connected to many radical centers to enlarge the L_{ij}^{AP} value. In such cases, large mixings of NBMOs are expected to generate on the benzene rings, leading to exchange interactions. Second, each radical center should not be connected to many benzene rings, because delocalized NBMOs generally provide a small L_{ij}^{AP} value due to their small NBMO coefficients.

In the present article, perfectly planar systems with no environment molecules were assumed for the simplicity. However, a deviation from the planar structure do not have primary importance for designing new materials, because it is obvious that such an effect always weakens the ferromagnetic properties due to a weakened π-conjugation. Indeed, the deviation effects can be easily estimated by considering overlaps between adjacent p-orbitals perpendicular to the molecular plane. Furthermore, exchange interactions through-space between environment and a target molecule is expected to be secondary effects in π-conjugated systems, because such through-space interactions generally show smaller effects than through-bond exchange interactions.

Although our new method was applied only to the periodic system, the method is also effective for predicting the ferromagnetism in aperiodic systems. In such cases, the L_{ij}^{AP} value is easily estimated by counting up the number of the NBMO mixings in the system.

4. Conclusion

The analytical prediction (AP) method can predict the high-spin stability ΔE(L−H) of nondisjoint closed-type benzyl radical oligomers without any direct quantum calculations. This method is composed of three procedures: (1) predicting the shape of NBMOs, (2) counting the mixings between NBMOs, and (3) formulating the relationship between the NBMO mixings and the number of radical centers (N). Although our AP method is based on the Hückel NBMO concept for simplicity, it was confirmed that the AP method well predicts the high-spin stability evaluated by ab initio MO calculations including the post-HF method or DFT method. That is to say, our new treatment simulates higher level calculations when considering the electron correlation effects.

The AP method has the potential for predicting ferromagnetic properties in large systems. For example, formulating the NBMO mixings for periodic polymers enables qualitative prediction of the ferromagnetic properties of polymers including those involving considerably large radical units. Furthermore, the AP method can also be applied to the prediction of the ferromagnetic properties of periodic and aperiodic polymers by directly counting all the mixings between NBMOs over the whole system.

Acknowledgment. This work was supported by a grant-in-aid from the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT) and by the Research and Development Applying Advanced Computational Science and Technology of the Japan Science and Technology Agency (ACT-JST). The calculations were performed on systems in our laboratory including Linux PCs.

Note Added after ASAP Publication. This article was released ASAP on March 30, 2006, with the incorrect Received Date. The correct version was posted on April 13, 2006.

References

(1) Kinoshita, M. In *Handbook of Organic Conductive Molecules and Polymers*; Nalva, H. S., Ed.; Wiley: New York, 1997; Vol. 1, Chapter 15.

(2) *Molecular Magnetism*; Ito, K., Kinoshita, M., Eds.; Kodansha, Gordon and Breach: Tokyo and Amsterdam, 2000.

(3) Rajca, A. *Chem. Rev.* **1994**, *94*, 871−893.

(4) Nishide, H.; Miyasaka, M.; Tsuchida, E. *Angew. Chem., Int. Ed. Engl.* **1998**, *37*, 2400−2402.

(5) Rajca, A.; Rajca, S.; Wongsriratanakul, J. *J. Am. Chem. Soc.* **1999**, *121*, 6308−6309.

(6) Rajca, A.; Wongsriratanakul, J.; Rajca, S. *Science* **2001**, *294*, 1503−1505.

(7) Rajca, A.; Wongsriratanakul, J.; Rajca, S. *J. Am. Chem. Soc.* **2004**, *126*, 6608−6626.

(8) Rajca, S.; Rajca, A.; Wongsriratanakul, J.; Butler, P.; Choi, S. *J. Am. Chem. Soc.* **2004**, *126*, 6972−6986.

(9) *Magnetic Properties of Organic Materials*; Lahti, P. M., Ed.; Marcel Dekker: New York, 1999.

(10) Sato, K.; Shimoi, D.; Takui, T.; Hattori, M.; Hirai, K.; Tomioka, H. *Synth. Met.* **2001**, *121*, 1816−1817.

(11) Oda, N.; Nakai, T.; Sato, K.; Shimoi, D.; Kozaki, M.; Okada, K.; Takui, T. *Synth. Met.* **2001**, *121*, 1840−1841.

(12) Michinobu, T.; Inui, J.; Nishide, H. *Org. Lett.* **2003**, *5*, 2165−2168.

(13) Fukuzaki, E.; Nishide, H. *J. Am. Chem. Soc.* **2006**, *128*, 996−1001.

(14) Mitani, M.; Mori, H.; Takano, Y.; Yamaki, D.; Yoshioka, Y.; Yamaguchi, K. *J. Chem. Phys.* **2000**, *113*, 4035−4051.

(15) Dietz, F.; Tyutyulkov, N. *Chem. Phys.* **2001**, *264*, 37−51.

(16) Huai, P.; Shimoi, Y.; Abe, S. *Phys. Rev. Lett.* **2003**, *90*, 207203.

(17) Hagiri, I.; Takahashi, N.; Takeda, K. *J. Phys. Chem. A* **2004**, *108*, 2290−2304.

(18) Borden, W. T.; Davidson, E. R. *J. Am. Chem. Soc.* **1977**, *99*, 4587−4594.

(19) Borden, W. T. *Mol. Cryst. Liq. Cryst.* **1993**, *232*, 195−218.

(20) Fang, S.; Lee, M.-S.; Hrovat, D. A.; Borden, W. T. *J. Am. Chem. Soc.* **1995**, *117*, 6727−6731.

(21) Aoki, Y.; Imamura, A. *Int. J. Quantum Chem.* **1999**, *74*, 491−502.

(22) Coulson, C. A.; Rushbrooke, G. S. *Proc. Camp. Phil. Soc.* **1940**, *36*, 193−200.

(23) Longuet-Higgins, H. C. *J. Chem. Phys.* **1950**, *18*, 283−291.

(24) Dewar, M. J. S. *J. Am. Chem. Soc.* **1952**, *74*, 3357−3363.

(25) *Gaussian 03, Revision C.02*; Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian, Inc.: Wallingford, CT, 2004.

(26) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.

(27) Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. *Chem. Phys. Lett.* **1989**, *157*, 200−206.

(28) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.

CT060008K

JCTC Journal of Chemical Theory and Computation

Ab Initio Simulations of the (101) Surfaces of Potassium Dihydrogenphosphate (KDP)

Damien J. Carter,[†,‡] Andrew L. Rohl,*[,†,§] and Julian D. Gale[†]

Nanochemistry Research Institute, Department of Applied Chemistry, Curtin University of Technology, GPO Box U1987, Perth, Western Australia 6845, Australia, School of Physics, The University of Sydney, Sydney, New South Wales, Australia, and iVEC, The hub of advanced computing in Western Australia, 26 Dick Perry Avenue, Technology Park, Kensington, Western Australia, Australia

Received January 9, 2006

Abstract: We have used density functional calculations to examine the (101) surfaces of KDP, under vacuum, nitrogen, and aqueous conditions, and these simulations are found to agree well with nanoscale experimental studies demonstrating that the density functional calculations are providing a good description of the surfaces of this complex inorganic salt.

Wolfgang Pauli is famously quoted as suggesting "God made solids ... but surfaces were the work of the Devil".[1] This statement reflects the immense difficulties faced by experimentalists when characterizing surfaces. As a result of this paucity of experimental data, computer simulations have been at the forefront of surface research. In the field of inorganic salts, one of the major triumphs was the prediction of extensive surface relaxation on the basal plane of alumina first by interatomic potential based calculations[2] and then by quantum mechanical simulations.[3−5] These predictions were confirmed by experiment very much later.[6,7] Similarly, good agreement has been obtained on the (001) surface of MgO. However, attention of the modeling community has turned to surfaces of complex salts, and yet to our knowledge there has been almost no direct validation of quantum mechanical simulations with experimental data for complex inorganic surfaces. This has been caused by both the lack of experimental data available, due to their nonconducting nature, and the complexity of the calculations. Recent advances in parallel efficiency and linear scaling algorithms now make it possible to simulate larger and more complex systems using quantum mechanics.[8] KDP is an example of a complex inorganic solid, and the structure of the (101) surface has been well characterized in a number of recent experimental studies.[9−12]

There has only been a small number of ab initio molecular modeling studies reported for KDP. The bulk structure of KDP has been examined in several recent studies[13−15] using density functional theory (DFT). The only reported theoretical study of the surfaces of KDP has been by Stack et al.[16] They used DFT to study the (100) surface of KDP, in particular relating to the adsorbing/detaching of growth units. Surface X-ray diffraction (SXRD) studies have examined the relaxation of atoms on the (101) surfaces of KDP, under aqueous conditions.[9,10] Recent SXRD studies on KDP surfaces in aqueous solution[11,12] suggest that there are several icelike ordered water layers on the surface. Thus we have used DFT calculations to examine relaxations of the (101) surface of KDP and compared our results to these recent experimental studies.

At room temperature, KDP forms a paraelectric phase in the tetragonal $I\bar{4}2d$ space group.[17] KDP undergoes a phase transition at 122 K to a ferroelectric (FE) phase in the orthorhombic *Fdd*2 space group.[17] For easier comparison to the paraelectric phase, we used the body-centered C_{2v}[19] I-setting space group as described by Baur.[18]

Density functional calculations on KDP were carried out using the SIESTA code.[19,20] The generalized gradient approximation has been employed for all calculations, using the functional of Perdew, Burke, and Ernzerhof.[21] Core electrons have been represented by norm-conserving pseudopotentials of the form proposed by Troullier and Martins.[22]

* Corresponding author phone: +618 92667317; fax: +618 92662300; e-mail: a.rohl@curtin.edu.au

† Curtin University of Technology

‡ The University of Sydney.

§ iVEC, 'The hub of advanced computing in Western Australia'.

10.1021/ct060012w CCC: $33.50
Published on Web 04/06/2006

Table 1. Surface Energies for the Two Surface Terminations of the (101) Surface of KDP

surface termination	surface energy (J/m^2)
$H_2PO_4^-$	0.80
K^+	0.54

The valence electron configurations used were $1s^1$, $2s^22p^3$, $2s^22p^4$, $3s^23p^3$, and $3p^64s^1$ for H, N, O, P, and K, respectively. A double-ζ basis set with polarization functions was used for all atoms except oxygen, which had a triple-ζ basis set with polarization functions. The Brillouin zone was sampled using a $3 \times 3 \times 3$ Monkhorst-Pack **k**-grid. The localized basis set in SIESTA consists of numerical atomic orbitals, which are radially confined to an extent that induces an energy shift in each orbital of 0.01 Ry. Hartree and exchange-correlation energies were evaluated on a uniform real space grid of points with a defined maximum kinetic energy of 200 Ry.

The bulk structure of KDP was relaxed at zero Kelvin, and then the (101) surface was generated using the GDIS program.[23] There are two possible ways to cut the (101) surface: one produces a cation (K^+) terminated surface and the other produces a dihydrogenphosphate ($H_2PO_4^-$) terminated surface. Surface slabs were created using a $1 \times 1 \times 4$ supercell, with a vacuum gap of approximately 26 Å. Relaxed surface configurations were generated by keeping the cell vectors fixed and only allowing the atom positions to relax, with calculations being run for both possible surface terminations. These calculations were then repeated using no vacuum gap. Surface energies were calculated for each surface termination using the following expression

$$E_{\text{surf}} = (E_{\text{slab}} - E_{\text{equiv bulk}})/2A$$

where E_{surf} is the surface energy, E_{slab} is the energy of the surface slab with a vacuum gap, $E_{\text{equiv bulk}}$ is the energy of the surface slab with no vacuum gap, and A is the surface area. E_{surf} is divided by two because each slab has two equivalent surfaces.

The use of $E_{\text{equiv bulk}}$, rather than an integer number of E_{bulk}, allows for maximum cancellation of errors relating to sampling and convergence for the different sized simulation cells. Two nitrogen molecules were added to the relaxed surface, and the system was relaxed to generate the nitrogenated surface configuration. A similar approach was taken to generate the hydrated surface configuration, by adding 12 water molecules onto the relaxed vacuum surfaces, and the system was then relaxed again.

The relaxed bulk structure for KDP matched closely to that from previous calculations, with the cell volume within 1.5% of the experimental value. Surface energies were calculated for each surface termination of the (101) surface, and the results are reported in Table 1. From the surface energy calculations, the cation (K^+) terminated (101) surface appears significantly more stable than the $H_2PO_4^-$ terminated surface. SXRD studies[9,10] of KDP have experimentally determined that the (101) surface is indeed cation terminated.

The reason the cation terminated surface is the most stable can be rationalized in terms of the relative surface area of the two surface terminations. In a crystal with isotropic interactions, the surface energy is proportional to the surface area of the exposed surface. The greater the surface area, the more energy that is required to create it. Using the GDIS program,[23] we calculated the molecular surface (the surface defined by the van der Waals spheres of the surface atoms) of the two surface terminations. The reported areas were 114 Å^2 for the cation terminated surface and 157 Å^2 for the anion terminated surface. Thus if the interactions within the KDP crystal were isotropic, we expect the anion terminated surface to have a surface energy 1.37 times higher than the cation terminated surface. The calculated surface energy of the anion terminated face is actually 1.48 times higher, suggesting that this simple model is a good approximation. In short, the cation terminated surface is more stable because it is much less corrugated than the anion terminated surface.

Table 2. Translational Relaxations of Species at the (101) Surface of KDP

	translational relaxations (Å)	
	K^+	$H_2PO_4^-$
Experimental Studies		
nitrogenated[a]	-0.13 ± 0.03	$+0.12 \pm 0.04$
aqueous solution[b]	$+0.10 \pm 0.05$	$+0.04 \pm 0.05$
Theoretical Calculations[c]		
vacuum	-0.25 (-0.27,-0.23)	$+0.12$ ($+0.18$,$+0.06$)
nitrogenated	-0.27 (-0.29,-0.25)	$+0.14$ ($+0.06$,$+0.22$)
hydrated	$+0.14$ ($+0.17$,$+0.11$)	$+0.08$ (0.00,$+0.16$)

[a] Reference 12. [b] References 9 and 10. [c] The average relaxation is reported for the theoretical calculations, with actual relaxations reported in parentheses.

The surface relaxations of the cation terminated surface of KDP was then measured. We used the approach taken in the experimental studies of de Vries et al.,[9−10] where a $H_2PO_4^-$ ion was treated essentially as a fixed group, and movements of the central atom in this group (phosphorus) were measured. The displacements of atoms (in the z direction) have been measured by comparing the positions in the relaxed and unrelaxed surfaces. These are reported in Table 2, with a negative (−) displacement referring to inward relaxation and a positive (+) displacement referring to an outward relaxation. Two relaxation values are reported for each particular surface species. This is because there are two symmetry independent $H_2PO_4^-$ and K^+ species on each surface, because of the hydrogen ordering in the orthorhombic phase at zero Kelvin. As the differences in interatomic distances between the orthorhombic and tetragonal phases are small, we believe that the average of our values should match closely with the experimental value, without recourse to the exceedingly compute expensive density functional molecular dynamics calculations that would be required to include the effect of temperature. The relaxed vacuum, nitrogenated, and hydrated (101) surfaces are illustrated in Figures 1−3, respectively.

The calculated vacuum and nitrogenated surface trends match the experimental trend found under nitrogen conditions, with the K^+ ion displacing toward the surface and the $H_2PO_4^-$ displacing away from it. De Vries et al.[10] have examined the surfaces of KDP under high vacuum, but the

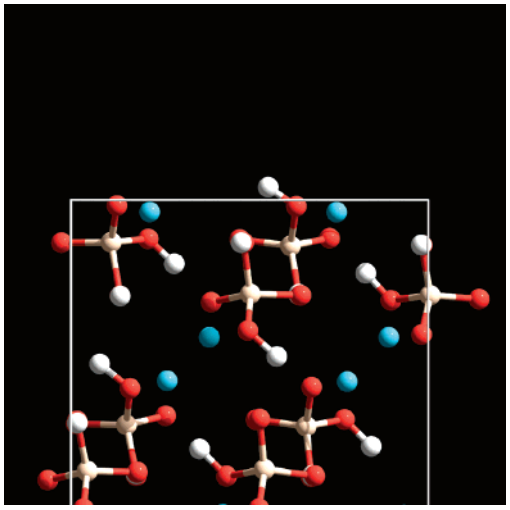


Figure 1. Relaxed vacuum (101) surface of KDP. Note: for this image and subsequent images, red = oxygen, orange = phosphorus, white = hydrogen, dark blue = nitrogen, and light blue = potassium.

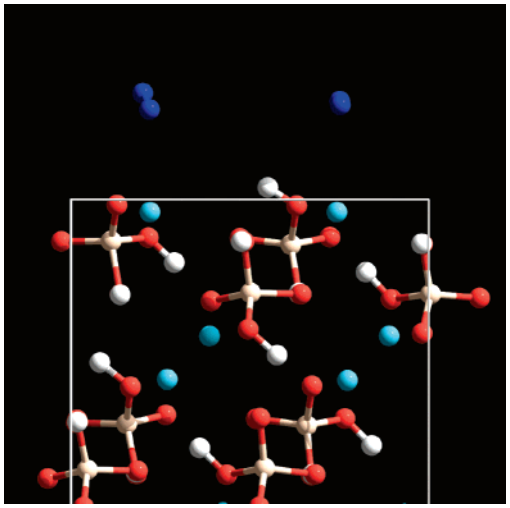


Figure 2. Relaxed nitrogenated (101) surface of KDP.

surface quality was found to change over time making it impossible to accurately determine the surface relaxation.

Comparing the experimental and nitrogenated surface relaxations, the magnitudes of relaxation compare favorably although vary slightly for K^+, with calculations producing an relaxation of −0.27 Å, while experiments reported a value of −0.13 Å.

We then performed calculations on a hydrated surface of KDP, using explicit water molecules, to produce a surface that would closely match the aqueous conditions reported in experiments. For the hydrated surface, the calculated surface trend matches that found experimentally under aqueous conditions, with both the K^+ and $H_2PO_4^-$ ions having an outward relaxation. The magnitudes of relaxation are also quite similar, although slightly larger in the calculations. When we examined the hydrated surface, we found that there were a number of strong hydrogen bonds formed between the $H_2PO_4^-$ ions and the surface water molecules, with typical hydrogen bond lengths of 1.5−1.9 Å. Density functional calculations are known to overestimate

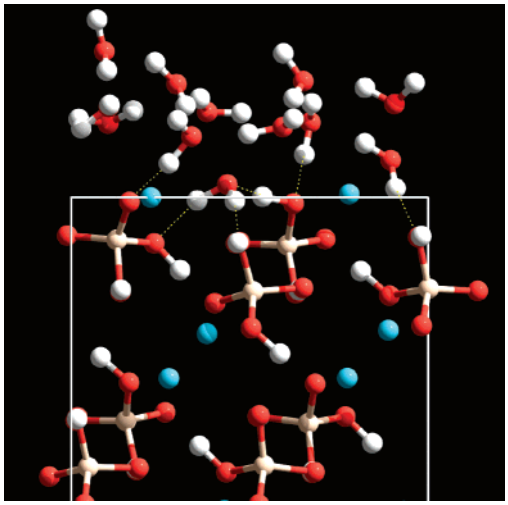


Figure 3. Relaxed hydrated (101) surface of KDP. Hydrogen bonds between water molecules and the surface are indicated with dashed lines.

the strength of hydrogen bonds,[24] so this could explain why the $H_2PO_4^-$ ions have a slighly larger outward relaxation in our calculations. There is potentially a range of configurations for water on the surface, and ideally molecular dynamics calculations should be run. However, for the present paper, we have examined two configurations and taken the lowest in energy (the energy differnce was small). These calculations were carried out at zero Kelvin, while the experimental values are measured at room temperature. This could also account for some of the differences in magnitudes.

Overall, the good match between calculated and experimental results suggests that density functional calculations can provide accurate simulations of the structures of complex inorganic surfaces.

References

(1) Toroczkai, Z.; Williams, E. D. Nanoscale Fluctuations at Solid Surfaces. *Phys. Today* **1999**, *December*, 24−28.

(2) Mackrodt, W. C.; Davey, R. J.; Black, S. N.; Docherty, R. The morphology of α-Al_2O_3 and α-Fe_2O_3: The importance of surface relaxation. *J. Cryst. Growth* **1987**, *80*, 441−446.

(3) Causa, M.; Dovesi, R.; Pisani, C.; Roetti, C. Ab initio characterization of the (0001) and (1010) crystal faces of α-alumina. *Surf. Sci.* **1989**, *215*, 259−271.

(4) Mackrodt, W. C. Classical and quantum simulation of the surface properties of α-Al_2O_3. *Philos. Trans. R. Soc. London, Ser. A* **1992**, *341*, 301−312.

(5) Manassidis, I.; De Vita, A.; Gillan, M. J. Structure of the (0001) surfaces of α-Al_2O_3 from first principles calculations. *Surf. Sci. Lett.* **1993**, *285*, L517−521.

(6) Ahn, J.; Rabalais, J. W. Composition and structure of the Al_2O_3 {0001}-(1 × 1) surface. *Surf. Sci.* **1997**, *388*, 121−131.

(7) Guénard, P.; Renaud, G.; Barbier, A.; Gautier-Soyer, M. Determination of the α-Al_2O_3(0001) Surface Relaxation and Termination by Measurements of Crystal Truncation Rods. *Surf. Rev. Lett.* **1998**, *5*, 321−324.

(8) Ordejon, P. Linear scaling ab-initio calculations in nanoscale materials with SIESTA. *Phys. Status Solidi B* **2000**, *217*, 335−356.

(9) De Vries, S. A.; Goedtkindt, P.; Bennett, S. L.; Huisman, W. J.; Zwanenburg, M. J.; Smilgies, D. M.; De Yoreo, J. J.; van Enckevort, W. J. P.; Bennema, P.; Vlieg, E. Surface Atomic Structure of KDP Crystals in Aqueous Solution: An Explanation of the Growth Shape. *Phys. Rev. Lett.* **1998**, *80*, 2229−2232.

(10) De Vries, S. A.; Goedtkindt, P.; Huisman, W. J.; Zwanenburg, M. J.; Feidenhans'l, R.; Bennett, S. L.; Smilgies, D. M.; Stierle, A.; De Yoreo, J. J.; Van Enckevort, W. J. P.; Bennema, P.; Vlieg, E. X-ray diffraction studies of potassium dihydrogen phosphate (KDP) crystal surfaces. *J. Cryst. Growth* **1999**, *205*, 202−214.

(11) Reedijk, M. F.; Arsic, J.; Hollander, F. F. A.; De Vries, S. A.; Vlieg, E. Liquid Order at the Interface of KDP Crystals with Water: Evidence for Icelike Layers. *Phys. Rev. Lett.* **2003**, *90*, 066103.

(12) Reedijk, M. F. Ordering at solid−liquid interfaces. Ph.D. Thesis, University of Nijmegen, The Netherlands, 2003.

(13) Koval, S.; Kohanoff, J.; Migoni, R. L.; Bussmann-Holder, A. Interplay between proton ordering and ferroelectric polarization in H-bonded KDP-type crystals. *Comput. Mater. Sci.* **2001**, *22*, 87−93.

(14) Zhang, Q.; Chen, F.; Kioussis, N.; Demos, S. G.; Radousky, H. B. *Ab initio* study of the electronic and structural properties of the ferroelectric transition in KH_2PO_4. *Phys. Rev. B* **2001**, *65*, 024128.

(15) Lin, Z.; Wang, Z.; Chen, C.; Lee, M. H. Mechanism of linear and nonlinear optical effects of KDP and urea crystals. *J. Chem. Phys.* **2003**, *118*, 2349−2356.

(16) Stack, A. G.; Rustad, J. R.; De Yoreo, J. J.; Land, T. A.; Casey, W. H. The growth morphology of the {100} surface of KDP (Archerite) on the molecular scale. *J. Phys. Chem. B* **2004**, *108*, 18284−18290.

(17) Nelmes, R. J.; Tun, Z.; Kuhs, W. F. A compilation of accurate structural parameters for KDP and DKDP, and a users' guide to their crystal structure. *Ferroelectrics* **1987**, *71*, 125−141.

(18) Baur, W. H. Reconstruction of local atomic environments in the disordered hydrogen-bonded crystal structures of paraelectric ammonium dihydrogen phosphate and potassium dihydrogen phosphate. *Acta Crystallogr., Sect. B* **1973**, *29*, 2726−2731.

(19) Ordejon, P.; Artacho, E.; Soler, J. M. Self-consistent order-N density-functional calculations for very large systems. *Phys. Rev. B* **1996**, *53*, R10441.

(20) Soler, J. M.; Artacho, E.; Gale, J. D.; Garcia, A.; Junquera, J.; Ordejon, P.; Sanchez-Portal, D. The SIESTA method for *ab initio* order-*N* materials simulation. *J. Phys.: Condens. Matter* **2002**, *14*, 2745−2779.

(21) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.

(22) Troullier, N.; Martins, J. L. Efficient Pseudopotentials for Plane-Wave Calculations. *Phys. Rev. B* **1991**, *43*, 1993−2006.

(23) Fleming, S. D.; Rohl, A. L. GDIS: A Visualization Program for Molecular and Periodic Systems. *Z. Kristallogr.* **2005**, *220*, 580−584.

(24) Koch, W.; Holthausen, M. C. *A Chemist's Guide to Density Functional Theory*, 2nd ed.; Wiley-VCH Verlag GmbH: Weinheim, 2001; p 231.

CT060012W

JCTC Journal of Chemical Theory and Computation

Theoretical DFT Study on the Interaction of NO and Br_2 with the Pt(111) Surface

Nurbosyn U. Zhanpeisov* and Hiroshi Fukumura

Department of Chemistry, Graduate School of Science, Tohoku University, Sendai 980-8578, Japan

Received December 7, 2005

Abstract: Density functional calculations were performed at the B3LYP level using combined basis sets for the NO and bromine interactions with the Pt(111) surface mimicked by the two-layer Pt_{10} cluster model. It explains well an attractive bonding interaction not only for bromine and Pt(111) but also for all three adsorption modes of NO on the Pt(111) surface. In accordance with the experimental observations, the calculations predict that the first peak in the IR spectra appears at around 1515 cm^{-1} at the initial stage of low NO coverage, while it would shift to 1707 cm^{-1} at high NO coverage. The bonding of NO on the 3-fold hollow fcc and hcp sites of Pt(111) proceeds via predominant back-donation interactions, while for the on-top adsorption, both the donation and back-donation interactions become equally important. Energetic criteria show also that the STM tip (made from Pt and Ir alloys) immersed into a bromine solution may contain only dissociated bromine atoms that bind strongly with the surface Pt atoms. As a result, the ν_{BrBr} stretching vibration mode for the bromine molecule may not be seen in the IR spectra because of its dissociation into adsorbed atoms. This leads to an appearance of a blue shifted band centered at ca. 202 cm^{-1}.

Introduction

Interaction of NO with metal surfaces and, in particular, with Pt and Pd has been the subject of many experimental and theoretical investigations because of considerable interest in the field of heterogeneous catalysis and various catalytic processes.[1−12] One of the important examples is a reduction of toxic NO_x from flue gases emitted by stationary and mobile sources, automobiles, coal-fired power plants, electric power generators, nitric acid factories, etc.[6,13,14] Since NO_x are the most toxic environmental pollutants, the control of NO emissions is a major challenge and the most highly desirable goal. Despite general agreement on a molecularly adsorption feature of NO on a Pt surface at low temperatures, there are still controversial discussions on the nature of target interactions and stable adsorption sites on the Pt(111) surface. Recently, Nakatsuji et al.[6] have shown that only the dipped admolecule model (DAM) can properly describe both the energetics and the vibrational modes for the NO adsorbed on the Pt(111) surface, while they have ruled out the adequacy in applying the so-called cluster model approach to mimic such kinds of interactions. At the same time, Trout et al.[7] have claimed that the periodic slab calculations can be also well suited in describing these interactions including such features as the order of stability for the adsorption complexes and their adsorption energies when considering the different adsorption modes realized on Pt(111). However, these latter two quantities, i.e., the relative stability order and adsorption energies, differ strongly from those predicted by Nakatsuji et al.[6]

Another aspect related to the Pt surface is studies on the adsorption of bromine or iodine species at different Pt facets that has been widely investigated under ultrahigh vacuum, at normal atmospheric pressure conditions as well as from aqueous solutions.[15−18] At temperatures as low as 25 K, bromine is dissociatively adsorbed on Pt(111). The strong interaction between bromine radicals and the Pt surface accounts for Br_2 dissociative adsorption. These adsorbed bromine atoms are supposed to be only transiently mobile,

* Corresponding author phone: 81-22-795-6568; e-mail: nurbosyn@orgphys.chem.tohoku.ac.jp.

10.1021/ct050308m CCC: $33.50
Published on Web 04/07/2006

while weakly bound Br_2 molecules would be adsorbed near or preexisting Br atoms that would act as preferential sites for Br_2 dissociation. Under atmospheric pressure conditions, the exposure of clean Pt(111) surface to bromine vapors results in the formation of about half of the monolayer coverage of bromine atoms while being neutral in character for the adsorbed bromine atoms. Earlier Solomon et al.[19] studied iodine molecule adsorption on steps of the Pt surface under vacuum and atmospheric pressure conditions using experimental LEED, AES, and thermal desorption spectroscopies. In contrast to smooth Pt(111), they found multiple phase domains of (3×3) or $(\sqrt{3}\times\sqrt{3})\times R30°$ and no special preferences for iodine adsorption on these step positions. Ertl and co-workers[20] have also shown that bromine adlayers formed by flame annealing of platinum single-crystal surfaces and quenching in bromine vapor, studied by STM and cyclovoltammetry, form protective layers which, similar to iodine, prevent contamination of the surface and lead to well-defined surfaces. On Pt(111), a (4×4) structure is observed, while on Pt(100), a disordered structure is observed. Both structures are compatible with densely packed adlayers of bromine, while all three low index surfaces including Pt-(110) yield air-stable bromine adlayers.

Since the discovery of powerful STM techniques in 1982,[21] to attain their main goals, scientists are driven not only to understand how atoms locate at surface interfaces and how they behave with the surface electrons but also to understand deeply the STM tip enhanced spectroscopic properties of adsorbed molecules and highly ordered structures as well as polymerization processes on different substrates.[22−24] Using this technique one may effectively initiate chemical reactions of interest, manipulate the oxidation or reduction processes at distinct active surface sites, etc. This methodology is also important in creating new microchips for optoelectronic devices, computers, etc. Also, because of considerable interest in the modification of STM tips through contacts with bromine or iodine solutions, theoretical considerations may further supplement a deep understanding of those unique properties at target surfaces and their interfaces. Moreover, they can be expected to be equally complimentary for the experimental investigations. The question that arises as to what extent this halogen modification of the STM tip can lead is of primary interest because the tip enhanced IR or Raman spectra is the origin to obtain additional information not only on surface structures but also on the quality of surfaces as well as on their environment.

In this paper, we report on the theoretical results obtained for the interaction of the NO and bromine with Pt(111) using the cluster approach. In contrast to the cited above most recent publications,[6,7] we will show that the cluster approach can be equally applied to describe reasonably the nature of the target interactions as well as to predict theoretical spectroscopic properties of adsorbed molecules. In particular, the IR spectra of NO acted as a probe molecule, and bromine on Pt(111) can be well explained.

Method and Models

Density functional theory (DFT) calculations were performed using the Gaussian03 program packages.[25] Geometry optimizations were carried out with the use of Becke's three-parameter hybrid method with the Lee, Yang, and Parr (B3LYP) gradient-corrected correlation functional[26] and the combined standard basis sets. The latter basis sets include the standard 6-31G* ones for the N and O atoms and Lanl2dz basis sets for all the Pt as well as Br atoms. As a result, they would be denoted as the B3LYP/(Lanl2dz + 6-31G*) level for the NO on the Pt(111) and B3LYP/Lanl2dz level for bromine on Pt(111), respectively. The Pt(111) surface was modeled by the two-layer Pt_{10} cluster model shown in Figure 1. The geometry optimizations were carried out taking into account symmetry restrictions and followed by estimation of vibration fundamentals via performing harmonic frequency calculations.

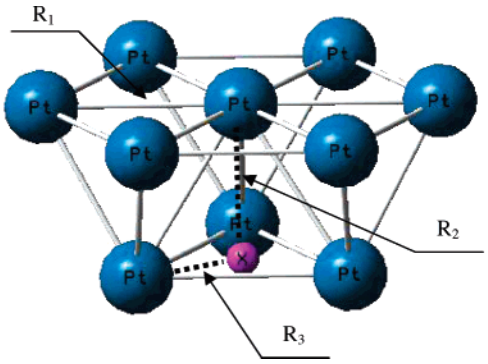


Figure 1. The two-layer Pt_{10} cluster model mimicking the Pt-(111) surface and the definition of the three independent variables.

Results and Discussion

The ground electronic state of the isolated Pt atom is $^3D_3$ with the electronic configuration of [Xe]: $4f^{14}5d^96s^1$. Because of this triplet ground state, it is not evident how the overall spin state would be for the cluster consisting of 10 Pt atoms. Trout et al.[7] have excluded in their calculations magnetic states and have performed only nonmagnetic slab calculations for the combined systems of the NO adsorbed on Pt because the chemisorbed NO is found to be nonspin-polarized according to previous studies.[2] However, this is not necessarily true either for the isolated Pt cluster with a finite size or for the combined adsorption complexes of NO considered here. To clarify this issue, we have first determined the optimal spin state for the isolated two-layer Pt_{10} cluster model. Table 1 lists the results of these calculations. Note that for each spin state the three independent structural parameters shown in Figure 1 were optimized. As is clear the closed-shell singlet spin state is highly unfavorable for the Pt_{10} cluster by energy (about 30 kcal/mol) than the optimal spin state one containing eight unpaired electrons with a multiplicity equal to 9. The spin densities for these 8 electrons are distributed over all the Pt atoms while being less localized on the central Pt site. Its spin equals only to 0.276. In addition, this state is free from spin contaminants because its $/S^2/$ value after spin annihilation matches exactly with the theoretical value of 20.

Moreover, the central Pt atom is the only site with a positive charge of $0.274e^-$ after conventional Mulliken population analysis, while the rest of the other Pt atoms in

Table 1. Multiplicity (M), Relative Energy (E_{rel}, kcal/mol), Value of $/S^2/$ before (after) Spin Annihilation, and Geometry (R_1, R_2, R_3, All in Å) of the Bare Pt_{10} Cluster

M	E_{rel}[a]	$/S^2/$	R_1	R_2	R_3
1	29.7		2.707	2.201	1.621
3	7.3	4.03 (5.44)	2.709	2.218	1.645
5	5.0	7.03 (6.35)	2.712	2.218	1.641
7	1.7	12.51 (12.04)	2.725	2.187	1.623
9	0	20.10 (20.00)	2.720	2.203	1.682
11	14.7	30.10 (30.00)	2.723	2.205	1.688

[a] Total energy of the bare cluster with multiplicity equal to nine is taken as an internal reference. The positive sign corresponds to lower stability.

the cluster are all slightly negatively charged. Evidently, the obtained high-spin state for the Pt_{10} cluster should be taken into account when considering the adsorption complexes of NO and bromine because they would strongly affect their adsorption energetics.

Thus, we have first examined the adsorption complexes of the NO molecule on the Pt_{10} cluster. Because the isolated NO molecule has one unpaired electron in its π^*-MO and radical in nature, the interaction with the high-spin Pt_{10} cluster leads to effectively coupling of these electrons. Therefore, the bonding between the NO and the Pt_{10} cluster can be governed by the donation and/or back-donation interactions. Nakatsuji et al.[6] have pointed out that for this target interaction the back-donation is more important and dominant and that the interaction can be only explained within the DAM model, while the cluster approach will fail to explain this bonding phenomenon. However, here we have some doubt about this conclusion. Below we will show that both donation and back-donation interactions are equally important and depend on the distinct adsorption modes. Moreover, the cluster model can be also considered as a complementary tool to reliably explain both the energetics of the NO adsorption and spectroscopic properties of the adsorbed NO on Pt(111).

Table 2 lists the adsorption energies, molecular properties, and main geometric parameters for the NO adsorption at the most common three active sites of Pt(111) mimicked by the Pt_{10} cluster model, while the stick and ball optimized geometries of adsorption complexes are shown in Figure 2. These adsorption modes correspond to the bonding of NO on the top site as well as on both the 3-fold hollow fcc and hcp sites. The latter two sites differ only by the absence or presence of the Pt atom of the second layer just below the center of the top 3-fold hollow site. Note that this applied Pt_{10} cluster model is the minimal one that contains all the sets of active sites on Pt(111). However, one should keep in mind that the use of such a kind of finite cluster model has well-documented disadvantages in describing extended surfaces.[10] Nevertheless, the applied Pt_{10} cluster model is well suited to the broad survey undertaken here, being computationally much less demanding than those based on the periodic slab approach.

As in the case of the isolated cluster model, for each adsorption mode considered here we have selectively determined the optimal spin state for the combined system. Accordingly, the optimal spin state for each adsorption complex together with their $/S^2/$ values are shown in Table 2. An analysis of these data shows the following peculiarities:

(i) The applied cluster model can explain the attractive bonding interaction between the NO and the Pt(111) surface

Table 2. Adsorption Energy (E_{ads}, kcal/mol), Multiplicity (M), Value of $/S^2/$ before (after) Spin Annihilation, Charges (Q_{NO} and Q_{Pt}, e^-) on NO and the Central Pt Atom, Spin Density (ρ_{NO}, e^-) on NO, Bond Lengths (R_{Pt-N} and R_{N-O}, Å), NO Stretching Frequency (ν, cm^{-1}), and HOMO and HOMO−LUMO Gap (Δ, au) for the Adsorption Complexes of NO on Pt(111) Mimicked by the Two-Layer Pt_{10} Cluster Model as Calculated at the B3LYP/(Lanl2dz + 6-31G*) Level

	on-top	fcc	hcp
E_{ads}[a]	19.1	14.8	14.1
	18.2	14.2	13.6
	19.3	16.8	16.1
M	8	6	6
$/S^2/$	15.82 (15.75)	8.92 (8.76)	8.86 (8.75)
Q_{NO}	+0.03	−0.46	−0.44
Q_{Pt}	−0.14	+0.16	+0.18
ρ_{NO}	+0.07	−0.06	−0.07
R_{Pt-N}	2.052	2.104 (1.999)[b]	2.104 (2.001)[b]
R_{N-O}[c]	1.167	1.205	1.204
ν_{NO}[d]	1707	1515	1530
HOMO	−0.2073	−0.2059	−0.2049
Δ	0.0654	0.0530	0.0585

[a] The values of adsorption energies at the second and third rows were estimated at the B3LYP/(Lanl2dz + 6-311++G**)//B3LYP/(Lanl2dz + 6-31G*) and B3LYP/(Lanl2dz + cc-pVTZ)//B3LYP/(Lanl2dz + 6-31G*) levels of theory, respectively. [b] The other two equal Pt−N bond distances formed by the N atom and the two nearest edge Pt atoms are shown in parentheses. [c] The N−O bond distance in the gas phase is calculated to be equal to 1.159 Å. [d] The ν_{NO} stretching frequency is scaled down by using a scale factor of 0.97.[30]

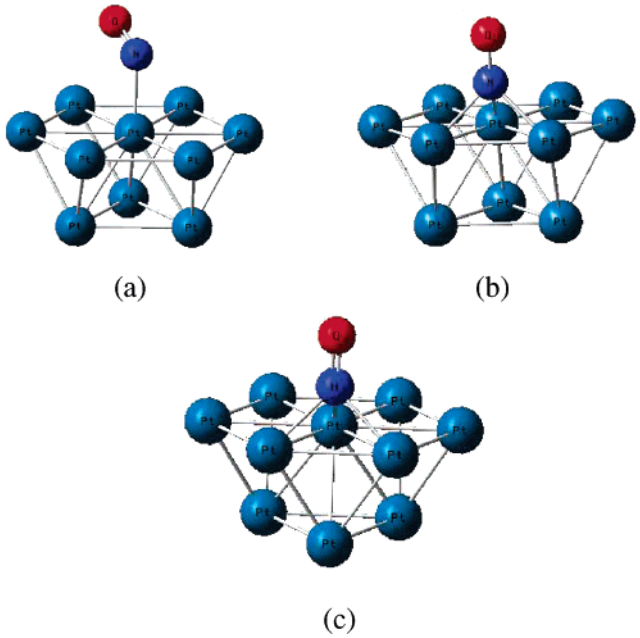


Figure 2. The optimized adsorption complexes of NO on the Pt(111) surface presented in stick and ball structures: (a) on-top, (b) 3-fold hollow fcc, and (c) 3-fold hollow hcp adsorption complexes.

for all three adsorption modes. This finding is in contrast with the conclusions of ref 6, where highly repulsive interactions have been observed for all the adsorption modes of NO on a similar Pt_{10} cluster model. For example, the adsorption of NO on the on-top, 3-fold hollow fcc and hcp sites was found to be highly unstable for 32.5, 35.3, and 36.9 kcal/mol, respectively.[6] Two factors, i.e., basis set inconsistency and low spin state for the Pt_{10} cluster applied, might be the reason for the latter repulsive interactions found in ref 6. Moreover, we have also found that the bonding mechanism is not one and the same for all three adsorption modes. For the on-top adsorption mode, both the donation and back-donation is equally important, with a slight preference only for the donation of the electron density from the NO molecule to the Pt(111) surface. The net electron density transfer amounts to $0.03e^-$. For the other two 3-fold hollow fcc and hcp sites, the back-donation is strictly predominant as can be seen from the Mulliken population analysis (see Table 2). As a result, the latter back-donation lowers the overall spin state for these adsorption modes. However, in accordance with the results of other previous studies,[2,7] the adsorbed NO molecule is nonspin-polarized because of negligibly small spin density on the N and O atoms for all three adsorption complexes.

We have also performed additional calculations for population analysis and atomic charge assignments that are based on natural bond orbital (NBO) analysis and charges determined by the fit to the electrostatic potential at points selected according to the CHelpG scheme.[25] The latter CHelpG charges support qualitatively the above findings; however, they are highly dependent on the radii of the centers used in fitting those potentials. NBO analysis shows that the d-population on the central Pt atom is smaller for ca. 0.10 e^-, while its 6p-AO earns ca. $0.13e^-$ as compared to other Pt sites in the case of the on-top adsorption mode of NO. The electronic configurations for the N and O atoms become $[He]2s^{1.64}2p^{3.16}$ and $[He]2s^{1.72}2p^{4.43}$, respectively. Conversely, the adsorption of NO on the 3-fold hollow fcc site decreases the d-population on the three Pt atoms of interest by 0.17, 0.14, and 0.14 e-, while electronic configurations of the N and O atoms look like $[He]2s^{1.44}2p^{3.59}$ and $[He]2s^{1.72}2p^{4.50}$, respectively. As is clear, a strong population of 2p-AO of the N site of NO is observed. A similar picture has been also observed for the adsorption of NO on the 3-fold hollow hcp site. Thus, the back-donation interaction favors the NO binding at high-coordination fcc and hcp sites, while the donation term is consistently more favorable for the on-top coordination.

(ii) The adsorption energies estimated in a conventional way lie in close proximity to each other with a slight preference (about 5 kcal/mol) for the on-top adsorption mode. The least stable mode is the adsorption on the 3-fold hollow hcp site, which is about 0.7 kcal/mol less strong than that of the 3-fold hollow fcc site. Thus, the order of stability is expressed as on-top $>$ fcc $\approx$ hcp and correlates with the highest occupied molecular orbital (HOMO) level. Additional single-point calculations were performed at the same B3LYP level, but using more larger and extended basis sets did not alter the above obtained stability order (see the second and third rows for the adsorption energy, Table 2). Also, the adsorption energies estimated at the B3LYP/(Lanl2dz + 6-31G*) level lie within 1−2 kcal/mol as compared to those estimated at higher levels of the theory for each selected adsorption mode of the NO on Pt(111). However, this obtained stability order of adsorption complexes is in contrast with the results of other theoretical studies[6,7] but supports well the results of an experimental study obtained by Matsumoto et al.[27] According to the latter experimental interpretation, NO occupies the three sites sequentially depending on the NO coverage. First NO adsorbs at the 3-fold hollow fcc site, next at the top site, and last at the 3-fold hollow hcp site.

The above findings of ours need additional remarks. A traditional picture proposed by Gland and Sexton[1b] for the interaction of NO with Pt(111) points out that at low coverages bridge sites are initially occupied by NO molecules. This erroneous ascription has been, however, later corrected by other studies[1c,3,27] so that it turned out to be 3-fold hollow fcc sites. Gland and Sexton[1b] claimed also that the bridge NO species leaves the bridge sites and moves to the on-top sites as the NO coverage increases that is based on the disappearance of the vibrational band at 1490 cm^{-1} at high NO coverage in the reflection absorption infrared spectroscopy measurements (RAIRS). However this mode remains in the high-resolution electron energy loss spectroscopy measurements (HREELS) even at the saturation coverage. Therefore, the site transfer model proposed by Sexton and Gland[1b] is insufficient and is not applicable to explain the discrepancy between the results of RAIRS and HREELS. Moreover, the previous theoretical calculations[3,6,7] have shown also that the 3-fold hollow fcc site is the most preferable adsorption site, so that there are no reasons for the move of the adsorption complex from the 3-fold hollow fcc site to on-top sites, at least at coverages lower than 0.25 multilayers. In addition, why the on-top species exists is irrespective of the relative instability as compared to the above 3-fold hollow fcc and hcp site species. The latter picture is even highly pronounced in ref 6, where the relative stability of the adsorption complexes on the 3-fold hollow fcc and the hcp sites are ca. 3.5 times larger than that of the on-top site that even rules out the formation of the latter adsorption complexes. Thus, our results obtained above are in line with the interpretation proposed by Matsumoto et al.,[27] and the move from the 3-fold hollow fcc site to the on-top site is driven by a slight preference in the adsorption energy for the latter on-top sites.

(iii) The adsorption of NO leads to the elongation of the N−O bond length as compared to the isolated gas-phase value. This increase in the bond distance is larger for the adsorption of NO on the 3-fold hollow fcc and hcp sites being about 0.05 Å. Note also that the two latter adsorption complexes of NO on 3-fold hollow sites result in the formation of three nonequivalent Pt−N bonds because of the shift of the center of mass of NO from the center of the respective hollow sites to the nearest edge site. The nearest distance from the metal surface amount to 1.291 and 1.293 Å for the adsorption complexes on the 3-fold hollow fcc and hcp sites, respectively. Note that the most stable on-top

adsorption mode of NO observed in the present study results in the formation of a tilted structure, the tilt angle being equal to 59.5°. This finding is in line with the results of other investigations on the adsorption of NO on Pt(111)[3,6,7] as well as on Pd(111).[28] Also, the geometry parameters that characterize these three adsorption complexes of NO (i.e., the N−O and Pt−N bond distances, see Table 2) are in good agreement with the observed data of the NEXAFS[5] and LEED[29] experiments.

(iv) The stretching frequency of NO shows a red shift during the adsorption as compared to that of the isolated gas-phase value. For the on-top adsorption complex, it appears at 1707 cm^{-1}, while for the 3-fold hollow fcc and hcp sites they are centered at 1515 and 1530 cm^{-1}, respectively. Because the adsorption energies differ slightly as well as these three adsorption sites equally available, we may conclude that at a low NO coverage, the adsorption takes place on 3-fold hollow sites, while the increase in the NO coverage will result on adsorption on the top sites. Accordingly, at the initial stage of a low NO coverage, the first peak would appear at around 1515 cm^{-1}, while it would shift to 1707 cm^{-1} at high NO coverage. These findings are in complete accordance with the observed experimental data (see, for example, ref 6).

Following the symmetry of the Pt_{10} cluster model, we have considered also some representative structures on extended models. The first one is obtained from Pt_{10} by an increase in Pt atoms at the top layer only (the Pt_{22} cluster model containing 19 and 3 atoms in two layers, respectively), while the second cluster corresponds to the Pt_{31} one containing atoms from the next shell in both layers (19 and 12 atoms at each layers, respectively, see Figure 3). Because of the size of these clusters, single point calculations at the similar B3LYP/Lanl2dz level were carried out to estimate an optimal spin state for these two cluster models at fixed geometries found for the Pt_{10} cluster (at its optimal high spin state). It was found that the energetically most preferable states are those containing 16 and 20 unpaired electrons for the Pt_{22} and Pt_{31} cluster models, respectively. The on-top adsorption mode of NO on the Pt_{22} cluster results in close geometrical parameters and adsorption energy as those found for the case of the minimal Pt_{10} cluster model. For example, the Pt−N and N−O bond distances are equal to 2.000 and 1.159 Å, respectively, while its adsorption energy amounts to 24.6 kcal/mol, an increase of 5.5 kcal/mol as compared to that found for the on-top adsorption on the Pt_{10} cluster model.

Let us move to consider the interaction of the bromine with the Pt(111) surface. As we have noted before, this issue is important, at least in relation to the STM tip modification and in understanding the nature of the target interaction. Note that in STM experiments the most common tip is made from a tungsten polycrystalline wire, while others are based on transition metals usually made from Pt and iridium alloys. To the first approximation, one may consider just a very dense Pt(111) structure as the tip and clarify its interaction with the bromine. Based on the above results obtained for NO interactions with the same Pt_{10} cluster model, we have considered adsorption modes for the bromine molecule and the atom on the top position of Pt(111) as well as the dissociative adsorption form of Br_2. Table 3 lists the results of these calculations, while Figure 4 shows their optimized

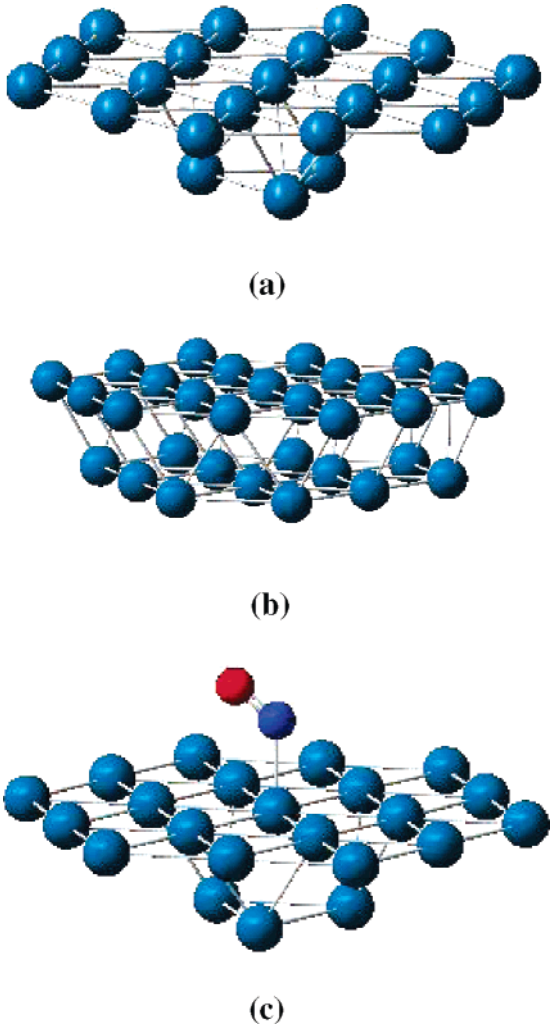


Figure 3. The extended cluster models of Pt_{22} (a) and Pt_{31} (b) as well as the optimized structure mimicking on-top adsorption mode of NO on the Pt_{22} cluster (c).

Table 3. Adsorption Energy (E_{ads}, kcal/mol), Multiplicity (M), Value of $/S^2/$ before (after) Spin Annihilation, Charges (Q_{Br} and Q_{Pt}, e^-) on the Br and Central Pt Atoms, Spin Density (ρ_{Br}, e^-) on the Adsorbate, Bond Lengths (R_{Pt-Br} and R_{Br-Br}, Both in Å), and HOMO (au) for the On-Top Adsorption Complexes of Br_2 and Br as Well as Dissociative Adsorption of Br_2 on Pt(111) Mimicked by the Two-Layer Pt_{10} Cluster Model

	on-top Br_2	on-top Br	dissociative Br_2
E_{ads}	20.4	42.5	72.9
M	9	8	9
$/S^2/$	20.10 (20.00)	15.81 (15.75)	20.07 (20.00)
Q_{Br}	+0.03 (−0.27)[a]	+0.04	+0.05 (+0.04)[b]
Q_{Pt}	+0.08	−0.24	−0.24
ρ_{Br}	+0.02 (+0.03)[a]	+0.03	+0.04 (+0.12)[b]
R_{Pt-Br}	2.720	2.570	2.568 (2.663)[b]
R_{Br-Br}	2.749		
HOMO	−0.2226	−0.2167	−0.2240

[a] The value in parentheses is for the end bromine atom. [b] The value in parentheses is for the Br atom that bonds with the nearest two edge Pt atoms.

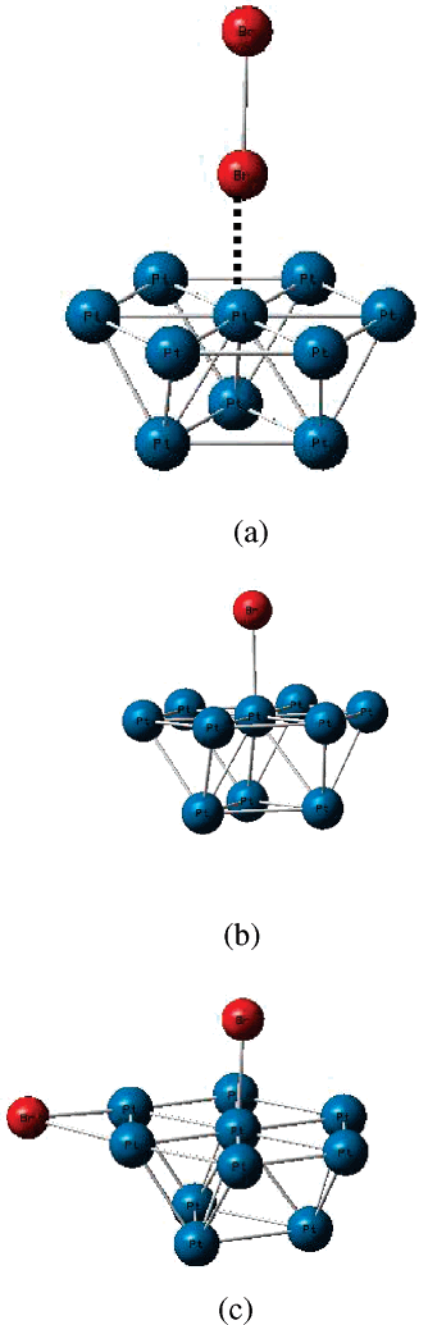


Figure 4. The optimized adsorption complexes of (a) molecular Br_2 and (b) atomic Br on the top position and (c) dissociative adsorption of Br_2 on-top as well as the two edge Pt atoms of the Pt(111) surface.

geometries. An analysis of these data shows that the molecular adsorption of Br_2 on the top position of Pt(111) is twice as weak as that of the atomic Br adsorption. Moreover, the molecular adsorption is much less profitable as compared to its dissociative adsorption form. In the latter dissociative adsorption form, the two Br atoms are initially attached to the on-top and to one of the edge Pt atoms. However, the geometry optimization leads to the structure shown in Figure 4c where one of the Br atoms makes in-plane symmetric two bonds with the two nearest edge Pt atoms. Its overall adsorption energy amounts to 72.9 kcal/mol as compared to the isolated noninteracting Pt_{10} cluster

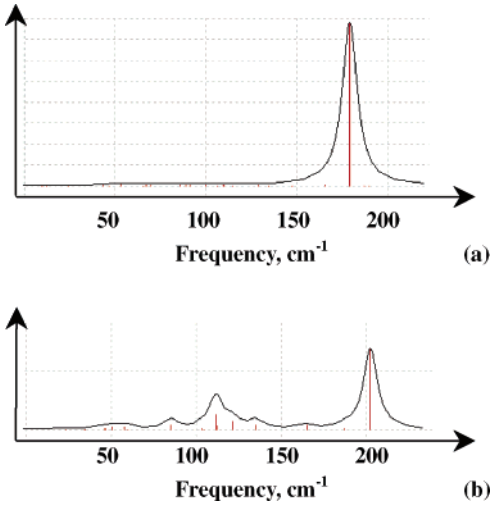


Figure 5. Theoretical IR spectra of bromine adsorbed on the Pt(111) surface: (a) molecular Br_2 adsorption and (b) atomic Br adsorption.

and the Br_2 molecule. Interesting to note is that both molecular and dissociative adsorption of Br_2 did not alter the optimal spin state of the bare Pt_{10} cluster, while an atomic Br adsorption lowers it, as is found before in the case of the on-top adsorption of NO. Based on these energetic criteria we conclude that the STM tip immersed into a bromine solution may contain only dissociated bromine atoms that are strongly bound with the surface Pt atoms. This finding is in complete agreement with the results of experimental investigations.[15,17] As a result, one may not see the ν_{BrBr} stretching vibration of the adsorbed Br_2 molecule (see Figure 5a) because it will be blue shifted due to the appearance of the ν_{PtBr} stretching vibration centered at around 202 cm^{-1} (see Figure 5b). Note that the theoretical IR spectrum has been simulated by the Lorentzian broadening of the calculated discrete vibrations with a half width of 20 cm^{-1}.

Summary and Conclusions

The obtained results of the density functional calculations at the B3LYP level using the above combined basis sets can be concluded as follows:

1. The adsorption energies of NO on different sites of the Pt(111) surface lie in close proximity to each other, while the relative stability order of adsorption complexes can be expressed as on-top > fcc ≈ hcp that correlates directly with the HOMO level. At low NO coverage, the first peak is predicted to appear at around 1515 cm^{-1}, while it would shift to 1707 cm^{-1} at high NO coverage in accordance with experimental observations.

2. An atomic Br as well as the dissociative adsorption of Br_2 proceed highly favorably as compared to the molecular adsorption of Br_2 on the top position of the Pt(111) surface that is the least favorable one. Both the molecular and dissociative adsorption of Br_2 did not alter the optimal spin state of the bare Pt_{10} cluster, while an atomic Br adsorption lowers it for one.

3. The STM tip immersed into a bromine solution may contain only dissociated bromine atoms that strongly bind with the surface Pt atoms. As a result, the ν_{BrBr} stretching

vibration of the adsorbed Br_2 molecule cannot be seen because of a blue shift and are centered at ca. 202 cm^{-1}.

Acknowledgment. We gratefully acknowledge supercomputing resources provided by the Information Synergy Center of Tohoku University. The present work was supported by a Grant-in-Aid from the Ministry of Education, Science, and Culture of Japan (16072203).

References

(1) (a) Schmatloch, V.; Kruse, N. *Surf. Sci.* **1992**, *270*, 488−494. (b) Gland, J. L.; Sexton, B. A. *Surf. Sci.* **1980**, *94*, 355−368. (c) Hayden, B. F. *Surf. Sci*. **1983**, *131*, 419−432.

(2) (a) Hammer, B. J. *J. Catal.* **2001**, *199*, 171−176. (b) Hammer, B. *Phys. Rev. Lett*. **1999**, *83*, 3681−3684. (c) Hammer, B.; Norskov, J. K. *Phys. Rev. Lett*. **1997**, *79*, 4441−4444.

(3) Aizawa, H.; Morikawa, Y.; Tsuneyuki, S.; Fukutani, K.; Ohno, T. *Surf. Sci.* **2002**, *514*, 394−403.

(4) Burch, R.; Daniells, S. T.; Hu, P. *J. Chem. Phys*. **2002**, *117*, 2902−2908.

(5) Esch, F.; Greber, T.; Kennou, S.; Siokou, A.; Ladas, S.; Imbihl, R. *Catal. Lett*. **1996**, *38*, 165−170.

(6) Nakatsuji, H.; Matsumune, N.; Kuramoto, K. *J. Chem. Theory Comput*. **2005**, *1*, 239−247.

(7) Tang, H.; Trout, B. L. *J. Phys. Chem. B* **2005**, *109*, 17630−17634.

(8) Brown, W. A.; King D. A. *J. Phys. Chem. B* **2000**, *104*, 2578−2595.

(9) Ge, Q.; King, D. A. *Chem. Phys. Lett*. **1998**, *285*, 15−20.

(10) Koper, M. T. M.; van Santen, R. A.; Wasilevski, S. A.; Weaver, M. J. *J. Chem. Phys*. **2000**, *113*, 4392−4407.

(11) Shelef, M.; Graham, G. W. *Catal. Rev. Sci.. Eng.* **1994**, *36*, 433−457.

(12) Taylor, K. C. *Catal. Rev. Sci. Eng.* **1993**, *35*, 457−481.

(13) Shelef, M. *Chem. Rev.* **1995**, *95*, 209−225.

(14) Zhanpeisov, N. U.; Higashimoto, S.; Anpo, M. *Int. J. Quantum Chem*. **2001**, *84*, 677−685.

(15) Ferro, S.; de Battisti, A. *J. Appl. Electrochem*. **2004**, *34*, 981−987.

(16) Swamy, K.; Hanesch, P.; Sandl, P.; Bertel, E. *Surf. Sci*. **2000**, *466*, 11−29.

(17) Xu, H.; Harrison, I. *J. Phys. Chem. B* **1999**, *103*, 11233−11236.

(18) Orts, J. M.; Gomez, R.; Feliu. *J. Electroanal. Chem*. **1999**, *467*, 11−19.

(19) Solomon, T.; Wieckowski, A.; Rosasco, S. D.; Hubbard, A. T. *Surf. Sci*. **1984**, *147*, 241−251.

(20) Bittner, A. M.; Wintterlin, J.; Beran, B.; Ertl, G. *Surf. Sci*. **1995**, *335*, 291−299.

(21) (a) Binnig, G.; Rohrer, H.; Gerber, Ch.; Weibel, E. *Phys. Rev. Lett*. **1982**, *49*, 57−61. (b) Binnig, G.; Rohrer, H. *Rev. Mod. Phys*. **1987**, *59*, 615−625.

(22) (a) van Hove, M. A.; Cerda, J.; Sautet, P.; Bocquet, M. L.; Salmeron, M. *Prog. Surf. Sci*. **1997**, *54*, 315−329. (b) Bocquet, M. L.; Cerda, J.; Sautet, P. *Phys. Rev. B* **1999**, *59*, 15437−15445. (c) Sautet, P. *Chem. Rev.* **1997**, *97*, 1097−1116; (d) Carlisle, C. I.; King, D. A.; Bocquet, M. L.; Cerda, J.; Sautet, P. *Phys. Rev. Lett*. **2000**, *84*, 3899−3902.

(23) Hofer, W. A.; Foster, A. S.; Schluger, A. L. *Rev. Mod. Phys*. **2003**, *75*, 1287−1332.

(24) Drakova, D. *Rep. Prog. Phys*. **2001**, *64*, 205−290.

(25) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, *Revision C.02*; Gaussian, Inc.: Wallingford, CT, 2004.

(26) (a) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098−3100. (b) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.

(27) Matsumoto, M.; Tatsumi, N.; Fukutani, K.; Okano, T. *Surf. Sci*. **2002**, *513*, 485−500.

(28) Hansen, K. H.; Sljivancanin, Z.; Hammer, B.; Laegsgaard, E.; Besenbacher, F.; Stensgaard, I. *Surf. Sci*. **2002**, *496*, 1−9.

(29) Materer, N.; Barbieri, A.; Gardin, D.; Starke, U.; Batteas, J. D.; Vanhove, M. A.; Somorjai, G. A. *Surf. Sci*. **1994**, *303*, 319−332.

(30) Because of evident facts originating from either the incomplete treatment of electroncorrelation or neglect of mechanical anharmonicity or basis set truncation effects, the calculated harmonic frequencies are generally overestimated.[31,32] The B3LYP method overestimates also the calculated harmonic frequencies, and, for this reason, the scaling factors have been proposed in the literature.[32,33]

(31) Scott, A. P.; Radom, L. *J. Phys. Chem*. **1996**, *100*, 16502−16513.

(32) (a) Halls, M. D.; Velkovski, J.; Schlegel, H. B. *Theor. Chem. Acc*. **2001**, *105*, 413−421. (b) Halls, M. D.; Raghavachari, K. *J. Phys. Chem. B* **2004**, *108*, 19388−19391.

(33) (a) Zhanpeisov, N. U.; Martra, G.; Ju, W. S.; Matsuoka, M.; Coluccia, S.; Anpo, M. *J. Mol. Catal. A: Chem.* **2003**, *201*, 237−246. (b) Zhanpeisov, N. U.; Nishio, S.; Fukumura, H. *Int. J. Quantum Chem.* **2005**, *105*, 368−375.

CT050308M

JCTC Journal of Chemical Theory and Computation

Theoretical Studies on the Nonlinear Optical Properties of Octupolar Tri-*s*-triazines

Wenxu Zheng,[†] Ning-Bew Wong,*[,†] Wai-Kee Li,[‡] and Anmin Tian*[,§]

Department of Biology and Chemistry, City University of Hong Kong, Kowloon, Hong Kong, Department of Chemistry, The Chinese University of Hong Kong, Shatin, Hong Kong, and Faculty of Chemistry, Sichuan University, Chengdu, Sichuan 610064, P. R. China

Received December 8, 2005

Abstract: The first-, second-, and third-order static and frequency-dependent polarizabilities of a series of octupolar tri-*s*-triazines have been investigated by using the ab initio coupled perturbed Hartree−Fock (CPHF) method. Effects of substitution have also been considered. The results show that α, β, and γ values for octupolar tri-*s*-triazines are much larger than those for *s*-triazine in both static and frequency-dependent cases. Attaching groups containing π systems such as azide and ethenyl to the tri-*s*-triazine molecule results in a significant increase of first-, second-, and third-order polarizabilities. Our calculations suggest that the octupolar tri-*s*-triazines may be prospective candidates for nonlinear optical materials.

Introduction

Molecules possessing large nonlinear optical (NLO) properties have been intensively investigated for potential applications in the areas of integrated optics and other sections of materials science.[1−3] As is well-known, second-order NLO compounds show great promise in their application in photonic devices such as high-speed electrooptic (E−O) modulators and switches, owing to their large E−O coefficients, ultrafast response times, and ease of processing into thin films.[4,5] On the other hand, third-order NLO properties provide the basic means for light controlling with light in all optics, as in optical bistability and phase conjugation. Furthermore, the understanding of all-optical switching, modulating, and computing devices is important in modern optical technology, and NLO materials with large third-order nonlinear susceptibilities are indispensable for such devices. For example, if a material has a high third-order nonlinear susceptibility, it can generate a strong conjugate wave in optical phase-conjugate wave generation devices, and the length of the resonator for reaching the bistable region in optical bistable devices can be much reduced. Our present research is focused on the development of NLO molecules with large nonlinearity, high thermal stability, and good processability.[6,7]

The traditional materials with enhanced molecular second-order NLO coefficients are those with extended π-systems and significant molecular dipoles. Although there are many examples of dipolar molecules exhibiting extremely large hyperpolarizabilities, several limitations prevent them from being a candidate for NLO materials, such as the NLO efficiency-transparency tradeoff, and the proclivity of dipolar molecules to adopt centrosymmetric packing in the solid state.[8] Octupolar molecules, which combine nonlinear optical properties with a strict cancellation of all vectorlike observable dipole moment,[8−11] overcome these problems due to the presence of a 3-fold symmetry axis in octupolar 1,3,5-substituted aromatic ring systems. Such a 3-fold axis may lead to better transparency characteristics,[11] and the lack of a molecular dipole enhances the prospects of noncentrosymmetric crystal packing.

In the 1990s, Zyss and co-workers introduced octupolar molecules for second-order nonlinear optics.[8,9] Since the development of hyper-Rayleigh scattering techniques, the hyperpolarizabilities of octupolar molecules, such as tricyanomethanide ion, tetraorganotin compounds, and sym-

* Corresponding author e-mail: bhnbwong@cityu.edu.hk (N.-B.W.) and suqcp@mail.sc.cninfo.net (A.T.).

[†] City University of Hong Kong.

[‡] The Chinese University of Hong Kong.

[§] Sichuan University.

10.1021/ct050312y CCC: $33.50
Published on Web 04/07/2006

metrically substituted benzenes and triazines as well as tertiary amines, have been evaluated experimentally.[12−17] In 1999, Humphrey et al. first reported data exploring the third-order NLO potential of an organometallic octupolar compound.[18] In 1995, Ray and Das measured the first hyperpolarizabilities of some symmetrically substituted triazines and compared them with those of the corresponding symmetrically substituted benzenes. It was found that the octupolar triazines have quadratic polarizabilities higher than the corresponding octupolar benzenes.[19] In 2002, Wu investigated the first- and third-order polarizabilities of a series of octupolar heteroaromatic-substituted triazines using ab initio methods. It was revealed that these triazines have high polarizabilities.[20] In a recent study, we calculated the geometric and electronic structures of some symmetrically substituted tri-*s*-triazines[21] and found that the octupolar tri-*s*-triazines molecules possess high thermal stability. In this work, we report NLO data for these octupolar compounds using ab initio molecular orbital method.

Methods

The geometries of all symmetrically substituted tri-*s*-triazines studied in this work were optimized at the B3LYP/aug-cc-pVDZ level of theory using the software package Gaussian 98.[22] The calculations for the NLO properties were performed using the HONDO 99.9 set of programs.[23] Static polarizabilities as well as the first- and second-order hyperpolarizability components were calculated analytically using the coupled perturbed Hartree−Fock (CPHF) method.[24] The Buckingham type expansion[25] was used in the CPHF calculation:

$$E = E^{(0)} - \mu_i^{(0)} F_i - \frac{1}{2}\alpha_{ij}F_iF_j - \frac{1}{6}\beta_{ijk}F_iF_jF_k - \frac{1}{24}\gamma_{ijkl}F_iF_jF_kF_l - \cdots$$

where E is the energy of a molecule under the electric field F, $E^{(0)}$ is the unperturbed energy of a free molecule, F_i is the vector component of the electric field in the ith direction, and μ_i, α_{ij}, β_{ijk}, and γ_{ijkl} denote the dipole moment, linear polarizability, and first- and second-order hyperpolarizabilities, respectively. Each of the subscripts i, j, k, and l represents the index of the Cartesian axes x, y, and z. To provide additional useful information to experimentalists, we also calculated the frequency-dependent polarizabilities and hyperpolarizabilities using the time-dependent CPHF method.[26] The 1064 nm line of Nd:YAG laser is used in our calculation since the second-harmonic generation (SHG) experiments are usually performed at this fundamental wavelength.

It is well-known that splitting valence basis sets with additional diffuse and polarization functions can be employed to correct the underestimation, especially for the γ-tensor components. Hurst et al.[24] reported that the 6-31G functions plus one p and one d function (6-31G plus lpld) is a reliable basis set which yields quality β and γ values of polyene systems, even though it is not a very big basis set. In this work, we selected Pople's standard 6-31++G(d,p) basis set in the CPHF calculations. The reliability of this method can be justified since it has been applied to a variety of molecules and has achieved good results.[20,27] So far *s*-triazine derivatives have not been extensively investigated for NLO applications. Only a few important *s*-triazine derivatives have been synthesized and their second harmonic generation properties measured. But these derivatives are too large to be dealt with the CPHF/6-31++G(d,p) method. Instead, we calculated the (hyper)polarizability of benzene and compared our computational results with the available experimental data. The calculated α (64.5 au) is very close to the experimental value of 71.5 au.[28] For the calculated γ (13 241 au, dcKerr 632.8 nm), we also find satisfactory agreement with the measured value (12 800 au, dcKerr 632.8 nm).[29] Hence the theoretical methods adopted in this work should be reliable.

The calculated tensors were transformed to mean scalar values by applying the following equations.

$$\alpha = \frac{1}{3}(\alpha_{xx} + \alpha_{yy} + \alpha_{zz})$$

$$\langle\gamma\rangle = \frac{1}{5}(\gamma_{xxxx} + \gamma_{yyyy} + \gamma_{zzzz} + 2\gamma_{xxyy} + 2\gamma_{xxzz} + 2\gamma_{yyzz})$$

Zyss et al.[8,9,11,30] have developed and discussed the tensorial nature of the second-order polarizability for octupolar molecules. Within their framework, the second-order polarizability tensor can be decomposed in two components $\beta_{J=1}$ and $\beta_{J=3}$, the dipolar and octupolar contributions, respectively.

$$\beta = \beta_{J=1} \otimes \beta_{J=3}$$

$$||\beta||^2 = ||\beta_{J=1}||^2 + ||\beta_{J=3}||^2$$

For 2-D planar systems, the following expressions apply:

$$||\beta_{J=1}||^2 = 3/4[(\beta_{xxx} + \beta_{xyy})^2 + (\beta_{yyy} + \beta_{yxx})^2]$$

$$||\beta_{J=3}||^2 = 1/4[(\beta_{xxx} - 3\beta_{xyy})^2 + (\beta_{yyy} - 3\beta_{yxx})^2]$$

Results and Discussion

Static First-, Second-, and Third-Order Polarizabilities. The octupolar tri-*s*-triazines studied in this work shown in Figure 1 are derived from symmetrical substitution of hydrogen atoms by various groups in tri-*s*-triazine. The *s*-triazine molecule is also displayed for comparison. The geometric structures of these molecules were discussed in detail in our previous work.[21] Tables 1−3 list the calculated values of the static first-, second-, and third-order polarizabilities and their individual components for all six tri-*s*-triazines shown in Figure 1. The tensor components listed in these tables have the following approximate relationships:

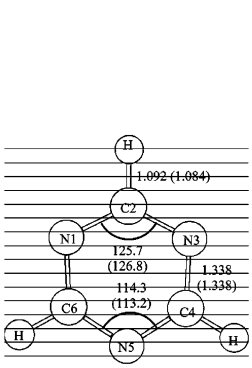
$$\alpha_{xx} = \alpha_{yy}$$

$$\beta_{xxy} = -\beta_{yyy}$$

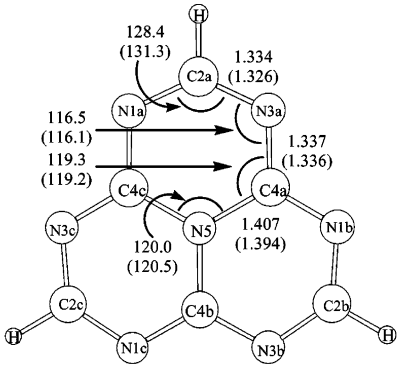
$$\gamma_{xxxx} = \gamma_{yyyy} = 3\gamma_{xxyy}$$

$$\gamma_{xxzz} = \gamma_{yyzz}$$

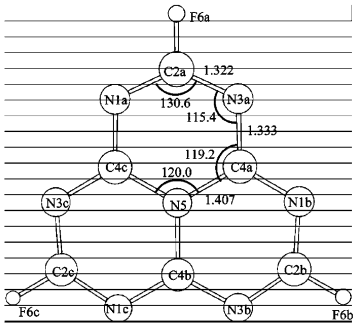
In all the molecules investigated in this work, the axial component α_{zz} of α along the C_{3v} axis is smaller than the



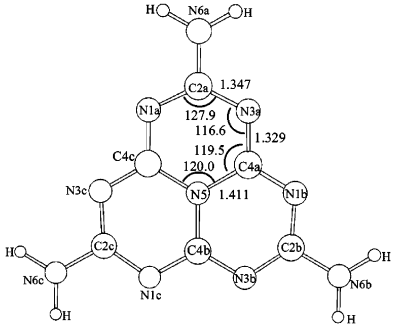
s-triazine



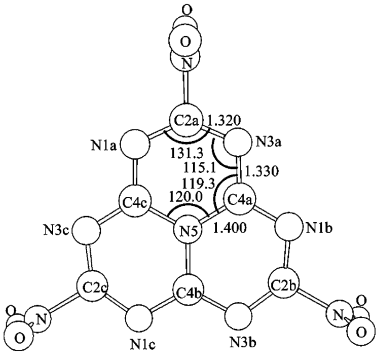
tri-*s*-triazine



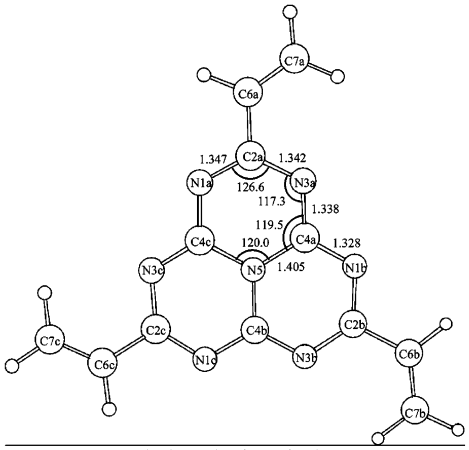
trifluoro-tri-*s*-triazine



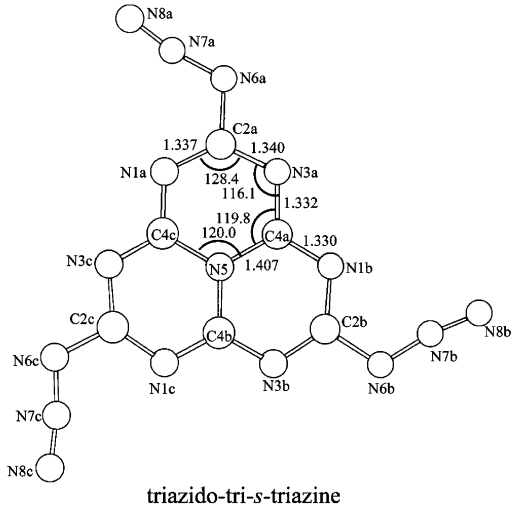
triamino-tri-*s*-triazine



trinitro-tri-*s*-triazine



triethenyl-tri-*s*-triazine



triazido-tri-*s*-triazine

Figure 1. Optimized molecular structures for *s*-triazine and octupolar tri-*s*-triazines with bond lengths in Å and angles in degrees. Experimental data for *s*-triazine and tri-*s*-triazine are given in parentheses.

Table 1. Static First-Order Polarizabilities (Å^3) and Their Individual Components of *s*-Triazine and Octupolar Tri-*s*-triazines

molecule	α_{xx}	α_{yy}	α_{zz}	α
s-triazine	8.3265	8.3265	4.3104	6.9878
tri-*s*-triazine	19.5108	19.5108	7.7516	15.5911
trifluoro-tri-*s*-triazine	19.3357	19.3357	7.6744	15.4486
triamino-tri-*s*-triazine	25.2204	25.2204	9.9322	20.1243
trinitro-tri-*s*-triazine	26.1252	26.1200	17.7568	23.3340
triethenyl-tri-*s*-triazine	37.9127	37.9127	13.3907	29.7387
triazido-tri-*s*-triazine	36.7575	36.7574	11.3723	28.2958

nonaxial components α_{xx} and α_{yy}. As expected, the polarizability α of tri-*s*-triazine is larger (by a factor of 2.2) than that of *s*-triazine. The average α of the parent molecule tri-*s*-triazine is 15.5911 Å^3. Upon substitution, the average α decreases slightly to 15.4486 Å^3 in the fluoro derivative but increases significantly to 20.1243, 23.3340, 28.2958, and 29.7387 Å^3 for the amino, nitro, azido, and ethenyl derivatives, respectively. It is noted that, when the substituent contains π system such as azide and ethenyl, the α value of the tri-*s*-triazine derivative is nearly twice as that of the parent molecule. The fluorine atom plays the role of π-acceptor and drains part of the π-electron cloud from the rings, which in turn decreases the degree of π-charge delocalization. Hence the fluoro-substituted tri-*s*-triazine has the smallest α value.

When we turn to the second-order polarizability β, we notice that the $||\beta||$ value for the parent tri-*s*-triazine molecule is 459.2260 $\times$ 10^{-32} esu, which is almost 8 times larger than that of *s*-triazine (64.6887 $\times$ 10^{-32} esu). By symmetry, the nonzero tensors of β in tri-*s*-triazine and its fluoro and amino derivatives, all with D_{3h} symmetry, are β_{yyy} and β_{xxy}. However, for the nitro, ethenyl, and azide derivatives, all with C_{3h} symmetry, the nonzero β tensors are β_{xxx}, β_{yyy}, β_{xxy}, and β_{xyy}. With tri-*s*-triazine as a reference, the $||\beta||$ value increases slightly in fluoro and nitro derivatives but increases by about 3−5 times in amino, azido, and ethenyl derivatives. This shows that the substituent containing π system enhances the second-order polarizability $||\beta||$ drastically.

The averaged third-order polarizabilities $<\gamma>$ of the *s*-triazine, tri-*s*-triazine, and its derivatives are listed in Table 3. Using *s*-triazine as a reference, the $<\gamma>$ value of tri-*s*-triazine increases by about 2-fold. However, comparing with the parent tri-*s*-triazine, trifluoro-tri-*s*-triazine decreases by about 25%, while the other derivatives experience an increase of 1 order of magnitude.

In summary, when compared with the parent molecule, all substituted tri-*s*-triazines, except the fluoro derivative, experience an increase, sometimes even by several orders of magnitude, for the static first-, second-, and third-order polarizabilities.

Frequency-Dependent First-, Second-, and Third-Order Polarizabilities. From a qualitative point of view, tuning up the frequency of the external electric field induces an increase of delocalizability, allowing a stronger electron mobility in the aromatic systems. This simple model leads to a practical explanation of the frequency dispersion pattern found in tri-*s*-triazines.

The frequency-dependent α and β values are reported in Table 4, where the laser frequency 1064 nm was used as a standard. For polarizability α, comparing the last column of Table 1 with the second column of Table 4, the effect of frequency dispersion is very small. But for second-order polarizability β, comparing the last column of Table 2 with the third and fourth columns of Table 4, frequency dispersion enhances the β value by 10% and more. In particular, the increase is more pronounced when electron-donating or π-bonding substituents are involved. Also, the increase for SHG β $(-2\omega; \omega, \omega)$ is more prominent than that for electrooptic pockels effect (EOPE) β $(-\omega; 0, \omega)$. Compared with the static value, for the ethenyl derivative, β $(-2\omega; \omega,$

Table 2. Static Second-Order Polarizabilities (10^{-32} esu) and Their Individual Components of *s*-Triazine and Octupolar Tri-*s*-triazines

molecule	β_{xxx}	β_{yyy}	β_{xxy}	β_{xyy}	$\|\|\beta\|\|$
s-triazine		32.3444	−32.3443		64.6887
tri-*s*-triazine		−229.6120	229.6133		459.2260
trifluoro-tri-*s*-triazine		−282.5875	282.5871		565.1744
triamino-tri-*s*-triazine		−632.9766	632.9786		1265.9562
trinitro-tri-*s*-triazine	−248.8180	265.8763	−265.9148	248.6477	728.1550
triethenyl-tri-*s*-triazine	276.1715	−1018.2450	1018.2460	−276.1698	2110.0658
triazido-tri-*s*-triazine	162.1427	−1019.7733	1019.7743	−162.1436	2065.1679

Table 3. Static Third-Order Polarizabilities (10^{-39} esu) and Their Individual Components of *s*-Triazine and Octupolar Tri-*s*-triazines

molecule	γ_{xxxx}	γ_{yyyy}	γ_{zzzz}	γ_{xxyy}	γ_{xxzz}	γ_{yyzz}	$<\gamma>$
s-triazine	2657.8618	2657.8703	1730.1931	885.9296	676.2007	676.2029	2304.5183
tri-*s*-triazine	6036.2210	6036.6554	2478.1237	2012.2301	1241.5368	1241.5300	4708.3188
trifluoro-tri-*s*-triazine	4638.8088	4639.2272	1873.7772	1546.4137	845.5892	845.5909	3525.4001
triamino-tri-*s*-triazine	13385.3187	13386.4202	3664.2631	4462.1743	1706.9594	1706.9679	9237.6411
trinitro-tri-*s*-triazine	8403.6906	8401.1645	951.2307	2800.4181	952.1211	952.0732	5433.4621
triethenyl-tri-*s*-triazine	44783.0373	44783.2771	10353.5549	14927.7331	4283.2492	4283.2457	29381.6605
triazido-tri-*s*-triazine	31072.2844	31071.7690	2688.5543	10357.2748	1757.4668	1757.4676	18515.4052

Table 4. Frequency-Dependent ($\lambda = 1064$ nm) First- (Å^3) and Second-Order (10^{-32} esu) Polarizabilities of *s*-Triazine and Octupolar Tri-*s*-triazines

molecule	$\alpha(-\omega, +\omega)$	$\|\beta\|(-2\omega; \omega, \omega)$	$\|\beta\|(-\omega; 0, \omega)$
s-triazine	7.0462	71.0365	66.6595
tri-*s*-triazine	15.7928	512.9644	529.8788
trifluoro-tri-*s*-triazine	15.6171	616.0259	581.22425
triamino-tri-*s*-triazine	20.3787	1435.5371	1318.2964
trinitro-tri-*s*-triazine	23.6296	813.2688	754.5485
triethenyl-tri-*s*-triazine	30.3280	2644.1821	2265.4672
triazido-tri-*s*-triazine	28.7613	2437.2169	2176.6892

Table 5. Frequency-Dependent ($\lambda = 1064$ nm) Third-Order (10^{-39} esu) Polarizabilities of *s*-Triazine and Octupolar Tri-*s*-triazines

molecule	$\gamma(-3\omega; \omega, \omega, \omega)$	$\gamma(-2\omega; 0, \omega, \omega)$	$\gamma(-\omega; \omega, \omega, -\omega)$	$\gamma(-\omega; 0, 0, \omega)$	$\gamma(0; 0, \omega, -\omega)$
s-triazine	2857.1075	2555.7084	2470.8334	2382.8672	2378.8438
tri-*s*-triazine	7003.7833	5329.7833	5164.3859	4899.9057	4889.4697
trifluoro-tri-*s*-triazine	4502.7955	3918.5818	3828.9645	3649.5615	3647.7562
triamino-tri-*s*-triazine	12844.0528	10663.8378	10336.4123	9690.8679	9728.7243
trinitro-tri-*s*-triazine	8334.8171	6094.7383	5972.8161	5646.8307	5659.3228
triethenyl-tri-*s*-triazine	55269.1883	38363.5500	35562.8818	31994.5338	32094.7608
triazido-tri-*s*-triazine	30792.1488	23025.3981	21757.3120	19873.1236	19952.7164

Table 6. Natural Atomic Charges of the Atoms in Tri-*s*-triazine Ring for Octupolar Tri-*s*-triazines

	tri-*s*-triazine					
		trifluoro-	triamino-	trinitro-	triethenyl-	triazido-
N1	−0.55	−0.57	−0.61	−0.52	−0.55	−0.55
C2	0.36	0.91	0.66	0.65	0.47	0.64
N3	−0.55	−0.57	−0.61	−0.51	−0.56	−0.59
C3a	0.68	0.69	0.69	0.70	0.69	0.69
N4	−0.55	−0.57	−0.61	−0.51	−0.55	−0.55
C5	0.36	0.91	0.66	0.63	0.47	0.64
N6	−0.55	−0.57	−0.61	−0.51	−0.56	−0.59
C6a	0.68	0.69	0.69	0.70	0.69	0.69
N7	−0.55	−0.57	−0.61	−0.51	−0.55	−0.55
C8	0.36	0.91	0.66	0.64	0.47	0.64
N9	−0.55	−0.57	−0.61	−0.51	−0.56	−0.59
C9a	0.68	0.69	0.69	0.70	0.69	0.69
N9b	−0.47	−0.48	−0.51	−0.46	−0.48	−0.49
total	−0.65	0.90	−0.12	0.49	−0.33	0.08

ω) is increased by 25%. For the azide and amino derivatives, the corresponding enhancements are 18% and 13%, respectively. For $\beta(-\omega; 0, \omega)$ values, a similar trend is detected. This means that, as the static β values increase, so do the frequency dispersion effects.

For the third-order polarizabilities of tri-*s*-triazines, five different optical processes are considered: third-harmonic generation (THG) $\gamma(-3\omega; \omega, \omega, \omega)$, electric-field-induced second-harmonic generation EFISH $\gamma(-2\omega; 0, \omega, \omega)$, degenerate four-wave mixing DFWM $\gamma(-\omega; \omega, \omega, -\omega)$, electric-field-induced Kerr effect EFIKE $\gamma(-\omega; 0, 0, \omega)$, and electric-field-induced optical rectification EFIOR $\gamma(0; 0, \omega, -\omega)$. The calculated values of the third-order polarizability γ for the aforementioned processes are tabulated in Table 5, where only the averaged scalar values $<\gamma>$ are listed. As can be seen from this table, the effect of frequency dispersion on γ is very strong, with the trend $\gamma(-3\omega; \omega, \omega, \omega) > \gamma(-2\omega; 0, \omega, \omega) > \gamma(-\omega; \omega, \omega, -\omega) > \gamma(0; 0, \omega, -\omega) \approx \gamma(-\omega; 0, 0, \omega)$ for all studied molecules. Again, as expected, the derivatives containing π-bonding substituents, such as ethenyl and azide groups, have the largest frequency-dependent third-order polarizability. Also, as the static γ values increase, so do the frequency dispersion affects. Taking triethenyl-tri-*s*-triazine as an example, $\gamma(-3\omega; \omega, \omega, \omega)$, $\gamma(-2\omega; 0, \omega, \omega)$, $\gamma(-\omega; \omega, \omega, -\omega)$, $\gamma(0; 0, \omega, -\omega)$, and $\gamma(-\omega; 0, 0, \omega)$ increase by 88%, 31%, 21%, 9%, and 9%, respectively.

Table 7. HOMO and LUMO Energies and the Their Gaps for Octupolar Tri-*s*-triazines

molecule	E_{HOMO} (eV)	E_{LUMO} (eV)	$\Delta E_{LUMO-HOMO}$ (eV)
tri-*s*-triazine	−7.32	−3.40	3.92
trifluoro-tri-*s*-triazine	−8.14	−3.64	4.50
triamino-tri-*s*-triazine	−6.44	−1.52	4.92
trinitro-tri-*s*-triazine	−8.91	−4.85	4.06
triethenyl-tri-*s*-triazine	−6.84	−3.19	3.65
triazido-tri-*s*-triazine	−7.48	−3.34	4.14

Ground-State Net Charges and Energy Levels. The ground-state net charge distribution of the tri-*s*-triazines studied in this work was analyzed by means of the natural bond orbital (NBO) model.[31−34] The natural atomic charges of the atoms in the tri-*s*-triazine ring are listed in Table 6. The results indicate that the total net charge of the fused ring for the tri-*s*-triazine molecule is −0.65 and that all substituents decrease the magnitudes of negative charges of the ring. These results imply that there is partial electron transfer from the rings to the substituents. This transfer is more pronounced for the fluoro and nitro derivatives, where the total net charges of the ring become positive. As a result of the charge transfer, the α, β, and γ values in triamino-tri-*s*-triazine, trinitro-tri-*s*-triazine, triethenyl-tri-*s*-triazine, and triazido-tri-*s*-triazine molecules are much greater than that

Table 8. Static First-Order Polarizabilities (Å^3) and Their Individual Components of Triethenyl-tri-*s*-triazine, Triazido-tri-*s*-triazine, and Their Configurations

molecule	α_{xx}	α_{yy}	α_{zz}	α
triethenyl-tri-*s*-triazine	37.9127	37.9127	13.3907	29.7387
triethenyl-tri-*s*-triazine-30[a]	31.4639	19.9353	36.4698	29.2897
triethenyl-tri-*s*-triazine-60[a]	32.5031	18.5031	33.4127	28.1396
triethenyl-tri-*s*-triazine-90[a]	28.5112	24.4644	29.3246	27.4334
triazido-tri-*s*-triazine	36.7575	36.7574	11.3723	28.2958
triazido-tri-*s*-triazine-30[b]	35.4912	13.3622	35.4744	28.1093
triazido-tri-*s*-triazine-60[b]	32.4198	17.3844	32.3598	27.3880
triazido-tri-*s*-triazine-90[b]	30.5213	24.2214	25.7043	26.8157

[a] Three N1−C2−C6−C7 dihedrals are 30°, 60°, and 90°, respectively. [b] Three N1−C2−N6−N7 dihedrals are 30°, 60°, and 90°, respectively.

Table 9. Static Third-Order Polarizabilities (10^{-39} esu) and Their Individual Components of Triethenyl-tri-*s*-triazine, Triazido-tri-*s*-triazine, and Their Configurations

molecule	γ_{xxxx}	γ_{yyyy}	γ_{zzzz}	γ_{xxyy}	γ_{xxzz}	γ_{yyzz}	γ
triethenyl-tri-*s*-triazine	44783.0373	44783.2771	10353.5549	14927.7331	4283.2492	4283.2457	29381.6605
triethenyl-tri-*s*-triazine-30[a]	27273.2289	11475.0252	80337.7657	8989.0430	11870.8883	5359.7104	26330.1226
triethenyl-tri-*s*-triazine-60[a]	2649.5188	6143.9387	27637.2076	3350.2628	8351.7971	3401.2846	18007.4708
triethenyl-tri-*s*-triazine-90[a]	11981.5824	7753.9653	13130.8486	4020.2897	4645.3995	3893.5861	11596.9894
triazido-tri-*s*-triazine	31072.2844	31071.7690	2688.5543	10357.2748	1757.4668	1757.4676	18515.4052
triazido-tri-*s*-triazine-30[b]	28777.3275	3398.8258	28823.6385	1954.6754	9620.8997	1897.6983	17589.2677
triazido-tri-*s*-triazine-60[b]	21667.6432	4080.1100	21437.4695	2005.1426	7297.3843	2146.9824	14016.8483
triazido-tri-*s*-triazine-90[b]	14865.2099	6947.8452	8648.6290	3463.8451	3606.3176	3639.6542	10376.2636

[a] Three N1−C2−C6−C7 dihedrals are 30°, 60°, and 90°, respectively. [b] Three N1−C2−N6−N7 dihedrals are 30°, 60°, and 90°, respectively.

observed in tri-*s*-triazine molecule, while, in the trifluoro-tri-*s*-triazine molecule, the β value increases and both the α and γ values decrease.

The energies of HOMO and LUMO and the gaps between them for the molecules investigated are summarized in Table 7. When compared with the parent molecule, the HOMO and LUMO energy levels decrease in fluoro and nitro derivatives, while they increase in amino and ethenyl substitutions. The energy gap decreases only in triethenyl-tri-*s*-triazine, in which the substituent is a very effective π-donating system. Owing to this small energy gap, the triethenyl-tri-*s*-triazine molecule possesses the largest α, β, and γ values.

π-Conjugation Effects. From the previous analysis, it should be noted that the substituents with π-donating group could enhance the (hyper)polarizabilities of tri-*s*-triazines dramatically. We now examine the role of π-conjugation between the substituent R group and the parent tri-*s*-triazine ring. In general, the most stable conformation of triethenyl-tri-*s*-triazine and triazido-tri-*s*-triazine adopts a planar geometry in order to achieve conjugation between the R group and the parent ring. Examination of the molecular orbitals reveals that the π orbital of the ring is composed of $2p_z$ orbitals of nitrogen and carbon atoms on the ring, and the π orbitals of ethenyl and azido groups are also composed of $2p_z$ orbitals of the atoms in the groups. The conjugation between the groups and the parent ring extends the delocalization of the electrons in the molecule. Hence triethenyl-tri-*s*-triazine and triazido-tri-*s*-triazine have the largest (hyper)-polarizabilities.

It is possible that this conjugation will be destroyed to a large extent when rotation around the dihedral angles N1−C2−C6−C7 of triethenyl-tri-*s*-triazine and N1−C2−N6−N7 of triazido-tri-*s*-triazine deviate significantly from 0°. When we rotated the three equivalent dihedral angles N1−C2−C6−C7 of triethenyl-tri-*s*-triazine and N1−C2−N6−N7 of triazido-tri-*s*-triazine to 30°, 60°, and 90° and calculated the (hyper)polarizabilities of these configurations, the results listed in Tables 8 and 9 were obtained. As can be seen from these tables, the (hyper)polarizabilities decrease with the increase of the dihedrals, which weakens the conjugation between the substituents and the ring. Even under such conditions, the (hyper)polarizabilities of all these configurations are still much larger than those of other substituted tri-*s*-triazines.

Conclusions

We have reported an ab initio study on the first-, second-, and third-order static and frequency-dependent polarizabilities of a series of octupolar tri-*s*-triazines. Our calculations show that octupolar tri-*s*-triazines have much larger α, β, and γ values than *s*-triazine, both in static and frequency-dependent cases. The π-containing substituents such as azide and ethenyl result in a significant increase of first-, second-, and third-order polarizabilities. Based on our results, these octupolar tri-*s*-triazines can be classified as a new family of NLO materials.

Acknowledgment. This work was supported by the Research Grants Council of Hong Kong (Account no. 9040979 (CityU 102404)), the National Science Foundation of China (Account no. 20373045), and Special Research Foundation of Doctoral Education of Chinese University (Account no. 20020610024).

References

(1) *Nonlinear Optical Properties of Organic Molecules and Crystals*; Chemla, D. S., Zyss, J., Eds.; Academic Press: New York, 1987.

(2) Prasad, P. N.; Williams, D. J. *Introduction to Nonlinear Optical Effect in Molecules and Polymers*; Wiley: New York, 1990.

(3) *Organic Nonlinear Optical Materials*; Bosshard, Ch., Sutter, K., Pretre, Ph., Hulliger, J., Flörsheimer, M., Kaatz, P., Günter P., Eds.; Overseas Publishers Association: Amsterdam, 1995.

(4) (a) Robinson, B. H.; Dalton, L. R.; Harper, A. W.; Ren, A.; Wang, F.; Zhang, C.; Todorova, G.; Lee, M.; Aniszfeld, R.; Garner, S.; Chen, A.; Steier, W. H.; Houbrecht, S.; Persoons, A.; Ledoux, I.; Zyss, J.; Jen, A. K.-Y. *Chem. Phys.* **1999**, *245*, 35. (b) Marder, S. R.; Kippelen, B.; Jen, A. K.-Y.; Peyghambarian, N. *Nature* **1997**, *388*, 845. (c) Ghebremichael, F.; Kuzyk, M. G.; Lackritz, H. S. *Prog. Polym. Sci.* **1997**, *22*, 1147. (d) Marks, T. J.; Ratner, M. A. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 155. (e) Burland, D. M.; Miller, R. D.; Walsh, C. A. *Chem. Rev.* **1994**, *94*, 31.

(5) (a) *Photonic Polymer Systems*; Wise, D. L., Wnek, G. E., Trantolo, D. J., Cooper, T. M., Gresser, J. D., Eds.; Marcel Dekker: New York, 1998. (b) *Polymers for Second-Order Nonlinear Optics*; Lindsay, G. A., Singer, K. D., Eds.; ACS Symposium Series No. 601; American Chemical Society: Washington, DC, 1995. (c) Prasad, P.; Williams, D. *Introduction to Nonlinear Optical Effects in Molecules and Polymers*; John Wiley and Sons: New York, 1991.

(6) (a) Ermer, S.; Lovejoy, S. M.; Leung, D. S.; Warren, H.; Moylan, C. R.; Twieg, R. J. *Chem. Mater.* **1997**, *9*, 1437. (b) Jen, A. K.-Y.; Rao, P.; Wong, K. Y.; Drost, K. J. *J. Chem. Soc., Chem. Commun.* **1993**, 90. (c) Lindsay, G. A.; Henry, R. A.; Yee, R. Y.; Chafin, A. P.; Seltzer, M. D.; Hill, R.; Knoesen, A. *Proc. SPIE-Int. Soc. Opt. Eng.* **1994**, *2143*, 88.

(7) (a) Prasad, P. N.; William, D. J. *Introduction to Nonlinear Optical Effects in Molecules and Polymers*; John Wiley: New York, 1991. (b) Ouder, J. L.; Le Person, H. *Opt. Commun.* **1975**, *15*, 258.

(8) Zyss, J. *Nonlinear Opt.* **1991**, *1*, 3.

(9) Zyss, J.; Ledoux, I. *Chem. Rev.* **1994**, *94*, 77.

(10) Brédas, J. L.; Meyers, F.; Pierce, B. M.; Zyss, J. *J. Am. Chem. Soc.* **1992**, *114*, 4928.

(11) Joffre, M.; Yarron, D.; Silbey, R. J.; Zyss, J. *J. Chem. Phys.* **1992**, *97*, 5607.

(12) Verbiest, T.; Clays, K.; Persoons, A.; Meyers, F.; Brédas, J. L. *Opt. Lett.* **1993**, *18*, 525.

(13) Verbiest, T.; Clays, K.; Samyn, C.; Wolv, J.; Reinhoudt, D.; Persoons, A. *J. Am. Chem. Soc.* **1994**, *116*, 9320.

(14) Lequan, M.; Branger, C.; Simon, J.; Thami, T.; Chauchard, E.; Persoons, A. *Adv. Mater.* **1994**, *6*, 851.

(15) Stadler, S.; Feiner, F.; Braüchle, C.; Brandle, S.; Gompper, R. *Chem. Phys. Lett.* **1995**, *245*, 292.

(16) Ray, P. C.; Das, P. K. *Chem. Phys. Lett.* **1995**, *244*, 153.

(17) Stadler, S.; Braüchle, C.; Brandl, S.; Gompper, R. *Chem. Mater.* **1996**, *8*, 414.

(18) McDonagh, A. M.; Humphrey, M. G.; Samoc, M.; Davies, B. L.; Houbrechts, S.; Wada, T.; Sasabe, H.; Persoons, A. *J. Am. Chem. Soc.* **1999,** *121*, 1405.

(19) Ray, P. C.; Das, P. K. *Chem. Phys. Lett.* **1995**, *244*, 153.

(20) Zhu, W.; Wu, G. S. *J. Phys. Chem. A* **2002**, *106*, 7216.

(21) Zheng, W. X.; Wong, N.-B.; Wang, W. Z.; Zhou, G.; Tian, A. M. *J. Phys. Chem. A* **2004**, *108*, 97.

(22) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. m. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, *Revision A.11*; Gaussian, Inc.: Pittsburgh, PA, 2001.

(23) Dupuis, M.; Marquez, A.; Davidson, E. R. *HONDO 99.9 from CHEMStation*; IBM Corporation: Neighborhood Road, Kingston, NY 12401, 1999.

(24) Hurst, G. J. B.; Dupuis, M.; Clementi, E. *J. Chem. Phys.* **1988**, *89*, 385.

(25) Buckingham, A. D. *Adv. Chem. Phys.* **1967**, *12*, 107.

(26) Dupuis, M.; Karna, S. *J. Comput. Chem.* **1991**, *12*, 487.

(27) (a) Barbosa, A. G. H.; Nascimento, M. A. C. *Chem. Phys. Lett.* **2001**, *343*, 15. (b) Hamada, T. *J. Phys.* Chem. **1996**, *100*, 19344. (c) Karna, S. P.; Dupuis, M. *J. Chem. Phys.* **1990**, *92*, 7418. (d) Bader, M. M.; Hamada, T.; Kakuta, A. *J. Am. Chem. Soc.* **1992**, *114*, 6475. (e) Nilar, S. H.; Pluta, T. S. *J. Am. Chem. Soc.* **1995**, *117*, 12603. (f) Dory, M.; Beudels, L.; Fripiat, J. G.; Delhakke, J.; Andre, J. M.; Dupuis, M. *Int. J. Quantum Chem.* **1992**, *42*, 1577.

(28) Bogaard, M. P.; Orr, B. J. *MTP International Review of Science, Physical Chemistry, Series Two; Vol. 2. Molecular structure and properties*; Butterworth: London, 1975.

(29) Shelton, D. P.; Rice, J. E. *Chem. Rev.* **1994**, *94*, 3.

(30) Zyss, J. *J. Chem. Phys.* **1993**, *98*, 6583.

(31) Carpenter, J. E.; Weinhold, F. *J. Mol. Struct.* (*Theochem*) **1988**, *169*, 41.

(32) Reed, A. E.; Curtiss, L. A.; Weinhold, F. *Chem. Rev.* **1988**, *88*, 899.

(33) Foster, J. P.; Weinhold, F. *J. Am. Chem. Soc.* **1980**, *102*, 7211.

(34) Reed, A. E.; Weinstock, R. B.; Weinhold, F. J. *Chem. Phys.* **1985**, *83*, 735.

CT050312Y

JCTC Journal of Chemical Theory and Computation

Combining Quantum Mechanics Methods with Molecular Mechanics Methods in ONIOM

Thom Vreven,*[,†,‡] K. Suzie Byun,[‡] István Komáromi,[‡,§] Stefan Dapprich,[‡] John A. Montgomery, Jr.,[†] Keiji Morokuma,*[,‡] and Michael J. Frisch[†]

Gaussian, Inc., 340 Quinnipiac Street, Building 40, Wallingford, Connecticut 06492, and Cherry Emerson Center for Scientific Computation and Department of Chemistry, Emory University, Atlanta, Georgia 30322

Received November 26, 2005

Abstract: The purpose of this paper is 2-fold. First, we present several extensions to the ONIOM(QM:MM) scheme. In its original formulation, the electrostatic interaction between the regions is included at the classical level. Here we present the extension to electronic embedding. We show how the behavior of ONIOM with electronic embedding can be more stable than QM/MM with electronic embedding. We also investigate the link atom correction, which is implicit in ONIOM but not in QM/MM. Second, we demonstrate some of the practical aspects of ONIOM(QM:MM) calculations. Specifically, we show that the potential surface can be discontinuous when there is bond breaking and forming closer than three bonds from the MM region.

1. Introduction

Hybrid methods allow the combination of two or more computational techniques in one calculation and make it possible to investigate the chemistry of very large systems with high precision. The region of the system where the chemical process takes place, for example bond breaking or bond formation, is treated with an appropriately accurate method, while the remainder of the system is treated at a lower level. The most common class of hybrid methods is formed by the QM/MM methods, which combine a quantum mechanical (QM) method with a molecular mechanics (MM) method.[1−4] The ONIOM (our Own N-layer Integrated molecular Orbital molecular Mechanics) scheme is more general in the sense that it can combine any number of molecular orbital methods as well as molecular mechanics methods.[5−13] Hybrid methods in general have been the subject of a number of recent reviews.[14−22]

A variety of QM/MM schemes have been reported in the literature. Although the main concepts are similar, the methods differ in a number of details. One major distinction is in the treatment of covalent interaction between the two regions. The resulting dangling bonds in the QM calculation need to be saturated, and the simplest approach is to use link atoms.[2,23] These are usually hydrogen atoms but can, in principle, be any atom that mimics the part of the system that it substitutes. Link atoms are used in a large proportion of QM/MM implementations as well as our ONIOM scheme. The main alternative to link atoms is the use of frozen orbitals, which can through parametrization and use of *p* and *d* orbitals describe a more accurate charge density than link atoms.[1,24−26] Although the few studies that directly compare link atom methods with frozen orbital methods show that both schemes perform well,[25,27] it appears generally accepted that the latter can provide a better description of the boundary. However, due to the required parametrization of the frozen orbitals, they lack the flexibility and generality of link atoms. In our implementation we exclusively use link atoms because we consider generality one of the key aspects of our ONIOM scheme, and it is not feasible to parametrize frozen orbitals for every possible method combination. It must be noted that to some extent also the accuracy of link atoms can be improved by parametrization. The electrone-

* Corresponding author fax: 203 284 2520; e-mail: thom@gaussian.com.
† Gaussian, Inc.
‡ Emory University.
§ Current address: Thrombosis and Haemostasis Research Group of the Hungarian Academy of Sciences at the University of Debrecen, H-4032 Debrecen, Hungary.

10.1021/ct050289g CCC: $33.50
Published on Web 04/14/2006

gativity can be modified through the use of a shift operator[28] or, in the case of semiempirical QM methods, the parameters involving link atoms can be adjusted.[29] We have not included this limited parametrization in our current method but may do so in the future. Besides frozen orbitals and link atoms, several QM/MM methods use pseudopotentials to handle the covalent interactions between the QM and MM regions.[30,31]

The second main difference between various QM/MM methods is the way the electrostatic interaction between the two layers is treated.[32] In its simplest form, the interaction between the QM and MM region is completely described by MM style terms. This includes the electrostatic interaction, which is then evaluated as the interaction of the MM partial charges with partial (point) charges assigned to the atoms in the QM region. This approach is usually referred to as *classical* or *mechanical embedding*. In the second approach, the charge distribution of the MM region interacts with the actual charge distribution of the QM region. In this case, the partial charges from the MM region are included in the QM Hamiltonian, which provides a more accurate description of the electrostatic interaction and, in addition, allows the wave function to respond to the charge distribution of the MM region. This approach is referred to as *electronic embedding*. In the original version of ONIOM, the QM region and MM region interact via mechanical embedding.

In this paper we present several issues that are specifically related to QM/MM and QM/QM/MM methods within the ONIOM framework, where we focus on aspects that are different in most other QM/MM schemes. We will discuss partitioning restrictions that follow from the fact that the ONIOM energy expression is in the form of an extrapolation, which we refer to as the *cancellation problem*. Further, we discuss the placement of the link atoms in the MM model system calculation. Finally, we extended our methods to include electronic embedding. For a discussion of the geometry optimization methods in our package we refer to a separate series of papers.[33−36] In the following sections we will first introduce the details of the ONIOM method. We will not provide a comprehensive overview of QM/MM methods but do present a 'generic QM/MM method' that contains all the components that we need for comparison to ONIOM. From here on we will use the term ONIOM to denote the combination of QM methods with MM methods in the ONIOM framework and QM/MM to denote the 'generic QM/MM method'. The ONIOM scheme has been implemented in the Gaussian suite of programs.[37] Some of the developments, however, are only available in the private development version of the program,[38] and it must be noted that a first version of electronic embedding in ONIOM was independently implemented by István Komáromi.

Finally, we want to clarify some of the misconceptions about our methods that occasionally seem to arise.[39] The ONIOM method as it is currently implemented is the latest incarnation of a series of hybrid methods developed by Morokuma and co-workers. This series includes IMOMM (Integrated Molecular Orbital + Molecular Mechanics), which combines a Molecular Orbital (MO) method with a MM method, and IMOMO (Integrated Molecular Orbital + Molecular Orbital), which combines two MO methods into one calculation. IMOMM and IMOMO are *not* a subset of ONIOM. The link atom position in ONIOM is obtained with a scale factor, while in IMOMM and IMOMO the link atom is placed at a specified distance from the atom to which it is connected. We consider the link atom treatment an intrinsic aspect of the ONIOM method.

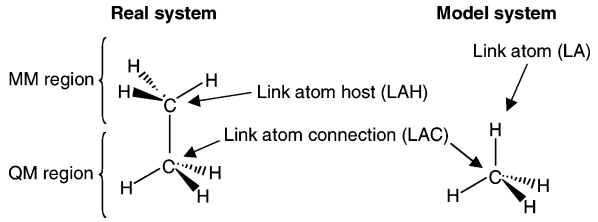


Figure 1. ONIOM terminology using ethane as an example.

2. Theory

2.1. MM Force Fields. An example of a typical force field, in this case Amber,[40] is of the form

$$E^{\text{total}} = \sum_{\text{bonds}} K_r(r - r_{\text{eq}})^2 + \sum_{\text{angles}} K_\theta(\theta - \theta_{\text{eq}})^2 + \sum_{\text{dihedrals}} \frac{V_n}{2}[1 + \cos(n\phi - \gamma)] + \sum_{i<j}\left[s_{ij}^{\text{VdW}}\left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}\right) + s_{ij}^{q}\frac{q_i q_j}{\epsilon r_{ij}}\right] \quad (1)$$

The first three terms describe the *bonded interactions*, formed by all the (chemical) bonds, angles, and dihedrals (including out-of-plane deformations) that are present in the system. The number of bonded terms scales linearly with the size of the system. The last term describes the *nonbonded interaction* between each pair of atom in the system. The van der Waals interaction, $(A_{ij}/r_{ij}^{12} - B_{ij}/r_{ij}^{6})$, and the Coulomb interaction, $q_i q_j/\epsilon r_{ij}$, are scaled by factors s_{ij}^{VdW} and s_{ij}^{q}, respectively. The factors only differ from unity when the centers i and j are separated by three bonds or fewer, and the argument for using them is that the van der Waals and electrostatic interactions are already included in the bonded terms. The number of nonbonded terms scales quadratically with the size of the system and would be the bottleneck in MM calculations if computed as written in eq 1. Most implementations, however, use either distance based cutoffs or linear scaling methods for the evaluation of the nonbonded interaction. In our implementation we use cutoffs and a boxing algorithm to evaluate with the van der Waals interaction[35] and the Fast Multipole Method to deal with the electrostatic interaction.[35,41−44]

2.2. ONIOM and QM/MM Energy Expressions. In a two-layer ONIOM(QM:MM) calculation, the total energy of the system is obtained from three independent calculations:

$$E^{\text{ONIOM}} = E^{\text{real,MM}} + E^{\text{model,QM}} - E^{\text{model,MM}} \quad (2)$$

The *real* system contains all the atoms and is calculated only at the MM level. The *model* system contains the part of the system that is treated at the QM level. Both QM and MM calculations need to be carried out for the model system. In Figure 1 we illustrate the terminology using ethane, where we include one methyl group in the QM region and the

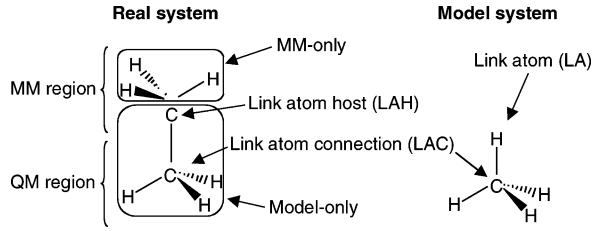


Figure 2. QM/MM terminology using ethane as an example.

remainder in the MM region. Because there is bonded interaction between the two regions, the model system includes a hydrogen link atom to saturate the open valence. The atoms that occur in both the model system and the real system have identical geometrical coordinates. The link atom (LA) is placed on the line that connects the center to which it is connected (the *Link Atom Connection*, LAC) with the atom that it substitutes (the *Link Atom Host*, LAH). The LAC-LA distance is obtained by scaling the original LAC-LAH distance with a constant factor, *g*, which is chosen so that a (chemically) reasonable LAC-LAH distance yields a reasonable LAC-LA distance.[5]

$$\mathbf{q}_{\mathrm{LA}} = \mathbf{q}_{\mathrm{LAC}} + g(\mathbf{q}_{\mathrm{LAH}} - \mathbf{q}_{\mathrm{LAC}}) \tag{3}$$

We will now relate the ONIOM(QM:MM) expression to the generic QM/MM scheme. The latter can be written as

$$E^{\mathrm{QM/MM}} = E^{\mathrm{MM-only,MM}} + E^{\mathrm{model,QM}} + E^{\mathrm{MM-only*model-only,MM}} \tag{4}$$

In Figure 2 we illustrate the QM/MM terminology. We assume that the position of the link atom is the same as in the ONIOM scheme. $E^{\mathrm{MM-only,MM}}$ is the MM energy of the part of the system that only involves MM atoms (thus excluding the LA's). $E^{\mathrm{MM-only*model-only,MM}}$ describes the interaction between the QM region and the MM region and contains all the MM terms that have at least one center in the MM-only region and at least one center in the model-only region. The QM/MM equation reflects a different approach than the ONIOM equation: Equation 4 is a summation scheme. It adds the QM energy of the QM region, the MM energy of the MM region, and the MM interaction energies between the two regions. The ONIOM expression 2, on the other hand, is cast as an extrapolation scheme. Note that in the ONIOM expression all three subcalculations are on 'complete systems', whereas in the QM/MM expression two of the terms are on partial systems, which is the reason for only ONIOM allowing the combination of QM methods with QM methods.

In the ONIOM scheme, most of the MM terms in the model system exist in the real system as well and cancel exactly in the full expression. The difference between the real system and model system MM calculations, the *S*-value, describes the contribution from the MM region, which includes both the energy of the MM region as well as the interaction between the QM region and the MM region.

$$S^{\mathrm{MM}} = E^{\mathrm{real,MM}} - E^{\mathrm{model,MM}} \tag{5}$$

S^{MM} plays the same roll as the $E^{\mathrm{MM-only,MM}} + E^{\mathrm{MM-only*model-only,MM}}$ terms in the QM/MM expression. In fact, when no bonded interactions are present between the regions, S^{MM} is identical to $(E^{\mathrm{MM-only,MM}}+E^{\mathrm{MM-only*model-only,MM}})$, and also the ONIOM and QM/MM energies become identical.

When bonded interactions are present in the system, the ONIOM and QM/MM functions are not identical. The terms involving LA in the model system calculation are not identical to the terms involving LAH in the real system calculation and do not cancel. The difference between the terms describes the difference between the LA and the LAH and can be interpreted as the MM extrapolation (or correction) of the link atoms in the model system QM calculation to the corresponding LAH atoms in the real system. In the generic QM/MM scheme, this extrapolation (or any other way to correct the QM link atom) is not present, although several QM/MM methods deal with this issue in other ways.

It is clear that when the ONIOM scheme is applied in its original formalism, as in eq 2, the interaction between the QM and MM regions is included via the MM calculations and therefore follows the *mechanical embedding* formalism. For this reason we also presented the generic QM/MM method in its mechanical embedding form. Later we will present *electronic embedding* formalisms for both ONIOM and QM/MM.

2.3. Derivatives and Three-Layer ONIOM. Derivatives with respect to geometrical coordinates and other properties can be obtained with the ONIOM formalism.[5] For example, the gradient is written as

$$\frac{E^{\mathrm{ONIOM}}}{\partial \mathbf{q}} = \frac{E^{\mathrm{real,MM}}}{\partial \mathbf{q}} + \frac{E^{\mathrm{model,QM}}}{\partial \mathbf{q}^{\mathrm{M}}_{\mathrm{QM}}}\mathbf{J}_{\mathrm{QM}} - \frac{E^{\mathrm{model,MM}}}{\partial \mathbf{q}^{\mathrm{M}}_{\mathrm{MM}}}\mathbf{J}_{\mathrm{MM}} \tag{6}$$

q is a vector that contains the Cartesian coordinates of the real system, and $\mathbf{q}^{\mathrm{M}}_{\mathrm{QM}}$ and $\mathbf{q}^{\mathrm{M}}_{\mathrm{MM}}$ are the Cartesian coordinates of the QM and MM model systems, respectively. The Jacobian **J** projects the gradients of the link atoms onto the link atom host (LAH) and connection (LAC) coordinates. Because the positions of the link atoms is a function of the geometry of the real system, there are no additional (or fewer) degrees of freedom in the ONIOM scheme, and the potential function is well-defined. ONIOM can therefore be used in standard geometry optimization schemes and almost every other technique for the investigation of potential energy surfaces.

In principle, ONIOM can handle any number of layers, although the current implementation is limited to three. This facilitates QM/QM/MM calculations, which can for a given accuracy handle much larger QM regions than regular QM/MM methods. The combined QM/QM region in ONIOM-(QM:QM:MM) is obtained in a similar manner as in ONIOM(QM:QM), using three independent QM subcalculations. This is conceptually different from the QM/QM/MM implementation using CDFT, in which only two QM subcalculations are carried out, and the wave functions are directly coupled.[45] The expression for ONIOM(QM-high: QM-low:MM), where QM-high and QM-low denote a high-

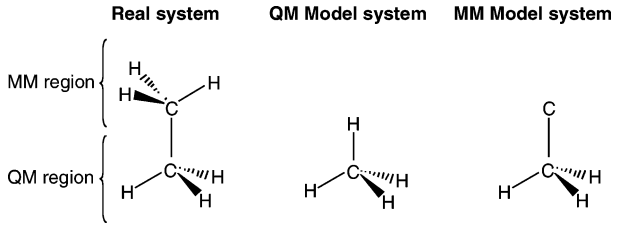


Figure 3. Alternative link atom placement in the MM model system in ONIOM.

Table 1: Vibrational Frequencies (cm^{-1}) of Ethane with ONIOM, B3LYP, and the Generic QM/MM

mode	B3LYP	QM/MM	ONIOM
CH_3-CH_3 twisting	312	317	316
asymmetric C−C−H bending (2)	832	903	972
C−C stretching	**1010**	**786**	**955**
symmetric C−C−H bending (2)	1235	1098	1168
asymmetric CH_3 umbrella	1433	1387	1468
symmetric CH_3 umbrella	1454	1474	1562
symmetric H−C−H bending (2)	1532	1414	1414
asymmetric H−C−H bending (2)	1537	1534	1549

level QM method and a low-level QM method, respectively, contains five terms:

$$E^{\mathrm{ONIOM3}} = E^{\mathrm{real,MM}} + E^{\mathrm{intermediate,QM-low}} - E^{\mathrm{intermediate,MM}} + E^{\mathrm{model,QM-high}} - E^{\mathrm{model,QM-low}} \quad (7)$$

Intermediate denotes the intermediate model system. Gradients and properties can be obtained for a three-layer system in the same way as for a two-layer system.

2.4. Link Atom Placement and Link Atom Extrapolation. When we first presented the ONIOM formalism, we assumed all the geometrical coordinates of the two model system calculations to be identical. The link atom scale factors are then the same for the QM and MM model systems, and also $\mathbf{J}_{\mathrm{QM}}$ and $\mathbf{J}_{\mathrm{MM}}$ in eq 6 are identical. In the most recent implementation we generalized the link atom placement and allow different scale factors g for the QM model system and the MM model system. The initial reason for lifting the restriction was to be able to minimize the error resulting from the link atom. In ONIOM(QM:MM) calculations, however, we can use this mechanism for a different purpose, which we illustrate in Figure 3. The QM model system calculation is obtained in the usual way with 'chemically reasonable scale factors', but the MM model system is obtained with unit scale factors, and $\mathbf{J}_{\mathrm{MM}}$ is a unit matrix. In addition, the MM atom type of the link atom is kept the same as the LAH in the real system. Carrying out an ONIOM(QM:MM) calculation in this way has the advantage that no MM parametrization for the link atoms is required and that therefore any system for which MM parameters are available can also be treated with ONIOM(QM:MM). However, close inspection of the MM terms in the ONIOM energy expression shows that all the MM terms from the model system occur in the real system as well and therefore cancel. Since the terms involving LAH and LA are now identical, the extrapolation of the LA to the LAH is no longer present. In fact, the ONIOM(QM:MM) expression

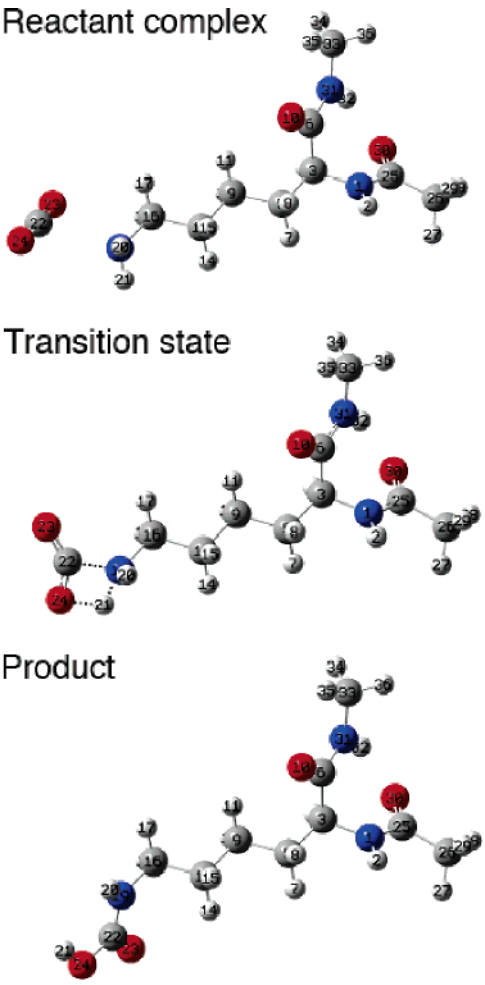


Figure 4. Carboxylation of lysine.

with unit scale factors for the MM model system becomes in this case identical to the QM/MM expression 4! Since the advantages of ONIOM(QM:MM) over QM/MM are removed with unit scale factors, we do not favor this.

The effect of the different link atoms placement is illustrated by the vibrational analysis of the ethane example from Figures 1−3. We calculated the vibrational spectrum at the B3LYP/6-31G(d) level of theory and compared this to the spectrum obtained with ONIOM and QM/MM. As outlined above, the latter is identical to ONIOM with the unit scale factor for the MM model system link atom. We used B3LYP/6-31G(d) for the QM method and the Amber force field for the MM method. In Table 1 we show the frequencies. The C−C stretching frequency is in bold. It is clear that the ONIOM value is much closer to the B3LYP reference value than the QM/MM value is. This is due to the link atom correction at the MM level of theory, which specifically improves the C−C stretch mode because it involves the border between the MM and QM region. Some of the other values are quite similar for QM/MM and ONIOM but differ significantly from B3LYP. These are modes that are either fully or partially located in the MM region and are therefore (partially) described by the MM parameters.

2.5. Cancellation Problem. We will now turn to a technical issue that needs to be taken into account when

Table 2: *S*-Values (Hartrees) for the Three Partitionings of the Carboxylation of Lysine

	reactant connectivity			product connectivity			
	$E^{MM,Real}$(TS)	$E^{MM,Model}$(TS)	*S*(TS,Reac)	$E^{MM,Real}$(TS)	$E^{MM,Model}$(TS)	*S*(prod)	ΔS(TS)
			Partitioning 1				
Coulomb	−0.07858	−0.07807	−0.00051	−0.21806	−0.21072	−0.00735	0.00683
van der Waals	7.91222	7.89104	0.02118	0.10302	0.09556	0.00746	0.01372
stretching	0.06245	0.06144	0.00102	0.10449	0.10348	0.00102	0.00000
bending	0.01353	0.01008	0.00345	0.05744	0.05509	0.00235	0.00110
torsion	0.00825	0.00000	0.00825	0.00755	0.00000	0.00755	0.00070
out-of-plane	0.00002	0.00000	0.00002	0.00005	0.00004	0.00002	0.00000
total	**7.91788**	**7.88448**	**0.03340**	**0.05449**	**0.04345**	**0.01105**	**0.02235**
			Partitioning 2				
Coulomb	−0.07858	−0.06837	−0.01021	−0.21806	−0.20726	−0.01080	0.00059
van der Waals	7.91222	7.90447	0.00775	0.10302	0.09522	0.00780	−0.00005
stretching	0.06245	0.06113	0.00132	0.10449	0.10317	0.00132	0.00000
bending	0.01353	0.01170	0.00183	0.05744	0.05562	0.00183	0.00000
torsion	0.00825	0.00076	0.00748	0.00755	0.00006	0.00748	0.00000
out-of-plane	0.00002	0.00000	0.00002	0.00005	0.00004	0.00002	0.00000
total	**7.91788**	**7.90969**	**0.00819**	**0.05449**	**0.04685**	**0.00764**	**0.00054**
			Partitioning 3				
Coulomb	−0.07858	−0.06416	−0.01442	−0.21806	−0.20364	−0.01442	0.00000
van der Waals	7.91222	7.90432	0.00790	0.10302	0.09512	0.00790	0.00000
stretching	0.06245	0.06130	0.00115	0.10449	0.10334	0.00115	0.00000
bending	0.01353	0.01182	0.00171	0.05744	0.05574	0.00171	0.00000
torsion	0.00825	0.00077	0.00748	0.00755	0.00007	0.00748	0.00000
out-of-plane	0.00002	0.00000	0.00002	0.00005	0.00004	0.00002	0.00000
total	**7.91788**	**7.91405**	**0.00383**	**0.05449**	**0.05066**	**0.00383**	**0.00000**

studying chemical reactivity. Standard (nonreactive) MM methods cannot handle bond breaking and forming. As a general rule, relative energies only make sense when the structures are computed with the same atom types and connectivity. In other words, the potential function 1 needs to be continuous, which means that the list of angles, bends, etc. and the associated parameter cannot change. For example, consider the carboxylation of the lysine amino acid residue (Figure 4). The reactant and product have a different connectivity, and therefore the reaction energy cannot be computed with MM methods. Somewhere during the reaction (in the TS region), the connectivity changes from 'reactant' to 'product'. This introduces a discontinuity in the potential function, which is clearly wrong; for a method to describe bond breaking and forming correctly, the potential energy surface must be continuous.

In contrast to MM methods, ONIOM and QM/MM are able to describe bond breaking and forming. We illustrate this with the carboxylation example. The bond between H21 and O24 is broken in the reactant complex and formed in the product. In a MM calculation there is therefore a stretching term for this bond in the product but no stretching term in the complex. In a pure MM calculation, this results in a discontinuity in the potential surface. However, the ONIOM energy expression contains two MM terms, $E^{real,MM}$ and $E^{model,MM}$, and the stretching term for H21 and O24 occurs in both. Because these two H21−O24 stretching terms enter the energy expression with different sign, they cancel exactly and have no contribution to the ONIOM energy. The result is that the energy does not depend on whether the bond between H21 and O24 is considered being broken or formed in the MM subcalculations. In other words, the ONIOM energy is continuous, despite the changes in the connectivity during the reaction.

We can present the issue raised in the previous paragraph in a more formal way. For the ONIOM surface to be continuous, the MM contribution to the energy must be continuous. The MM contribution is defined as the *S*-value, eq 5. Using the *S*-value, the ONIOM energy can be written as

$$E^{ONIOM} = E^{model,QM} + S^{MM} \tag{8}$$

For the *S*-value to be continuous, it must be independent of the definition of the connectivity in the reaction center. For example, when we take the transition state structure of the carboxylation reaction, the *S*-value calculated with the connectivity as in the reactant structure must be the same as the *S*-value calculated with the connectivity as in the product structure:

$$\Delta S^{MM} = S^{MM}(\text{TS,reactant connectivity}) - S^{MM}(\text{TS,product connectivity}) \Rightarrow \text{must be zero}$$

In Table 2 we show the S^{MM} and ΔS^{MM} values for the transition state of the carboxylation reaction, for the three partitionings in Figure 5. We calculated the *S*-value with both the reactant connectivity and the product connectivity and show the difference ΔS^{MM} in the final column. Besides the total MM values (in bold), we also show the values broken down for each type of MM term.

In the table we see that the reactant *S*-value and product *S*-value differ for both partitioning 1 and partitioning 2. Thus

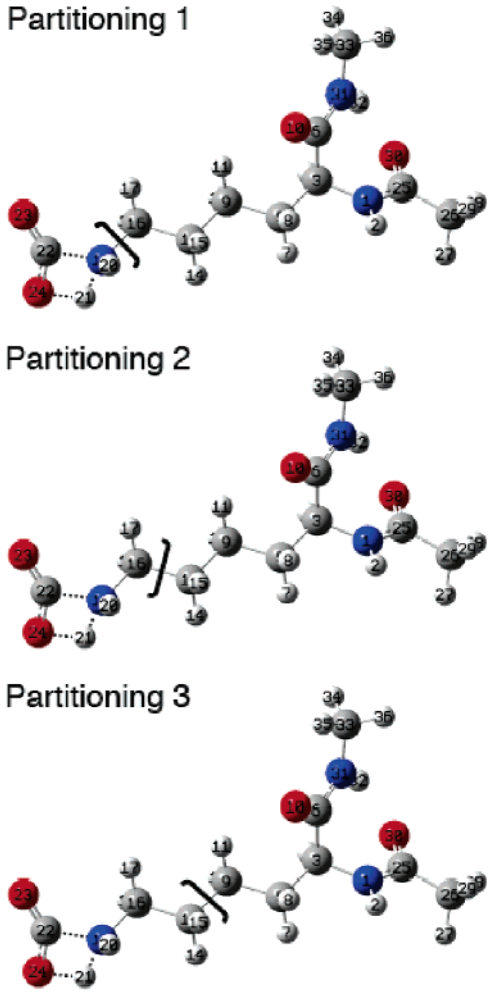


Figure 5. Partitionings for the carboxylation of lysine.

ΔS is not zero, and the result depends on the definition of the connectivity. This is clearly not what we expect from a correct partitioning, which we refer to as the *cancellation problem*.

How is it possible that ΔS is not zero, despite the only changes in the connectivity being completely in the QM region? From Table 2 we see that for partitioning 1, ΔS is zero for the stretching terms. Indeed, the stretching terms that are affected by the change in connectivity are identical and therefore cancel exactly in the real system and in the model system. However, the torsional terms, for example, have a nonzero ΔS for partitioning 1. Consider the torsional term for atoms H17−N19−H21−O24. This term does not exist in the model system, so its value does not cancel in the ONIOM expression. But whether this term is included in the energy expression (via the $E^{\mathrm{MM,Real}}$ term) depends on whether the bond between N19 and H21 is considered broken or formed. If the bond is considered broken, the term will not enter the energy expression. If the bond is considered formed, it will enter the ONIOM energy expression. The result is that for partitioning 1, the total ONIOM energy depends on the connectivity and is not continuous. Therefore, partitioning 1 is wrong and should not be used. Looking at the table, partitioning 2 still produces a small energy difference, which results from the nonbonded terms that involve the link atom. This difference is so small that in practice this partitioning can be used. Only the third partitioning generates exactly the same energies with both connectivity sets. Of course, this 'cancellation problem' is directly the result of MM terms extending over more than one bond. Since the 'largest' MM terms extend over three bonds (dihedral angle terms), the cancellation problem will never occur when there are more than three bonds between the MM region and any bond that is formed or broken during the reaction. However, avoiding the cancellation problem is only one requirement for a reliable potential. In most cases it will be necessary to include many more centers in the QM region in order to obtain meaningful results.

To summarize, when bond breaking and forming is part of the process in the QM region, it is safest to have the MM region at least three bonds away from the bond breaking/formation to avoid the cancellation problem. However, depending on the exact parameters involved in the MM terms in the border region, there are instances where one or two bond separations are sufficient. To test whether this is possible for a particular problem, once can apply the tests using the *S*-value as described in the paragraphs above. Although we discussed the problem with the cancellation using the ONIOM expression, it will occur with the generic QM/MM expression as well. The $E^{\mathrm{MM-only*model-only,MM}}$ term in eq 4 contains all the MM terms that have at least one center in the MM-only region and at least one center in the model-only region. Again, which terms are included in $E^{\mathrm{MM-only*model-only,MM}}$ depends on the connectivity, and the potential surface may be discontinuous when bond breaking and formation takes place close to the MM region.

2.6. Electronic Embedding. In QM/MM methods, there are two choices for dealing with the electrostatic interactions between the QM layer and the MM layer. The first, *classical embedding* or *mechanical embedding*, treats the cross-region electrostatic interactions at the molecular mechanics level. The second, *electronic embedding*, incorporates the cross-region electrostatic interaction in the QM Hamiltonian. The latter avoids the approximation of the QM charge distribution by point charges and allows the wave function to be polarized by the charge distribution of the MM region. From the original formulation, as outlined in the previous sections, it follows that ONIOM uses mechanical embedding by default. In this section we will present the modification of the ONIOM scheme to include electronic embedding. We also present the electronic embedding version of the generic QM/MM method and compare the two schemes in the results section.

To illustrate the different embedding approaches, we use the deprotonation of Histidine in Figure 6 as an example. The ONIOM expression contains two molecular mechanics terms, of which $E^{\mathrm{model,MM}}$ includes the electrostatic interaction for the QM region, while $E^{\mathrm{real,MM}}$ includes the electrostatic interaction for the *full* system. The latter includes the electrostatic interactions between atoms within the MM region, atoms within the QM region, and atoms in the QM region with atoms in the MM region. Electrostatic interactions between atoms that are separated by three bonds or less are scaled according to the MM force field definition, because they are (partially) implicit in the stretch, bend, and

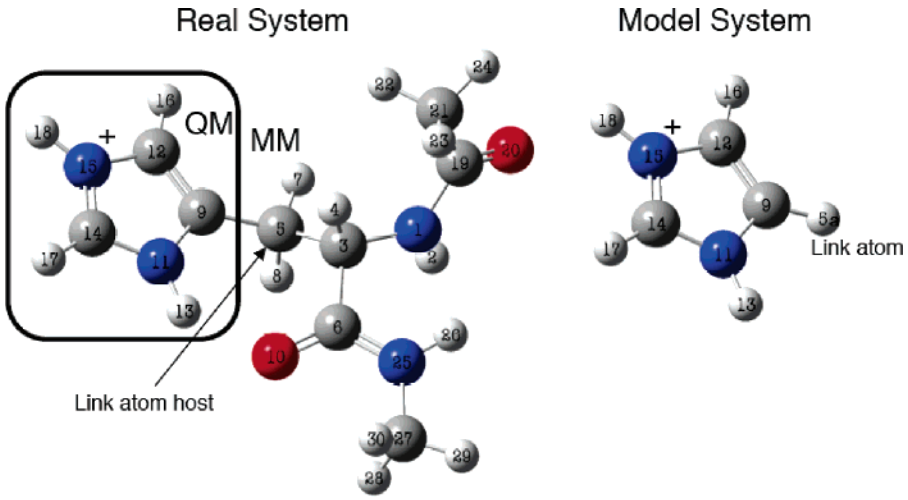


Figure 6. Real system, QM/MM partitioning, and model system for the (H18) deprotonation of histidine.

Table 3: Inclusion of Specific Interactions in the Standard (Mechanical Embedding) ONIOM(QM:MM) Energy Expression

centers	included in MM model system?	included in MM real system?	in QM model?
N1−C6	no (centers not in model)	no (2 bond separation)	no
N1−N25	no (centers not in model)	scaled (3 bond separation)	no
N1−C27	no (centers not in model)	yes	no
C3−C9	no (C3 not in model)	no (2 bond separation)	no
C3−C12	no (C3 not in model)	scaled (3 bond separation)	no
C3−N15	no (C3 not in model)	yes	no
C6−C9	no (C6 not in model)	scaled (3 bond separation)	no
C6−N11	no (C6 not in model)	yes	no
C6−C14	no (C6 not in model)	yes	no
C6−H17	no (C6 not in model)	yes	no
H16−C9	no (2 bond separation)	no (2 bond separation)	yes
H16−N11	scaled (3 bond separation)	scaled (3 bond separation)	yes
H16−H13	yes	yes	yes
C5−N11	no (C5 not in model system)	no (2 bond separation)	no
C5−C14	no (C5 not in model system)	scaled (3 bond separation)	no
C5−H17	no (C5 not in model system)	yes	no
H5a−N11	no (2 bond separation)	no (H5a not in real system)	yes
H5a−C14	scaled (3 bond separation)	no (H5a not in real system)	yes
H5a−H17	yes	no (H5a not in real system)	yes

torsional terms (For example, Amber uses a factor of zero for one and two bond separated electrostatic interactions and a factor of 1/1.2 for three bond separated interactions.). Using Figure 6, we show in Table 3 a number of specific interactions to illustrate which terms are included.

Because the interactions between QM and QM atoms (for example between H16 and H13) are included in *both* the $E^{\text{model,MM}}$ and the $E^{\text{real,MM}}$ terms, they cancel in the ONIOM energy expression. It follows that the only electrostatic interactions at the MM level that are retained in eq 1 are those between MM atoms and MM atoms and those between MM atoms and QM atoms. Hence, the electrostatic interaction between the two layers is described by the MM component of the energy expression, which is referred to as mechanical embedding. The point-charge electrostatic interactions involving LA and LAH are a special case. The interactions of the QM atoms with both LAH (C5) and LA (H5a) are retained but with different value, sign, and position. Their difference represents the extrapolation of the hydrogen link atom in the QM calculation to the carbon atom as it is in the real system, similar to the bonded terms involving LA and LAH as discussed in the previous sections.

We will first discuss how the generic QM/MM scheme is extended to electronic embedding. The QM/MM energy expression 4 is modified to

$$E^{\text{QM/MM-EE}} = E^{\text{MM-only,MM}} + E_{\text{v}}^{\text{model,QM}} + E_{\text{noQ}}^{\text{MM-only*model-only,MM}} \quad (9)$$

where

$$\hat{H}_{\text{v}}^{\text{model,QM}} = \hat{H}^{\text{model,QM}} - \sum_i \sum_N \frac{s_N q_N}{r_{iN}} + \sum_J \sum_N \frac{Z_J s_N q_N}{r_{JN}} \quad (10)$$

N, J, and i refer to the atoms from the MM region, atoms from the QM region, and electrons, respectively. The subscript noQ indicates that the electrostatic terms are excluded. The scaling factor s_N is used to avoid overpolarization of the wave function due to large charges close to the QM region. Usually s_N is zero for charges less than three

Table 4: Inclusion of Specific Interactions in the Generic Electronic Embedding QM/MM Energy Expression, with the Charges on C3 and C5 Scaled to Zero in the Model System Calculations[a]

centers	included in MM model system?	included in MM real system?	in QM model?
N1−C6	no (centers not in model)	no (2 bond separation)	no
N1−N25	no (centers not in model)	scaled (3 bond separation)	no
N1−C27	no (centers not in model)	yes	no
C3−C9	*no (C3 scaled to zero)*	*no (excluded)*	*no (C3 scaled)*
C3−C12	*no (C3 scaled to zero)*	*no (excluded)*	*no (C3 scaled)*
C3−N15	*no (C3 scaled to zero)*	*no (excluded)*	*no (C3 scaled)*
C6−C9	*yes*	*no (excluded)*	*yes*
C6−N11	*yes*	*no (excluded)*	*yes*
C6−C14	*yes*	*no (excluded)*	*yes*
C6−H17	*yes*	*no (excluded)*	*yes*
H16−C9	no (2 bond separation)	no (2 bond separation)	yes
H16−N11	scaled (3 bond separation)	scaled (3 bond separation)	yes
H16−H13	yes	yes	yes
C5−N11	*no (C5 scaled to zero)*	*no (excluded)*	*no (C5 scaled)*
C5−C14	*no (C5 scaled to zero)*	*no (excluded)*	*no (C5 scaled)*
C5−H17	*no (C5 scaled to zero)*	*no (excluded)*	*no (C5 scaled)*
H5a−N11	*no (excluded)*	no (H5a not in real system)	yes
H5a−C14	*no (excluded)*	no (H5a not in real system)	yes
H5a−H17	*no (excluded)*	no (H5a not in real system)	yes

[a] Entries in italics are different from those in Table 3.

bonds away from the QM region, and unit for the remaining charges. The scaling factor also avoids 'overcounting'. For example, for the C3−C5−C9 angle there are molecular mechanics bending (C3−C5−C9) and stretching terms (C3−C5 and C5−C9) in $E_{\mathrm{noQ}}^{\mathrm{MM-only*model-only,MM}}$ (see Table 4). The electrostatic and van der Waals interactions between these three centers are implicit in these bending and stretching terms and are therefore excluded from the list of nonbonded interactions in MM calculations. However, if the partial charge of C3 were included in the QM Hamiltonian, the electrostatic interaction of this center with the charge density of C9 would be fully included via the term $E_{\mathrm{v}}^{\mathrm{model,QM}}$. The electrostatic interaction would then be included twice: once through the C3−C5−C9 bending term and once through the inclusion of the charge on C3 in the Hamiltonian. We refer to this as 'overcounting', and scaling the charge on C3 to zero ensures that it does not take place. This solution, however, is not satisfactory, because scaling the charge on C3 to zero results in the exclusion of the electrostatic interaction of this center with all the other QM atoms (C14, N15, etc.). Since these interactions should in fact be included, we refer to this as 'undercounting'. In short, standard QM/MM will always result in either undercounting or overcounting, which is ultimately the result of the incompatibility between the QM charge density and the MM atom centered charges.

Some QM/MM implementations deal with the overpolarizaton and overcounting problem by using delocalized charges instead of point charges,[46,47] or by redistributing the charges close to the QM region.[48] For electronic embedding in ONIOM we follow an approach that differs from standard QM/MM schemes. Following the spirit of ONIOM, we perform the model system calculations on the same system, which includes the charges that come from the MM region. In the QM model system these point charges are then incorporated in the Hamiltonian, while in the MM model system calculation they are evaluated at the classical level. Furthermore, only the model system calculations are modified, while the real system MM term remains identical to that in the ONIOM-ME (ONIOM-Mechanical Embedding) expression 1. The expression for ONIOM-EE (ONIOM-Electronic Embedding) becomes

$$E^{\mathrm{ONIOM-EE}} = E_{\mathrm{v}}^{\mathrm{model,QM}} + E^{\mathrm{real,MM}} - E_{\mathrm{v}}^{\mathrm{model,MM}} \qquad (11)$$

where

$$E_{\mathrm{v}}^{\mathrm{model,MM}} = E^{\mathrm{model,MM}} + \sum_{J}\sum_{N}\frac{q_J s_N q_N}{r_{JN}} \qquad (12)$$

The QM calculation in eq 11 is identical to that in eq 9. Because the model systems must be identical, we use the same scale factor s_{N} in both the QM and the MM model system calculations.

It must be noted that the use of delocalized or redistributed charges, as in other QM/MM schemes, is not mutually exclusive with the ONIOM implementation of electronic embedding. Applying both methods simultaneously might provide a superior scheme. Furthermore, from a practical point of view, a major difference between the QM/MM and ONIOM electronic embedding schemes is that in the latter the user must specify charges for the QM region.

In Table 5 we give again the specific electrostatic interactions in the MM terms, with the entries in italic being different from the corresponding Table 3 for ONIOM-ME. We will now compare the specific electrostatic interactions between standard QM/MM-EE and ONIOM-EE. With QM/MM-EE, we saw that the interaction between C3 and N15 is undercounted because the charge on C3 is scaled to zero in the QM calculation, and the C3−N15 interaction is excluded from the MM calculations. In the ONIOM-EE scheme, however, this interaction is still present in the $E^{\mathrm{real,MM}}$

Table 5: Inclusion of Specific Interactions in the Electronic Embedding ONIOM(QM:MM) Energy Expression, with the Charges on C3 and C5 Scaled to Zero in the Model System Calculations[a]

centers	included in MM model system?	included in MM real system?	in QM model?
N1−C6	no (centers not in model)	no (2 bond separation)	no
N1−N25	no (centers not in model)	scaled (3 bond separation)	no
N1−C27	no (centers not in model)	yes	no
C3−C9	*no (C3 scaled to zero)*	no (2 bond separation)	*no (C3 scaled)*
C3−C12	*no (C3 scaled to zero)*	scaled (3 bond separation)	*no (C3 scaled)*
C3−N15	*no (C3 scaled to zero)*	yes	*no (C3 scaled)*
C6−C9	*yes*	scaled (3 bond separation)	*yes*
C6−N11	*yes*	yes	*yes*
C6−C14	*yes*	yes	*yes*
C6−H17	*yes*	yes	*yes*
H16−C9	no (2 bond separation)	no (2 bond separation)	yes
H16−N11	scaled (3 bond separation)	scaled (3 bond separation)	yes
H16−H13	yes	yes	yes
C5−N11	*no (C5 scaled to zero)*	no (2 bond separation)	*no (C5 scaled)*
C5−C14	*no (C5 scaled to zero)*	scaled (3 bond separation)	*no (C5 scaled)*
C5−H17	*no (C5 scaled to zero)*	yes	*no (C5 scaled)*
H5a−N11	no (2 bond separation)	no (H5a not in real system)	yes
H5a−C14	scaled (3 bond separation)	no (H5a not in real system)	yes
H5a−H17	yes	no (H5a not in real system)	yes

[a] Entries in italics are different from those in Table 3.

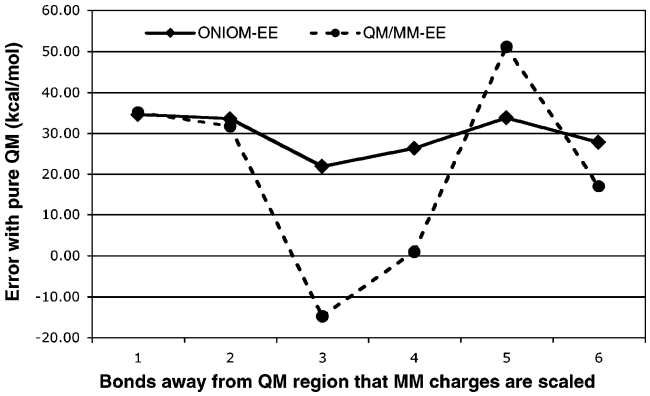


Figure 7. Error of ONIOM-EE and QM/MM-EE as a function of the charges scaled based on distance from the QM region.

term. The electrostatic interaction between C3 and N15 is therefore still included in the ONIOM expression, albeit at the classical level. Another example is the interaction between C6 and C9. The MM torsional term C9−C5−C3−C6 already contains implicitly part of the electrostatic interaction, which is why the three-bond separated nonbonded interactions are scaled. However, the C6−C9 electrostatic interaction is without scaling included in the QM term in equation $E_v^{model,QM}$. In QM/MM-EE this leads to overcounting, but in ONIOM-EE this particular interaction is included fully in the model system MM term $E_v^{model,MM}$ and scaled in the real system MM term $E^{real,MM}$. The difference between the MM terms can be regarded as a correction to the overcounting at the QM level. In other words, in ONIOM-EE the overcounting or undercounting introduced at the QM level is always corrected automatically at the classical level. This follows naturally from the ONIOM expressions and does not require any of the corrections that generic QM/MM-EE schemes need.

In Figure 7 we show the error of the ONIOM-EE and QM/MM-EE calculations on the deprotonation of histidine, compared to full QM calculations. Our focus is not on the absolute performance of the hybrid schemes but rather on the difference in behavior between ONIOM and the generic QM/MM method. The graph shows the error as a function of the number of bonds away from the QM region that have the charges scaled to zero. The charge on C5 is always scaled to zero. At $x = 1$, only charges one bond out (C5) are scaled. At $x = 2$, charges up to two bonds out (C5, H7, H8, and C3) are scaled, and so forth.

Even though the QM/MM-EE has the smallest absolute error (at x=4), its behavior is much more erratic than ONIOM-EE. The reason is that, in this particular example, the sign of the charges in the MM region is alternating with each step further away from the QM region and that therefore the total charge included in the set of (nonzero) point charges changes strongly with each increasing x-value. In ONIOM-EE the total charge is always the same, because the error at the QM level is corrected at the MM level, and the graph is much less erratic.

At $x = 6$, all the charges in the MM region are scaled to zero in the model system calculations, and ONIOM-EE becomes identical to ONIOM-ME. We see that the result is quite similar to the ONIOM-EE results, which indicates that the partial charges assigned to the QM region describe the real QM charge density quite well. Note that ONIOM-EE and QM/MM-EE are not identical to each other when all the charges are scaled to zero, at $x = 6$. In that case, QM/MM-EE completely ignores all the electrostatic interactions between the two layers, while QM/MM-EE becomes ONIOM-ME and still includes the electrostatic interactions at the classical level, through the $E^{real,MM}$ calculation.

Finally, we want to stress that the example above is intended primarily to demonstrate that correct charge balancing follows naturally from the electronic embedding version of ONIOM. Our generic QM/MM implementation is crude in many ways, and the unfavorable comparison to ONIOM-

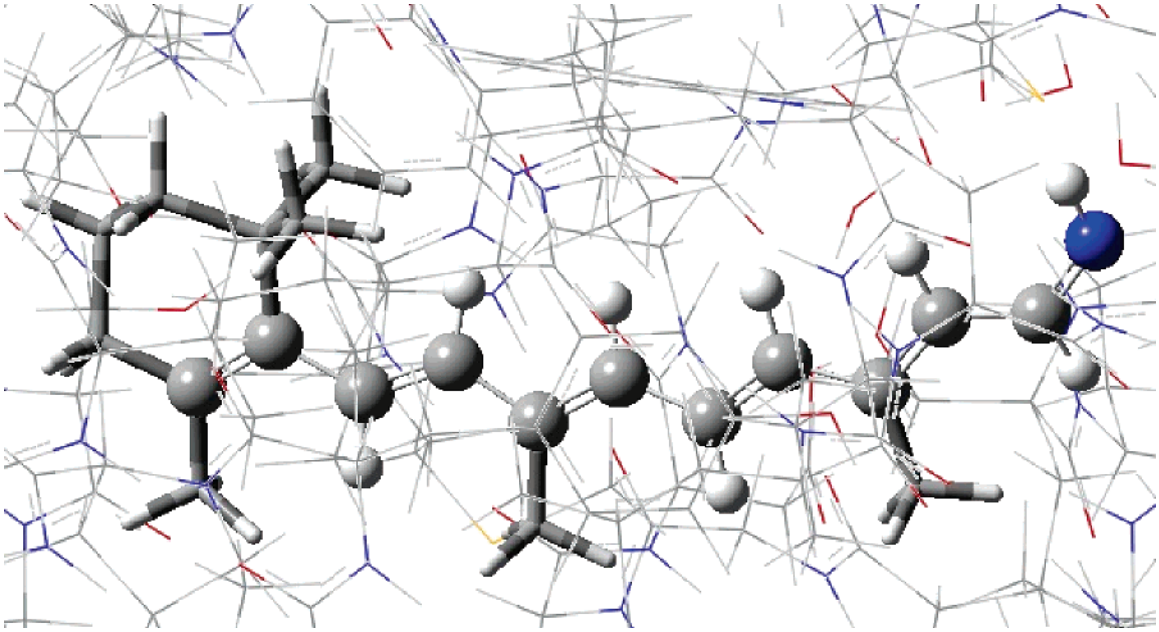


Figure 8. ONIOM3 partitioning of bR. Ball and stick represents B3LYP, tube HF, and wireframe the Amber molecular mechanics force field.

Table 6: Excitation Energies of BR for Various Partitionings and Method Combinations

	real (kcal/mol)	model-high (kcal/mol)	model-low (kcal/mol)	ONIOM (kcal/mol)	time (min)
HF:Amber-EE	81.45			81.45	39
B3:Amber-EE (benchmark)	58.25			58.25	318
B3:HF:Amber-EE	69.70	62.80	71.24	61.26	109
B3:HF:Amber-EEx	72.33	60.78	69.62	63.50	102

EE in Figure 7 is not typical for state-of-the-art QM/MM implementations, which have incorporated other methods for dealing with the charge-balancing. These methods, however, are complimentary with ONIOM, and using both approaches simultaneously may provide the superior QM/MM scheme.

2.7. Three-Layer Electronic Embedding. In ONIOM-(QM-high:QM-low:MM)-EE, we need to decide which charges need to be included in each of the four QM model system calculations. In the calculations on the intermediate model system, we always include the MM charge density. In the calculation on the small model system, we have two choices. First, we can carry out the calculations on the small model system with the same charge cloud as the intermediate model system calculations. We denote this as ONIOM3-EE. Note that in this case, only the partial charges of the MM layer are included and not the charges associated with the intermediate layer. Second, we can carry out the calculations on the small model system without any charges, assuming that the effect of the charge cloud is accurately included via the calculations on the intermediate model system. This is denoted as ONIOM3-EEx.

$$E^{\mathrm{ONIOM3-EE}} = E_{\mathrm{v}}^{\mathrm{model,QM-high}} + E_{\mathrm{v}}^{\mathrm{intermediate,QM-low}} - E_{\mathrm{v}}^{\mathrm{model,QM-low}} + E^{\mathrm{real,MM}} - E_{\mathrm{v}}^{\mathrm{intermediate,MM}} \quad (13)$$

$$E^{\mathrm{ONIOM3-EEx}} = E^{\mathrm{model,QM-high}} + E_{\mathrm{v}}^{\mathrm{intermediate,QM-low}} - E^{\mathrm{model,QM-low}} + E^{\mathrm{real,MM}} - E_{\mathrm{v}}^{\mathrm{intermediate,MM}} \quad (14)$$

It depends on the problem whether ONIOM3-EE or ONIOM3-EEx is the best choice. The QM-high layer can be far away from the MM layer (with the QM-low layer acting as buffer). If the MM density is included in the model system calculations, following ONIOM3-EE, the QM-high layer 'feels' the full charge density. However, in reality the charges from the MM region will be screened by the QM-low layer. If this screening is important, ONIOM3-EE will be unphysical. Excluding the MM charges from the model system calculation, in ONIOM3-EEx, is not always correct either, especially when the QM-low method cannot describe accurately the effect of the MM charge distribution on the QM-high layer.

To illustrate the three-layer version of ONIOM with electronic embedding, we use the $S_0 \rightarrow S_1$ excitation in bacteriorhodopsin (bR) from ref 11. We include the complete chromophore in the intermediate model system, while the chromophore without the substituents (only the Schiff base polyene) is in the small model system (Figure 8). The levels of theory used are B3LYP/6-31G(d) for QM-high and HF/3-21G for QM-low and TD-B3LYP and TD-HF for the excitation energies. We calculated the HF:Amber and B3LYP:HF:Amber excitations on the geometries optimized at the specific level of theory.

In Table 6 we show the excitation energies for the three-layer ONIOM calculations as well as two-layer ONIOM calculations. In the latter, the full chromophore is included in the QM region. The two-layer B3LYP:Amber calculation then forms the benchmark to which we compare the results.

The results are clear. B3LYP:HF:Amber does always a better job than the two-layer HF:Amber. In addition, for this example it does appear necessary to include the charge cloud also in the smallest model system, although the difference is small. In the three-layer calculation, we obtain virtually identical results, for about a third of the CPU time as needed in the two-layer calculation.

3. Conclusions

We presented several developments of the ONIOM(QM:MM) method and compared the behavior and performance to a generic QM/MM scheme. The electronic embedding version of ONIOM appears to be more stable than QM/MM with electronic embedding. However, we implemented only the simple charge zeroing version of the latter, and QM/MM might compare more favorably to ONIOM when more sophisticated methods for avoiding overpolarization are implemented, such as delocalized charges or charge redistribution. Combining these methods with the ONIOM scheme may present a superior scheme for electronic embedding. Furthermore, we showed that the link atom correction that is implicit in ONIOM and omitted from simple QM/MM schemes can improve the results significantly.

We analyzed in detail the occurrence of discontinuities in the ONIOM(QM:MM) energy when bond breaking and formation is involved. The 'better safe than sorry' approach is to ensure that there are at least three bonds between the MM region and the bonds in the QM region that are being broken or formed. In some cases it may be possible to have only one or two bonds, but careful testing is required to ensure that the potential is correct.

To conclude, we feel that the extrapolation nature of ONIOM provides elegant and implicit solutions to some problems, especially related to charge balancing, that require ad-hoc corrections in most other QM/MM schemes. Some of the solutions that follow from the ONIOM method can be incorporated in other QM/MM schemes. In turn, many of the promising developments in other QM/MM methods can be formulated in the ONIOM framework as well.

Acknowledgment. K.M. acknowledges support in part by a grant (CHE-0209660) from the U.S. National Science Foundation and by a grant from Gaussian, Inc.

References

(1) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227−249.

(2) Maseras, F.; Morokuma, K. *J. Comput. Chem.* **1995**, *16*, 1170−1179.

(3) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700−733.

(4) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1986**, *7*, 718−730.

(5) Dapprich, S.; Komáromi, I.; Byun, K. S.; Morokuma, K.; Frisch, M. J. *J. Mol. Struct. (THEOCHEM)* **1999**, *461*−*462*, 1−21.

(6) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357−19363.

(7) Vreven, T.; Mennucci, B.; da Silva, C. O.; Morokuma, K.; Tomasi, J. *J. Chem. Phys.* **2001**, *115*, 62−72.

(8) Karadakov, P. B.; Morokuma, K. *Chem. Phys. Lett.* **2000**, *317*, 589−596.

(9) Rega, N.; Iyengar, S. S.; Voth, G. A.; Schlegel, H. B.; Vreven, T.; Frisch, M. J. *J. Phys. Chem. B* **2004**, *108*, 4210−4220.

(10) Humbel, S.; Sieber, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 1959−1967.

(11) Vreven, T.; Morokuma, K. *Theor. Chem. Acc.* **2003**, *109*, 125−132.

(12) Vreven, T.; Morokuma, K. *J. Comput. Chem.* **2000**, *21*, 1419−1432.

(13) Hopkins, B. W.; Tschumper, G. S. *J. Comput. Chem.* **2003**, *24*, 1563.

(14) Warshel, A. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 425−443.

(15) Ujaque, G.; Maseras, F. *Structure Bonding* **2004**, *112*, 117−149.

(16) Gao, J. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH Publishers: New York, 1996; Vol. 7, pp 119−301.

(17) Froese, R. D. J.; Morokuma, K. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Allinger, N. L., Kollman, P. A., Clark, T., Schaefer, H. F., III, Gasteiger, J., Schreiner, P. R., Eds.; Wiley: Chichester, 1998; Vol. 2, pp 1244−1257.

(18) Gao, J. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Allinger, N. L., Kollman, P. A., Clark, T., Schaefer, H. F., III, Gasteiger, J., Schreiner, P. R., Eds.; Wiley: Chichester, 1998; Vol. 2, pp 1257−1263.

(19) Merz, K. M.; Stanton, R. V. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Allinger, N. L., Kollman, P. A., Clark, T., Schaefer, H. F., III, Gasteiger, J., Schreiner, P. R., Eds.; Wiley: Chichester, 1998; Vol. 4, pp 2330−2343.

(20) Ruiz-López, M. F.; Rivail, J. L. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Allinger, N. L., Kollman, P. A., Clark, T., Schaefer, H. F., III, Gasteiger, J., Schreiner, P. R., Eds.; Wiley: Chichester, 1998; Vol. 1.

(21) Tomasi, J.; Pomelli, C. S. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Allinger, N. L., Kollman, P. A., Clark, T., Schaefer, H. F., III, Gasteiger, J., Schreiner, P. R., Eds.; Wiley: Chichester, 1998; Vol. 4, pp 2343−2350.

(22) Vreven, T.; Morokuma, K. *Ann. Rep. Comput. Chem.*, in press.

(23) Derat, E.; Bouquant, J.; Humbel, S. *J. Mol. Struct. (THEOCHEM)* **2003**, *632*, 61−69.

(24) Thery, V.; Rinaldi, D.; Rivail, J. L.; Maigret, B.; Ferenczy, G. G. *J. Comput. Chem.* **1994**, *15*, 269−282.

(25) Reuter, N.; Dejaegere, A.; Maigret, B.; Karplus, M. *J. Phys. Chem. A* **2000**, *104*, 1720−1735.

(26) Gao, J.; Amara, P.; Alhambra, C.; Field, M. J. *J. Phys. Chem. A* **1998**, *102*, 4714−4721.

(27) Nicoll, R. M.; Hindle, S. A.; MacKenzie, G.; Hillier, I. H.; Burton, N. A. *Theor. Chem. Acc.* **2000**, *106*, 105−112.

(28) Koga, N.; Morokuma, K. *Chem. Phys. Lett.* **1990**, *172*, 243−248.

(29) Antes, I.; Thiel, W. *J. Phys. Chem. A* **1999**, *103*, 9290−9295.

(30) Zhang, Y. K.; Lee, T. S.; Yang, W. T. *J. Chem. Phys.* **1999**, *110*, 46−54.

(31) DiLabio, G. A.; Hurley, M. M.; Christiansen, P. A. *J. Chem. Phys.* **2002**, *116*, 9578−9584.

(32) Bakowies, D.; Thiel, W. *J. Phys. Chem.* **1996**, *100*, 10580−10594.

(33) Vreven, T. Manuscript in preparation.

(34) Vreven, T. Manuscript in preparation.

(35) Vreven, T.; Frisch, M. J.; Kudin, K. N.; Schlegel, H. B.; Morokuma, K. *Mol. Phys.* **2006**, *104*, 701−714.

(36) Vreven, T.; Morokuma, K.; Farkas, Ö.; Schlegel, H. B.; Frisch, M. J. *J. Comput. Chem.* **2003**, *24*, 760−769.

(37) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03 Revision C.02 ed.*; Gaussian, Inc.: Wallingford, CT, 2004.

(38) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian Development Version Revision C.01 ed.*; Gaussian, Inc.: Wallingford, CT, 2004.

(39) Swart, M. *Int. J. Quantum Chem.* **2002**, *91*, 177−183.

(40) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.

(41) Kudin, K. N.; Scuseria, G. E. *Chem. Phys. Lett.* **1998**, *283*, 61−68.

(42) Strain, M. C.; Scuseria, G. E.; Frisch, M. J. *Science* **1996**, *271*, 51.

(43) Greengard, L. *The rapid evaluation of potential fields in particle systems*; MIT Press: Cambridge, MA, 1988.

(44) Greengard, L.; Rokhlin, V. *J. Comput. Phys.* **1987**, *73*, 325.

(45) Strajbl, M.; Hong, G.; Warshel, A. *J. Phys. Chem. B* **2002**, *106*, 13333−13343.

(46) Amara, P.; Field, M. J. *Theor. Chem. Acc.* **2003**, *109*, 43−52.

(47) Das, D.; Eurenius, K. P.; Billings, E. M.; Sherwood, P.; Chatfield, D. C.; Hodoscek, M.; Brooks, B. R. *J. Chem. Phys.* **2002**, *117*, 10534−10547.

(48) Lin, H.; Trulahr, D. G. *J. Phys. Chem. A* **2005**, *109*, 3991−4004.

CT050289G

JCTC Journal of Chemical Theory and Computation

Rotational *g* Tensors Calculated Using Hybrid Exchange-Correlation Functionals with the Optimized Effective Potential Approach

Ola B. Lutnæs,[†] Andrew M. Teale,[‡] Trygve Helgaker,[†] and David J. Tozer*[,‡]

Department of Chemistry, University of Oslo, P.O. Box 1033, Blindern, N-0315 Oslo, Norway, and Department of Chemistry, University of Durham, South Road, Durham, DH1 3LE United Kingdom

Received February 2, 2006

Abstract: The calculation of rotational *g* tensors using density functional theory (DFT) with hybrid exchange-correlation functionals is considered. A total of 143 rotational *g* tensor elements in 58 molecules (67 isotopic combinations) are calculated using three standard hybrid functionals. Tensor elements determined using an uncoupled approach with orbitals and eigenvalues calculated from the multiplicative optimized effective potential (OEP) constitute a significant improvement over those determined in the conventional coupled manner with a nonmultiplicative exchange-correlation operator. Relative to experimental results, mean absolute errors are reduced by a factor of 2; mean errors and standard deviations are reduced by more than a factor of 3. The results are also an improvement over those determined using a generalized gradient-approximation functional optimized for magnetic response properties. The influence of orbital exchange is investigated for a representative subset of molecules, yielding an optimal amount near 0.3. Rotational *g* tensors are also determined from coupled-cluster electron densities using a combined DFT/wave-function approach. Being substantially more expensive, they do not offer a notable improvement on the pure DFT values from OEP-based hybrid calculations.

1. Introduction and Background

A rotating molecule acquires a magnetic moment proportional to its angular momentum. In an external magnetic field, the Zeeman interaction between this magnetic moment and the external magnetic induction causes a shift in the rotational energy levels, as observed by molecular beam[1] and microwave Zeeman experiments.[2,3] In atomic units (as used throughout this paper), the energy shift is conventionally expressed as

$$\Delta E = -\mu_N \mathbf{B}^T \mathbf{g} \mathbf{J} \tag{1}$$

where μ_N is the nuclear magneton, **B** is the external magnetic induction, **J** is the angular momentum of the molecule with respect to its center of mass, and **g** is the dimensionless 3×3 rotational *g* tensor. The *g* tensor can be evaluated as the second derivative of the electronic energy *E* with respect to **B** and **J**:

$$\mathbf{g} = -\frac{1}{\mu_N} \frac{\partial^2 E(\mathbf{B},\mathbf{J})}{\partial \mathbf{B} \partial \mathbf{J}}\bigg|_{\mathbf{B},\mathbf{J}=0} \tag{2}$$

The calculation of *g* tensors is thus similar to that of other singlet second-order magnetic response properties such as nuclear magnetic resonance (NMR) shielding constants, spin-rotation constants, and magnetizabilities. Indeed, in the center-of-mass coordinate system,[4] the *g* tensor can be determined directly from the magnetizability

$$\boldsymbol{\xi} = -\frac{\partial^2 E(\mathbf{B})}{\partial \mathbf{B}^2}\bigg|_{\mathbf{B}=0} \tag{3}$$

* Corresponding author fax: 0191 384 4737; e-mail: D.J.Tozer@Durham.ac.uk.
[†] University of Oslo.
[‡] University of Durham.

10.1021/ct060038n CCC: $33.50
Published on Web 04/15/2006

using the relation

$$\mathbf{g} = -\frac{2}{\mu_N}\xi_{CM}^{para}\,\mathbf{I}_{nuc}^{-1} + \frac{1}{2\mu_N}\sum_K Z_K\,(\mathbf{R}_K^T\mathbf{R}_K\mathbf{I}_3 - \mathbf{R}_K\mathbf{R}_K^T)\mathbf{I}_{nuc}^{-1} \quad (4)$$

where ξ_{CM}^{para} is the paramagnetic contribution to the magnetizability calculated with the gauge origin at the center of mass, $\mathbf{I}_{nuc}$ is the moment-of-inertia tensor, and Z_K is the charge of nucleus K at position $\mathbf{R}_K$ relative to the center of mass. This relation arises because the paramagnetic contributions to the g tensor and to the magnetizability (with the gauge origin at the center of mass) are proportional to each other, $\mathbf{g}^{para} = -2\xi_{CM}^{para}\mathbf{I}_{nuc}^{-1}/\mu_N$, as follows by comparing the first-order electronic perturbations generated by the external magnetic field $\mathbf{B}$ and by the rotationally induced magnetic field $-2\mathbf{I}_{nuc}^{-1}\mathbf{J}$. The paramagnetic contribution ξ_{CM}^{para} is straightforwardly obtained from the total magnetizability (calculated with London orbitals) ξ^{LAO}, by subtracting the diamagnetic (expectation value) contribution ξ_{CM}^{dia}. For further details regarding the theory and computation of rotational g tensors, see ref 4.

The high accuracy obtained in the experimental determination of rotational g tensors makes them an excellent test case for quantum-chemical theories. Indeed, a wide range of electronic-structure methods have been used to calculate g tensors, including Hartree−Fock (HF) theory;[4−7] multiconfigurational self-consistent field (MCSCF) theory;[8−15] Møller−Plesset theory to second,[14,16,17] third,[14,16] and fourth orders[14] (MP2, MP3, and MP4); linearized coupled-cluster doubles (L-CCD) theory;[16] coupled-cluster singles-and-doubles (CCSD) theory;[14] the second-order polarization propagator approximation (SOPPA);[14,18−20] the coupled-cluster polarization propagator approximation (CCSDPPA);[21−24] the SOPPA using CCSD amplitudes [SOPPA(CCSD)];[14,25] full configuration-interaction (FCI) theory;[26−28] and density functional theory (DFT).[13,29,30] DFT is particularly attractive because of its low computational cost. Recently, Wilson et al.[30] presented an extensive assessment of DFT rotational g tensors, using a range of exchange-correlation functionals. They considered the local density approximation (LDA), the Becke−Lee−Yang−Parr (BLYP)[31] generalized gradient approximation (GGA), and the Becke−3-parameter−Lee−Yang−Parr (B3LYP)[32] hybrid functional. They also considered the Keal−Tozer (KT2) GGA functional,[33] which was specifically designed to yield good quality magnetic response parameters. In particular, this functional has been shown to provide improved magnetizabilities compared to conventional approximations.[33] For rotational g tensors, Wilson et al.[30] observed that the quality improved in the order LDA < BLYP < B3LYP < KT2, reflecting the close relationship between the rotational g tensor and the magnetizability (see eq 4).

For the LDA and GGA functionals, the Kohn−Sham orbitals $\{\varphi_p(\mathbf{r})\}$ and eigenvalues $\{\epsilon_p\}$—the key quantities in the calculation of DFT response properties—are determined from the Kohn−Sham equations

$$\left[-\frac{1}{2}\nabla^2 + v_s(\mathbf{r})\right]\varphi_p(\mathbf{r}) = \epsilon_p\,\varphi_p(\mathbf{r}) \quad (5)$$

where $v_s(\mathbf{r})$ is a multiplicative one-electron potential

$$v_s(\mathbf{r}) = v_{ext}(\mathbf{r}) + v_J(\mathbf{r}) + v_{xc}(\mathbf{r}) \quad (6)$$

comprising external, Coulomb, and exchange-correlation components, respectively. The LDA and GGA magnetic Hessian, which determines the imaginary component of the orbital response, is diagonal, and so NMR shielding and spin-rotation constants as well as magnetizabilities and rotational g tensors are determined in an uncoupled manner. The whole approach is rigorously Kohn−Sham theory.

By contrast, for hybrid functionals such as B3LYP, which combines a GGA functional with an amount ξ of exact orbital exchange, the standard implementation in all widely used programs is not strictly Kohn−Sham theory. Specifically, rather than containing a multiplicative exchange-correlation potential as in eq 5, the orbital equations involve the scaled nonmultiplicative Hartree−Fock exchange operator

$$\left[-\frac{1}{2}\nabla^2 + v_s(\mathbf{r})\right]\varphi_p(\mathbf{r}) - \xi\int\frac{\rho_1(\mathbf{r},\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,\varphi_p(\mathbf{r}')\,d\mathbf{r}' = \epsilon_p\,\varphi_p(\mathbf{r}) \quad (7)$$

where $\rho_1(\mathbf{r},\mathbf{r}')$ is the one-particle density matrix and the $v_{xc}(\mathbf{r})$ part of $v_s(\mathbf{r})$ (see eq 6) is the multiplicative potential associated with the GGA component of the functional. The resulting magnetic Hessian is nondiagonal, and so magnetic properties are determined in a coupled manner. This was the approach used for the B3LYP rotational g tensor calculations of Wilson et al.[30]

There has recently been much interest[34−42] in determining uncoupled second-order magnetic properties from hybrid functionals using a rigorous Kohn−Sham equation of the form of eq 5. The high-quality results obtainable from such an approach were first demonstrated for shielding constants—Wilson and Tozer[34] determined multiplicative potentials from electron densities with the Zhao−Morrison−Parr (ZMP) approach[43] and used the resulting orbitals and eigenvalues to determine uncoupled shieldings. When hybrid functional DFT densities were used, the ZMP results for a series of small, highly correlated molecules were two to three times more accurate than those of conventional hybrid theory; the results were also a significant improvement on the uncoupled values of conventional GGA theory. Cohen et al.[37] subsequently demonstrated that essentially identical shieldings are obtained when the potentials are determined using the rigorous optimized effective potential (OEP) approach.[44] Similar improvements have been observed for chemical shifts[35] and magnetizabilities,[36] and so analogous improvements are also anticipated for rotational g tensors. See refs 38−42 for more recent studies using alternative potential constructions.

The aim of the present study is to quantify the improvement in rotational g tensors. Specifically, we shall use the OEP method to determine multiplicative potentials associated with the B3LYP, B97-2,[45] and B97-3[46] hybrid functionals and then use the orbitals and eigenvalues from a rigorous Kohn−Sham equation of the form of eq 5 to determine uncoupled rotational g tensors. The results will be compared

with the conventional g tensors, determined in a coupled manner with a nonmultiplicative exchange-correlation operator as in eq 7, as well as with the uncoupled results from the KT2 GGA functional.[33] All results will be compared with experimental values, mentioning, where possible, the effect of zero-point vibrational corrections. We shall also investigate the influence of orbital exchange and determine rotational g tensors directly from ab initio coupled-cluster electron densities using the Wu−Yang (WY) approach.[47] We commence by summarizing the OEP and WY approaches and providing some computational details in section 2. The results of our calculations are presented in section 3. Conclusions are presented in section 4.

2. Theory and Computational Details

The OEP method[44] is the rigorous approach for handling orbital-dependent (e.g., hybrid) functionals in Kohn−Sham theory. The aim of the OEP method is to find the multiplicative potential, $v_s(\mathbf{r})$, in a rigorous Kohn−Sham equation of the form of eq 5, that minimizes the total electronic energy for the chosen exchange-correlation functional. We have implemented the OEP method, using the approach of Yang and Wu,[48,49] in the CADPAC quantum-chemistry code.[50] Specifically, the potential is written

$$v_s(\mathbf{r}) = v_{ext}(\mathbf{r}) + v_0(\mathbf{r}) + \sum_t b_t\, g_t(\mathbf{r}) \tag{8}$$

where $v_{ext}(\mathbf{r})$ is the external potential, $v_0(\mathbf{r})$ is a fixed reference potential, and the final term is an expansion in an auxiliary basis of Gaussian functions $g_t(\mathbf{r})$. The only unknown parameters are the coefficients $\{b_t\}$, which are determined by a direct minimization of the total electronic energy with respect to these parameters, using an approximate Newton scheme.[49] For the reference potential, we use the Fermi−Amaldi potential constructed using the Hartree−Fock density of the system,

$$v_0(\mathbf{r}) = \left(1 - \frac{1}{N}\right) v_J^{HF}(\mathbf{r}) \tag{9}$$

where N is the number of electrons and $v_J^{HF}(\mathbf{r})$ is the Coulomb potential of the Hartree−Fock density.

While the OEP method provides the rigorous scheme for orbital-dependent functionals in DFT, the constrained-search formulation[51] provides an alternative method for determining multiplicative exchange-correlation potentials. In this method, the noninteracting kinetic energy is minimized subject to the constraint that the orbitals yield a supplied electron density. Wu and Yang have proposed a particular method[47] for such calculations, denoted WY, which is computationally similar to the OEP method and formally equivalent to the ZMP method. We have also implemented this approach in CADPAC. For a given input density $\rho_{in}(\mathbf{r})$, the solution to the problem can be recast as the unconstrained maximization of the functional

$$W_s[\Phi_{det}, v(\mathbf{r})] = 2\sum_i^{N/2}\left\langle \varphi_i \middle| -\frac{1}{2}\nabla^2 \middle| \varphi_i \right\rangle + \int d\mathbf{r}\, \{v_{ext}(\mathbf{r}) + v_0(\mathbf{r})\}\{\rho(\mathbf{r}) - \rho_{in}(\mathbf{r})\} + \int d\mathbf{r} \sum_t b_t\, g_t(\mathbf{r})\, \{\rho(\mathbf{r}) - \rho_{in}(\mathbf{r})\} \tag{10}$$

with respect to $\{b_t\}$, where the orbitals are obtained from a Kohn−Sham equation of the form of eq 5 with the potential defined as in eq 8 and $\rho(\mathbf{r})$ is the density constructed from these orbitals. The maximization is carried out using a quasi-Newton algorithm similar to the one used for the OEP approach. For further details, see ref 47.

The WY method can be applied to a DFT density associated with a hybrid exchange-correlation functional. As demonstrated by Cohen et al.,[37] the resulting WY shielding constants are very close to those obtained by applying the OEP method to the same functional. However, the WY method has the important advantage that it does not require explicit knowledge of the exchange-correlation functional associated with the input density. In the present work, we shall therefore also calculate g tensors using potentials determined from high-quality Brueckner-doubles (BD) coupled-cluster densities.

A key feature of the OEP and WY schemes is that $v_s(\mathbf{r})$ is expressed analytically through eq 8. Once optimal $\{b_t\}$ coefficients have been determined, it is therefore straightforward to reconstruct this potential in any electronic-structure program. Specifically, we here determine these coefficients using CADPAC and then read the values into a modified version of the DALTON program,[52] which assembles the potential using the same geometry, reference potential, and orbital and auxiliary basis sets. The resulting Kohn−Sham equations are then solved [by a single diagonalization because $v_s(\mathbf{r})$ does not depend on the orbitals], and the orbitals and eigenvalues are used to compute the uncoupled rotational g tensors.

The main advantage of this two-code approach is that it allows us to use rotational London atomic orbitals in the g-tensor evaluation to ensure fast basis-set convergence and gauge-independent results,[4] noting that such orbitals are essential for molecules of the size considered in this work and that they are not available in the CADPAC program. The approach also allows the enhanced property calculations of DALTON to be exploited in future investigations. Our implementation was checked by confirming that the converged one-electron eigenvalues in DALTON agreed with those of the CADPAC OEP/WY calculation, to within numerical integration error. Property calculations were checked by comparing single-origin magnetizability calculations in DALTON and CADPAC. The London-orbital calculations were checked by confirming gauge-origin invariance and by comparing large-basis single-origin magnetizabilities with London-orbital magnetizabilities for small molecules.

We closely follow the study of Wilson et al.,[30] using the same aug-cc-pVTZ basis[53−55] (but with Cartesian rather than spherical functions) and the same molecules, omitting the largest four because of computational limitations. The 58

Table 1. Molecules Considered in This Study

carbon monoxide [CO], carbon sulfide [CS], carbon selenide [CSe], hydrogen cyanide [HC^{15}N], fluoro cyanide [FC^{15}N], chloro cyanide [ClC^{15}N], bromo cyanide [BrC^{15}N], carbonyl sulfide [OCS, OC^{34}S, O^{13}CS], nitrous oxide [^{15}N^{15}NO, ^{14}N^{14}NO], carbonyl selenide [OC^{80}Se, OC^{76}Se], methylidene phosphine [HCP, DCP], hydrogen boron sulfide [HBS], fluoro acetylene [FCCH], chloro acetylene [^{35}ClCCH, ^{37}ClCCH], bromo acetylene [^{79}BrCCH, ^{81}BrCCH], ammonia [^{15}NH$_3$], trifluoromethane [CHF$_3$], methylisocyanide [CH$_3{}^{14}$NC], acetonitrile [CH$_3$C^{15}N, CH$_3$C^{14}N, CD$_3$C^{14}N], fluoromethane [CH$_3$F], chloromethane [CH$_3$Cl], acrolein [CH$_2$CHCHO], propene [CH$_2$CHCH$_3$], propynal [HCCCHO], dimethyl ether [CH$_3$OCH$_3$], dimethylsulfane [CH$_3$SCH$_3$], acetaldehyde [CH$_3$CHO], formaldehyde [H$_2$CO], thioformaldehyde [H$_2$CS], formic acid [HCOOH], formamide [HCONH$_2$], glycoaldehyde [CH$_2$OHCHO], methyl formate [HCOOCH$_3$], ketene [H$_2$CCO], difluoromethane [CH$_2$F$_2$], carbonic difluoride [F$_2$CO], formyl fluoride [HFCO], fluoroethene [CFHCH$_2$], 1,1-difluoroethene [CF$_2$CH$_2$], *cis*-difluoroethene [CFHCFH], fluoroethane [CFH$_2$CH$_3$], trifluoroethene [CF$_2$CFH], ozone [O$_3$], sulfur dioxide [SO$_2$], difluorooxide [F$_2$O], hypofluoros acid [HOF], water [H$_2$O], hydrogendisulfide [H$_2$S], methylenecyclopropane [C$_4$H$_6$], cyclopropene [C$_3$H$_4$], aziridine [C$_2$H$_5$N], oxirane [C$_2$H$_4$O], thiirane [C$_2$H$_4$S], cyclopropenone [C$_3$H$_2$O], methylcyclopropene [C$_4$H$_6$], cyclobutene [C$_4$H$_6$], oxetane [C$_3$H$_6$O], β-propiolactone [C$_3$H$_4$O$_2$]

molecules considered (67 isotopic combinations, 143 tensor elements) are listed in Table 1. Rotational g tensors are known to be sensitive to geometry,[13] although Wilson et al. observed that the relative performance of different DFT functionals was the same whether optimized or experimental geometries were used. We only consider experimental geometries here, taken from ref 30 apart from ozone, for which we use $r = 1.2717$ Å and $\theta = 116.78°$.[56] With the exception of ammonia and thioformaldehyde, we compare results with the experimental g tensors compiled by Wilson et al.[30] For ammonia, we found experimental results with smaller error bars;[57] for thioformaldehyde, we used -5.2602 for the dominant diagonal element of the g tensor, noting that Wilson et al. used an incorrect value of -5.6202 because of a typographical error in the original experimental paper.[58]

We commenced by performing extensive investigations into the choice of the auxiliary Gaussian basis functions $g_t(\mathbf{r})$ for the expansion of the potential in eq 8 and the value of the filter in the singular-value decomposition (SVD) procedure used to construct the inverse Hessian in the optimization schemes. For the auxiliary basis, we considered both even-tempered basis sets and the use of the primary orbital basis. We found that the primary orbital basis was not sufficiently large, with deviations of more than 2% from converged results for simple molecules. For the even-tempered auxiliary basis sets, s, p, d, and f functions were placed on all atom centers, with exponents λ^{n_s}, λ^{n_p}, and so forth, where λ is a universal base and n_s, n_p, and so forth are (negative and positive) integers that define the range of the functions. After some investigations, a base of $\lambda = 2$ was used for all functions with the range chosen such that the exponents span those of the orbital basis, with the lowest exponent multiplied by 0.1. Because very large exponents led to a spurious structure in the OEP potential, functions with exponents larger than 2^6 were discarded, noting that the removal of the high exponents affected the g tensors less than a change in the orbital basis from triple- to quadruple-ζ. For the SVD filter, an analysis of the shape of the potentials and the dependence of the highest occupied molecular orbital eigenvalue led us to conclude that a value of 10^{-4} was appropriate. Once again, the effect of changing the filter was smaller than the orbital-basis incompleteness error.

To test our auxiliary basis set, we used the OEP procedure to calculate rotational g tensors for the KT2 GGA functional. Because the KT2 Kohn−Sham equation already takes the form of eq 5, the OEP procedure should, in the limit of a complete auxiliary basis, yield the same solution. A comparison of conventional uncoupled KT2 g tensors with those determined using the OEP orbitals and eigenvalues therefore provides a measure of the quality of the auxiliary basis set. Calculations were performed on a representative subset (specifically, the molecules considered in section 3.1). For the 15 tensor components, the average deviation from conventional KT2 values was 0.1%, with a maximum deviation of 0.3%.

3. Results and Discussion

Table 2 lists the mean absolute error (MAE), mean error (ME), percentage mean absolute error (PMAE), and standard deviation (SD) for the calculated rotational g tensors of the molecules in Table 1, compared with experimental values. Following Wilson et al.,[30] we present the error analysis both including and omitting the ozone molecule, because of its significant multireference character. Whereas the values denoted B3LYP, B97-2, and B97-3 refer to g tensors determined in a conventional, coupled manner using orbitals and eigenvalues from eq 7, those denoted O-B3LYP, O-B97-2, and O-B97-3 are the new values, determined in an uncoupled manner, using orbitals and eigenvalues from eq

Table 2. Mean Absolute Error (MAE), Mean Error (ME), Percentage Mean Absolute Error (PMAE), and Standard Deviation (SD) for Rotational g Tensor Elements, Relative to Experimental Values

	including ozone				excluding ozone			
	MAE	ME	PMAE	SD	MAE	ME	PMAE	SD
KT2	0.0082	−0.0043	10.9	0.0203	0.0078	−0.0048	11.1	0.0197
B3LYP	0.0153	−0.0133	11.3	0.0636	0.0107	−0.0087	11.2	0.0347
B97-2	0.0125	−0.0084	8.8	0.0583	0.0079	−0.0037	8.7	0.0224
B97-3	0.0149	−0.0107	9.2	0.0735	0.0091	−0.0049	9.1	0.0297
O-B3LYP	0.0077	−0.0037	12.0	0.0189	0.0070	−0.0046	12.2	0.0162
O-B97-2	0.0057	0.0013	7.1	0.0165	0.0048	0.0003	7.2	0.0120
O-B97-3	0.0064	0.0018	6.8	0.0214	0.0049	0.0003	6.9	0.0120

Table 3. Rotational g Tensor Components for the Ozone Molecule

	g_{xx}	g_{yy}	g_{zz}
KT2	−2.9230	−0.2313	−0.0792
B3LYP	−3.6369	−0.2657	−0.0783
B97-2	−3.6351	−0.2635	−0.0751
B97-3	−3.7980	−0.2729	−0.0754
O-B3LYP	−2.8729	−0.2259	−0.0782
O-B97-2	−2.8508	−0.2226	−0.0752
O-B97-3	−2.7752	−0.2195	−0.0754
exptl.	−2.9877	−0.2295	−0.0760
exptl.[a]	−2.9169	−0.2277	−0.0746

[a] Experimental results with calculated zero-point vibrational contributions (from ref 13) removed.

5 with the multiplicative OEP exchange-correlation potential. The results of all individual calculations, together with experimental data, are available in the Supporting Information.

For hybrid functionals, the uncoupled OEP approach gives significantly improved g tensors compared with those of the conventional approach. When all of the molecules are included, mean absolute errors reduce by at least a factor of 2 in moving from B3LYP, B97-2, and B97-3 to O-B3LYP, O-B97-2, and O-B97-3, respectively, while mean errors and standard deviations reduce by more than a factor of 3. The percentage mean absolute error reduces for the latter two functionals. The percentage errors are particularly large for tensor elements in molecules such as dimethyl ether and fluorethane. When ozone is excluded from the analysis, the improvement is more modest, reflecting the fact that ozone is particularly poorly described by conventional calculations. This is to be expected because the method has a close relationship (because of the fraction of orbital exchange) to coupled Hartree−Fock theory, which is very poor because of the multireference nature of this molecule. By contrast, ozone is not an especially challenging case for the OEP-based calculations. To quantify this, Table 3 presents rotational g tensors for ozone, compared with experimental values, both including and excluding zero-point vibrational contributions. For the g_{zz} component, which is described well by the conventional hybrid calculations, the OEP calculations yield essentially the same results. For the g_{xx} and g_{yy} components, which are in error with the conventional approach, the OEP calculations are a notable improvement. The improvement in g_{xx} is particularly pronounced.

With all molecules included, the quality as measured by the MAE can be summarized as

$$\text{B3LYP} < \text{B97-3} < \text{B97-2} < \text{KT2} < \text{O-B3LYP} < \text{O-B97-3} < \text{O-B97-2} \quad (11)$$

with the top method, O-B97-2, exhibiting MAE, ME, PMAE, and SD values of 0.0057, 0.0013, 7.1%, and 0.0165, respectively. Figure 1 quantifies the correlation between the O-B97-2 and experimental g tensors; the slope and intercept of the linear trendline are 1.0102 and −0.0002, respectively, with $R^2 = 0.9993$. When O_3 is excluded, there is little to choose between O-B97-2 and O-B97-3 as the top method. Whether ozone is included or not, the KT2 functional, which

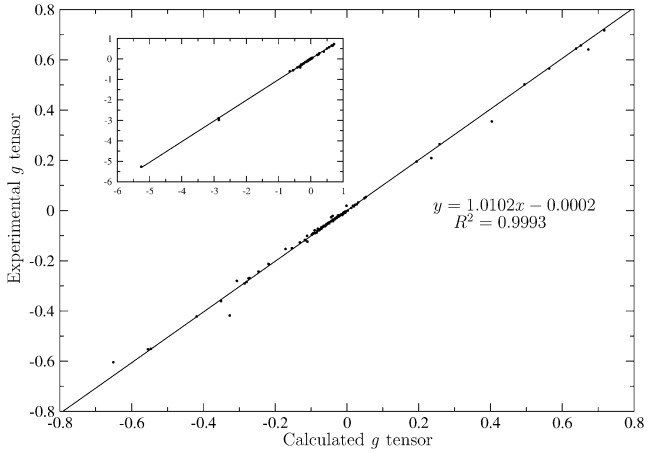


Figure 1. Correlation of O-B97-2 rotational g tensors with experimental values. The inset includes the entire range of g tensors.

was found to be the most accurate DFT method in the study of Wilson et al.,[30] is outperformed by the OEP-based hybrid functionals.

The errors are particularly small for the 22 linear molecules in the data set. For this subset, O-B97-3 gives the lowest MAE of just 0.0011, followed by O-B97-2 (0.0013), KT2 (0.0017), O-B3LYP (0.0018), B97-2 (0.0040), B97-3 (0.0049), and B3LYP (0.0050). The smallest errors are approaching the experimental error bars. Considering individual linear molecules, the OEP-based hybrid values are a particular improvement on the conventional hybrid results for CO, CS, CSe, and HCN. Other g-tensor elements for which the OEP results represent a sizable improvement are the large components in thioformaldehyde, formaldehyde, and sulfur dioxide. Indeed, it was the analysis of our OEP results that highlighted the error in the quoted thioformaldehyde experimental number of refs 3 and 30. After ozone, the most significant error with O-B97-2 occurs for the largest component of the ketene g tensor. The O-B97-2 value (−0.3267) is well-above the experimental value (−0.4182) and is even less accurate than the conventional B97-2 result (−0.3601). The discrepancy is evident in Figure 1. We have confirmed that this discrepancy does not arise from inadequacies in the orbital or auxiliary basis sets. It is interesting to note that the KT2 functional is also relatively poor for ketene (−0.3433).

It is pertinent to comment on the effect of zero-point vibrational corrections (ZPVCs) on rotational g tensors. The experimental data used in the above comparison with DFT do not include ZPVCs; they contain vibrational as well as electronic contributions. Vibrational corrections are scarce, although we are aware[13,15] of MCSCF values for NH_3, H_2O, HOF, and O_3 and a Hartree−Fock value for H_2S. To examine the importance of such corrections, we have computed errors for these molecules, comparing them with corrected as well as uncorrected experimental data. For O-B3LYP, O-B97-2, and O-B97-3, inclusion of the ZPVC reduces the MAE, by an average of 6%. For B3LYP, B97-2, and B97-3, by contrast, the correction increases the error, by an average of 14%. For all six methods, the PMAEs actually increase slightly because of the small g-tensor components, although

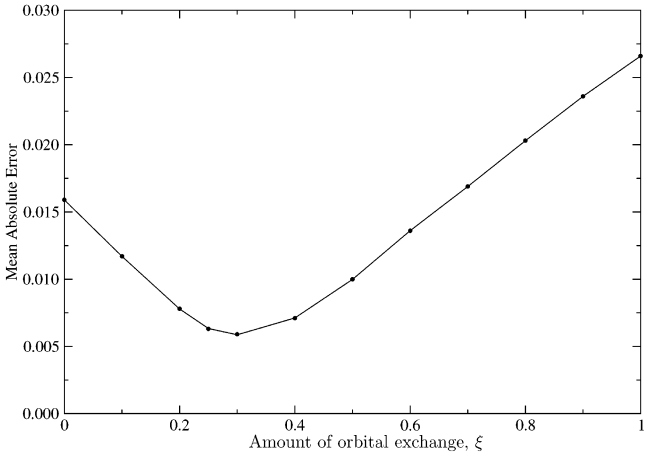


Figure 2. Mean absolute error for a subset of eight molecules (see text) as a function of the amount of orbital exchange, ξ.

the average increase is smaller for the OEP-based methods. Explicit values for ozone are presented in Table 3. This analysis, albeit based on limited data, supports our conclusions regarding the improved quality of predictions from OEP-based hybrid calculations.

3.1. Influence of Orbital Exchange. In ref 59, we investigated the influence of the amount of orbital exchange ξ (see eq 7) on uncoupled shielding constants determined using orbitals and eigenvalues associated with a multiplicative exchange-correlation potential. Using the ZMP exchange-correlation potential, we found that an amount $0.2 \leq \xi \leq 0.3$ was optimal, consistent with the amount used in common hybrid functionals such as B3LYP. By contrast, Arbuznikov and Kaupp[40] used the localized Hartree−Fock (LHF) approach[60] [equivalent to the common energy denominator approximation (CEDA)[61]] and instead found that an amount nearer 0.6 was optimal. These different optimal amounts of exact exchange reflect differences between the LHF and ZMP potentials. The reader is referred to refs 38 and 62 for further discussion on the differences between these and other potentials.

In the present investigation, we have used the OEP procedure to determine orbitals and eigenvalues for the series of functionals in ref 59, with different amounts of exchange $0 \leq \xi \leq 1$. From these potentials, we have determined uncoupled rotational g tensors for the representative subset of molecules CO, CSe, OCS, N_2O, NH_3, HOF, H_2O, and SO_2. Figure 2 presents the mean absolute error relative to experimental results as a function of ξ. In line with the NMR-shielding observations of ref 59, the best results are obtained with $\xi \approx 0.3$, which is consistent with previous observations highlighting the similarity of the ZMP and OEP potentials.[37,62]

3.2. *g* Tensors Calculated from Coupled-Cluster Densities. The results discussed so far have been pure DFT results. We complete this study by instead determining rotational g tensors directly from ab initio electron densities using the WY approach. Specifically, we calculate BD coupled-cluster relaxed densities and input them [$\rho_{in}(\mathbf{r})$ in eq 10] into our implementation of the WY procedure in CADPAC, which yields an optimal set of expansion parameters $\{b_t\}$. These are then transferred to DALTON as in the OEP calculations, and the uncoupled g tensor is evaluated. The results are denoted WY(BD).

We have considered the same subset of molecules as that in section 3.1; all results are presented in the Supporting Information. The conventional hybrids give mean absolute errors from 0.0127 (B97-2) to 0.0139 (B97-3), whereas the OEP errors are between 0.0068 (O-B97-3) and 0.0084 (O-B3LYP). The WY(BD) mean absolute error is slightly lower, at 0.0066. The WY(BD) results are of good quality but, of course, require the evaluation of the coupled-cluster relaxed density. The modest gain in accuracy is therefore associated with a significant increase in computational cost; the method is not practical for larger molecules. We also note that the WY(BD) results are poor for the ozone molecule. If this molecule is included in the subset, then the OEP methods are an improvement over WY(BD).

4. Conclusions

In line with previous studies of NMR shielding constants,[34,37] chemical shifts,[35] and magnetizabilities,[36] we have demonstrated that significantly improved rotational g tensors can be obtained with hybrid functionals by an uncoupled evaluation, using orbitals and eigenvalues from a rigorous Kohn−Sham equation, eq 5. Compared with the conventional approach, which uses a coupled formalism and nonmultiplicative exchange-correlation operator, eq 7, mean absolute errors are reduced by at least a factor of 2, whereas mean errors and standard deviations are reduced by more than a factor of 3. When ozone is omitted, the improvement is more modest, reflecting the fact that ozone is particularly poorly described by conventional calculations. Notably, the results are also an improvement over those from the KT2 GGA functional, which was developed with an emphasis on a high-quality magnetic response and which provided the highest quality in the study of Wilson et al.[30] The OEP calculations are more computationally demanding than those of the KT2 functional, requiring an implementation of the OEP methodology, although they do exhibit the favorable scaling of DFT. We have also demonstrated that, for a representative subset of molecules, optimal results are obtained with an amount of orbital exchange $\xi \approx 0.3$. When g tensors are determined from coupled-cluster electron densities, the improvement is only modest, at significant additional computational expense.

This study highlights the high-quality rotational g-tensor predictions possible within DFT; further investigations are under way. We are also using the ideas of this paper to investigate spin-rotation constants, which are related to shielding constants in the same way that g tensors are related to magnetizabilities. Results will be reported in a forthcoming publication.

Acknowledgment. We are grateful to D. J. D. Wilson for providing input geometry files for DALTON and to P. Salek for implementation of the B97-2 and B97-3 functionals in DALTON. This work has received support from the EPSRC (A.M.T. studentship) and the Norwegian National Research Council through a Strategic University Program in Quantum Chemistry grant (Grant No. 154011/420) and

through a grant of computer time from the program of Supercomputing. O.B.L. acknowledges the facilities provided by the University of Durham, during his six month visit.

Supporting Information Available: All calculated rotational *g* tensors and experimental data. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Ramsey, N. F. *Molecular Beams*; Clarendon Press: Oxford, U. K. ,2000.

(2) Flygare, W. H.; Benson, C. H. *Mol. Phys.* **1971**, *20*, 225−250.

(3) Flygare, W. H. *Chem. Rev.* **1974**, *74*, 653−687.

(4) Gauss, J.; Ruud, K.; Helgaker, T. *J. Chem. Phys.* **1996**, *105*, 2804−2812.

(5) Stevens, R. M.; Pitzer, R. M.; Lipscomb, W. N. *J. Chem. Phys.* **1963**, *38*, 550−560.

(6) Kelly, H. M.; Fowler, P. W. *Chem. Phys. Lett.* **1993**, *206*, 568−573.

(7) Ruud, K.; Helgaker, T. *Chem. Phys. Lett.* **1997**, *264*, 17−23.

(8) Ogilvie, J. F.; Cheah, S.-L.; Lee, Y.-P.; Sauer, S. P. A. *Theor. Chem. Acc.* **2002**, *108*, 85−97.

(9) Åstrand, P.-O.; Ruud, K.; Mikkelsen, K. V.; Helgaker, T. *Mol. Phys.* **1997**, *92*, 89−96.

(10) Åstrand, P.-O.; Ruud, K.; Mikkelsen, K. V.; Helgaker, T. *J. Chem. Phys.* **1999**, *110*, 9463−9468.

(11) Ruud, K.; Helgaker, T.; Jørgensen, P. *J. Chem. Phys.* **1997**, *107*, 10559−10606.

(12) Ruud, K.; Vaara, J.; Lounila, J.; Helgaker, T. *Chem. Phys. Lett.* **1998**, *297*, 467−474.

(13) Mohn, C. E.; Wilson, D. J. D.; Lutnæs, O. B.; Helgaker, T.; Ruud, K. *Adv. Quantum Chem.* **2005**, *50*, 77−90.

(14) Sauer, S. P. A. *Adv. Quantum Chem.* **2005**, *48*, 319−334.

(15) Ruud, K.; Åstrand P.-O.; Taylor, P. R. *J. Chem. Phys.* **2000** *112*, 2668−2683.

(16) Cybulski, S. M.; Bishop, D. M. *J. Chem. Phys.* **1997**, *106*, 4082−4090.

(17) Cybulski, S. M.; Bishop, D. M. *J. Chem. Phys.* **1993**, *100*, 2019−2026.

(18) Oddershede, J.; Sabin, J. R. *Chem. Phys.* **1988**, *122*, 291−296.

(19) Sauer, S. P. A.; Špirko, V.; Paidarová, I.; Oddershede, J. *Chem. Phys.* **1994**, *184*, 1−3.

(20) Sauer, S. P. A.; Špirko, V.; Oddershede, J. *Chem. Phys.* **1991**, *153*, 189−200.

(21) Sauer, S. P. A.; Ogilvie, J. F. *J. Phys. Chem.* **1994**, *98*, 8617−8621.

(22) Ogilvie, J. F.; Oddershede, J.; Sauer, S. P. A. *Chem. Phys. Lett.* **1994**, *228*, 183−190.

(23) Geertsen, J.; Scuseria, G. E. *J. Chem. Phys.* **1989**, *90*, 6486−6490.

(24) Sauer, S. P. A.; Oddershede, J.; Geertsen, J. *Mol. Phys.* **1992**, *76*, 445−465.

(25) Sauer, S. P. A. *Chem. Phys. Lett.* **1996**, *260*, 271−279.

(26) Ruud, K.; Åstrand, P.-O.; Helgaker, T.; Mikkelsen, K. V. *THEOCHEM* **1996**, *388*, 231−235.

(27) Åstrand, P.-O.; Ruud, K.; Mikkelsen, K. V.; Helgaker, T. *Chem. Phys. Lett.* **1997**, *271*, 163−166.

(28) Bak, K. L.; Sauer, S. P. A.; Oddershede, J.; Ogilvie, J. F. *Phys. Chem. Chem. Phys.* **2005**, *7*, 1747−1758.

(29) Wilson, P. J.; Amos, R. D.; Handy, N. C. *J. Mol. Struct.* **2000**, *506*, 335−342.

(30) Wilson, D. J. D.; Mohn, C. E.; Helgaker T. *J. Chem. Theory Comput.* **2005**, *1*, 877−888.

(31) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098−3100. Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.

(32) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623−11627. Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.

(33) Keal, T. W.; Tozer, D. J. *J. Chem. Phys.* **2003**, *119*, 3015−3024.

(34) Wilson, P. J.; Tozer, D. J. *Chem. Phys. Lett.* **2001**, *337*, 341−348.

(35) Allen, M. J.; Keal, T. W.; Tozer, D. J. *Chem. Phys. Lett.* **2003**, *380*, 70−77.

(36) Wilson, P. J.; Tozer, D. J. *J. Mol. Struct.* **2002**, *602*−*603*, 191−197.

(37) Cohen, A. J.; Wu, Q.; Yang, W. *Chem. Phys. Lett.* **2004**, *399*, 84−88.

(38) Teale, A. M.; Tozer, D. J. *Chem. Phys. Lett.* **2004**, *383*, 109−114.

(39) Hieringer, W.; Della Sala, F.; Gorling, A. *Chem. Phys. Lett.* **2004**, *383*, 115−121.

(40) Arbuznikov, A. V.; Kaupp, M. *Chem. Phys. Lett.* **2004**, *386*, 8−16.

(41) Arbuznikov, A. V.; Kaupp, M. *Chem. Phys. Lett.* **2004**, *391*, 16−21.

(42) Arbuznikov, A. V.; Kaupp, M. *Int. J. Quantum Chem.* **2005**, *104*, 261−271.

(43) Zhao, Q.; Morrison, R. C.; Parr, R. G. *Phys. Rev. A* **1994**, *50*, 2138−2142.

(44) Sharp, R. T.; Horton, G. K. *Phys. Rev.* **1953**, *90*, 317. Talman, J. D.; Shadwick, W. F. *Phys. Rev. A* **1976**, *14*, 36−40.

(45) Wilson, P. J.; Bradley, T. J.; Tozer, D. J. *J. Chem. Phys.* **2001**, *115*, 9233−9242.

(46) Keal, T. W.; Tozer, D. J. *J. Chem. Phys.* **2005**, *123*, 121103.

(47) Wu, Q.; Yang, W. *J. Chem. Phys.* **2003**, *118*, 2498−2509.

(48) Yang, W.; Wu, Q. *Phys. Rev. Lett.* **2002**, *89*, 143002.

(49) Wu, Q.; Yang, W. *J. Theor. Comput. Chem.* **2003**, *2*, 627−638.

(50) Amos, R. D. *CADPAC 6.5, The Cambridge Analytic Derivatives Package*; Cambridge University: Cambridge, England, 1998.

(51) Levy, M. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 6062−6065.

(52) *DALTON*, release 2.0; 2005. See http://www.kjemi.uio.no/software/dalton/dalton.html (accessed June 2005).

(53) Kendall, R. A.; Dunning, T. H., Jr. *J. Chem. Phys.* **1992**, *96*, 6796−6806.

(54) Woon, D. E.; Dunning, T. H., Jr. *J. Chem. Phys.* **1993**, *98*, 1358−1371.

(55) Wilson, A. K.; Woon, D. E.; Peterson, K. A.; Dunning, T. H. *J. Chem. Phys.* **1999**, *110*, 7667−7676.

(56) Tanaka, T.; Morino, Y. *J. Mol. Spectrosc.* **1970**, *33*, 538−551.

(57) Hüttner, W.; Frank, U. E.; Majer, W.; Mayer, K.; Špirko, V. *Mol. Phys.* **1988**, *64*, 1233−1249.

(58) Rock, S. L.; Flygare, W. H. *J. Chem. Phys.* **1972**, *56*, 4723−4728.

(59) Wilson, P. J.; Tozer, D. J. *J. Chem. Phys.* **2002**, *116*, 10139−10147.

(60) Della Sala, F.; Görling, A. *J. Chem. Phys.* **2001**, *115*, 5718−5732.

(61) Gritsenko, O. V.; Baerends, E. J. *Phys. Rev. A* **2001**, *64*, 042506.

(62) Teale, A. M.; Tozer, D. J. *Phys. Chem. Chem. Phys.* **2005**, *7*, 2991−2998.

CT060038N

JCTC Journal of Chemical Theory and Computation

Energy-Represented Direct Inversion in the Iterative Subspace within a Hybrid Geometry Optimization Method

Xiaosong Li*

Department of Chemistry, University of Washington, Seattle, Washington 98195-1700

Michael J. Frisch

Gaussian, Inc., Wallingford, Connecticut 06492

Received November 8, 2005

Abstract: A geometry optimization method using an energy-represented direct inversion in the iterative subspace algorithm, GEDIIS, is introduced and compared with another DIIS formulation (controlled GDIIS) and the quasi-Newton rational function optimization (RFO) method. A hybrid technique that uses different methods at various stages of convergence is presented. A set of test molecules is optimized using the hybrid, GEDIIS, controlled GDIIS, and RFO methods. The hybrid method presented in this paper results in smooth, well-behaved optimization processes. The optimization speed is the fastest among the methods considered.

I. Introduction

Geometry optimization is an essential part of computational chemistry. Any theoretical investigation that involves calculations of transition structures, barrier heights, heats of reaction, or vibrational spectra requires searches for one or more minima or saddle points on a potential energy surface (PES). Computational methods are applied to large systems of ever-increasing size. Biomolecules, polymers, and nanostructures with hundreds to thousands of atoms are often difficult to optimize because of excessive degrees of freedom. Any decrease in the computational cost and increase in the general stability of geometry optimization would be welcome.

A variety of algorithms for geometry optimization are widely used in computational chemistry (for some recent reviews, see refs 1 and 2). Geometry optimization methods can be broadly classified into two major categories. First-order methods use only analytic first derivatives to search for stationary points; second-order methods use analytic first and second derivatives, assuming a quadratic model for the potential energy surface and a Newton−Raphson step for the minima search

$$\Delta x = -\mathbf{H}^{-1}\mathbf{g} \tag{1}$$

where $\mathbf{g}$ is the gradient (first derivative) and $\mathbf{H}^{-1}$ is the inverse Hessian (second derivative). While second-order optimization schemes need fewer steps to reach convergence than first-order methods,[3] this approach can quickly become very expensive with increasing system size because the explicit computation of the Hessian scales as $O(N^4)-O(N^5)$, where N is a measure of the system size. Quasi-Newton methods are intermediate between the first- and second-order approaches. An initial estimate of the Hessian is obtained by some inexpensive method. Subsequently, the Hessian is updated using the first derivatives and displacements, by methods such as BFGS,[4−7] SR1,[8] and PSB.[9,10] The quasi-Newton approach is comparable in computational cost to first-order methods and in convergence speed to second-order methods.

The use of a Newton−Raphson step when the PES is far away from a quadratic region can lead to overly large step sizes in the wrong direction. The stability of a Newton−Raphson geometry optimization can be enhanced by controlling the step size using techniques such as rational function optimization (RFO)[11,12] and the trust radius model.[1−3,13−16] To reduce the number of iterations required to reach convergence, a least-squares minimization scheme is used: direct inversion in the iterative subspace (DIIS).[17,18] The DIIS approach is efficient in both converging the wave function[17−21] and optimizing the geometry.[22,23] It extrapolates/interpolates a set of vectors $\{\mathbf{R}_i\}$ by minimizing the errors in the least-

* Corresponding author. Fax: 206-685-8665. E-mail: li@chem.washington.edu.

10.1021/ct050275a CCC: $33.50
Published on Web 04/18/2006

squares sense (for a leading reference, see ref 17):

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & \cdots & a_{1,N} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{N,1} & \cdots & a_{N,N} & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix}$$

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,N} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{N,1} & \cdots & a_{N,N} & 1 \\ 1 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_N \\ \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \text{ and } \sum c_i = 1 \qquad (2)$$

where λ is the Lagrangian multiplier. The coefficients minimize a working function, which can be an error function[17,18,21−23] or an energy function.[19,20] Fock matrix and nuclear positions are often chosen as the vectors for self-consistent field (SCF)[17,18,20,21] and geometry[22,23] optimizations, respectively. When the working function is an error function, the matrix **A** is usually defined as the product of error vectors, $a_{i,j} = e_i e_j^{\mathrm{T}}$, where the error vector can be a quasi-Newton step for geometry optimization[22,23] or the commutator [**F**, **P**] for SCF convergence.[17,18] Solving eq 2 leads to a set of DIIS coefficients c_i that are used to obtain a new vector, $\mathbf{R^*} = \sum_i c_i \mathbf{R}_i$, which has the minimum functional value within the search space. Direct solution of eq 2 often leads to unproductive oscillations when optimizing large systems.[20,23] Farkas and Schlegel have introduced some controls to ensure a downhill DIIS extrapolation/interpolation for geometry optimization (GDIIS).[23] Scuseria and co-workers have introduced a stable and efficient alternative DIIS formalism for SCF convergence, energy-DIIS (SCF-EDIIS).[19,20] The SCF-EDIIS algorithm minimizes an energy function and enforces coefficients to be positive definite.

This paper extends the energy-represented direct inversion in the iterative subspace algorithm from SCF convergence to geometry optimization (GEDIIS). A hybrid geometry optimization technique that utilizes the advantages of different methods of various stages of the optimization is then illustrated and tested.

II. Methodology and Benchmarks

Optimizations are carried out using the development version of the Gaussian series of programs[24] with the addition of the geometry optimization algorithms using GEDIIS and the hybrid method presented here. For all methods, the geometry optimization is considered converged when the root-mean-square (RMS) force is less than 1×10^{-5} au, the RMS geometry displacement is less than 4×10^{-5} au, the maximum component of the force vector is less than 1.5×10^{-5} au, and the maximum component of the geometry displacement is less than 6×10^{-5} au. All calculations are carried out with redundant internal coordinates. No symmetry operations or reorientations are imposed. Starting geometries are available upon request (li@chem.washington.edu).

A. Geometry Optimization Using Energy-Represented Direct Inversion in the Iterative Subspace (GEDIIS). As in all DIIS-based schemes, the GEDIIS formalism yields a new vector **R*** constructed from a linear combination of N previously computed vectors $\mathbf{R}_i$

$$\mathbf{R^*} = \sum_{i=1}^{N} c_i \mathbf{R}_i, \qquad \sum_{i=1}^{N} c_i = 1 \qquad (3)$$

In GEDIIS, $\mathbf{R}_i$'s are geometries and c_i's minimize an energy function. The energy of the new structure **R*** can be approximated to first order with an expansion at $\mathbf{R}_i$

$$E(\mathbf{R^*}) = E(\mathbf{R}_i) + (\mathbf{R^*} - \mathbf{R}_i)\mathbf{g}_i \qquad (4)$$

where $E(\mathbf{R}_i)$ and $\mathbf{g}_i$ are the energy and gradient of structure $\mathbf{R}_i$, respectively. Multiplying both sides of eq 4 by c_i, and summing over N points, we obtain

$$E(\mathbf{R^*}) = \sum_{i=1}^{N} c_i [E(\mathbf{R}_i) + \sum_{j=1}^{N} c_j \mathbf{R}_j \mathbf{g}_i - \mathbf{R}_i \mathbf{g}_i] \qquad (5)$$

A simple algebraic manipulation leads to the GEDIIS working function

$$E(\mathbf{R^*}) = \sum_{i=1}^{N} c_i E(\mathbf{R}_i) - \frac{1}{2}\sum_{i,j=1}^{N} c_i c_j [\mathbf{R}_i(\mathbf{g}_i - \mathbf{g}_j) + \mathbf{g}_i(\mathbf{R}_i - \mathbf{R}_j)] \qquad (6)$$

or

$$E(\mathbf{R^*}) = \sum_{i=1}^{N} c_i E(\mathbf{R}_i) - \frac{1}{2}\sum_{i,j=1}^{N} c_i c_j (\mathbf{g}_i - \mathbf{g}_j)(\mathbf{R}_i - \mathbf{R}_j)$$

where

$$\sum_{i=1}^{N} c_i = 1$$

The quadratic term in eq 6 represents the variations in energy for changes of geometric coordinates and gradients within the search space. Equation 6 is formally identical to Scuseria's SCF-EDIIS equation for wave-function optimization.[20] In both the GEDIIS and SCF-EDIIS formalisms, the energy function is minimized directly with respect to the expansion coefficients, c_i's. The working functions in GEDIIS, SCF-DIIS,[17,18] SCF-EDIIS,[20] and GDIIS[22,23] are all quadratic with respect to c_i, and minimizations are therefore performed in a least-squares sense. The comparison of this technique with Pulay's DIIS[17,18] has been discussed extensively in the literature.[19,20]

The set of DIIS coefficients, c_i's, that minimize the energy function, eq 6, are used to construct a new geometry **R*** from a linear combination of known points (eq 3). The resulting geometry **R*** is, to a first-order approximation, associated with the optimal energy within the search space. However, **R*** is not necessarily the final minimum structure. The distance from **R*** to the minimum structure can be approximated by a second-order Newton step. In this work, we use a RFO step for the second-order correction

$$\mathbf{R}_{N+1} = \mathbf{R^*} + \Delta\mathbf{R} = \sum_{i=1}^{N} c_i \mathbf{R}_i - \sum_{i=1}^{N} c_i \mathbf{g}_i (\mathbf{H} - \xi)^{-1} \qquad (7)$$

where the parameter ξ is optimized using the RFO approach

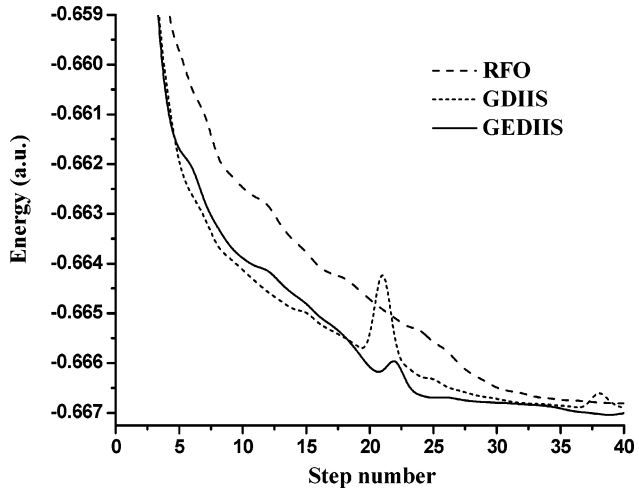


Figure 1. Energy vs step number for geometry optimizations on taxol at AM1.

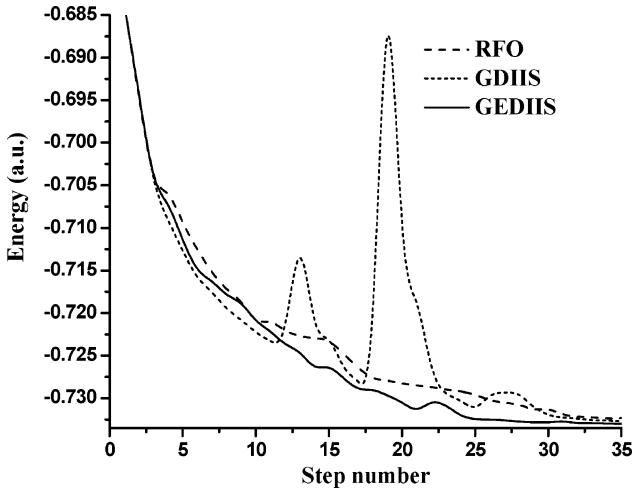


Figure 2. Energy vs step number for geometry optimizations on For−$(Ala)_{10}$−NH_2 at AM1.

(note that the Hessian is a constant in a quadratic approximation).[11,12] With no constraint on the sign of c_i, eq 7 is essentially an extrapolation; when $c_i > 0$, it becomes an interpolation step. When the molecular geometry is far from convergence, extrapolations can lead to erroneously large steps away from the optimized geometry. To ensure optimization stability, an enforced interpolation constraint, $c_i > 0$, is added into eq 2 (for detailed machinery, see ref 20). As a result, GEDIIS searches for a new geometry in the region close to the local potential energy surface, that is, interpolations only.

We update the geometric Hessian with first derivatives: a weighted combination of BFGS and SR1[23] with the square root of the Bofill[25] weighting factor

$$\mathbf{H}_i = \mathbf{H}_{i-1} + \varphi\Delta\mathbf{H}_i^{\mathrm{SR1}} + (1-\varphi)\Delta\mathbf{H}_i^{\mathrm{BFGS}}$$

$$\Delta\mathbf{H}_i^{\mathrm{SR1}} = -\frac{(\mathbf{H}_{i-1}\Delta\mathbf{R}_{i-1} - \Delta\mathbf{g}_i)(\mathbf{H}_{i-1}\Delta\mathbf{R}_{i-1} - \Delta\mathbf{g}_i)^{\mathrm{T}}}{(\mathbf{H}_{i-1}\Delta\mathbf{R}_{i-1} - \Delta\mathbf{g}_i)^{\mathrm{T}}\Delta\mathbf{R}_{i-1}}$$

$$\Delta\mathbf{H}_i^{\mathrm{BFGS}} = \frac{\Delta\mathbf{g}_i\Delta\mathbf{g}_i^{\mathrm{T}}}{\Delta\mathbf{R}_{i-1}^{\mathrm{T}}\Delta\mathbf{g}_i} - \frac{\mathbf{H}_{i-1}\Delta\mathbf{R}_{i-1}\Delta\mathbf{R}_{i-1}^{\mathrm{T}}\mathbf{H}_{i-1}}{\Delta\mathbf{R}_{i-1}^{\mathrm{T}}\mathbf{H}_{i-1}\Delta\mathbf{R}_{i-1}}$$

$$\varphi = \sqrt{\varphi^{\mathrm{Bofill}}} = \sqrt{\frac{[(\mathbf{H}_{i-1}\Delta\mathbf{R}_{i-1} - \Delta\mathbf{g}_i)^{\mathrm{T}}\Delta\mathbf{R}_{i-1}]^2}{(\mathbf{H}_{i-1}\Delta\mathbf{R}_{i-1} - \Delta\mathbf{g}_i)^{\mathrm{T}}(\mathbf{H}_{i-1}\Delta\mathbf{R}_{i-1} - \Delta\mathbf{g}_i)\Delta\mathbf{R}_{i-1}^{\mathrm{T}}\Delta\mathbf{R}_{i-1}}}$$

$$\Delta\mathbf{g}_i = \mathbf{g}_i - \mathbf{g}_{i-1}$$

$$\Delta\mathbf{R}_i = \mathbf{R}_i - \mathbf{R}_{i-1} \qquad (8)$$

Equation 8 has been successfully used for large-molecule geometry optimization.[26]

Figures 1 and 2 show energy profiles for the first 40 optimization steps of taxol and For−$(Ala)_{10}$−NH_2 using RFO, GDIIS, and GEDIIS methods at the AM1 level of theory. The GEDIIS method reduces, or even eliminates, large energy oscillations, which often appear in GDIIS at the early stage of optimization because of large extrapolation steps. The convergence behavior of the GEDIIS approach is generally smooth and well-behaved. Note that GEDIIS shows small bumps at places similar to those in GDIIS, indicating that it is taking approximately the same path but in a more controlled fashion. This enhanced stability is a result of the better representation of error vectors and the enforced DIIS interpolation. The SCF-EDIIS method for wave-function optimization also shows enhanced stability.[20] Because the DIIS algorithm collects information from previously computed points for a better picture of the local potential surface than that obtained from a single point, as in the RFO method, the potential energies at the same optimization step are generally in the order $E_{\mathrm{GEDIIS}} < E_{\mathrm{GDIIS}} < E_{\mathrm{RFO}}$.

B. Mixed Geometry Optimization Scheme. The simple RFO, GDIIS, and GEDIIS methods each perform best in different stages of the geometry optimization. Even though GEDIIS exhibits fast and smooth behavior at the early stage of geometry optimization, the first-order energy-represented **A** matrix has no knowledge of the curvature of the local potential surface, so the problems of gradient-based methods are still present in a direct implementation of the GEDIIS method presented here. In a well-behaved quadratic region near convergence, first-derivative-based optimizations such as GEDIIS are generally slower than those using both first and second derivatives. To take advantage of the fast convergence of Hessian-based methods near minima, GDIIS using RFO steps as error vectors can be used instead of GEDIIS when the optimization process is close to convergence. Another critical problem for methods using DIIS minimization schemes results from interpolation where the optimization scope is limited to neighboring potential surfaces only: optimizations can easily be trapped in undesired shallow potential wells. This is particularly troublesome for optimizing large molecules with many degrees of freedom. By contrast, RFO is essentially a single extrapolation step that is better able to overcome shallow potential wells. Therefore, it is a reasonable strategy to use RFO as a *preoptimizer* or *potential well selector* followed by DIIS-based optimization methods.

We thus combine methods to gain a fast hybrid geometry optimization scheme: (1) The geometry optimization begins with the RFO method to use the Hessian-based quadrature

Table 1. Comparison of the Computational Costs for Quasi-Newton RFO, Controlled GDIIS, GEDIIS, and Hybrid Optimization Methods on Select Baker Test Molecules

		RFO		GDIIS		GEDIIS		hybrid	
	E^a (au)	T^b	N^c	T^b	N^c	T^b	N^c	T^b	N^c
disilyl ether	−657.881 31	1.28	13	1.18	12	1.10	11	1.00	10
1,3,5-trisilaceyclohexane	−990.064 87	1.10	12	0.93	10	1.07	12	1.00	11
2-hydroxybicyclopentane	−270.492 14	1.14	15	0.99	13	1.00	13	1.00	13
achtar10	−363.046 87	1.27	18	1.06	15	1.00	14	1.00	14
benzidine	−574.008 39	1.08	12	1.00	11	1.01	11	1.00	11
pterin	−580.647 03	1.00	12	1.01	12	1.01	12	1.00	12
histidine	−548.755 26	1.01	17	1.00	17	1.00	17	1.00	17
caffeine	−680.376 96	1.00	11	1.00	11	0.86	9	1.00	11
menthone	−467.143 10	1.00	14	1.00	14	0.95	13	1.00	14
acthcp	−852.314 18	1.05	17	1.00	16	1.00	16	1.00	16
histamine H^+	−360.593 73	1.06	16	0.92	14	0.92	14	1.00	15
hydrazobenzene	−573.970 61	1.00	20	1.10	22	1.41	28	1.00	20

[a] Final energy. [b] Total CPU time with respect to that of the hybrid method. [c] Number of optimization steps.

Table 2. Comparison of the Computational Costs for Quasi-Newton RFO, Controlled GDIIS, GEDIIS, and Hybrid Optimization Methods

	histidine (20 atoms)	hydrazobenzene (26 atoms)	taxol (113 atoms)	For−$(Ala)_{10}$−NH_2 (106 atoms)	For−$(Ala)_{20}$−NH_2 (206 atoms)
	MP2/6-311G(d)	B3LYP/6-31G(d)	PBEPW91/3-21G	HF/STO-3G	AM1
			Final Energy		
RFO	−545.663 46	−573.970 61	−2911.140 32	−2593.884 16	−1.425 66
GDIIS	−545.663 46	−573.970 61	−2911.140 33	−2593.882 62	−1.425 66
GEDIIS	−545.663 46	−573.970 61	−2911.140 33	−2593.882 58	−1.425 66
hybrid	−545.663 46	−573.970 61	−2911.140 33	−2593.884 16	−1.425 66
			Number of Optimization Steps		
RFO	21	20	135	94	107
GDIIS	21	22	136	111	92
GEDIIS	16	28	125	67	88
hybrid	19	20	105	79	91
			Relative CPU Times[a]		
RFO	1.10	1.00	1.22	1.11	1.13
GDIIS	1.10	1.10	1.30	1.47[b]	1.02
GEDIIS	0.85	1.41	1.19	0.81[b]	0.96
hybrid	1.00	1.00	1.00	1.00	1.00

[a] Total CPU time with respect to that of the hybrid method. [b] Minima higher in energy than that optimized by the hybrid method.

to avoid shallow potential wells; (2) the optimization algorithm switches to fast and smoothly converging GEDIIS when the root-mean-square force of the latest point is smaller than 10^{-2} au; (3) when the root-mean-square RFO step of the latest point is less than 2.5×10^{-3} au, GDIIS[23] is used until convergence. This hybrid method takes advantage of the three methods considered here—the ability to overcome shallow potential wells by RFO, smooth optimization by GEDIIS, and fast convergence of RFO−DIIS near the minimum. The switching criteria were optimized from tests on the Baker set (Table 1).[27] Figure 3 illustrates optimizations using RFO, GDIIS, GEDIIS, and the hybrid method for For−$(Ala)_{10}$−NH_2 at the HF/STO-3G level of theory. Both GDIIS and GEDIIS converge to a minimum that is about 1.5×10^{-3} au higher in energy than that by RFO. The smallest harmonic frequency for the DIIS minimum is only about 7 cm^{-1}, indicating a very flat potential well. The hybrid optimization switches to GEDIIS from RFO after the first 20 steps and switches to GDIIS after another 26 steps. Convergence is reached within 79 steps, significantly faster than the 94 steps by RFO. While the DIIS methods remain at a higher energy minimum, the hybrid method overcomes

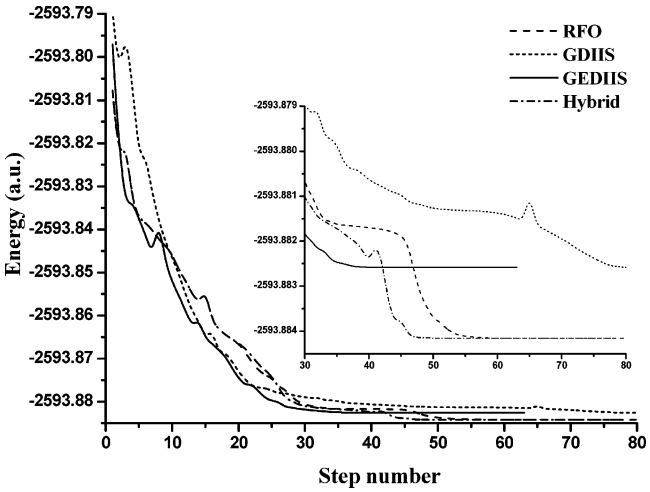


Figure 3. Energy vs step number for geometry optimizations on For−$(Ala)_{10}$−NH_2 at HF/STO-3G.

the shallow potential well and optimizes For−$(Ala)_{10}$−NH_2 to the RFO minimum.

Table 2 compares final energies and computational costs by the RFO, GDIIS, GEDIIS, and hybrid optimization methods on select molecules at various levels of theory. The computational costs are presented as the relative CPU times. The hybrid method introduced in this paper is the fastest among the popular geometry optimization methods considered here. Most importantly, the optimization behavior of the hybrid method is consistently smooth and fast through optimizations of the Baker set and large biochemical molecules tested in this paper.

III. Conclusion

This paper presents a geometry optimization method using GEDIIS. The GEDIIS method minimizes an energy representation of the local potential energy surface in the vicinity of previously computed points (gradients and geometries) as a least-squares problem. The enforced interpolation in GEDIIS leads to enhanced stability.

A hybrid geometry optimization algorithm is proposed that takes into account the problems and advantages of different optimization methods. The hybrid method starts off with RFO as a *preoptimizer* and switches to GEDIIS when a certain convergence threshold is met. Near the minimum, GEDIIS switches to RFO−DIIS. This takes advantage of RFO's ability to overcome shallow potential wells, smooth optimization by GEDIIS, and fast convergence of RFO−DIIS near the minimum. Optimizations of test molecules with the hybrid method are shown to be smooth, reliable, and fast.

Acknowledgment. The authors are grateful for discussions with Professor Schlegel at Wayne State University.

References

(1) Schlegel, H. B. Geometry Optimization on Potential Energy Surfaces. In *Modern Electronic Structure Theory*; Yarkony, D. R., Ed.; World Scientific: Singapore, 1995; p 459.

(2) Schlegel, H. B. Geometry Optimization. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Allinger, N. L., Kollman, P. A., Clark, T., Schaefer, H. F., III, Gasteiger, J., Schreiner, P. R., Eds.; Wiley: Chichester, U. K., 1998; Vol. 2, p 1136.

(3) Fletcher, R. *Practical Methods of Optimization*; Wiley: Chichester, U. K., 1981.

(4) Broyden, C. G. *J. Inst. Math. Appl.* **1970**, *6*, 76.

(5) Fletcher, R. *Comput. J. (Switzerland)* **1970**, *13*, 317.

(6) Goldfarb, D. *Math. Comput.* **1970**, *24*, 23.

(7) Shanno, D. F. *Math. Comput.* **1970**, *24*, 647.

(8) Murtagh, B.; Sargent, R. W. H. *Comput. J. (Switzerland)* **1972**, *13*, 185.

(9) Powell, M. J. D. *Nonlinear Programing*; Academic: New York, 1970.

(10) Powell, M. J. D. *Math. Program.* **1971**, *1*, 26.

(11) Banerjee, A.; Adams, N.; Simons, J.; Shepard, R. *J. Phys. Chem.* **1985**, *89*, 52.

(12) Simons, J.; Nichols, J. *Int. J. Quantum Chem.* **1990**, *24*, 263.

(13) Murray, W.; Wright, M. H. *Practical Optimization*; Academic: New York, 1981.

(14) *Nonlinear Optimization*; Powell, M. J. D., Ed.; Academic: New York, 1982.

(15) Dennis, J. E.; Schnabel, R. B. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*; Prentice Hall: Upper Saddle River, New Jersey, 1983.

(16) Scales, L. E. *Introduction to Nonlinear Optimization*; Macmillam: Basingstoke, England, 1985.

(17) Pulay, P. *Chem. Phys. Lett.* **1980**, *73*, 393.

(18) Pulay, P. *J. Comput. Chem.* **1982**, *3*, 556.

(19) Cancès, E.; Le Bris, C. *Int. J. Quantum Chem.* **2000**, *79*, 82.

(20) Kudin, K. N.; Scuseria, G. E.; Cancès, E. *J. Chem. Phys.* **2002**, *116*, 8255.

(21) Li, X.; Millam, J. M.; Scuseria, G. E.; Frisch, M. J.; Schlegel, H. B. *J. Chem. Phys.* **2003**, *119*, 7651.

(22) Csaszar, P.; Pulay, P. *J. Mol. Struct.* **1984**, *114*, 31.

(23) Farkas, Ö.; Schlegel, H. B. *Phys. Chem. Chem. Phys.* **2002**, *4*, 11.

(24) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Ehara, M.; Toyota, K.; Hada, M.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Kitao, O.; Nakai, H.; Honda, Y.; Nakatsuji, H.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Cammi, R.; Pomelli, C.; Gomperts, R.; Stratmann, R. E.; Ochterski, J.; Ayala, P. Y.; Morokuma, K.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, *J. A. Development Version Rev. D01 ed.*; Gaussian, Inc.: Pittsburgh, PA, 2005.

(25) Bofill, J. M. *J. Comput. Chem.* **1994**, *15*, 1.

(26) Farkas, Ö.; Schlegel, H. B. *J. Chem. Phys.* **1999**, *111*, 10806.

(27) Baker, J. *J. Comput. Chem.* **1993**, *14*, 1085.

CT050275A

JCTC Journal of Chemical Theory and Computation

Transition Networks for the Comprehensive Characterization of Complex Conformational Change in Proteins

Frank Noé,[†,‡] Dieter Krachtus,[†,‡] Jeremy C. Smith,[†] and Stefan Fischer*[,‡]

Computational Molecular Biophysics, Interdisciplinary Center for Scientific Computing (IWR), University of Heidelberg, Im Neuenheimer Feld 368, 69120 Heidelberg, Germany, and Computational Biochemistry, Interdisciplinary Center for Scientific Computing (IWR), University of Heidelberg, Im Neuenheimer Feld 368, 69120 Heidelberg, Germany

Received June 24, 2005

Abstract: Functionally relevant transitions between native conformations of a protein can be complex, involving, for example, the reorganization of parts of the backbone fold, and may occur via a multitude of pathways. Such transitions can be characterized by a transition network (TN), in which the experimentally determined end state structures are connected by a dense network of subtransitions via low-energy intermediates. We show here how the computation of a TN can be achieved for a complex protein transition. First, an efficient hierarchical procedure is used to uniformly sample the conformational subspace relevant to the transition. Then, the best path which connects the end states is determined as well as the rate-limiting ridge on the energy surface which separates them. Graph-theoretical algorithms permit this to be achived by computing the barriers of only a small number out of the many subtransitions in the TN. These barriers are computed using the Conjugate Peak Refinement method. The approach is illustrated on the conformational switch of Ras p21. The best and the 12 next-best transition pathways, having rate-limiting barriers within a range of 10 kcal/mol, were identified. Two main energy ridges, which respectively involve rearrangements of the switch I and switch II loops, show that switch I must rearrange by threading Tyr32 underneath the protein backbone before the rate-limiting switch II rearrangement can occur, while the details of the switch II rearrangement differ significantly among the low-energy pathways.

1. Introduction

Conformational changes are critical to the function of many proteins. Well-known examples of functional transitions include the rearrangement of subunits in the hemoglobin tetramer upon oxygen binding,[1] the lever-arm motion in myosin during muscle contraction,[2] and the molecular switch in Ras p21 (see Figure 2) that signals cell division.[3,4] Such functional changes in conformation are often complex, involving the rearrangement of backbone segments or the packing at domain interfaces. Understanding the mechanism of these transitions is particularly challenging, because the nature and order of their subtransitions are difficult to predict and may, in principle, occur in different ways.

X-ray crystallography and nuclear magnetic resonance spectroscopy can provide atomic-detail structures for stable end states of conformational transitions and sometimes long-lived intermediates. However, the transitions themselves are difficult to characterize experimentally because, although the time required for a complete structural change can be relatively long (μs or longer), the transition states involved

* Corresponding author e-mail: stefan.fischer@iwr.uni-heidelberg.de.
† Computational Molecular Biophysics.
‡ Computational Biochemistry.

10.1021/ct050162r CCC: $33.50
Published on Web 03/11/2006

are very short-lived. Computer simulation can help to gain insight into these processes.

Because complex conformational transitions usually occur on long time scales, they are not accessible to unbiased molecular dynamics (MD) simulation with presently available computing power. Consequently, alternative computational approaches must be used. Variations of molecular dynamics have been proposed to overcome this time scale problem. For example, multiple time-step methods[5] are quite successful in certain multiple-time scale contexts. However, they do not achieve sufficient speedup for the present purposes.[6,7] Other methods bias the underlying energy potential[8] or reduce the dimension of the conformational space.[9] These methods face the difficulty that a good guess of the energy surface along the whole transition must in principle be known a priori, which is usually not possible for complex transitions in proteins. Steered and Targeted Molecular Dynamics[10,11] incorporate a constraint into the energy function that directs the system toward the desired product structure. While these methods are successful in cases where the reaction follows a pathway that is compatible with these constraints,[12] they lead to unnatural structures and unrealistic energy barriers in other cases.[13] A further variant is conformational flooding,[14] which approximates the local shape of the underlying energy surface explored by an MD trajectory by computing its principal components and then escapes the local energy minimum by adding a multivariate Gaussian function to the energy function the form of which depends on these principal components. Although this method allows the trajectory to overcome high energy barriers, it is designed to explore yet unknown new states rather than performing a transition into a predefined state.

Pathway methods are a different approach to simulating molecular transitions. Starting from an initial guess, the transition pathway is allowed to relax on the energy surface by constrained molecular dynamics[15,16] or by local minimization methods.[17−21] These methods have been applied successfully in cases where the transition does not involve too complex rearrangements of the protein, such that a number of reasonable initial guesses of the pathway can easily be formulated.[22−24] This is particularly the case when the conformational distance between the transition end states is small. In contrast, when the transition involves rearrangements of the protein fold, a guess for the initial path is more difficult to make. Moreover, such transitions can follow multiple pathways, as the energy landscape is likely to include broad energy ridges with many saddle-points of similar energies. Therefore, the determination of a single reaction pathway (even if it is the lowest-energy one) does not yield a comprehensive description of the transition.[13]

To represent multiple pathways, the *transition network* approach may be used. Transition networks are a discrete and simplified representation of configurational space and encode the possible transition pathways in a network of subtransitions. Each subtransition occurs between two conformations that are relatively close in conformational space. Each conformation in the network can be reached and left through at least one, but usually several, subtransitions. Each subtransition has an associated energy barrier that can be used to determine an associated rate constant or a mean passage time (i.e. "cost"). See Figure 1 for an illustration.

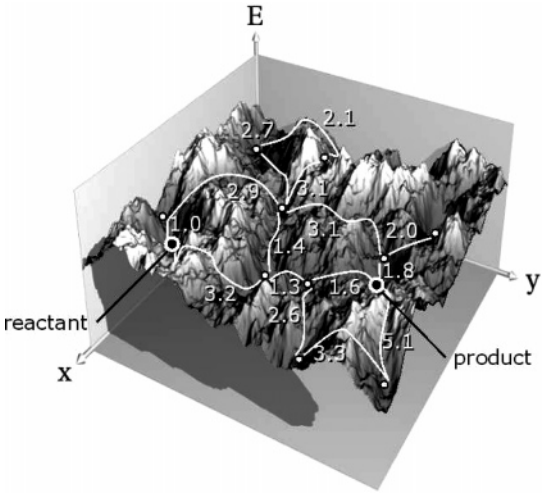


Figure 1. Transition network on a schematic two-dimensional energy surface. The network vertices (white bullets) correspond to low-energy intermediates between the reactant and product end states of the transition (black bullets). The network edges (white lines) correspond to subtransitions between the vertices and are associated with weights (white numbers), which are the rate-limiting energy barriers along each subtransition.

The construction of transition networks is documented in a large number of studies which have addressed the analysis of energy surfaces by mapping its local minima and saddle points.[25−46] These stationary points can be generated by local optimization starting from conformational ensembles that are generated by high-temperature molecular dynamics,[29,33,36,39,47] by a mode-following guided parallel search starting from a deep initial minimum,[35,42,48] or by Discrete Path Sampling (DPS).[43,45,49] The kinetics between groups of stationary points may be recovered using Master-Equation dynamics (MED),[28,29,32,36,39,35,38−40,42,43,45,46] Kinetic Monte Carlo (KMC),[45] or, again, by Discrete Path Sampling (DPS).[43,45,49] Typical applications of the above methodology are the rearrangement of atomic or molecular clusters[29−31] and the rearrangement or folding of peptides[27,28,32,33,36,35,38,40,43,45] and of model proteins.[34,42,50]

The applicability of the above approaches to complex transitions between native conformations of a protein is limited by two main difficulties. The first involves the generation of the minima which serve as TN vertices: It is a priori unclear how a conformational ensemble can be generated that adequately covers the volume of conformational space that is relevant for the transition. In particular, the direct manipulation of the backbone torsion angles or high-temperature dynamics are likely to disrupt the native structure, while search-based procedures may get lost in the huge number of possibly distant low-energy minima. Discrete Path Sampling is likely to be successful in identifying a connected channel between the end states, but it is unclear how it can identify a collection of considerably different channels. The second problem involves the computation of energy barriers. The determination of global properties of the network, such as the kinetics or the optimal path between

two end states,[51] requires the barriers of the subtransitions in the network to be known. Dense transition networks for complex macromolecular transitions typically have so many edges and the computation of each subtransition barrier is so CPU-demanding that the computation of all subtransition barriers cannot be afforded.

In the present contribution we address these two problems. In section 3.1 we present a procedure for efficiently sampling the relevant degrees of freedom of a complex transition between two native conformations of a protein so as to yield a representative set of low-energy conformers. Then, in section 3.2 we present a graph-theoretical approach that allows for determining global network properties (such as the best transition pathway or the rate-limiting ridge dividing the energy surface into reactant and product basins) based on the computation of only a limited subset of all subtransition barriers.

The methods introduced here are used to identify likely pathways and the order of events in an example system, the molecular switch of Ras p21 (section 4). This switch plays an important role in the signal transduction pathways that control proliferation, differentiation, and metabolism.[3,4] The conformational transition, which occurs in the GDP-bound state, involves a complex rearrangement of the backbone fold around the nucleotide binding site (Figure 2). The complexity of the transition suggests that it may occur via many different pathways, thus making it an excellent case for testing the present transition network approach. Among the questions that have been raised by a previous study of this conformational transition[13] and that are addressed here are as follows: (1) Is the rearrangement of switch I characterized by the side chain of Tyr32 threading underneath the backbone or by moving it through the solvent (see Figure 2)? (2) Is there a coupling between the switch I and switch II transitions, i.e. is the relative order of events in the two switch regions strictly defined? (3) Is there a well-defined unfolding pathway of switch II?

The present study reports on methodological advances which permit the generation and analysis of a comprehensive set of pathways for a complex conformational transition between two native conformations of a protein. The methodology is applicable to complex conformational changes in many other proteins whose functional time scale and complexity precludes the use of direct simulation.

2. Theory

Transition networks (TN) can in principle be used to model any dynamical system that can be appropriately described by a (possibly large) number of states and interstate transition rules. Here, we focus on TN for molecular systems and in particular for conformational changes in proteins. A TN is a discrete model which abstracts dynamic properties from the full system and captures its relevant kinetic behavior. Formally, a TN is a weighted graph, $\mathcal{G} = (\mathcal{V},\mathbf{X}^S,\mathbf{E}^S,\mathcal{E},\mathbf{X}^{TS},\mathbf{E}^{TS})$, where the list of vertices, $\mathcal{V} = (1,...,|\mathcal{V}|)$, represents the stable states of the protein, $\mathbf{X}^S = (\mathbf{x}_1^S,...,\mathbf{x}_{|\mathcal{V}|}^S)$ are the corresponding configuration vectors, and $\mathbf{E}^S = (E_1^S,..., E_{|\mathcal{V}|}^S)$ are corresponding state energies. The list of edges, $\mathcal{E} = ((u, v),...,(w, y))$, specifies between which pairs of states direct subtransitions are considered. $\mathbf{X}^{TS} = (\mathbf{x}_{uv}^{TS},..., \mathbf{x}_{wy}^{TS})$ are the configuration vectors of the corresponding transition states and $\mathbf{E}^{TS} = (E_{uv}^{TS},..., E_{wy}^{TS})$ are the associated transition state energies.

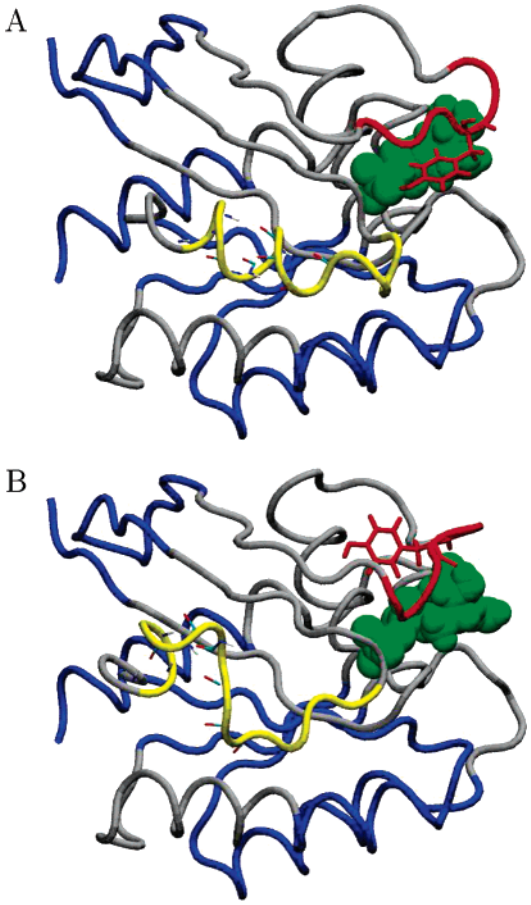


Figure 2. The conformational switch in Ras p21. (A) The GTP-bound and (B) GDP-bound conformers of the Ras p21 transition. The blue regions are very similar in the crystallographic end states and were kept fixed during the simulations. During the transition, switch I (residues 30−35, in red) rearranges such that Tyr32 (shown in red) is repositioned on the opposite site of the backbone and opens the nucleotide binding site to prepare for the release of GDP (shown in green van der Waals spheres). Switch II (residues 61−71, in yellow) unfolds from a helix to a coil structure.

Each TN vertex, v, corresponds to a region R_v of the configurational space, containing a group of geometrically similar molecular configurations. What is appropriate as a definition of "group" depends on the application. For the current work, each given vertex v corresponds to an *attraction basin*, i.e. the set of configurations that can be mapped to the same local minimum $\mathbf{x}_v^S$ on the potential energy surface $U(\mathbf{x})$ by a direct minimization.[25,52] Each vertex, v, is associated with a state energy E_v^S, generally, the free energy of the basin R_v. Depending on the complexity of the system used, different approximations to the free energy may be necessary. For systems where all basins can be mapped, it has been shown that free energies calculated from harmonic approximations to the potential in each basin are able to reproduce thermodynamic properties.[42,43]

The TN edges represent subtransitions between pairs of neighboring vertices. To any given edge (u, v), there is an associated transition state configuration, $\mathbf{x}_{uv}^{TS}$. The energy E_{uv}^{TS} associated with the edge is generally the free energy of the transition state.

The absolute height of the vertex and edge energies can be shifted by subtracting an arbitrary constant value E_0 without affecting the results. To avoid numerical problems when using exponentials of E_{uv}^{TS}, it is desirable to choose E_0 in such a way as to keep E_{uv}^{TS} small.

Figure 1 shows a schematic representation of a transition network.

2.1. Best Paths. Given the weighted graph $\mathcal{G}$, one can search for the "best" path connecting two particular vertices v_R and v_P (e.g. corresponding to experimentally determined "reactant" and "product" structures). For this, consider a subtransition along one edge between two vertices, $u \rightarrow v$, with associated energies E_u^S and E_{uv}^{TS}. We can express the flux from state u into state v, k_{uv}, as a product of the probability of being in state u, p_u, with the rate constant for the transition $u \rightarrow v$, k'_{uv}:[53]

$$k_{uv} = p_u k'_{uv} \tag{1}$$

Using free energies for vertices and edges, the probability p_u at equilibrium is given by the ratio of the partition functions for u and the full phase space:

$$p_u = \frac{\exp(-E_u^S/k_B T)}{\sum_{w=1}^{|\mathcal{V}|} \exp(-E_w^S/k_B T)} \tag{2}$$

Using the phenomenological form of transition state theory,[54] and assuming that all subtransitions have a similar dynamic prefactor, γ, the rate constant can be expressed as

$$k'_{uv} = \gamma \exp\frac{-(E_{uv}^{TS} - E_u^S)}{k_B T} \tag{3}$$

where k_B is Boltzmann's constant, and T is the temperature. Substituting eqs 2 and 3 into eq 1, we see that the equilibrium flux for the transition $u \rightarrow v$ is proportional to the Boltzmann weight of E_{uv}^{TS}:

$$k_{uv} \propto \exp\left(-\frac{E_{uv}^{TS}}{k_B T}\right) \tag{4}$$

The mean time between two subsequent transition events from u to v, τ_{uv}, is given by the inverse of the flux: $\tau_{uv} = k_{uv}^{-1}$. We take the *edge cost*, $\bar{w}_{uv}$, as proportional to τ_{uv}, setting the proportionality factor to unity:

$$\bar{w}_{uv} = \exp\left(\frac{E_{uv}^{TS}}{k_B T}\right) \tag{5}$$

For a path connecting vertices $v_1 = v_R$ and $v_m = v_P$ via a series of m vertices, $P = (v_1, v_2, ..., v_m)$, travelling over edges $((v_1, v_2), ..., (v_{m-1}, v_m))$, the best path is defined as that which minimizes the cumulative edge costs

$$C(P) = \sum_{k=1}^{m-1} \bar{w}_{v_k v_{k+1}} \tag{6}$$

This definition of a best path is similar to the previously proposed notion of the continuous pathway with "maximum flux" or "minimum resistance".[15,55] To determine the best path in practice, the edge energy vector $\mathbf{E}^{TS}$ is transformed into a cost-vector $\bar{\mathbf{w}}$ using eq 5. $\bar{\mathbf{w}}$ has size $|\mathcal{E}|$ and assigns a cost $\bar{w}_{uv}$ to each edge (u, v) in $\mathcal{E}$. Subsequently, the Dijkstra algorithm[51] is used to identify a best path between the two end states through the weighted network defined by $(\mathcal{V}, \mathcal{E}, \bar{\mathbf{w}})$. This path minimizes the path cost $C(P)$ given in eq 6.

This best path furnishes a preliminary understanding of the transition,[13] and it may be used as a guess for a reaction coordinate for free energy calculations[56] or as a starting point for discrete path sampling.[49] However, it dominates the transition only if the barriers of alternative pathways are considerably higher. To obtain an idea of the number of different accessible pathways and their associated structures, it is useful to determine the set of k different pathways, $(P_1, P_2, ..., P_k)$ with costs $(C_1 \leq C_2 \leq ... \leq C_k)$, where P_1 is the path with the lowest cost, C_1, P_2 is the path with the second-lowest cost, C_2, etc. This so-called "k best path problem" is well-known in graph theory.[57] To precisely define it, one must define in which way two paths must differ in order to be treated as different. In a transition network, it is clearly not very meaningful to distinguish two pathways which differ only in two low-energy, non-rate-limiting barriers. Therefore, two paths are treated as different only if their rate-limiting steps (i.e. their highest-energy edges) do not coincide. The k best paths are determined in k steps: The second best path is found by using the Dijkstra algorithm after "blocking" the edge (u, v) associated with the highest energy barrier in the previously found best path (by setting its $E_{uv}^{TS} = \infty$). The third best path is found by blocking the highest edges of the best and second best paths, etc.

2.2. Energy Ridge. The collection of rate-liming transition states from all different (as defined above) paths from a defined reactant to a defined product belongs to a $(D-1)$-dimensional transition surface that divides the D-dimensional conformation space into a reactant and a product side. In terms of topography, this transition surface corresponds to an *energy ridge*, as illustrated in Figure 3. On a geographical landscape, it is analogous to a water-shed, i.e. the mountain ridge that separates water flows toward distinct oceans. The particular interest of the energy ridge is that it allows for a feeling to be quickly obtained for how degenerate the transition is, i.e., how many significantly different paths are likely to be accessible. For instance, if one transition state in the ridge has a significantly lower energy than the other transition states in the ridge, then the transition mechanism is dominated by a well-defined bottleneck. In contrast, if the ridge contains many different transition states with similar energies, the transition mechanism is not well defined.

In graph-theoretical terms, an energy ridge is a *cut*. The name "cut" stems from the fact that deletion of its edges

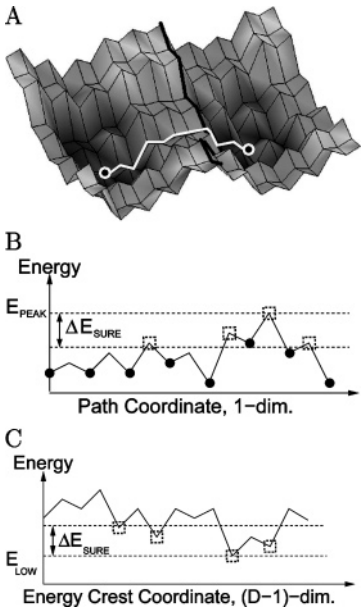


Figure 3. Schematic representation of the graph-theoretical concepts introduced in sections 2 and 3.3. (A) The best path (white line) connecting the transition end states (black bullets) and the energy ridge (black line) separating them. (B) Profile of vertex and edge energies along the best path through the network. The best path is requested to be correct in all edges with energies in the range $[E_{peak} - \Delta E_{sure}, E_{peak}]$ (indicated by squares). (C) Profile of the energy ridge cutting the TN into two conformational regions. The energy ridge is guaranteed to be correct in all edges with energies in the range $[E_{low}, E_{low} + \Delta E_{sure}]$.

dissociates the network into two disconnected subnetworks. Formally, the cut C is a set of M edges $C = \{(u_1, v_1),...,(u_M, v_M)\}$ with the property that each vertex u_i belongs to one set, U (e.g. "reactant side"), each vertex v_i belongs to another set, V (e.g. "product side"), and (U, V) partition the set of all vertices (i.e., $U \cup V = \mathcal{V}$ and $U \cap V = \varnothing$).

When the best and all next-best paths each have a dominant (rate-limiting) step, the energy ridge is identical to the cut whose total flux k_{UV} across it is minimal. k_{UV} is given by the sum of all localized fluxes $k_{u_iv_i}$ in the direction $u \rightarrow v$ across edges (u_i, v_i) in the cut

$$k_{UV} = \sum_{(u_i,v_i)\in C} k_{u_iv_i} \tag{7}$$

where $k_{u_iv_i}$ is the equilibrium flux from eq 4. By dismissing the proportionality constant, we obtain the normalized total flux, $k_{UV,0}$:

$$k_{UV,0} = \sum_{(u_i,v_i)\in C} \exp\left(-\frac{E^{TS}_{u_iv_i}}{k_B T}\right) \tag{8}$$

Note that the cut that minimizes $k_{UV,0}$ (the *rate-limiting cut*) and the cut associated with the topographic energy ridge are not always identical. For example, consider a case where the topographic ridge is very broad, its cut containing many edges of similar energy, whereas another cut contains only a single edge of slightly lower energy than those of the topographic ridge. Then the cut with the single edge has a lower $k_{UV,0}$ than the cut of the topographic ridge, because the many individual fluxes across the broad topographic ridge add up to a larger total flux. In the current context, however, this theoretical difference is not of importance.

The rate-limiting cut can be found by defining the vector of weights $\mathbf{w} = (w_{\mathcal{E}_1},..., w_{\mathcal{E}_{|\mathcal{E}|}})$, where for each edge (u, v) in the network $w_{uv} = \exp(-E^{TS}_{u_iv_i}/k_BT)$, and using the algorithm of Nagamochi and Ibaraki.[58] However, this algorithm is computationally expensive (scaling as $O(|\mathcal{V}|^3)$ or $O(|\mathcal{V}|^2 + |\mathcal{V}||\mathcal{E}| \log|\mathcal{V}|)$, depending on the implementation). Since the computation of the cut has to be repeated many times (see section 3.3.2), we used the topographical energy-ridge cut rather than the rate-limiting cut.

The topographical energy-ridge cut is determined by an algorithm that can be likened to flooding the energy landscape by stepwise filling up its basins. The ridge that last divides the reactant and product "lakes" before they become connected is the energy ridge. In the network, the edges which define the energy ridge are identified iteratively, starting from an edgeless network, $\mathcal{G}$, consisting only of the vertices, $\mathcal{V}$. In each iteration, a new edge $e \in \mathcal{E}$ is added to the network in order of increasing edge energy. At each iteration, the topology of $\mathcal{G}$ allows the identification of connected subgraphs (i.e. sets of vertices in which each vertex has at least one link to another vertex in the set). Each vertex is assigned an identifier that is unique for the connected subgraph it belongs to. The subgraph containing the reactant vertex is always assigned the identifier '0', while the subgraph containing the product vertex is always assigned '1'. Whenever an edge would be added that connects two vertices with identifiers '0' and '1', this edge is not added but marked as part of the energy ridge. The full ridge is determined when all edges have been iterated.

3. Methods

3.1. Efficient Sampling Procedure for Complex Conformational Changes in Proteins. In this section, a method is described for generating a representative sample of low-energy minima covering the conformational (sub)space relevant to a conformational transition. The method consists mainly of two stages: (1) generation of a sample of low-energy minima that are sparsely distributed over a large conformational subspace and (2) finding new low-energy minima between the minima found in (1) so as to densely map out the low-energy regions of conformational space.

In the first stage, the sampling can be limited to conformations that are likely to be relevant to the transition, thus avoiding sampling of the full conformational space (which would include, for example, mostly unfolded structures of the protein). For most transitions between native protein conformations, a good estimation can be made as to which structural regions are likely to require significant sampling. Small deformations in the remaining domains, which are structurally similar in the two end states, are then sampled by a simple interpolation between the end states. This partitioning and sampling procedure is described in sections 3.1.1−3.1.3 and is illustrated in Figure 4. Section 3.1.4 describes strategies for finding a uniformly dense set of low-

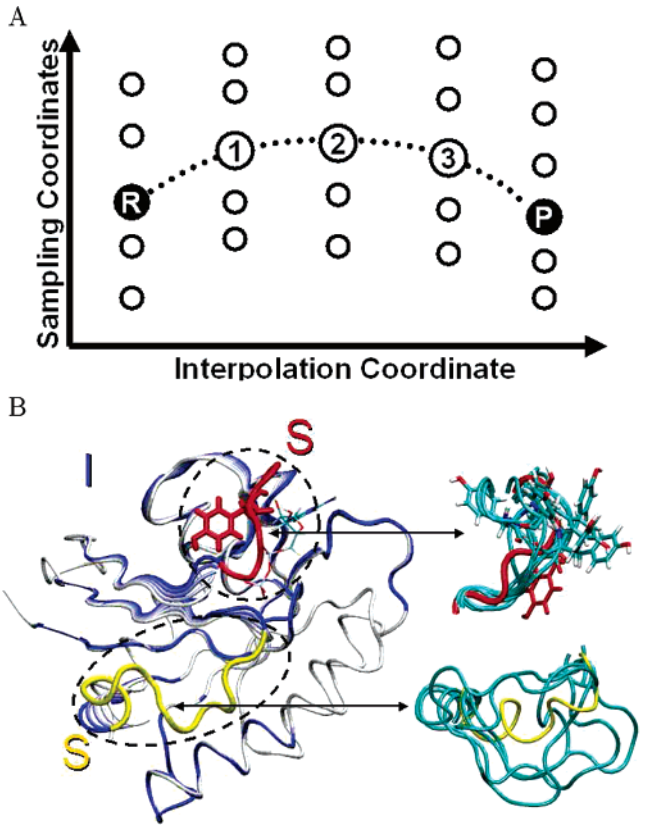


Figure 4. Illustration of the sampling procedure. (A) Some intermediate conformations are generated by interpolating the positions of a subset of protein atoms (the **I**-region) between their end state positions "R" and "P" (here 3 intermediate structures are generated, shown in circles). From each structure along the interpolation, a large set of conformations is generated by sampling the torsional angles of the **S**-region of the protein. The full set of conformations is defined by all combinations of the five (including both end states) conformations for the **I**-region with each sample of the **S**-region. (B) The interpolation (**I**) and sampling (**S**) regions in Ras p21. Left: The atoms of the **I**-region are interpolated between the transition end states (shown in white and dark blue), here producing three intermediates (shown in shades of blue). Right: The single-bond torsion angles of the **S**-region (switch I in red, switch II in yellow) are sampled uniformly (examples of several **S** conformations are overlaid). The **S** region encompasses switch I (residues 30−35) and switch II (residues 61−70).

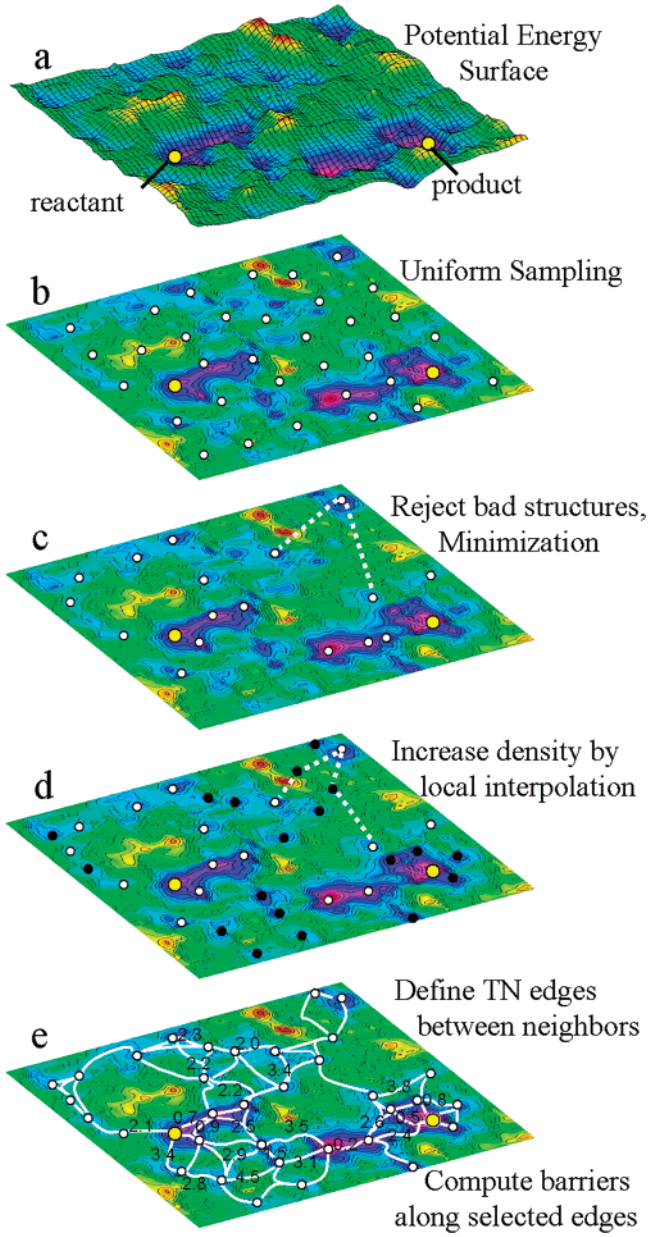


Figure 5. Overview of the steps in generating a transition network (TN). (a) The potential energy surface and the minimized reactant and product end states of the transition. (b) Conformers (white bullets) are uniformly spread over the part of conformational space that is relevant to the transition (see Figure 4 and section 3.1.3). (c) Structures without steric clashes are accepted (see Appendix A) and energy-minimized (see section 3.1.4). Interpolation between pairs of available conformers is used to explore nearby low-energy regions (dashed line). (d) The minimized interpolation intermediates (black bullets) increase the density of low-energy minima (see section 3.1.4). These minima form the TN vertices. (e) Pairs of neighboring minima are associated, forming the TN edges. The subtransitions of selected edges are computed by CPR, yielding the rate-limiting energy barriers for the TN edges (see section 3.2).

energy minima. Figure 5 gives a schematic overview of all the steps involved in the generation of the TN.

3.1.1. The Sampling (S) and Interpolation (I) Regions. Functionally relevant conformational changes in proteins are usually relatively local in the sense that most native contacts are preserved. While there might be complex rearrangements in certain regions, involving sometimes even refolding of parts of the backbone (such as in Ras p21), the remainder of the protein only deforms flexibly. This allows a sampling subspace with a considerably reduced dimensionality to be defined. Thus, the protein can be partitioned into interpolation (**I**) and sampling (**S**) regions. The changes in the **I** atoms are sampled by simply interpolating between the atomic positions in the two transition end states (see Figure 4a for an illustration). For each such interpolated structure, the rotable torsion angles of the **S**-region (including backbone ϕ/ψ torsions and single-bond side-chain torsions) are sampled

uniformly (Figure 4b). Finally, all degrees of freedom are locally relaxed during the energy minimizations that follow this combined **I**/**S** sampling procedure.

3.1.2. Interpolation of I-Atoms. To obtain a smooth variation of the positions of atoms in the **I** region near the boundary to the **S** region, the coordinates of the **I** region are generated by interpolating between the end states of the transition. For this, a combined interpolation procedure is used: First, so as to preserve the backbone fold, the backbone atoms are interpolated in Cartesian coordinates, and then the side-chain atoms are built onto the interpolated backbone, using internal coordinate values that are interpolated between the internal coordinates of the end states. This interpolation method has been shown to produce less distorted structures than Cartesian or internal coordinate interpolation alone.[13]

For practical convenience, the combined interpolation is done for all atoms of the protein (including the atoms of the **S** region). Because the **S** region has by definition very different conformations in the end states, the interpolated structures involve distorted internal coordinates in the **S** part of the backbone. To start the **S** sampling with reasonable values of the internal coordinates of the **S** region, each interpolated structure is energy minimized with positional harmonic constraints on the **I** atoms (force constant 1 kcal $mol^{-1}Å^{-1}$). In the example treated here, $n_{\text{interpol}} = 3$ interpolated structures of Ras p21 were generated in this way, yielding 5 structures along the interpolation including the end states.

3.1.3. Conformational Sampling of the S Region. For each of the structures along the interpolation between the two end states, many conformers of the **S** region are generated (Figure 5b). Sampling of the **S** region is performed uniformly in the space of flexible torsion angles, comprising the ϕ/ψ backbone and single-bond side-chain angles. The stiff internal degrees of freedom (i.e. bond lengths, valence angles and backbone ω angles) were not sampled here. The flowchart in Figure 6 summarizes the following algorithm for sampling backbone conformations.

If the sampling region is located at one of the termini of the polypeptide chain, there are no closure constraints on the backbone, allowing the ϕ/ψ angles to be sampled directly by setting them to random values. However, when the sampling region is within the polypeptide chain (as is the case for the switch I and switch II loops in Ras p21), this "free" sampling is not possible, as it would involve violation of the backbone closure (i.e., some backbone bond lengths and angles would not be preserved) or disruption of the native fold of the protein. Therefore, random backbone conformations are generated using a variant of the so-called window method.[59,60] This procedure allows backbone variation of a series of $r \geq 3$ consecutive residues (the "window") while preserving the position and orientation of the backbone at the boundaries of that window. Out of the windows' $2r$ torsion angles (ϕ and ψ), $2r - 6$ can be freely chosen and rotated randomly. The remaining six torsion angle values are determined by the window method (see ref 59 for a detailed description). In each sampling step, the location and length of the window in the **S** region and the rotated torsion angles are randomly chosen.

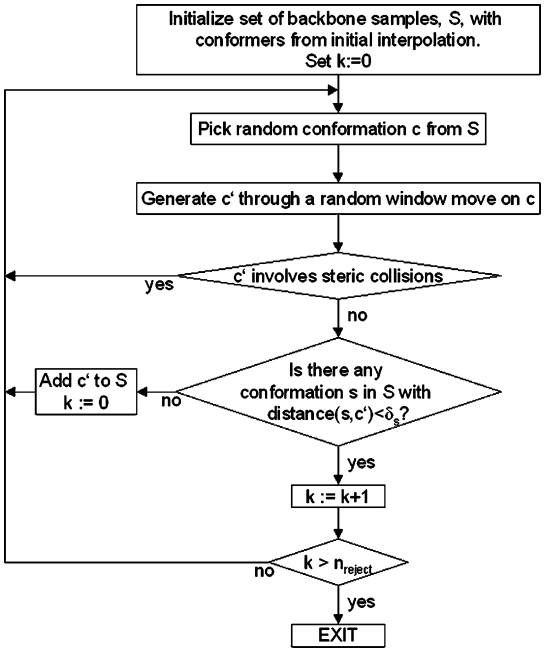


Figure 6. Flowchart for the backbone sampling procedure (section 3.1.3).

A backbone conformer is considered "valid" if it does not produce steric collisions. For this, it is checked whether the resulting backbone atoms and atoms whose positions are directly dependent on the backbone configuration (i.e., backbone O and H, C_β, and proline side chains) can be placed without collision among themselves and with the **I** region of the protein. Since a large number of **S** conformers have to be tested for collisions, an efficient strategy is used to perform these collision checks (see Appendix A).

To obtain a conformational sample that is approximately uniform, conformers that are valid (i.e., have no collisions) are "accepted" (i.e., added to a "conformational repository") only if they significantly differ from already-accepted conformers as measured by the ϕ/ψ dihedral RMS difference, which must exceed a chosen value, δ_s (see Table 2 for suggested parameter values). The choice of δ_s determines the density of sampling. It is set according to how finely the details of the transition should be probed. Backbone conformers are generated until the sampling density defined by δ_s has been reached. The criterion used here for this is that no "valid" structures are "accepted" anymore for a number of n_{reject} successive attempts. This yields a number of n_i^{back} backbone conformations for each interpolation step i ($i \in \{0,1,...,n_{\text{interpol}} + 1\}$, where $i = 0$ and $i = n_{\text{interpol}} + 1$ are associated with the end states and $1,...,n_{\text{interpol}}$ are the interpolated intemediates).

To obtain a complete conformation, the side chains of the **S** region are built onto a randomly picked backbone out of the n_i^{back} generated backbones, using randomly chosen single-bond torsion angles. The resulting conformer is accepted if it does not involve atom collisions (see Appendix A), giving for each interpolation step i a number n_i^{full} of sterically valid conformations of the **S** region. $n_i^{\text{full}} = k^{\text{side}} n_i^{\text{back}}$, where k^{side} is the desired average number of side chain conformers per backbone conformer. An efficient

Table 1: Frequently Used Symbols

symbol	meaning
$U(\mathbf{x})$	energy function
$\mathcal{V}=(1,...,\lvert\mathcal{V}\rvert)$	list of vertices in network; $\lvert\mathcal{V}\rvert$ is the number of vertices
$\mathrm{X}^S=(\mathbf{x}_1^S,...,\mathbf{x}_{\lvert\mathcal{V}\rvert}^S)$	conformers corresponding to vertices (here: minima on $U(\mathbf{x})$)
E_0	energy of the minimized reactant structure (see section 3.4)
$\mathbf{E}^S=(E_1^S,...,E_{\lvert\mathcal{V}\rvert}^S)$	vertex energies (here: $E_i^S=U(\mathbf{x}_i^S)-E_0$)
$\mathcal{E}=((u_1,v_1),...,(u_{\lvert\mathcal{E}\rvert},v_{\lvert\mathcal{E}\rvert}))$	list of edges in the network, each edge connecting two vertices of V; $\lvert\mathcal{E}\rvert$ is the number of edges
$\mathbf{X}^{TS}=(\mathbf{x}_{u_1v_1}^{TS},...,\mathbf{x}_{u_{\lvert\mathcal{E}\rvert}v_{\lvert\mathcal{E}\rvert}}^{TS})$	for each edge in $\mathcal{E}$, highest saddle point on minimum-energy path connecting the conformers $\mathbf{x}_{u_i}^S$, $\mathbf{x}_{v_i}^S$
$\mathbf{E}^{TS}=(\mathbf{E}_{u_1v_1}^{TS},...,\mathbf{E}_{u_{\lvert\mathcal{E}\rvert}v_{\lvert\mathcal{E}\rvert}}^{TS})$	edge (i.e. saddle-point) energies: $E_{u_iv_i}^{TS}=U(\mathbf{x}_{u_iv_i}^{TS})-E_0$
$\mathcal{G}=(\mathcal{V},\mathbf{X}^S,\mathbf{E}^S,\mathcal{E},\mathbf{X}^{TS},\mathbf{E}^{TS})$	transition network composed of vertices $\mathcal{V}$ connected by edges $\mathcal{E}$
$\mathbf{E}^{TS,min}=(\mathbf{E}_{u_1v_1}^{TS,\mathrm{min}},...,\mathbf{E}_{u_{\lvert\mathcal{E}\rvert}v_{\lvert\mathcal{E}\rvert}}^{TS,\mathrm{min}})$	lower bounds to the (yet unknown) edge energies
$\mathbf{E}^{TS,\mathrm{max}}=(\mathbf{E}_{u_1v_1}^{TS,\mathrm{max}},...,\mathbf{E}_{u_{\lvert\mathcal{E}\rvert}v_{\lvert\mathcal{E}\rvert}}^{TS,\mathrm{max}})$	upper bounds to the (yet unknown) edge energies
$w=(w_{u_1v_1},...,w_{u_{\lvert\mathcal{E}\rvert}v_{\lvert\mathcal{E}\rvert}})$	Boltzmann weights of the edge energies: $w_{u_iv_i}=\exp(-E_{u_iv_i}^{TS}/k_\mathrm{B}T)$
$\bar{\mathbf{w}}=(\bar{w}_{u_1v_1},...,\bar{w}_{u_{\lvert\mathcal{E}\rvert}v_{\lvert\mathcal{E}\rvert}})$	inverse Boltzmann weights of the edge energies: $w_{u_iv_i}=exp(E_{u_iv_i}^{TS}/k_\mathrm{B}T)$
$\bar{\mathbf{w}}^{\mathrm{min}}=(\bar{w}_{u_1v_1}^{\mathrm{min}},...,\bar{w}_{u_{\lvert\mathcal{E}\rvert}v_{\lvert\mathcal{E}\rvert}}^{\mathrm{min}})$	inverse Boltzmann weights of the lower edge energy bounds: $w_{u_iv_i}^{\mathrm{min}}=exp(E_{u_iv_i}^{TS,\mathrm{min}}/k_\mathrm{B}T)$
$\bar{\mathbf{w}}^{\mathrm{max}}=(\bar{w}_{u_1v_1}^{\mathrm{max}},...,\bar{w}_{u_{\lvert\mathcal{E}\rvert}v_{\lvert\mathcal{E}\rvert}}^{\mathrm{max}})$	same as $\bar{\mathbf{w}}^{\mathrm{min}}$, for upper bounds

Table 2: Parameters for the Sampling Algorithm

parameter	meaning	value
n_{interpol}	number of steps along the interpolation between the transition end states, including the end states ($n_{\mathrm{interpol}}\geq 2$)	5
δ_s	shortest permitted RMS distance between two accepted backbone conformers (here: in ϕ/ψ torsion angle space)	50°
n_{reject}	sampling has converged when n_{reject} newly generated conformers are successively rejected because they are closer than δ_s to already-accepted conformers	1000
k^{side}	average number of side chain conformers per backbone conformer.	10
E_{tol}	largest permitted interaction energy between each atom-pair, for a structure to be considered valid	20 kcal/mol
n_{min}	number of structures drawn from the conf. repository for minimization	15000
E_{low}	largest permitted energy difference above minimized reactant structure to accept a conformer	40 kcal/mol
$\delta_{\mathrm{min}}^{\mathrm{interpol}}$	shortest and longest permitted RMS distance between a pair of minima to generate additional conformers by interpolation between them	$\delta_{\mathrm{min}}^{\mathrm{interpol}}=0.75$ Å
$\delta_{\mathrm{max}}^{\mathrm{interpol}}$		$\delta_{\mathrm{max}}^{\mathrm{interpol}}=2$ Å
$\delta_{\mathrm{min}}^{\mathrm{connect}}$	shortest permitted RMS distance between any pair of minima $\mathbf{x}_u^S$, $\mathbf{x}_v^S$, to avoid redundancy in the TN	0.75 Å
$\delta_{\mathrm{max}}^{\mathrm{connect}}$	longest permitted RMS distance between any pair of minima $\mathbf{x}_u^S$, $\mathbf{x}_v^S$, to form an edge (u, v) in the TN	1.5 Å
$n_{\mathrm{max}}^{\mathrm{connect}}$	maximum number of neighbors for each vertex	20

protocol for building side chains on large **S** regions is described in Appendix B.

For Ras p21, the switch I and II **S** regions were sampled independently, using $\delta_\mathrm{s}=50°$ and $n_{\mathrm{reject}}=1000$. For each interpolation step, i ($i\in\{0,...,4\}$), this yielded $n_i^{\mathrm{back1}}\approx 30$ backbone conformers for switch I and $n_i^{\mathrm{back2}}\approx 10^4$ backbone conformers for switch II. An average of $k_{\mathrm{side}}=10$ side-chain conformations per backbone conformer were generated, yielding $n_i^{\mathrm{full1}}\approx 300$ and $n_i^{\mathrm{full2}}\approx 10^5$ collision-free conformations of switch I and II, respectively. Combining pairs of these switch I and II conformers yielded 3×10^7 fully built protein structures for each interpolation step i. Thus, the total number of collision-free and significantly different structures is $n^{\mathrm{full}}=1.5\times 10^8$, forming a large conformational repository from which structures can be drawn and further energy optimized. The conformations in this repository are distributed uniformly within the sterically accessible regions of the conformational subspace spanned by the torsional coordinates of **S** and the interpolation coordinate of **I**.

3.1.4. Constructing a Uniformly Dense Set of Low-Energy Minima. To obtain a representative collection of low-energy minima, a number of n_{min} conformers is drawn randomly from the conformational repository and energy-minimized on the potential $U(\mathbf{x})$ (see Figure 5c). Only minima which reach a low-energy region defined by $U(\mathbf{x})<E_{\mathrm{low}}$ are accepted, where E_{low} is a predefined constant. Minimization of many conformers is expensive, so it is desirable to reject structures early which are not likely to fall into low-energy minima. An efficient method to do this is proposed in Appendix C. For Ras p21, $n_{\mathrm{min}}=15\,000$ conformers were randomly retrieved from the conformational repository. Out of these, 189 reached the desired low-energy region below E_{low}, which was taken here as 40 kcal/mol above the energy of the minimized reactant structure (obtained by quenched molecular dynamics, see section 3.4). These were minimized to a gradient RMS of 10^{-3} kcal mol^{-1} Å^{-1}. They form a sparse set of low-energy conformations in the desired region of conformational space (see Table 4).

The density of conformers in the low-energy regions can, in principle, be increased by minimizing more structures from the conformational repository. Given the low yield of this approach (see above: $189/15\,000\approx 1.25\%$), this is com-

Table 3: Sampling of the **S** Regions in Ras P21

symbol	meaning (see section 3.1.3)	value
n_i^{back1}	number of backbone conformers (residues 30−35), for each interpolation step i ($i \in \{0,...,4\}$)	≈30
n_i^{back2}	same as n_i^{back1}, for residues 61−70	$\approx 10^4$
n_i^{full1}	number of fully built conformers with side chains (residues 30−35), for interpolation step i	≈300
n_i^{full2}	same as n_i^{full1}, for residues 61−70	$\approx 10^5$
n^{full}	total size of the conformational repository	$1.5 \cdot 10^8$

Table 4: Size and Density of the Network during Sampling

	minima[a]	accepted[b]	neighbors[c]
after first sampling[d]	15002	189	3
after increasing density[e]	35836	10831	267
TN vertices[f]	n/a	6242	117

[a] The total number of generated energy-minima. [b] The number of accepted minima with energy below E_{low}. [c] The average number of neighbors around accepted minima within a distance-range from $\delta_{\text{min}}^{\text{connect}}$ to $\delta_{\text{max}}^{\text{connect}}$. [d] After the sampling of the **I** and **S** regions (sections 3.1.2 and 3.1.3). [e] After increasing the density of low-energy minima (section 3.1.4). [f] After removing redundancy by not allowing neighbors closer than $\delta_{\text{min}}^{\text{connect}}$ (section 3.2).

putationally inefficient. Instead, additional conformers are built by interpolation between the already-found low-energy conformers. This can be done in various ways. The strategy used here was to select each pair of low-energy conformers separated by a distance in the range $\delta_{\text{min}}^{\text{interpol}} = 0.75$ Å and $\delta_{\text{max}}^{\text{interpol}} = 2$ Å (measured as Cartesian RMSD of the C_α atoms in the **S** region) and to generate an interpolation pathway between them using the method described in section 3.1.2. Two structures were generated, one-third and two-thirds of the way along each interpolation, respectively, and energy minimized as described in section 3.1.4 (see Figure 5d). This procedure was efficient in finding low-energy minima, increasing the number of conformers below E_{low} from 189 to 10 831 (see Table 4). This considerably increased the average number of neighbors for each minimum from 3 to 267 ("neighborhood" being defined by a cutoff distance $\delta_{\text{max}}^{\text{connect}}$, see section 3.2).

During minimization, it is possible that some conformers end up in similar minima. This produces conformational redundancy, which was subsequently removed. For this, minima were considered in the order of increasing energy, accepting only those minima whose nearest-neighbor distance to any already-accepted minimum was at least $\delta_{\text{min}}^{\text{connect}} = 0.75$ Å. This led to a final number of $|\mathcal{V}| = 6242$ diverse minima.

3.1.5. Verification of the Set of Minima. The available set of minima is approximately uniformly distributed in a conformational subspace which depends on the original definitions of **S** and **I**. There are two questions regarding the adequacy of this set of minima: (1) is the set dense enough and (2) are relevant parts of conformational space sampled.

The density of the set may be increased by reducing the parameter $\delta_{\text{min}}^{\text{connect}}$ and conducting further interpolations. Clearly, there is a tradeoff between the density of minima and the computational requirements. The important question

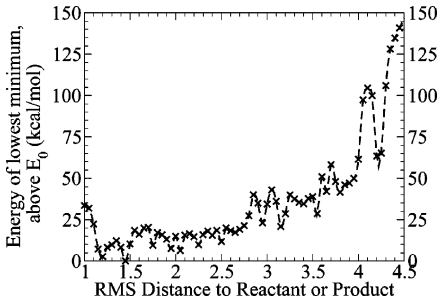


Figure 7. Dependence of energies of minima on the distance from the crystallographic end states. For each minimum, the distance to the reactant or product structure (whichever is closer) is calculated. The minima with distances between 1 and 4.5 Å were grouped according to their distances, each group being 0.1 Å wide. The lowest energy of the minima within each group is plotted versus the distance of that group. The plot shows that the energies increase considerably with increasing distance from the crystallographic end states, indicating that the relevant portions of conformational space have been sampled.

is how sensitive the analyses are to the density of minima. In the present study, density variation, within reason, has little effect as the CPR path calculations used to compute the edge energies ensure that no important intervening barriers are missed (see section 3.2). Moreover, the purpose of the present calculations is to generate a coarse-grained model of the Ras p21 energy function which is analyzed in the Results section based on qualitative properties of large sets of pathways. Local features of the network do not play a role in this analysis, and therefore the density of minima was not further increased.

A more critical question is whether there are important parts of the conformational space that are not sampled at all. A logical check of this is to examine whether any low-energy minima exist in regions of conformational space continguous with regions already explored. If so, energetically accessible pathways might exist that lead out of the available set of minima into other regions of the conformational space which were not included in the initial sampling. This can be checked by calculating the lowest-energy minima within shells that are increasingly distant from the reactant and product vertices and ensuring that the found set of minima defines an energetic basin that is unlikely to be left.

To examine this, we have analyzed all 35 836 minima that were generated (see Table 4) and computed their distances to the reactant or product structure (whichever of the two was closer). For each distance-window between 1 and 4.5 Å, the lowest energy of all minima within that window was recorded. The result, which is shown in Figure 7, shows that there is a strong increase in energy with increasing distance, reaching about 150 kcal/mol above E_0 at 4.5 Å. This result shows that the existence of low-energy exit pathways from the initial set of minima is unlikely, and therefore a sufficient volume of conformational space has been sampled.

3.2. Construction of the Ras p21 Transition Network. The final number of $|\mathcal{V}| = 6242$ diverse minima served as the vertices of the transition network (see Table 4).

Due to the large number of possible conformations of the system studied in this article, it is possible to map only a subset of the dynamically accessible minima, so that the computation of meaningful free energies is not feasible. Here, the contributions from local vibrations around the minima are neglected, and the potential energy $U(\mathbf{x}_v^S)$ is used directly for E_v^S. Nevertheless, the important free energy contributions from bulk solvent are accounted for in the calculation of $U(\mathbf{x})$ through a continuum solvent method (see section 3.4).

The absolute height of the vertex can be shifted by subtracting an arbitrary constant value E_0 without affecting the results. Thus, we define E_v^S as $E_v^S := U(\mathbf{x}_v^S) - E_0$. Here, E_0 is chosen as the minimized reactant energy.

The "reactant" and "product" vertices were redefined by selecting the lowest energy minima within the vicinity of the crystallographic reactant and product structures after quenched MD (see section 3.4). The "vicinity" was defined here to be within both a ϕ/ψ-RMSD of 50° and a Cartesian RMSD of 1.5 Å for the C_α-atoms of the switch regions. The resulting Cartesian RMSD over all C_α atoms between the crystal structures and the so-chosen reactant and product conformers was 1.4 and 1.5 Å, respectively.

Given the complete set of vertices, $\mathcal{V}$, the edges of the transition network are generated by defining connections according to distance-based criteria. Each vertex is connected to up to $n_{\max}^{\text{connect}}$ of its nearest neighbors that are within a distance $\delta_{\max}^{\text{connect}}$ (see Figure 5e). For Ras p21, $\delta_{\max}^{\text{connect}} = 1.5$ Å (measured as RMS distance between the C_α-atoms of the **S** region), and $n_{\max}^{\text{connect}} = 20$ were used. The resulting transition network had $|\mathcal{E}| = 47\,404$ edges and was fully connected (i.e. any given pair of vertices is connected by some pathway).

To determine the energy barrier associated with a given edge, (u, v), a Minimum Energy Path (MEP) between the two minima $\mathbf{x}_u^S$ and $\mathbf{x}_v^S$ corresponding to that edge was computed using the Conjugate Peak Refinement (CPR) method.[18] An initial guess for the path was generated by interpolation between the edge end structures, using the procedure described in section 3.1.2. Starting from this guess, CPR identifies all the first-order saddle points that are local energy maxima along the Minimum Energy Path. Here, the CPR calculation was stopped as soon as the highest (i.e. rate-limiting) of these saddle points along the edge was determined. Its structure is assigned to $\mathbf{x}_{uv}^{TS}$, the corresponding edge energy is taken as $E_{uv}^{TS} = U(\mathbf{x}_{uv}^{TS}) - E_0$ (see section 2).

3.3. Efficient Determination of Best Paths and Energy Ridges. Even though the subtransition pathways are short compared with a whole pathway between the transition end states, finding the highest saddle-point along an edge can still be very CPU intensive (here, the average time on a single 3 MHz CPU was about 2 h per subtransition). Therefore, it is computationally infeasible to do this for all edges in the network. The problem thus arises that global properties of the network (such as the best path or the dividing energy ridge) must be determined using incomplete information on the barriers. We solve this problem by devising a strategy such that only a small number of edge energies need to be computed to determine the best path and the energy ridge. The strategy relies on the introduction of lower and upper bounds, $E_{uv}^{TS,\min}$, $E_{uv}^{TS,\max}$ on each edge energy, which bracket the (yet unknown) true edge energy, E_{uv}^{TS}.

The "safe" lower bound for the edge energy is $E_{uv}^{TS,\min} = \max\{E_u^S, E_v^S\}$, because the energy of each barrier is at least as high as the highest of the two minima it connects. The safe upper bound is $E_{uv}^{TS,\max} = \infty$, but for numerical reasons it is taken as $E_{uv}^{TS,\max} = E_{uv}^{TS,\min} + M$, where M is a large but finite number (here, M=100 kcal/mol). A tighter upper bound could also be obtained by performing a very short (i.e. unconverged) CPR path refinement on the edge (u, v) and using the highest energy along the resulting path as $E_{uv}^{TS,\max}$. An alternative method to both the lower and upper bounds, based on statistical estimates, is used here (see Appendix D).

To implement lower and upper bounds, two graphs are defined, which have the same topology as the actual transition network: one using the lower bounds for the edge energies, $\mathcal{G}^{\min} = (\mathcal{V}, \mathbf{X}^S, \mathbf{E}^S, \mathcal{E}, \mathbf{E}^{TS,\min})$, and the other using the with upper bounds for the edge energies, $\mathcal{G}^{\max} = (\mathcal{V}, \mathbf{X}^S, \mathbf{E}^S, \mathcal{E}, \mathbf{E}^{TS,\max})$. The best paths (and energy ridges) through $\mathcal{G}^{\min}$ and $\mathcal{G}^{\max}$ can be computed, using the corresponding inverse Boltzmann weight vectors, $\bar{\mathbf{w}}^{\min}, \bar{\mathbf{w}}^{\max}$ and Boltzmann weight vectors $\mathbf{w}^{\min}, \mathbf{w}^{\max}$. This is addressed in the next sections.

3.3.1. Best Paths. The flowchart in Figure 8 summarizes the procedure for finding the best path. The best path through the transition network is found iteratively as follows: In each iteration, the best path through the graph $\mathcal{G}^{\min}$, $P_{\text{best}}^{\min} = (v_R, \ldots, v_P)$, is determined as described in section 2.1. The edge with the highest unknown energy along $P_{\text{best}}^{\min}$, (u, v), is identified, and its true energy E_{uv}^{TS} is computed by CPR. The network $\mathcal{G}^{\min}$ is updated by setting $E_{uv}^{TS,\min}$ to the true edge energy, E_{uv}^{TS} (the weights $\bar{\mathbf{w}}^{\min}$ are also updated accordingly). This procedure is repeated until all edge energies along the resulting best path have been computed, yielding the true best path.

A preliminary estimate of the energy barrier of the best path can be easily obtained. For this, in each iteration of the above algorithm, the best path is calculated on both graphs $\mathcal{G}^{\min}$ and $\mathcal{G}^{\max}$, yielding best paths $P_{\text{best}}^{\min}$ and $P_{\text{best}}^{\max}$, respectively. The rate-limiting barrier of the true best path is bounded by those of $P_{\text{best}}^{\min}$ and $P_{\text{best}}^{\max}$. During successive iterations, these bounds converge to the true value (see Figure 9).

Often, one is not interested in the details of how the best path travels in the low-energy regions, since it is the highest-energy edges along the whole path that are rate determining. Computation time can thus be saved if only the high-energy edges of the best path, i.e., those with barrier energies within a range ΔE_{sure} of the highest-energy barrier along the path, E_{peak}, are required to be correct (see Figure 3b). To achieve this, the computation proceeds as above until the energy of the rate-limiting barrier, E_{peak}, is identified. This is the case when $P^{\min}$ and $P^{\max}$ have the same value for E_{peak}. After that, whenever a barrier is computed whose energy E_{uv}^{TS} is below the threshold $E_{\text{peak}} - \Delta E_{\text{sure}}$, the transition networks are updated by setting $E_{uv}^{TS,\min} = E_{uv}^{TS,\max} := \max\{E_u, E_v\}$, i.e., as

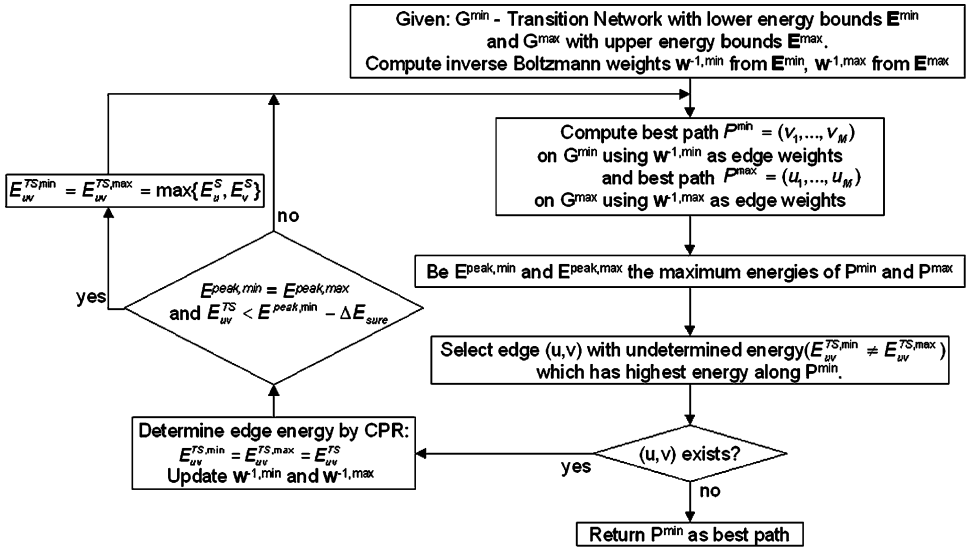


Figure 8. Flowchart for the best path finding procedure (section 3.3.1).

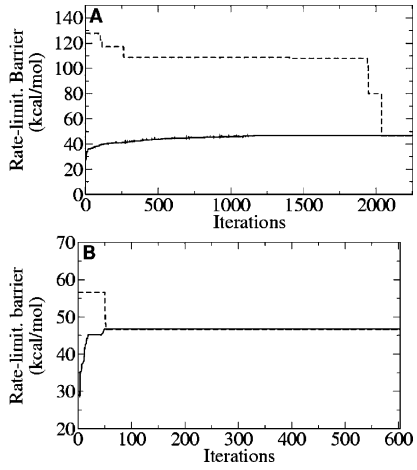


Figure 9. Convergence behavior of the best-path. The energy of the rate-limiting transition state along the best path $P_{\text{best}}^{\min}$ (solid line) or $P_{\text{best}}^{\max}$ (dashed line) is plotted as a function of iterations of the algorithm (see section 3.3.1). (A) Using "safe" bounds, i.e. setting the lower and upper bounds to 0 and ∞. (B) Using guessed bounds, i.e. setting the barrier bounds based on statistics (see Figure 13).

if the transition were barrierless. This saves computational effort as it makes sure that (u, v) is a part of the best path identified in the next Dijkstra computation, thereby avoiding to spend time in identifying paths avoiding (u, v) that might have a lower energy barrier than E_{uv}^{TS}.

The computing time can be drastically reduced, at the expense of possibly failing to identify the true best path, if the "safe" lower edge energy bound $E_{uv}^{TS,\min} = \max\{E_u, E_v\}$ is replaced with a statistical estimate (described in more detail in Appendix D). In this case the lower energy bound is not necessarily correct as it might overestimate the real barrier. That is, edges which are not included in the resulting best path and have been rejected based on their lower estimate $E_{uv}^{TS,\min}$ might in fact have a true edge energy $E_{uv}^{TS} < E_{uv}^{TS,\min}$. The maximum overestimation possible, $\text{err}_{\max}$, is given by the maximal difference found between any estimated lower bound and the corresponding safe lower bound:

$$\text{err}_{\max} = \text{MAX}[E_{uv}^{TS,\min} - \max\{E_u^S, E_v^S\}]_{\text{all pairs}(u,v)} \qquad (9)$$

Thus, $\text{err}_{\max}$, also gives the maximum possible error on the rate-limiting barrier of the path. If the graph $\mathcal{G}^{\max}$ is used to obtain a preliminary estimate of the energy barrier of the best path, this estimate is usually improved by also replacing the safe upper energy bound $E_{uv}^{TS,\max} = \infty$ by an statistical estimate. As the identification of the best path does not depend on $\mathcal{G}^{\max}$, an incorrect upper bound $E_{uv}^{TS,\max}$ cannot lead to a wrong result but may give a wrong upper bound for the preliminary estimate of the best path's energy barrier.

3.3.2. Energy Ridges. As defined in section 2.2, the energy ridge is the rate-limiting cut that divides the TN into a reactant and product side. The algorithm that computes the energy ridge while determining only a limited number of edge barriers uses a strategy similar to the one used to find the best-path. The energy-ridge cut is determined (as described in section 2.2) on the graph $\mathcal{G}^{\max}$ (i.e. using upper bounds for the yet unknown edge energies). The lowest-energy unknown edge, (u, v), in the resulting cut, $C^{\max}$, is computed by CPR, and $\mathcal{G}^{\max}$ is updated with the value of E_{uv}^{TS}. This process is repeated in successive iterations. When an energy-ridge cut $C^{\max}$ is identified whose edge barriers are all determined, it is identical to the true energy ridge.

In practice, it is sufficient to compute only the low-energy barriers of the energy ridge, since the higher-energy barriers are not populated. Thus, one is only interested in finding the energy barriers of the energy ridge that are up to an energy difference ΔE_{sure} above the energy of the lowest barrier in the ridge, E_{low} (see Figure 3c). To find these barriers, the algorithm given above proceeds until the value of E_{low} is identified. From this moment on, whenever a barrier is computed which has $E_{uv}^{TS} > E_{\text{low}} + \Delta E_{\text{sure}}$, $\mathcal{G}^{\max}$ is updated by setting $E_{uv}^{TS} = \infty$. This fools the algorithm so that it leaves these high-energy barriers in the ridge and thereby saves the computational cost of identifying the high-energy regions of the full ridge.

Finding the energy ridge depends on the upper bounds on the edge energies; therefore, the performance of the algorithm is considerably increased if estimates (see Appendix D) are used instead of infinite upper bounds. However, in contrast to the error involved in the determination of best paths (eq 9), no upper bound for the error from using such estimates can be derived here.

3.4. Test Case: Protein Model and Energy Function. The method is tested here on Ras p21. The conformational change in Ras occurs after the γ-phosphate of bound GTP is cleaved off. GTP hydrolysis can be catalyzed by the binding of a GTPase-activating protein (GAP)[61] or it can take place as a result of the weak intrinsic GTPase activity of Ras in absence of GAP. The conformational transition studied here is in the absence of GAP, as would occur after intrinsic hydrolysis of GTP. Crystallographic structures exist for both the GTP-bound (Protein Data Bank structure 5p21[62]) and GDP-bound states (1q21[63]) of Ras p21. The γ-phosphate was deleted from 5p21, to yield the reactant state. The 1q21 structure served as product state. The HBUILD facility in CHARMM[64] was used to place the missing hydrogens.

All calculations were performed using the extended-carbon potential function (PARAM19).[65] Contributions from bulk solvent to the free energy of the conformational substates were included with the Generalized Born model of continuum solvation, using version 2 of the Analytical Continuum Electrostatics (ACE) method.[66] Nonbonded interactions were smoothly brought to zero by multiplying them with a switching function between 8 and 12 Å.

The structure of a protein may be affected by the crystal environment. Therefore, both the reactant and the product structures were first relaxed using molecular dynamics simulations with ACE. For this 20 ps of heating were followed by 100 ps of equilibration and a 10 ns production run. One structure every 100 ps (making up 100 structures in total) was selected and energy minimized with ACE to a gradient RMS of 10^{-3} kcal mol^{-1} Å^{-1}. The structures with the lowest energies were selected as reactant and product structures. The potential energy of these structures was lower than that obtained by a direct minimization of 5p21 and 1q21 by 30−45 kcal mol^{-1}. Structurally, the differences compared to 5p21 and 1q21 were rather small, consisting mainly of exposed side-chain rearrangements, while the backbone fold of the switch regions was preserved. The RMS coordinate deviations from the directly minimized crystallographic end states were <1.8 Å for the nonfixed atoms (<2.4 Å for the switch regions).

To remove insignificant degrees of freedom, residues which were not involved in the conformational switch and whose atoms had similar positions in both end states were fixed (residues 1−4, 42−53, 77−95, 110−115, 124−143, 155−167), leaving 1001 atoms free to move. To obtain the same positions for the fixed atoms in the two end states, the product structure was oriented onto the reactant structure so as to minimize the RMS deviation between the fixed atom coordinate sets. Then, the reactant and product values of these coordinates were averaged. The averaged coordinates of the fixed atoms were used for all calculations. Furthermore, insignificant differences in the side chains of nonswitch regions were removed from the end states as described in ref 13. Finally, both end states were minimized to a gradient RMS of 10^{-3} kcal mol^{-1} Å^{-1}.

Table 5: Number of Edges Computed with CPR To Determine the Best Path[a]

ΔE_{sure}[b]	safe bounds[c]	guessed bounds[d]
∞[e]	2252	603
30	2246	589
25	2224	565
20	2208	505
15	2115	321
10	2069	212
5	2059	114
0	2059	106
path length[f]	23	24
energy barrier[g]	45.7	45.7

[a] Assuming no energy barrier has been previously computed. [b] Energy range below the highest barrier for which the barriers of the best path are to be determined (see Figure 3B). [c] Using $E_{uv}^{\min} = \max\{E_u, E_v\}$, $E_{uv}^{\max} = \infty$ as bounds on the unknown energy barriers. [d] Using statistical estimates (Appendix D) to guess the $E_{uv}^{\min}$ and $E_{uv}^{\max}$ bounds. [e] ∞ means that the whole best path with all its edges is to be determined. [f] Number of edges along the fully determined best path. [g] Rate-limiting energy barrier along the best path, in kcal/mol relative to the reactant.

4. Results and Discussion

4.1. Performance of Best Path Calculations. Best paths between the reactant and product structures of the Ras p21 conformational switch were computed using the iterative algorithm described in section 3.3.1. The performance was evaluated, first using safe values for the upper and lower bounds on the edge barriers (i.e. $E_{uv}^{TS,\min} = \max\{E_u, E_v\}$ and $E_{uv}^{TS,\max} = \infty$). Alternatively, statistical estimates for the bounds (described in Appendix D) were used. The partial computation of best paths, using different values for the energy interval ΔE_{sure} (see section 3.3.1 and Figure 3B) was also examined.

Table 5 shows how many edges need to be computed with CPR in order for the best path to be determined under these different conditions (starting the count from scratch for each setting). To determine the full best path ($\Delta E_{\text{sure}} = \infty$) using safe bounds on the energy barriers, 2252 edges had to be computed (only 5% of the total number of 47 404 edges). It was possible to reduce this number by a factor of 4 (to 603) when statistical estimates of the bounds were used. This faster convergence behavior is also apparent when comparing Figure 9A,B and demonstrates that the computation time can be greatly reduced by introducing a relatively small uncertainty. In the worst case, the error on the rate-limiting barrier resulting from the present estimates of $E_{uv}^{TS,\min}$ could have been as much as 5.25 kcal/mol (from eq 9). But in the present case, statistical estimation actually resulted in a best path with the same rate-limiting energy barrier as found when safe bounds were used. Moreover, except for one additional, insignificant low-energy step, the estimated best path is equal to the true best path.

The computational savings are even larger when only the highest-energy barriers of the best path are determined. The number of edges that need to be computed with CPR when

Table 6: Number of Edges Computed with CPR To Determine Ridge 2[a]

ΔE_{sure}[b]	safe bounds[c]	guessed bounds[c]
∞[d]	1092	667
20	1092	622
15	1092	509
10	897	383
5	862	293
0	805	214
ridge size[e]	174	162
energy barrier[f]	45.7	45.7

[a] The energy ridge of the switch II rearrangement, assuming no energy barrier has been previously computed. [b] Energy range above the lowest barrier for which ridge barriers are to be determined (see Figure 3C). [c] Same meaning as in Table 5. [d] ∞ means that all barriers of the ridge are determined. [e] Number of edges in the fully determined energy ridge. [f] Lowest edge barrier of the energy ridge, in kcal/mol relative to the reactant.

only the highest barrier along the best path is requested to be certain ($\Delta E_{sure} = 0$) in conjunction with statistical estimates on E_{uv}^{min} is 106, reducing the number of edge computations by a factor of 6 (from 603). This shows that statistical estimates help to quickly isolate the rate-limiting step of the reaction.

4.2. Performance of Energy Ridge Computations. The highest energy ridge in the TN (here: termed as ridge 2, as it is associated with rearrangements in the switch II) was computed with and without the use of statistical estimates for the barrier bounds to test the performance of the algorithm given in section 3.3.2. The results are shown in Table 6, where the counting is started from scratch for each setting, assuming that no energy barrier has been computed yet. 1092 energy barriers were computed to determine the full energy ridge ($\Delta E_{sure} = \infty$) with safe bounds, which amounts to $\approx$2% of the total number of 47 404 TN edges. By using statistical estimates, this number was reduced to 667. When only the lowest energy barrier of the ridge ($\Delta E_{sure} = 0$) was computed, the computational savings were comparatively slight (805 energy barriers were computed for this), but using both $\Delta E_{sure} = 0$ and statistical estimates reduced the number of computed barriers to 214. The "safe" and the "estimated" ridge 2 agree in their lower-energy edges (up to 5 kcal/mol above the lowest edge). For edges of higher energy, only about 25% of the edges in the "estimated" ridge belong to the safe energy ridge.

4.3. Structural Mechanism of Ras p21. Structural analysis shows that in the best path about half of the switch II helix first unfolds before the rearrangement of the switch I, in which Tyr32 passes underneath the backbone. Subsequently, the rearrangement of switch II completes. This latter step is rate limiting, having the highest potential energy barrier along the best path (E_{peak}=45 kcal/mol relative to the reactant). From the time scale of the Ras p21 conformational switch[67] (on the order of 10^4 s in absence of GAP), it follows that the highest free energy barrier along the path cannot exceed 23 kcal/mol.[13] This indicates that a significant contribution from entropy, due possibly to an increase in backbone flexibility, reduces the high enthalpic barrier found here.

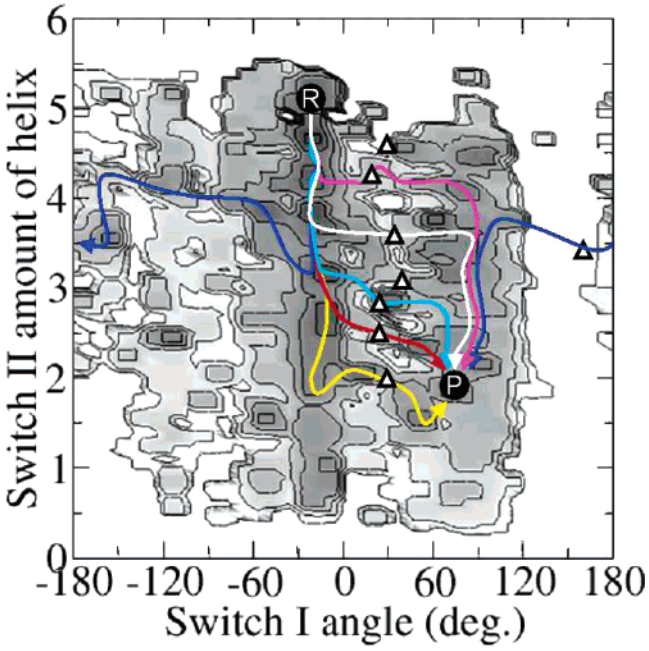


Figure 10. Two-dimensional representation of the potential energy surface of Ras p21. The horizontal axis measures the orientation of Tyr32 on the switch I loop (α = dihedral angle P_β,C_{32},N_{32},OH_{32}, in degrees). The vertical axis measures the helicity of switch II (number of α-helical H-bonds). The contour levels show the energy of the TN vertices (dark gray = 0−10 kcal/mol, light gray > 60 kcal/mol). Reactant and product structures are labeled as 'R' and 'P'. The best transition pathway is shown in white, and the next-best transition pathways (with a rate-limiting step up to 10 kcal/mol higher than the white) in yellow, red, magenta, and cyan. Triangles mark the rate-limiting transition state of the switch I rearrangement (corresponding to the lowest-energy points shown in Figure 11b,c) and belong to ridge 1 (see section 4.3). Ridge 1 can be split into two energy ridges: one along $\alpha \approx 30°$ (where Tyr32 passes underneath the backbone, Figure 11b) and another along $\alpha \approx 150°$ (where Tyr32 passes through the solvent, Figure 11c). The best path with Tyr32 moving through the solvent is shown in dark blue.

The next-best pathways (i.e. pathways having a different rate-limiting transition state, see section 2.1) with rate-limiting barriers within 10 kcal/mol above that of the best path were also computed. There are 12 such pathways in the present TN. The order of events in these pathways is similar to the events described above for the best path, i.e., switch II unfolds (to a varying degree) in the first part of the transition, and the subsequent switch I transition occurs with Tyr32 passing underneath the backbone (see Figure 10). The differences between these pathways are mainly in the precise order of events in the switch II rearrangement.

In the lowest best path, the passage of Tyr32 underneath the backbone is associated to an important barrier of about 25 kcal/mol. This raises the question whether the Tyr32 must necessarily pass underneath the backbone. An obvious alternative would be for Tyr32 to pass the other way (i.e. through the solvent). To better analyze the motion of Tyr32, the energy ridge corresponding to its reorientation (abbreviated as ridge 1, since it is the rate-limiting step of switch I

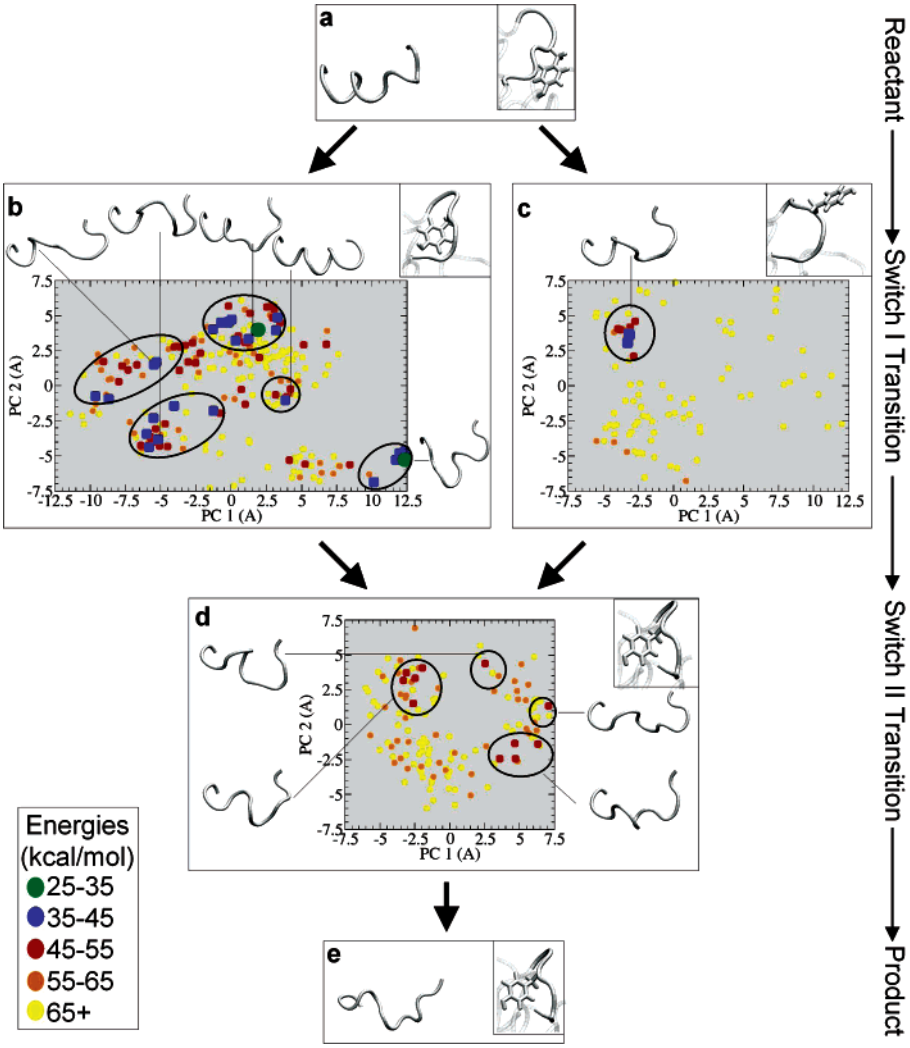


Figure 11. Two-dimensional projection of the energy ridges of the Ras p21 transition. Three major ridges were identified: two for the switch I rearrangement (both belonging to ridge 1) and one for the switch II rearrangement (ridge 2). Transition states from each ridge were projected on their two first principal components (computed from the C_α-coordinates). Each panel (b,c,d) shows one ridge and the corresponding conformation of the switch I loop (box in top right corner of each panel). The projected points cluster (ellipsoids) according to their different switch II conformations (typical backbone conformation shown for each cluster). The energy of each transition state is coded by color. (a) Reactant state: switch I has Tyr32 pointing to the 'right', switch II is a helix. From here, the conformational change proceeds through panels b or c. (b) Energy ridge of the switch I-transition (ridge 1), with Tyr32 passing underneath the backbone. There is a large variety of alternative switch II-conformations at this step of the transition. (c) Ridge 1 with Tyr32 moving through the solvent. (d) Energy ridge of the switch II-transition (ridge 2), which is globally rate-limiting. The transition of switch I is already completed and Tyr32 is pointing to the 'left'. Various isoenergetic ways for the switch II rearrangement coexist. (e) Product state: switch I is pointing to the 'left' and switch II helix has fully unfolded.

rearrangement) was determined. Tyr32 passes from an orientation where its side chain points toward GDP and $-30° < \alpha < -10°$, to an orientation where $60° < \alpha < 110°$. α is an artificial dihedral angle, defined over atoms P_β,C_{32},N_{32},-OH_{32} (P_β is the β-phosphorus of GDP). Ridge 1 was computed with $\Delta E_{sure} = 30$ kcal/mol and using safe values for the barrier bounds (see section 3.3). The resulting energy ridge consists of 92 transition states. In 11 of them, Tyr32 goes through the solvent, and the associated barrier is at least 40 kcal/mol, clearly indicating that passage underneath the backbone is the preferred mechanism. In Figure 10, ridge 1 appears split in two regions.

To visualize the two energy ridges, Figure 11 shows a two-dimensional projection of the transition states contained in ridge 1 and ridge 2. Ridge 1 was split into two sets: one set containing the transition states that involve the passage of Tyr32 underneath the backbone and the other set containing the transition states having Tyr32 passing through the solvent. In the case where passage of Tyr32 is underneath the backbone, there are 7 different transition states in ridge 1 up to 10 kcal/mol above the lowest transition state in ridge 1. These differ considerably in the amount of unfolding of the switch II helix: some still form a perfect helix, while in others the helix is fully unfolded (see Figure 11b). In the unfavorable case that Tyr32 passes through the solvent the conformation of the partially unfolded switch II helix is well defined, as can be seen from its similar structure in all next-higher transition states (Figure 11c).

After the switch I rearrangement has completed, the transition pathways must cross ridge 2, which contains the globally rate-limiting transition states. Ridge 2 contains 14 transition states within 10 kcal/mol above the lowest transition state in ridge 2 (which here is identical to the highest transition state in the lowest best path). These alternative transition states are highly scattered in Figure 11d, showing that the structure of switch II varies considerably. Thus, there are many different ways in which switch II can rearrange toward the product structure, and the coupling between switch I and II is weak enough to allow for many different orders of the conformational events in both switch regions. This means that the Ras p21 conformational switch is highly degenerate, thus confirming a significant entropic contribution to the free energy profile of the conformational switch.[13]

5. Conclusions

We have introduced here an efficient method for mapping the low-energy minima involved in a complex conformational transition in a protein. The method was shown to be effective in identifying minima belonging to very different conformational pathways. Furthermore, the resulting set of minima is dense in the low-energy regions.

A transition network is constructed to connect the available set of low-energy minima. The graph-theoretical methods have allowed to determine global properties of the network while only requiring computation of a small subset of the subtransition barriers in the network. When applied to the conformational switch of Ras p21, the globally best pathway connecting the transition end states and the energy ridge separating them could be determined while computing less than 5% of the total number of subtransitions in the network.

The energetically best pathway and the two main energy ridges of the Ras p21 switch give insight in the mechanism of the transition and provide answers to the three questions asked in the Introduction: (1) The rearrangement of switch I always occurs such that Tyr32 is threaded underneath the protein backbone. (2) This rearrangement of switch I must be finished before the rate-limiting rearrangement of switch II can start. (3) The order of conformational events in either switch I or II and the details of rearrangement in switch II vary substantially. This confirms that complex conformational transitions in proteins such as Ras may occur via multiple pathways.

The methodological advances presented here allow comprehensive analysis of the mechanism of complex transitions in proteins. To allow for comparison with certain experiments, it will be desirable to obtain free energy TN that allow calculation of thermodynamic and kinetic properties. This might be achieved by estimating vibrational free energies for the TN states[42,43] and merging vertices which are separated by low-energy barriers so as to account for intrastate configurational entropy.[32,42,43]

Acknowledgment. We thank Prof. G. Reinelt and M. Oswald for valuable discussions concerning the graph-theoretical aspects of this work. We kindly acknowledge Deutsche Forschungsgemeinschaft (DFG) for financial support.

Appendices: A. Checking Conformers for Steric Collisions

During the sampling procedures described in section 3.1.3, new conformers are validated by checking that they do not produce very high potential energies, mostly due to atom collisions (see Figure 5c). For each pair of atoms i, j, the criterion used is that the combined Lennard-Jones and Coulomb interaction energy should not exceed a tolerance value E_{tol} ($E_{\mathrm{tol}} = 20$ kcal mol^{-1} in this study). This check needs to be repeated so often that it is a computational bottleneck for the sampling method. Therefore, to avoid computing all pairwise interaction energies for each conformer, we first precompute the minimum distance, $d_{i,j}^{\mathrm{min}}$ allowed for each atom pair. This is obtained as the solution of

$$\epsilon_w\left[\left(\frac{\sigma_{ij}}{d_{i,j}^{\mathrm{min}}}\right)^{12}-\left(\frac{\sigma_{ij}}{d_{i,j}^{\mathrm{min}}}\right)^{6}\right]+\frac{q_i q_j}{4\pi\epsilon_r\epsilon_0 d_{i,j}^{\mathrm{min}}}-E_{\mathrm{tol}}=0 \quad (10)$$

where ϵ_w is the van der Waals well depth, σ_{ij} is the effective van der Waals radius for atoms i and j, q_i and q_j are the partial charges of atoms i and j, and $\epsilon_r\epsilon_0$ is the dielectric constant. The above equation is solved for $d_{i,j}^{\mathrm{min}}$ with Newton's root-finding method. For the E_{tol} used in this study, there was always a unique solution for $d_{i,j}^{\mathrm{min}}$. If smaller E_{tol} are used, eq 10 may have two solutions, in which case, the smaller solution must be used so as to ensure that $d_{i,j}^{\mathrm{min}}$ reflects the repulsive interaction. The resulting $d_{i,j}^{\mathrm{min}}$ values are stored. A given conformation is treated as valid if all nonbonded atom distances, $d_{i,j}$ (excluding 1−4 pairs) fulfill the criterion $d_{i,j}^{\mathrm{min}} \leq d_{i,j}$. The number of distance computations is kept small by embedding the protein coordinates in a lattice and computing distances only between atoms which have been changed in a given sampling step and atoms which are in the same or adjacent lattice cells.

B. Efficient Side-Chain Sampling Method

Given a set of backbone conformations that is uniformly distributed in ϕ/ψ-torsional space, a uniformly distributed set of full (backbone and side chain) conformations can be build by repeating following steps: (1) randomly selecting a backbone conformation, (2) building all side chains on this backbone conformation, using random torsion angles, and (3) accepting the conformation if it does not produce collisions. This trivial method is not very efficient in practice, first because some backbone conformers may never allow a given side chain to be built without collisions, and second because for a given backbone conformer it is unlikely that placing all side chains at once produces a conformation without collisions. Here, a more efficient method is used that consists of the following steps: (1) For each backbone conformation c, a weight w_c is computed which is equal to the probability that a set of noncolliding side chains can be built on this backbone, when a uniform distribution of side chain torsion angles is used. (2) A random backbone conformation is selected according to the probability $p_c = w_c/\sum_k w_k$. (3) Onto the selected backbone, each side chain is build by itself in a number of conformations that do not produce collisions with the backbone and the nonsampled

regions of the protein. (4) Side-chain conformations from step 3 are combined randomly to form a fully build protein conformation, which is accepted if is does not have any collisions. Steps 2−4 are repeated until a desired number of conformations have been generated.

The weight w_c is computed as follows: For each backbone conformation c, an acceptance probability $p_{c,i}$ for each side chain i is calculated by generating a large number of random rotamers for that side chain (in the absence of the other side chains of the **S** region) and counting the number of noncolliding rotamers. If any $p_{c,i} = 0$ (i.e. some side chain cannot be placed at all without producing collisions), then backbone conformation c is permanently rejected and $w_c = 0$. Otherwise, the probability q_c to find a noncolliding combination of the individually valid side-chain conformations is computed. This is done by generating a large number N_c of random combinations of valid side-chain rotamers and counting the number n_c of noncolliding combinations, $q_c = n_c/N_c$. The weight w_c is obtained as $w_c = q_c \prod_i p_{c,i}$.

C. Efficient Method for Early Rejection of High-Energy Minima

This describes a method for the early rejection of energy minima with $U(\mathbf{x}) > E_{\text{tol}}$ during the minimization. Early rejection is based on statistics that are collected during a number of preliminary full minimizations, correlating the energy difference between partially minimized and final structures with the gradient of the partially minimized structure.

For Ras p21, 100 samples were retrieved from the sample repository and minimized to a gradient RMS of 10^{-5} kcal mol^{-1} Å^{-1}. Each of these minimization trajectories delivered a series of gradients ($\mathbf{g}_0,\mathbf{g}_1,...,\mathbf{g}_n$) and associated potential energies (U_0, U_1,..., U_n), where the pair ($\mathbf{g}_n$, U_n) corresponds to the fully minimized structure. All pairs ($\mathbf{g}_i$, U_i) from all 100 minimization trajectories were used to derive correlation statistics between $\mathbf{g}_i$ and $\Delta U = U_i - U_n$, i.e., the energy difference from the fully minimized structure. These statistics, shown in Figure 12, were used to obtain for each range of gradient a corresponding value of ΔU that was higher than 90% of the ΔU's in that range. This yields an upper estimate of ΔU, given a certain gradient $\mathbf{g}$. This estimate was used to reject structures during minimizations if their minimum energy, predicted from this upper estimate, considerably exceeded an energy tolerance threshold: $U(\mathbf{x}) - \Delta U > E_0 + E_{\text{low}} + 10$ kcal mol^{-1}. E_0 is the minimized reactant energy, and E_{low} was set to 40 kcal mol^{-1} (see Table 2).

D. Barrier Estimation. A method is given for the statistical estimation of lower and upper bounds for the energy barriers of subtransitions. For this, one correlates available information on the edges $e = (u, v)$, such as distance between its vertices $\delta_{uv} = |\mathbf{x}_u - \mathbf{x}_v|$, with the computed energy barriers $B_{uv} = E^{TS}_{uv} - \max\{E^S_u, E^S_v\}$. Using a certain confidence interval, one obtains upper and lower estimates, $B^{\min}_{uv}(\delta_{uv})$ and $B^{\max}_{uv}(\delta_{uv})$, which are used to replace the strict edge-weight bounds by $\max\{E^S_u, E^S_v\} + B^{\min}_{uv}(\delta_{uv})$ and $\max\{E_u, E_v\} + B^{\max}_{uv}(\delta_{uv})$.

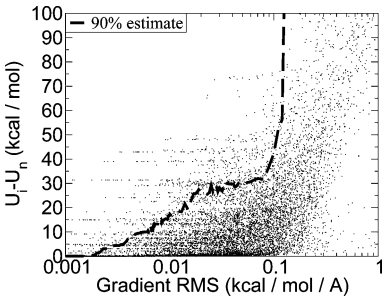


Figure 12. Using the gradient during minimizations to predict the expected energy at the minimum. Based on the minimizations of 100 different conformers, each minimization going through a series of intermediates with gradients ($\mathbf{g}_0,...,\mathbf{g}_n$) and energies ($U_0,..., U_n$), the difference between the energy of an intermediate and the final (minimum) energy, $U_i - U_n$ is plotted against the current gradient $\mathbf{g}_i$. 90% of the points are below the dashed line, which can be used to estimate how much more the energy may decrease during a minimization, based on the current gradient value, thus allowing nonpromising minimizations to be stopped early.

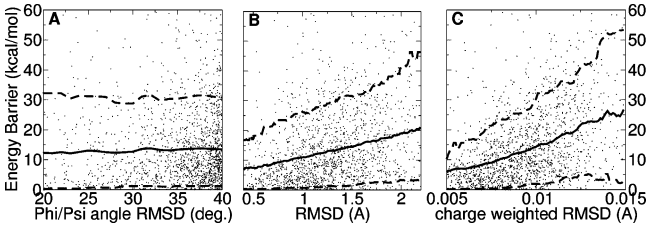


Figure 13. Predicting lower and upper bounds to the energy barrier of subtransitions. The energy barrier is plotted versus the distance between the end states of a given subtransition in Ras p21, using different distance metrics: (A) RMSD in ϕ/ψ-dihedral space of the **S**-regions, (B) all-atom RMSD in Cartesian space, and (C) same, but with each atomic distance weighted by the absolute atomic charge. Solid line: average barrier. 90% of the points lie below the upper dashed line, 10% below the lower dashed line. These were used as lower and upper estimates for the estimation of optimistic and pessimistic best paths (see Figure 9).

For Ras p21, after computing the first ∼2000 energy barriers, these barriers were correlated with the distance between the corresponding minima so as to yield a distance-dependent barrier estimate. Figure 13 shows a plot the first ∼2000 barriers against three different distance measures. The average value and the boundaries of a 90% confidence interval are given. Clearly, the ϕ/ψ-RMSD is not a useful measure here as it is not correlated with the energy barrier. The Cartesian RMSD gives a better correlation, while the charge-weighted RMSD, $d_C(\mathbf{x},\mathbf{y})$, defined as

$$d_C(\mathbf{x},\mathbf{y}) = \sqrt{\frac{\sum_{i=1}^{N}(\mathbf{x}_i - \mathbf{y}_i)^2 q_i^2}{N}}$$

where N is the number of atoms and q_i is the charge on atom i, here gives the best correlation of the three distance measures. The 90% confidence interval was used to derive $B^{\min}_{uv}(\delta_{uv})$ and $B^{\min}_{uv}(\delta_{uv})$.

References

(1) Olsen, K.; Fischer, S.; Karplus, M. A continuous path for the $T \rightarrow R$ allosteric transition of hemoglobin. *Biophys. J.* **2000**, *78*, 394A.

(2) Fischer, S.; Windshuegel, B.; Horak, D.; Holmes, K. C.; Smith, J. C. Structural mechanism of the recovery stroke in the myosin molecular motor. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6873.

(3) Coleman, M. L.; Marshall, C. J.; Olson, M. F. Ras and Rho GTPases in G1-Phase Cell-Cycle Regulation. *Nat. Rev. Mol. Cell Bio.* **1993**, *62*, 851.

(4) Vojtek, A. B.; Der, C. J. Increasing Complexity of the Ras Signaling Pathway. *J. Biol. Chem.* **1998**, *32*, 19925.

(5) Streett, W. B.; Tildesley, D. J.; Saville, G. Multiple time step methods in molecular dynamics. *Mol. Phys.* **1978**, *35*, 639.

(6) Peskin, C. S.; Schlick, T. Molecular dynamics by the backward: Euler's method. *Commun. Pure Appl. Math.* **1989**, *42*, 1001.

(7) Schlick, T.; Barth, E.; Mandziuk, M. Biomolecular dynamics at long time steps: Bridging the time scale gap between simulation and experimentation. *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 181.

(8) Sanz-Navarro, C. F.; Smith, R. Numerical calculations using the hyper-molecular dynamics method. *Comput. Phys. Comm.* **2001**, *137*, 206.

(9) Amadei, A.; Linssen, A. B.; Berendsen, H. J. C. Essential dynamics of proteins. *Proteins* **1993**, *17*, 412.

(10) Krammer, A.; Lu, H.; Isralewitz, B.; Schulten, K.; Vogel, V. Forced unfolding of the fibronectin type III module reveals a tensile molecular recognition switch. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 1351.

(11) Schlitter, J.; Engels, M.; Krüger, P. Targeted molecular dynamics: A new approach for searching pathways of conformational transitions. *J. Mol. Graphics* **1994**, *12*, 84.

(12) Böckmann, R.; Grubmüller, H. Nanoseconds molecular dynamics simulations of primary mechanical energy transfer steps in F_1-ATP synthase. *Nat. Struct. Biol.* **2002**, *9*, 196.

(13) Noé, F.; Ille, F.; Smith, J. C.; Fischer, S. Automated computation of low-energy pathways for complex rearrangements in proteins: Application to the conformational switch of ras p21. *Proteins* **2005**, *59*, 534.

(14) Grubmüller, H. Predicting slow structural transitions in macromolecular systems: conformational flooding. *Phys. Rev. E* **1995**, *52*, 2893.

(15) Huo, S.; Straub, J. E. The maxflux algorithm for calculating variationally optimized reaction paths for conformational transitions in many body systems at finite temperature. *J. Chem. Phys.* **1997**, *107*, 5000.

(16) Huo, S.; Straub, J. E. Direct computation of long time processes in peptides and proteins: Reaction path study of the coil-to-helix transition in polyalanine. *Proteins* **1999**, *36*, 249.

(17) Czerminski, R.; Elber, R. Self-avoiding walk between two fixed points as a tool to calculate reaction paths in large molecular systems. *Int. J. Quantum Chem.* **1990**, *24*, 167.

(18) Fischer, S.; Karplus, M. Conjugate peak refinement: an algorithm for finding reaction paths and accurate transition states in systems with many degrees of freedom. *Chem. Phys. Lett.* **1992**, *194*, 252.

(19) Jónsson, H.; Mills, G.; Jacobsen, K. W. Nudged Elastic Band Method for Finding Minimum Energy Paths of Transitions. In *Classical and Quantum Dynamics in Condensed Phase Simulations*; Berne, B. J., Ciccotti, G., Coker, D. F., Eds.; World Scientific: Singapore, 1998; pp 385−404.

(20) Henkelman, G.; Uberuaga, B. P.; Jonsson, H. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *J. Chem. Phys.* **2000**, *113*, 9901.

(21) Henkelman, G.; Jóhannesson, G.; Jónsson, H. Methods for Finding Saddle Points and Minimum Energy Paths. In *Progress on Theoretical Chemistry and Physics*; Schwartz, S. D., Ed.; Kluwer Academic: New York, 2000; pp 269−300.

(22) Fischer, S.; Michnick, S.; Karplus, M. A mechanism for rotamase catalysis by the fk506 binding protein (fkbp). *Biochemistry* **1993**, *32*, 13830.

(23) Bondar, A. N.; Elstner, M.; Suhai, S.; Smith, J. C.; Fischer, S. Mechanism of primary proton transfer in bacteriorhodopsin. *Structure* **2004**, *12*, 1281.

(24) Gruia, A. D.; Bondar, A. N.; Smith, J. C.; Fischer, S. Mechanism of a molecular valve in the halorhodopsin chloride pump. *Structure* **2005**, *13*, 617.

(25) Stillinger, F. H.; Weber, T. A. Hidden structure in liquids. *Phys. Rev. A* **1982**, *25*, 978.

(26) Stillinger, F. H. A topographic view of supercooled liquids and glass formation. *Science* **1995**, *267*, 1935.

(27) Czerminski, R.; Elber, R. Reaction path study of conformational transitions and helix formation in a tetrapeptide. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 6963.

(28) Czerminski, R.; Elber, R. Reaction path study of conformational transitions in flexible systems: Application to peptides. *J. Chem. Phys.* **1990**, *92*, 5580.

(29) Berry, R. S.; Breitengraser-Kunz, R. Topography and dynamics of multidimensional interatomic potential surfaces. *Phys. Rev. Lett.* **1995**, *74*, 3951.

(30) Wales, D. J. Structure, Dynamics, and Thermodynamics of Clusters: Tales from Topographic Potential Surfaces. *Science* **1996**, *271*, 925.

(31) Ball, K. D.; Berry, R. S.; Kunz, R. E.; Li, F.-Y.; Proykova, A.; Wales, D. J. From topographies to dynamics on multidimensional potential energy surfaces of atomic clusters. *Science* **1996**, *271*, 963.

(32) Becker, O. M.; Karplus, M. The topology of multidimensional potential energy surfaces: Theory and application to peptide structure and kinetics. *J. Chem. Phys.* **1996**, *106*, 1495.

(33) Levy, Y.; Becker, O. M. Effect of conformational constraints on the topography of complex potential energy surfaces. *Phys. Rev. Lett.* **1998**, *81*, 1126.

(34) Miller, M. A.; Wales, D. J. Energy landscape of a model protein. *J. Chem. Phys.* **1999**, *111*, 6610.

(35) Mortenson, P. N.; Wales, D. J. Energy landscapes, global optimization and dynamics of the polyalanine $Ac(ala)_8NHMe$. *J. Chem. Phys.* **2001**, *114*, 6443.

(36) Levy, Y.; Becker, O. M. Energy landscapes of conformationally constrained peptides. *J. Chem. Phys.* **2001**, *114*, 993.

(37) Brooks, C. L., III; Onuchic, J. N.; Wales, D. J. Taking a walk on a landscape. *Science* **2001**, *293*, 612.

(38) Levy, Y.; Jortner, J.; Becker, O. M. Dynamics of hierarchical folding on energy landscapes of hexapeptides. *J. Chem. Phys.* **2001**, *115*, 10533.

(39) Levy, Y.; Jortner, J.; Becker, O. M. Solvent effects on the energy landscapes and folding kinetics of polyalanines. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 2188.

(40) Mortenson, P. N.; Evans, D. A.; Wales, D. J. Energy landscapes of model polyalanines. *J. Chem. Phys.* **2001**, *117*, 1363.

(41) Levy, Y.; Becker, O. M. Conformational polymorphism of wild-type and mutant prion proteins: energy landscape analysis. *Proteins* **2002**, *47*, 458.

(42) Evans, D. A.; Wales, D. J. Free energy landscapes of model peptides and proteins. *J. Chem. Phys.* **2003**, *118*, 3891.

(43) Evans, D. A.; Wales, D. J. The free energy landscape and dynamics of met-enkephalin. *J. Chem. Phys.* **2003**, *119*, 9947.

(44) Wales, D. J.; Doye, J. P. K. Stationary points and dynamics in high-dimensional systems. *J. Chem. Phys.* **2003**, *119*, 12409.

(45) Evans, D. A.; Wales, D. J. Folding of the gb1 hairpin peptide from discrete path sampling. *J. Chem. Phys.* **2004**, *121*, 1080.

(46) Despa, F.; Wales, D. J.; Berry, R. S. Archetypal energy landscapes: Dynamical diagnosis. *J. Chem. Phys.* **2005**, *122*, 024103.

(47) Levy, Y.; Jortner, J.; Becker, O. M. Dynamics of hierarchical folding on energy landscapes of hexapeptides. *J. Chem. Phys.* **2001**, *115*, 10533.

(48) Miller, M. A.; Doye, J. P. K.; Wales, D. J. Energy landscapes of model polyalanines. *J. Chem. Phys.* **2002**, *117*, 1363.

(49) Wales, D. J. Discrete path sampling. *Mol. Phys.* **2002**, *100*, 3285.

(50) Apaydin, M. S.; Brutlag, D. L.; Guestrin, C.; Hsu, D.; Latombe, J.-C.; Varma, C. Stochastic Road map Simulation: An Efficient Representation and Algorithm for Analyzing Molecular Motion. *J. Comput. Bio.* **2003**, *10*, 257.

(51) Dijkstra, E. A note on two problems in connection with graphs. *Num. Math.* **1959**, *1*, 269.

(52) Stillinger, F. H.; Weber, T. A. Packing structures and transitions in liquids and solids. *Science* **1984**, *228*, 983.

(53) McQuarrie, D. A.; Simon, J. D. *Molecular Thermodynamics*; University Science Books: Sausalito, CA, 1999.

(54) Hänggi, P.; Talkner, P.; Borkovec, M. Reaction rate theory: Fifty years after kramers. *Rev. Mod. Phys.* **1990**, *62*, 251.

(55) Berkowitz, M.; Morgan, J. D.; McCammon, J. A.; Northrup, S. H. Diffusion-controlled reactions: A variational formula for the optimum reaction coordinate. *J. Chem. Phys.* **1983**, *79*, 5563.

(56) Elber, R. Reaction Path Studies of Biological Molecules. In *Recent Developments in Theoretical Studies of Proteins (Advanced Series in Physical Chemistry, Vol. 7)*; World Scientific: Singapore, 1996.

(57) Eppstein, D. Finding the *k* shortest paths. *SIAM J. Comp.* **1998**, *28*, 652.

(58) Nagamochi, H.; Ibaraki, T. Computing edge connectivity in multigraphs and capacitated graphs. *SIAM J. Discr. Math.* **1992**, *5*, 54.

(59) Hoffmann, D.; Knapp, E.-W. Polypeptide folding with off-lattice Monte Carlo dynamics: the method. *Eur. Biophysics J.* **1996**, *24*, 387.

(60) Mezei, M. Efficient Monte Carlo sampling for long molecular chains using local moves, tested on a solvated lipid bilayer. *J. Chem. Phys.* **2003**, *118*, 3874.

(61) Lowy, D. R.; Willumsen, B. M.; Function and regulation of ras. *Annu. Rev. Biochem.* **1993**, *62*, 851.

(62) Pai, E. F.; Krengel, U.; Petsko, G. A.; Goody, R. S.; Kabsch, W. and Wittinghofer, A. Refined crystal structure of the triphosphate conformation of H-ras p21 at 1.35 Å resolution: implications for the mechanism of GTP hydrolysis. *EMBO J.* **1990**, *9*, 2351.

(63) Tong, L. A.; de Vos, A.; Milburn, M. V.; Kim, S. H. Crystal structures at 2.2 Å resolution of the catalytic domains of normal ras protein and an oncogenic mutant complexed with GDP. *J. Mol. Biol.* **1991**, *217*, 503.

(64) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. Charmm: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **1983**, *4*, 187.

(65) Neria, E.; Fischer, S.; Karplus, M. Simulation of activation free energies in molecular systems. *J. Chem. Phys.* **1996**, *105*, 1902.

(66) Schaefer, M.; Karplus, M. A Comprehensive Analytical Treatment of Continuum Electrostatics. *J. Chem. Phys.* **1996**, *100*, 1578.

(67) Neal, S. E.; Eccleston, J. F.; Webb, M. R. Hydrolysis of GTP by $p21^{NRAS}$, the *NRAS* protooncoge ne product, is accompanied by a conformational change in the wild-type protein. Use of a single fluorescent probe at the catalytic site. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 3562.

CT050162R

JCTC Journal of Chemical Theory and Computation

Role of Solvent Dynamics in Stabilizing the Transition State of RNA Hydrolysis by Hairpin Ribozyme

Hwangseo Park*,† and Sangyoub Lee*,‡

Department of Bioscience and Biotechnology, Sejong University, 98 Kunja-Dong, Kwangjin-Ku, Seoul 143-747, Korea, and Department of Chemistry, Seoul National University, Seoul 151-747, Korea

Received December 1, 2005

Abstract: Structural and mechanistic studies of the hairpin ribozyme have been actively pursued over the last two decades to understand its catalytic strategy for RNA hydrolysis. Based on molecular dynamics simulations with the newly developed force field parameters for a vanadium−oxygen complex, we investigate the dynamic properties of the hairpin ribozyme in complex with a transition state analogue for the phosphodiester cleavage. The results indicate that the three nucleobases of the hairpin ribozyme (G8, A9, and A38) stabilize the negatively charged oxygen atoms in the transition state through the formation of five hydrogen bonds, which is consistent with the X-ray crystallographic data. In addition to the three catalytic nucleobases, several solvent molecules are also found to contribute to the catalytic action of the hairpin ribozyme by hydrogen bond stabilization of the negatively charged oxygens as well as by optimally positioning the catalytic nucleobases in the active site.

Introduction

The hairpin ribozyme catalyzes the reversible and sequence-specific cleavage of phosphodiester backbone of an RNA substrate.[1] This catalytic transesterification reaction is facilitated by the intramolecular attack of the neighboring 2′-hydroxy moiety on the phosphorus atom, leading to the formation of a cyclic phosphorane intermediate or transition state that precedes 2′,3′-cyclic phosphate and 5′-hydroxyl termini (Figure 1). In the past two decades, numerous studies with a variety of model systems have been carried out to clarify the underlying chemical principles of the catalytic cleavage of RNA.[2] It was suggested that although divalent metal ions, Mg^{2+} in particular, stabilize the structure of the hairpin ribozyme in a functionally active form, they are known to be nonessential in the catalytic mechanism, contributing at most a 10−20-fold to rate enhancement.[3−5] This indicates that the major role in the catalytic mechanism of the hairpin ribozyme should be played by the nucleobases themselves, possibly involving the general acid−base mechanism[6] to stabilize the unfavorable charges developing in the transition state.

A few years ago, high-resolution X-ray crystal structures of the hairpin ribozyme were reported in complex with a noncleavable substrate analogue or a transition state mimic including a pentavalent vanadium ion.[7] These structures suggested a model for ribozyme catalysis in which the nucleobases of G8, A9, and A38 serve as the electrophilic catalysts that stabilize the transition state by donating multiple hydrogen bonds. The crystal structures also revealed that A38 acts as a hydrogen bond donor to the leaving O5′ atom in its protonated form at N1 position. However, recent kinetic studies disproved the critical roles of the nucleobases, especially with regards to G8 in general acid-general base catalysis of the hairpin ribozyme.[8,9] Although the involvement of water molecule(s) in the catalytic mechanism of ribozymes has been proposed,[10,11] no ordered water molecule was detected near the active site of hairpin ribozyme in the X-ray crystal structures. Therefore, the precise role of water molecule(s) in the catalytic mechanism of the hairpin ribozyme still remains unclear.

* Correspondence may be addressed to either author. Phone: +82-2-3408-3766. Fax: +82-2-3408-3334. E-mail: hspark@sejong.ac.kr (H.P.); sangyoub@snu.ac.kr (S.L.).

† Sejong University.

‡ Seoul National University.

10.1021/ct0503015 CCC: $33.50
Published on Web 03/14/2006

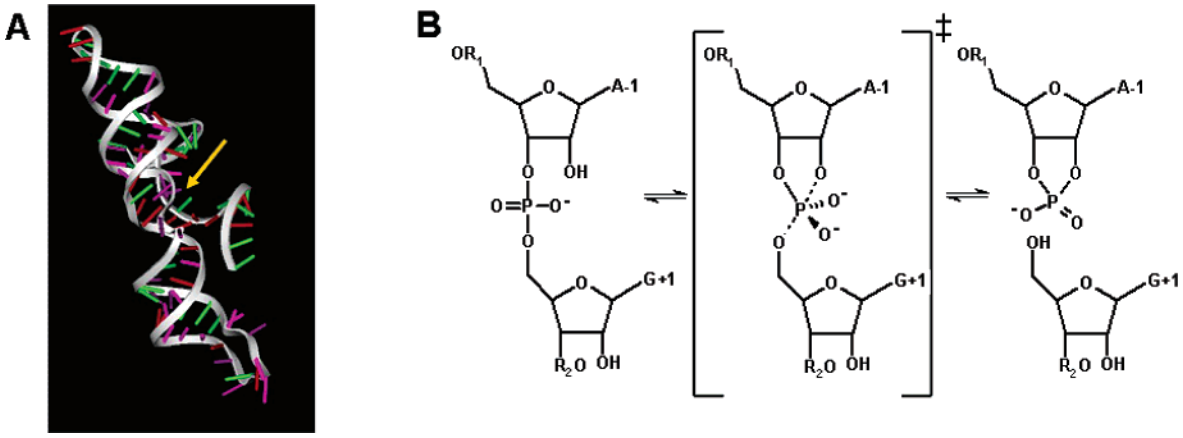


Figure 1. (A) Crystal structure of the hairpin ribozyme. Indicated by a yellow arrow is the catalytic site in which a substrate RNA is hydrolyzed. (B) Schematic of the reversible transesterification reaction of a substrate RNA catalyzed by the hairpin ribozyme.

In this paper, we report the dynamic properties of the hairpin ribozyme in complex with the vanadate transition state analogue based on molecular dynamics (MD) simulation in aqueous solution. We focus our interest on clarifying the roles of solvent molecules in maintaining the structure of the hairpin ribozyme and in stabilizing the transition state. Also presented are the potential parameters for the vanadium−oxygen complex mimicking the transition state of the phosphodiester cleavage, which are unavailable in the current force field database. The detailed structural and dynamic features found in this MD study are expected to provide new insight into the catalytic mechanism of the hairpin ribozyme.

Computational Methods

A. Force Field Design for the Vanadium−Oxygen Complex. To obtain the missing force field parameters, we used a simplified structural model for the hairpin ribozyme in complex with the transition state mimic. This model structure includes the two riboses of (A − 1) and (G + 1) residues and the bridging vanadium ion whose coordinates were extracted from the original X-ray structure. The two nucleobases were omitted for simplicity. We adopted the bonded approach proposed by Hoops et al.[12] to introduce explicit bonds between the central pentavalent vanadium ion and its oxygen ligands. In the derivation of the missing force field parameters, we followed the procedure suggested by Fox and Kollman[13] to be consistent with the standard AMBER force field.[14] The equilibrium bond lengths and bond angles involving the pentavalent vanadium ion were computed using the optimized structure of the vanadium−oxygen complex mimicking the transition state of the phosphodiester cleavage. For the force constant parameters associated with the vanadium ion, we used those for the sugar−phosphate backbone of RNA or DNA that are available in the force field database. The Lennard-Jones parameters of the backbone phosphorus atom were used for the vanadium ion. Geometry optimization of the vanadium−oxygen complex was performed at the B3LYP/6-31G* level of theory with the Gaussian98 program. Using the energy-minimized structure, atomic partial charges for the transition state mimic were also calculated at the RHF/6-31G* level of theory through the RESP method[15] to be consistent with the standard AMBER force field. We computed the potential parameters for the protonated adenine at N1 position by following the same procedure as for the vanadium−oxygen complex, which involves geometry optimization and charge fitting with the RESP method.

B. MD Simulations. As a starting structure of MD simulation with the AMBER7 program,[16] we used the X-ray crystal structure of the hairpin ribozyme in complex with a vanadate transition state analogue (PDB entry: 1M5O). In addition to the 37 Ca^{2+} ions in the original X-ray structure, 29 Na^{+} ions were added to neutralize the negative charges of the system. The all-atom model for the complex was then immersed in a rectangular box containing 11 891 TIP3P[17] water molecules. After 1000 cycles of energy minimization (500 steps for solvent molecules only followed by 500 steps for the entire system with no restraints) to remove bad steric contacts, we equilibrated the ribozyme-transition state mimic complex beginning with 20 ps equilibration dynamics of the solvent molecules at 300 K. The next step involved equilibration of the solute with a fixed configuration of the solvent molecules for 10 ps at 10, 50, 100, 150, 200, 250, and 300 K. These partial equilibrations were used to ensure the stability of the entire system although the full equilibration had been more conventional.[18] Then, the equilibration dynamics of the entire system was performed at 300 K for 100 ps. Following the equilibration procedure, 1.8 ns MD simulations were carried out with a periodic boundary condition in the NPT ensemble at 300 K using Berendsen temperature coupling[19] and constant pressure (1 atm) with isotropic molecule-based scaling. The SHAKE algorithm,[20] with a tolerance of 10^{-6}, was applied to fix all bond lengths involving hydrogen atom. We used a time step of 1.0 fs and a nonbond-interaction cutoff radius of 12 Å; the trajectory was sampled every 0.2 ps (200 step intervals) for analysis.

Results and Discussion

To extend the AMBER force field for modeling the vanadium−oxygen complex that mimics the transition state of the phosphodiester cleavage, we used the standard procedure that starts with the geometry optimization of the complex at the B3LYP/6-31G* level of theory. Figure 2 displays the structure of a local energy minimum whose input structure was extracted from the X-ray crystal structure of the hairpin ribozyme in complex with the vanadate transition

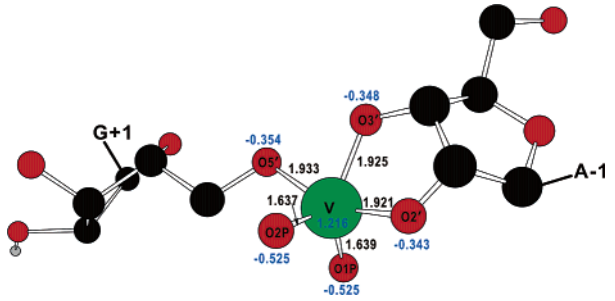


Figure 2. Optimized structure of the vanadium−oxygen complex mimicking the transition state of the phosphodiester cleavage. The calculated RESP atomic charges and the V−O coordination distances are given in e and Å, respectively. All hydrogen atoms bonded to carbon atoms are omitted for simplicity.

state mimic.[8] Also presented are the calculated RESP atomic charges and the coordination distances between the vanadium ion and its oxygen ligands. Consistent with the crystal structure, the five oxygen ligands are coordinated to the central vanadium ion to form a distorted trigonal bipyramidal geometry. The interatomic distances associated with the V−O coordination compare reasonably well with those in the crystal structure with a difference of 0.028 Å on average.

In Figure 2, it is noted that the RESP atomic charge of the pentavalent vanadium ion decreases from +5.000 e in the free state to +1.216 e in the coordination complex. On the other hand, the atomic charges of oxygen ligands become less negative by 0.145−0.271 e when compared to those in the absence of the vanadium ion. These changes reflect the redistribution of charges between the vanadium ion and its oxygen ligands upon the formation of the complex. We used these newly obtained atomic charges in the subsequent MD simulation of the hairpin ribozyme in complex with the transition state mimic because it is well-attested that the M^{n+} model for a metal ion is inadequate for maintaining the actual coordination geometry of a transition metal complex contained in biomolecules.[21]

We checked the reliability of current MD simulation by examining whether the structure of the hairpin ribozyme was maintained stable under the simulation condition described in the previous section. For this purpose, we calculated the root-mean-square-deviations from starting structure ($RMSD_{init}$) for all backbone heavy atoms of the hairpin ribozyme and for all heavy atoms of the transition state mimic as a function of simulation time, which are compared in Figure 3. The computed $RMSD_{init}$ values for the hairpin ribozyme remain within 2.5 Å, indicating that the conformation of hairpin ribozyme is maintained stable during the entire course of simulation. The $RMSD_{init}$ values of the transition state mimic fall into 1.5 Å and remain lower than those of the backbone atoms of the hairpin ribozyme. Judging from this difference in dynamic behavior, the movement of the transition state mimic around the active site seems to be insignificant when compared to the conformational change of the hairpin ribozyme.

Shown in Figure 4 is the representative MD trajectory snapshot of the hairpin ribozyme in complex with the transition state mimic. As in the crystal structure reported by Rupert et al.,[8] the nucleophilic 2′OH at the cleavage site

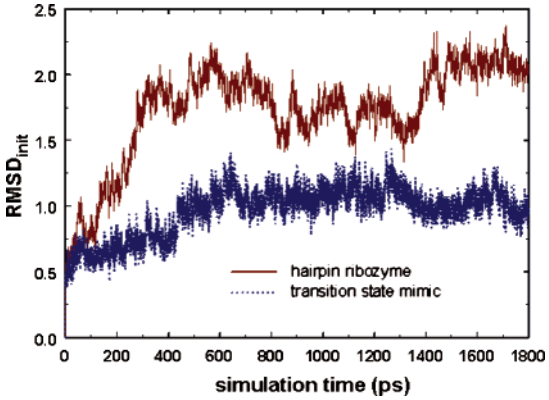


Figure 3. Comparative view of the time evolutions of the root-mean-square deviations for backbone heavy atoms of the hairpin ribozyme (red) and all heavy atoms of the transition state mimic (blue).

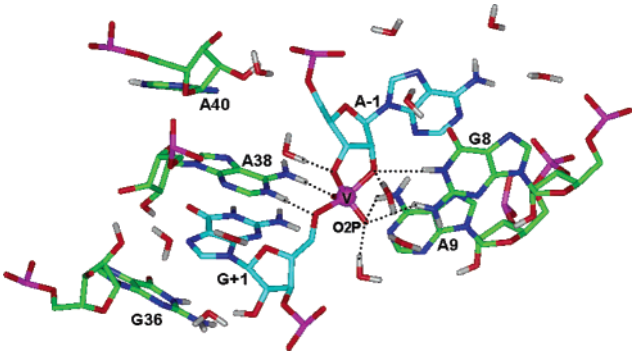


Figure 4. Representative MD trajectory snapshot of the hairpin ribozyme in complex with the vanadate transition state mimic including the water molecules found near the active site. Carbon atoms of the hairpin ribozyme and the transition state mimic are indicated in green and cyan, respectively. Each dotted line indicates a hydrogen bond.

(A − 1) of the transition state mimic, vanadium ion, and the 5′ oxygen of the leaving group (G + 1) reside in nearly optimal positions for an in-line S_N2-type reaction mechanism. It is also consistent with the X-ray structure that the three nucleobases of the hairpin ribozyme (G8, A9, and A38) donate five hydrogen bonds to stabilize the oxygen atoms coordinated to the vanadium ion. In this solution-phase structure, O2′ and O3′ atoms of the (A − 1) residue and O2P atoms of the (G + 1) residue are further stabilized by forming three hydrogen bonds with the water molecules that diffuse into the active site of the hairpin ribozyme from bulk solvent. This is a structural feature inconsistent with the X-ray crystal structure in which no ordered water molecule was found within the distance of 10 Å from the vanadium ion although the transition state mimic occupies a small portion of the active site volume and leaves sufficient space for additional molecules to get into the active site. Related with such a discrepancy, it can be argued that the access of water molecules to the active site may depend on the crystallization procedure. The presence of a hydrogen bond between the deprotonated O2′ atom of the (A − 1) residue and a water molecule supports the possibility that the role of proton acceptor for the nucleophilic oxygen of an RNA

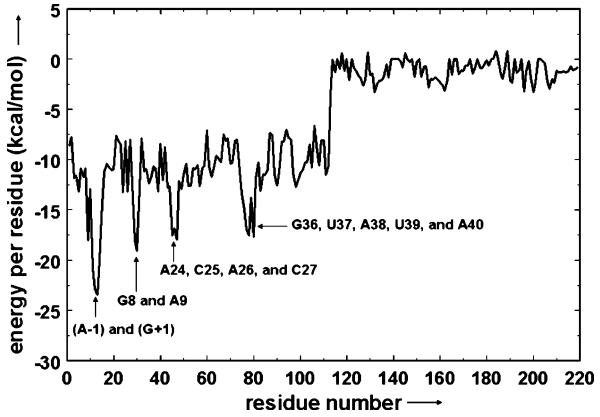


Figure 5. Interaction energy per residue for the average structure of the hairpin ribozyme in complex with the transition state mimic. Residues 1−21, 22−113, and 114−220 correspond to the transition state mimic, the hairpin ribozyme, and the solvent molecules found within 5 Å around the transition state mimic, respectively.

substrate should be played by water in the catalytic mechanism of the hairpin ribozyme.

To provide further evidence for strong interactions between the transition state mimic and the catalytic nucleobases and solvent molecules, we calculate the interaction energy per residue for the average structure of the complex. As indicated in Figure 5, the most significant contributions to the stabilization of the complex seem to come from the interaction between (A − 1) and (G + 1) residues of the transition state mimic and the catalytic nucleobases including G8, A9, C25, A24, C25, A26, C27, G36, U37, A38, U39, and A40. However, no solvent molecule is found that can stabilize the transition state mimic as comparable to the catalytic nucleobases. Possible reasons for this may include the small size of water molecule and the exchange of the water molecules donating hydrogen bonds to the transition state mimic.

To estimate dynamic stabilities of the hydrogen bonds relevant to the stabilization of the transition state mimic, we have calculated time evolutions of the associated interatomic distances, which are plotted in Figure 6. We note that the hydrogen bondings of the catalytic nucleobases to the transition state mimic are maintained stable during the entire course of simulation with 95% of residence time on average. Similarly, all three hydrogen bonds established between water molecules and the transition state mimic are maintained for more than 80% of simulation time if the distance limit for the O···H hydrogen bond is assumed to be 2.2 Å as suggested by Jeffrey.[22] In contrast to the obvious dynamic stabilities of the hydrogen bondings to the transition state mimic by the three catalytic nucleobases, however, we find that there are exchanges of the water molecules that act as a hydrogen bond donor with respect to the transition state mimic. This indicates the involvement of complex solvent dynamics in an optimal stabilization of the transition state by the hairpin ribozyme. It is noteworthy that as shown in Figure 7, the three O···H−O hydrogen bond angles between water molecules and the transition state mimic are also maintained stable during the entire course of simulation. This implies that a complex solvent dynamics would have an insignificant effect on the strengths of the three hydrogen bonds. Thus, the structural and dynamical features found in this MD study support the hypothesis that the hairpin ribozyme has evolved to maximize its hydrogen bonding interactions with the trigonal bipyramidal transition state.

Conclusions

We have developed the force field parameters appropriate for modeling the vanadate transition state mimic for the phosphodiester cleavage by the hairpin ribozyme. The hairpin ribozyme in complex with the transition state mimic shows a dynamic stability in aqueous solution, maintaining a square bipyramidal coordination geometry at the reaction center including the nucleophilic 2′OH at the cleavage site (A − 1), vanadium ion, and the 5′ oxygen of the leaving group (G + 1). This supports the S_N2-type reaction mechanism of the hairpin ribozyme. As in the crystal structure, the three nucleobases of the hairpin ribozyme (G8, A9, and A38) stabilize the negatively charged oxygen atoms in the transition state mimic through the formation of five hydrogen

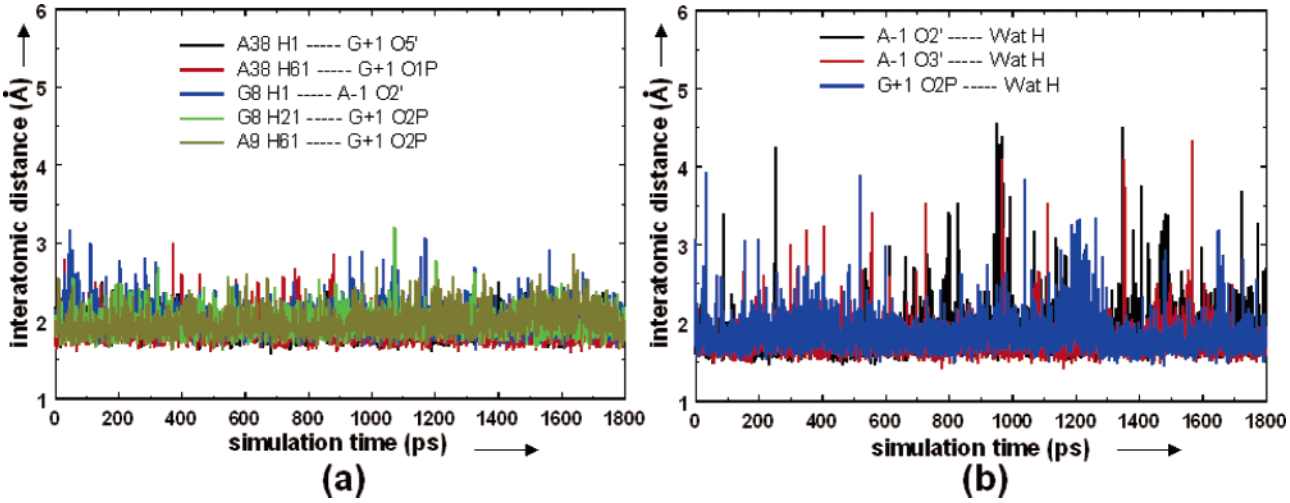


Figure 6. Time evolutions of the interatomic distances associated with hydrogen-bond interactions (a) between catalytic nucleobases of hairpin ribozyme and vanadate transition state mimic and (b) between solvent molecules and the transition state mimic.

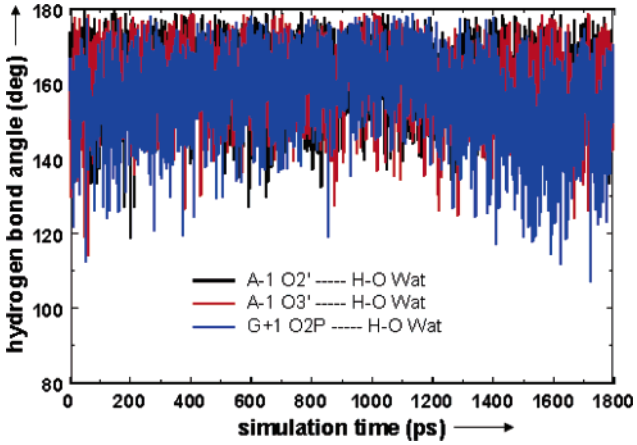


Figure 7. Time evolutions of the O···H−O angles associated with the hydrogen bonds between the transition state mimic and solvent molecules.

bonds. Several solvent molecules are also found to contribute to the catalytic action of the hairpin ribozyme by hydrogen bond stabilization of the negatively charged oxygens and by optimally positioning the catalytic nucleobases in the active site. The presence of a hydrogen bond between the deprotonated O2′ atom at the (A − 1) site and a solvent molecule supports the possibility that the nucleophilic oxygen of an RNA substrate would be deprotonated by water in bulk solvent during the catalytic action of the hairpin ribozyme.

Acknowledgment. This work was supported by Grant No. R01-2004-000-10354-0 from the Basic Research Program of the Korea Science & Engineering Foundation. The authors would also like to acknowledge the support from KISTI (Korea Institute of Science and Technology Information) under "The Seventh Strategic Supercomputing Support Program" with Dr. Sang Min Lee as the technical supporter. The use of the computing system of the Supercomputing Center is greatly appreciated.

Note Added after ASAP Publication. This article was released ASAP on March 14, 2006, with the incorrect Received Date. The correct version was posted on April 13, 2006.

Supporting Information Available: Force field parameters of the vanadium−oxygen complex used to represent the transition state mimic. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Fedor, M. J. Structure and function of the hairpin ribozyme. *J. Mol. Biol.* **2000**, *297*, 269−291.

(2) DeRose, V. J. Two Decades of RNA Catalysis. *Chem. Biol.* **2002**, *9*, 961−969.

(3) Curtis, E. A.; Bartel, D. P. The hammerhead cleavage reaction in monovalent cations. *RNA* **2001**, *7*, 546−552.

(4) O'Rear, J. L.; Wang, S.; Feig, A. L.; Beigelman, L.; Uhlenbeck, O. C.; Herschlag, D. Comparison of the hammerhead cleavage reactions stimulated by monovalent and divalent cations. *RNA* **2001**, *7*, 537−545.

(5) Nakano, S.; Proctor, D. J.; Bevilacqua, P. C. Mechanistic characterization of the HDV genomic ribozyme: assessing the catalytic and structural contributions of divalent metal ions within a multichannel reaction mechanism. *Biochemistry* **2001**, *40*, 12022−12038.

(6) Bevilacqua, P. C. Mechanistic considerations for general acid−base catalysis by RNA: revisiting the mechanism of the hairpin ribozyme. *Biochemistry* **2003**, *42*, 2259−2265.

(7) Rupert, P. B.; Massey, A. P.; Sigurdsson, S. Th.; Ferre-D'Amare, A. R. Transition state stabilization by a catalytic RNA. *Science* **2002**, *298*, 1421−1424.

(8) Kuzmin, Y. I.; Da Costa, C. P.; Fedor, M. J. Role of an active site guanine in hairpin ribozyme catalysis probed by exogenous nucleobase rescue. *J. Mol. Biol.* **2004**, *340*, 233−251.

(9) Kuzmin, Y. I.; Da Costa, C. P.; Cottrell, J. W.; Fedor, M. J. Role of an active site adenine in hairpin ribozyme catalysis. *J. Mol. Biol.* **2005**, *349*, 989−1010.

(10) Nakano, S.; Chadalavada, D. M.; Bevilacqua, P. C. General acid−base catalysis in the mechanism of a hepatitis delta virus ribozyme. *Science* **2000**, *287*, 1493−1497.

(11) Shih, I.; Been, M. D. Catalytic strategies of the hepatitis delta virus ribozymes. *Annu. Rev. Biochem.* **2002**, *71*, 887−917.

(12) Hoops, S. C.; Anderson, K. W.; Merz, K. M., Jr. Force field design for metalloproteins. *J. Am. Chem. Soc.* **1991**, *113*, 8262−8270.

(13) Fox, T.; Kollman, P. A. Application of the RESP methodology in the parameterization of organic solvents. *J. Phys. Chem. B* **1998**, *102*, 8070−8079.

(14) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.

(15) Bayly, C. A.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. A well behaved elestrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.* **1993**, *97*, 10269−10280.

(16) Case, D. A.; Cheatham, T. E., III; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005**, *26*, 1668−1688.

(17) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.

(18) Beveridge, D. L.; Dixit, S. B.; Barreiro, G.; Thayer, K. M. Molecular dynamics simulations of DNA curvature and flexibility. *Biopolymers* 2004, *73*, 380−403.

(19) Berendsen, H. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.

(20) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.

(21) Stote, R. H.; Karplus, M. Zinc binding in proteins and solution: a simple but accurate nonbonded representation. *Proteins* **1995**, *23*, 12−31.

(22) Jeffrey, G. A. *An Introduction to Hydrogen Bonding*; Oxford University Press: Oxford, 1997.

CT0503015

JCTC Journal of Chemical Theory and Computation

Theoretical Calculation of Hydrogen-Bonding Strength for Drug Molecules

Ming-Hong Hao*

Department of Medicinal Chemistry, Boehringer Ingelheim Pharmaceuticals, Inc., Ridgefield, Connecticut 06877

Received January 18, 2006

Abstract: Hydrogen bond is an important type of interaction between drug molecules and their receptors. We present here a computational method for accurately predicting the hydrogen-bonding strength for different acceptors with respect to a given donor or vice versa. The method is based on quantum chemistry DFT calculation of the interaction energy between hydrogen bond donors and acceptors. An excellent linear correlation is observed between the calculated hydrogen-bonding energies and the experimentally measured hydrogen-bonding constants log K_β on a variety of known hydrogen bond acceptors and donors. These results not only indicate the predictive power of this method but also shed light on factors that determine the magnitude of experimentally measured hydrogen-bonding constants for different acceptors with respect to a given donor, suggesting a primarily enthalpic contribution from hydrogen-bonding energy. The method can be used for evaluating the effects of steric interference and inhibitor binding geometry on hydrogen-bonding strength in drug design.

Introduction

The importance of the hydrogen bond to drug design is well recognized. Hydrogen bonds are not only crucial in dictating the orientation of an inhibitor binding in the receptor but also contribute importantly to binding affinity. Hydrogen bond capacity is an essential factor in the strategy of bioisosterism for drug design and optimization. When a bioisostere is used to replace an existing moiety of a compound, the replacement has to match the hydrogen-bonding characteristics of the parent and would preferably further improve upon compound properties including binding potency. There is strong evidence indicating that the strength of different hydrogen bond donors and/or acceptors varies significantly. The hydrogen-bonding constants of commonly encountered hydrogen-bonding groups have been measured to vary over more than 3 orders of magnitudes.[1] Furthermore, in the drug optimization process, it is often observed that electron-withdrawing or donating substituents have opposite effects on the activity of the compounds. Some of the underlying causes of such structure−activity relationships (SAR) could be traced to a modulation of the hydrogen-bonding interaction of the inhibitor with the receptor.[2] Clearly, the variations in hydrogen bond strength could be utilized in drug design. A fundamental understanding of hydrogen bond interactions will greatly facilitate such efforts.

In a comprehensive monograph on the subject,[3] Jeffrey and Saenger summarized the known characteristics of hydrogen bonds, some of which are particularly instructive to a theoretical investigation: (1) hydrogen bonds are not properties of atom pairs but are dependent on the pair of atom groups that forms the extensive donor and acceptor subunits;[3,1] (2) the major component of hydrogen-bonding interaction is electrostatic;[3−5] and (3) hydrogen bonds are soft interactions, and hydrogen bond lengths and angles fluctuate according to local environments.[3,6,7] Because of the importance of hydrogen bonds to drug design, much work has been done in the past on the theoretical modeling of hydrogen bonds for QSAR studies. The approaches span from simple indicator methods[8,9] to parametrization approaches using theoretically calculated properties, such as atomic charges,[10] molecular electrostatic potential,[11] LUMO

* Corresponding author phone: (203)791-6033; e-mail: mhao@rdg.boehringer-ingelheim.com. Corresponding author address: Boehringer Ingelheim Pharmaceuticals, Inc., 900 Ridgebury Rd., P.O. Box 368, Ridgefield, CT 06877.

10.1021/ct0600262 CCC: $33.50
Published on Web 04/05/2006

and HOMO properties,[10,12] atom polarizability,[13] and superdelocalizability,[14,15] to model hydrogen bond strength. However, due to the character of being the property of a group of atoms and the susceptibility to local environments, hydrogen bonds could not be modeled accurately by a general semiempirical or rule-based method because there are many exceptions, such as steric factors, to be accommodated by a finite set of rules. To treat complicated systems such as hydrogen bonding, the ab initio quantum chemistry method is an ideal approach since all electronic and steric effects are fully taken into account in such a treatment. Over the past decade, much work has been published on the ab initio study of hydrogen bonds.[16−30] However, few ab initio calculations have been carried out on hydrogen bond systems that are directly related to medicinal chemistry in drug design. Hydrogen bond systems encountered in drug design involve relatively large organic molecules that interact with protein targets. We explore here the application of ab initio calculations to problems of hydrogen bonding relevant to drug design.

In this work, an ab initio procedure for calculating the hydrogen bond strengths for molecules of interest to medicinal chemistry was investigated. The fundamental quantum chemistry theory for calculation of the interaction energy between two molecular species has been well established.[20,31] However, there are practical issues as to the accuracy of the calculations on specific systems and the relevance of the calculated properties to a practical process. The ab initio calculation directly yields the binding energy for the hydrogen-bonding interaction, but the strength of the hydrogen bond is determined by its free energy change. Furthermore, hydrogen bond strength, as a function of free energy, will also be influenced by the interaction of the inhibitor with water solvent. These are clearly complicated systems. However, experimental observations help to simplify the perspective considerably. There had been historically the postulation of a linear relationship between hydrogen bond free energy and enthalpy.[32,33] Abraham el al. provided clear experimental evidence supporting such a relationship.[1] They observed a linear correlation between hydrogen-bonding energy, measured by spectroscopic wavelength shift, and hydrogen bond free energy measured by the equilibrium constant of hydrogen bond formation for many hydrogen bond donors and acceptors. On the basis of these observations and being concerned with only the relative hydrogen bond strength among different acceptors with respect to a given donor, it is possible that the ab initio calculated bonding energy can yield information, even though indirectly, for hydrogen-bonding free energy. In this paper, the term hydrogen-bonding strength will be used to refer to either the thermodynamic stability or the interaction energy of a hydrogen bond in a different context, even though the term is strictly speaking only a function of the free energy.

The paper is organized as follows: the basic computational method is presented first, followed by an investigation into the sensitivity of the calculated results on hydrogen bond geometry such as separation and orientation. The main results of this work are the calculations of the energy for a variety of hydrogen bond acceptors with respect to a given donor and the subsequent comparison of these calculated results with experimental determined hydrogen-bonding constants that provide a validation of the present approach. Additional issues such as secondary interactions and effects of different donors are further investigated. In the Discussion section, comparison of the accuracy of the present method with earlier approaches will be made. Insights gained from this study will also be elaborated on.

Theory and Method

All quantum chemical calculations in this work were carried out with the program package Jaguar from Schrodinger, Inc.[34] The standard energy difference method[20,31] was used to calculate the energy of the hydrogen-bonding interaction between a donor and acceptor

$$\Delta E_{HB}(R) = E_{AD}(R) - E_A - E_D \quad (1)$$

where ΔE_{HB} is the energy of bonding interaction, R is the set coordinates that define the structure of the hydrogen-bonding complex, $E_{AD}(R)$ is the total energy of the complex, and E_A and E_D are the individual energies of the donor and the acceptor, respectively. Because the calculations involve a complex system on one hand and two subunits on the other hand, the standard molecular orbital method will use different basis functions for the complex and for the subunits, respectively, thus incurring a basis set superimposition error (BSSE)[20,31] in calculating the energy difference. To correct for the BSSE, the counterpoise procedure[31] was applied in the calculation of E_A and E_D in which the virtual orbitals of the other subunit were included in the basis functions for the subunit that is to be calculated. Two quantum chemistry methods were used in this work. The majority of the calculations were carried out using density function theory (DFT) with the B3LYP procedure and a large 6-31++G** basis set. A subset of the calculations was carried out with both DFT and the MP2 procedures to evaluate the effects of the computational methods on the calculated hydrogen bond energy. In the MP2 calculations, the 6-31G** basis set was used.

One of the major technical issues encountered in this work was to find a proper procedure for geometry optimization. For simpler hydrogen bond systems, such as when one subunit is a water molecule, a single constraint on the hydrogen bond angle would be sufficient to keep the donor and the acceptor in a proper geometry through the geometry optimization process.[35] However, for calculations involving relatively large donors and acceptors in which only one or two internal variable constraints were applied, geometry optimization by energy minimization often results in structures that are substantially different from the desired hydrogen bond conformation. The reason for such an outcome is that, when two relatively large molecules are free to move, energy minimization always brings secondary interactions, in addition to hydrogen bonding, into the total energy. The secondary interactions not only perturb the hydrogen bond geometry but also cause difficulty in separating hydrogen-bonding energy from other nonbonded contributions. In a realistic and complete inhibitor−receptor complex, the hydrogen bond interactions between the donor

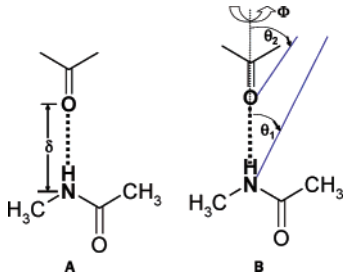


Figure 1. Variables used to define the hydrogen bond geometry between a donor and an acceptor. A: The separation between the donor and the acceptor. B: The variables defining the orientation of the acceptor with respect to the donor.

and the acceptor are not totally free to reach an absolute energy minimum but are subjected to numerous additional constraints at the binding site such as tethering from the protein structure. As a result, the hydrogen bond geometry in inhibitor−receptor complexes can be quite different from that produced from energy minimization of free subunits. We desired a procedure that computes hydrogen-bonding interaction at a specified geometry and can control the effects of secondary interaction contributions to hydrogen-bonding energy.

We devised a constrained energy minimization scheme for the calculation of hydrogen bond energy. The donor and the acceptor were first energy minimized separately with full geometry optimization. They were then brought together into a complex structure defined by one distance variable and three orientation variables. Figure 1 shows the definition of the separation variable, δ, and angular orientation variables, θ_1, θ_2, and Φ, used in this work. Without loss of generality, the donor is assumed to be fixed, and then the hydrogen bond geometry is specified by translation and rotation of the acceptor with respect to the donor. The separation between the donor and the acceptor is defined by the distance between the heavy atoms in the hydrogen bond. The orientation variable Φ describes the rotation of the acceptor around the N−O line (Figure 1) drawn between the heavy atoms of the donor and the acceptor; θ_1 defines the rotation of the acceptor away from the N−O line, with the rotation center being the heavy atom of the acceptor; θ_2 defines the rotation of the acceptor away from the N−O line, with the rotation center being the heavy atom of the donor. At a specified geometry, three atoms from the acceptor and three atoms from the donor are frozen during the geometry optimization process. These constrained atoms do not include the atoms that directly participate in the hydrogen bond. In this manner, geometry optimization can eliminate nonobvious steric problems in the system but will retain the specified hydrogen-bonding geometry. After the structure of the complex is geometrically optimized, the total energy of the complex system is first calculated, and then the energies of the donor and the acceptor are calculated at the optimized complex geometry, including the BSSE correction. To determine the energy minimum for a hydrogen bonding complex, we scanned over the separation and orientation variable space. The setup of different geometries for a given hydrogen bond system was done with a script program; the ab initio calculations on

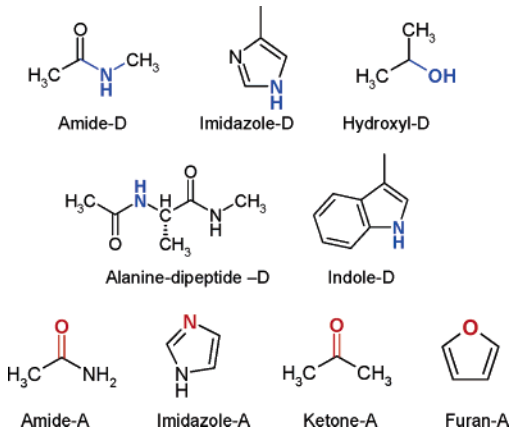


Figure 2. Chemical structures of the first group of donors and acceptors studied in this work. Donor atoms are labeled by blue color and acceptor atoms by red color.

multiple conformations were processed in parallel on a Linux cluster. A constrained geometry optimization job with the DFT method, run on one CPU of a Linux PC, took from 1 to 10 h, depending on the difficulty of the geometry optimization. A single point DFT calculation took less than 20 min on a Linux PC. MP2 calculations with geometry optimization would take a longer time.

Results

The most important hydrogen-bonding interactions to drug design are those between ligand and protein. We used an amide group to mimic the peptide hydrogen bond donor or acceptor. Hydroxyl, imidazole, and indole groups were used as surrogates for hydrogen-bonding donors from the protein side chains. Figure 2 shows a partial list of the chemical structures of the donors and acceptors studied in this work. The hydrogen bond donor and acceptor atoms are shown in blue and red colors, respectively. The labels for the different molecules will be used subsequently in describing the calculation results. We first focused on the dependency and sensitivity of the calculated hydrogen-bonding energy on the geometry of the hydrogen bond, using the amide-D (Figure 2) as donor and amide-A, imidazole-A, ketone-A, and furan-A, defined in Figure 2, as acceptors.

1. Hydrogen-Bonding Energy as a Function of Separation between Donor and Acceptor. Figure 3 shows the interaction energy, calculated with the DFT procedure, of four hydrogen bond complexes as a function of the separation. By definition, a larger negative energy indicates a stronger hydrogen-bonding interaction. Results presented in Figure 3 show that, for all four complexes, there is a sharp energy minimum at a separation of 3.0 Å. At larger separations, the interaction energy for all four complexes decays slowly, at a rate proportional to the inverse distance, resembling the behavior of electrostatic interactions. There are large differences in the interaction energy among the four hydrogen bond complexes: ranging from −6.53 kcal/mol for the imidazole acceptor to −2.13 kcal/mol for the furan acceptor. The minimum interaction energy for the amide and ketone acceptors is −5.44 and −4.06 kcal/mol, respectively. The order of hydrogen-bonding strength for these four acceptors by this calculation is imidazole > amide > ketone

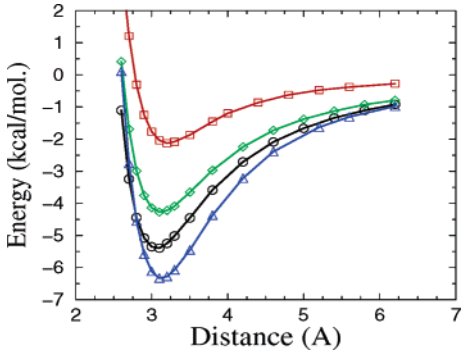


Figure 3. The interaction energy of a hydrogen bond as a function of the separation between the donor and the acceptor calculated with the DFT/B3LYP (6-31++G**) method. The donor is amide-D; the colored coded curves in blue, black, green, and red are for imidazole-A, amide-A, ketone-A, and furan-A as the acceptor, respectively. The molecular structures are given in Figure 2.

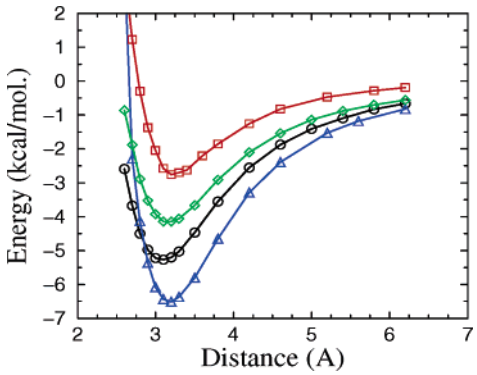


Figure 4. Hydrogen bond energy as a function of the separation between the donor and the acceptor calculated with the MP2 method. The labeling of the curves is identical to that of Figure 3.

> furan. Despite the large difference in the interaction energy, there is little difference in the energy-minimum distance between the four complexes. It was noticed that the energy-minimum distances of the hydrogen bonding calculated here appear slightly longer than the average hydrogen bond length observed from neutron diffraction measurements [Chapter 7 of ref 3]. However, they are within the range of calculated hydrogen bond lengths using different ab initio methods [Table 4.3 of ref 3, Chapter 2 of ref 20].

Quantitatively similar results for the interaction energy of the four hydrogen-bonding complexes were obtained from calculations with the MP2 method. Figure 4 shows the MP2 results. The minimum energy values calculated by the MP2 method are −6.51, −5.26, −4.15, and −2.76 kcal/mol for imidazole-A, amide-A, ketone-A, and furan-A, respectively. The order of hydrogen bond strengths for the four acceptors calculated by DFT and MP2 methods are identical. The calculated energy-minimum distances by the two methods are also similar. However, for furan-A, the weakest hydrogen bond acceptor of the four molecules, the minimum energy calculated by the MP2 method, is lower by 0.63 kcal/mol than that obtained by the DFT method. We observed that the MP2 method yielded, in our calculations, somewhat larger interaction energies for weaker hydrogen bond acceptors (such as the CH- - -O hydrogen bond) than the DFT method. However, MP2 calculations were considerably more time-consuming in geometry optimization on large complex systems. Since the calculation results with the two methods are similar for most common hydrogen bond acceptors, we chose to use the DFT method on all subsequent calculations in this work for convenience and consistency.

2. Directionality of Hydrogen-Bonding Interaction. Moving onto the directional properties of hydrogen bonds, we calculated the interaction energy for acceptors in different orientations with respect to the donor. Figure 5 shows sets of orientations between the imidazole acceptor and the amide donor generated by systematic variation of the variables of θ_1, θ_2, and Φ at a fixed separation of 3 Å between the two subunits. Figure 6 shows the hydrogen-bonding energy of these structures, where curve *a*, *b*, and *c* correspond to the geometric series shown in Figure 5, parts a−c, respectively. These results indicate that the directional property of a hydrogen-bonding interaction is anisotropic. Using the structure of the donor amide as the reference frame, i.e., fixing the position of amide donor, the orientation of the hydrogen bond of this system can vary in two orthogonal directions: i.e., the acceptor imidazole ring swings either within the plane or perpendicular to the plane of the amide structure. The ensemble of structures in Figure 5a is generated as a result of the imidazole ring swinging in the direction perpendicular to the structural plane of the amide donor. Curve *a* in Figure 6 shows the variation of the hydrogen-bonding energy as a function of such an orientational change. In this direction, the acceptor imidazole can rotate by ±45° from the normal direction, while the hydrogen-bonding energy changes by less than 0.5 kcal/mol. In contrast, when the imidazole ring swings within the structural plane of amide around angle θ_1, with the resulting conformations shown in Figure 5c, the hydrogen-bonding interaction weakens rapidly as shown by curve c of Figure 6. The directional dependency of the hydrogen bond energy as a function of angle θ_2, i.e., another orientational variation within the structural plane of the amide donor, and the changes of hydrogen bond energy are between the above two extremes, see the energy curve *b* of Figure 6 and in reference to the ensemble of structures shown in Figure 5b.

Figure 7 shows structural variations between the amide donor and the amide acceptor resulting from rotation of the acceptor around the angle Φ, with the angle θ_1 fixed at 0° and θ_2 at 15° and 30°, respectively. The separation between the donor and the acceptor was fixed at 3.0 Å. Figure 8 shows the calculated hydrogen-bonding energy for the two structural series of Figure 7. The peaks in the energy plots in Figure 8 correspond to structures in which the amide acceptor orients toward the right-hand side methyl group of the amide donor seen from the front of Figure 7. Significantly, as the acceptor tilts closer to the donor, i.e., when θ_2 changes from 15° to 30°, the energy peak increases quickly, as can be seen from a comparison of curves *a* and *b* of Figure 8. This result suggests that the primary reason for the weakening of the hydrogen-bonding interaction in these directions is steric interference between the donor and the acceptor.

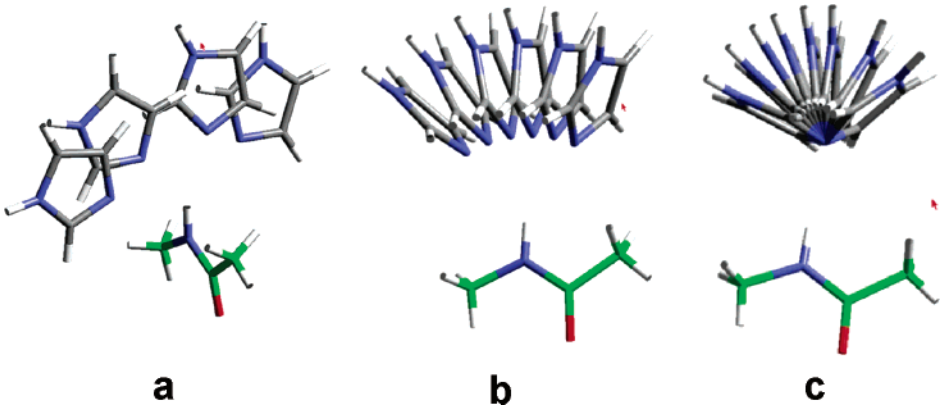


Figure 5. Different orientations between the imidazole acceptor and the amide donor at a fixed separation of 3 Å. a: Conformations generated from varying angle θ_2 with $\Phi = 90°$ and $\theta_1 = 0$. b: Conformations generated from varying angle θ_2 with $\Phi = 0$ and $\theta_1 = 0$. c: Conformations from a variation of angle θ_1 with $\Phi = 0$ and $\theta_2 = 0$.

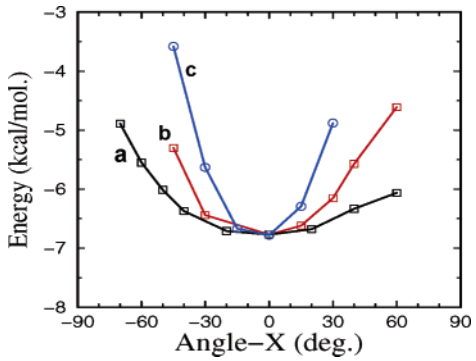


Figure 6. Hydrogen-bonding energy between the amide donor and the imidazole acceptor as a function of orientational variables. Curve *a* is for the conformational series of Figure 5a with the variable X = θ_2. Curve *b* is for the conformational series of Figure 5b with the variable X = θ_2. Curve *c* is for the conformational series of Figure 5c with the variable X = θ_1.

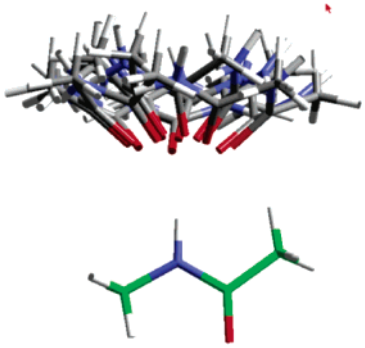


Figure 7. Orientation of the hydrogen bond between the amide donor and the amide acceptor at a constant separation of 3 Å. The conformations are generated by varying angle Φ with $\theta_1 = 0$ and $\theta_2 = 15°$ and 30°, respectively.

Combining the results from Figures 6 and 8, we observed that the hydrogen-bonding energy is rather insensitive to the orientation between the donor and the acceptor. There are rather broad ranges of directions in which acceptor (or donor) can tilt away from the normal direction of the hydrogen bond, by up to ±45°, while the energy changes less than 0.5 kcal/mol. However, the hydrogen-bonding energy could quickly deteriorate when steric interference arises between the donor and the acceptor. The directional anisotropy is determined by the specific chemical structure of the donor and the acceptor and has to be dealt with on a case-by-case basis. In contrast, hydrogen-bonding energy is much more sensitive to the separation between the donor and the acceptor: a 0.2 Å shift in separation can cause over a 1 kcal/mol change in energy. Based on these observations, in our subsequent calculations of hydrogen-bonding energies for different molecules, the initial geometry of the hydrogen bond was set up based on templates of analogous systems from X-ray crystal structures of inhibitor−protein complexes. The orientation between the donor and the acceptor was checked sparsely to ensure no steric interference. The separation was systematically scanned to locate the energy minimum of the hydrogen-bonding interaction.

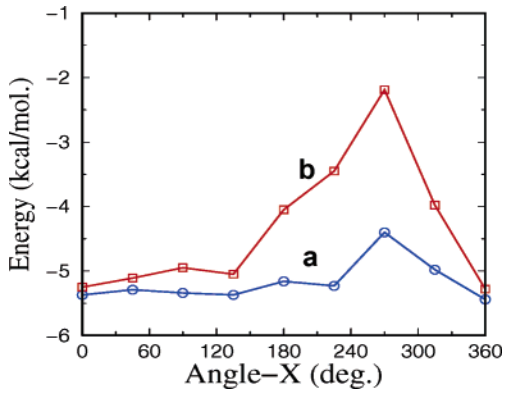


Figure 8. Interaction energy between the amide donor and the amide acceptor as a function of the hydrogen bond orientation at a constant separation of 3 Å. The hydrogen bond conformations are shown in Figure 7. The variable angle-X is Φ angle. Curves *a* and *b* correspond to angle $\theta_2 = 15°$ and 30°, respectively, while angle θ_1 is fixed at 0° and $\theta_2 = 30°$.

3. Comparison of the Hydrogen-Binding Energy of Different Acceptors. Having investigated the calculation procedure, we turned to the validation of the computational procedure by comparison of the calculation results with experimental data. Abraham et al.[1] determined the hydrogen-bonding constants for a large number of donors and acceptors. Hydrogen bond donors are generally confined only to a few types of NH or OH groups. In contrast, there is a large

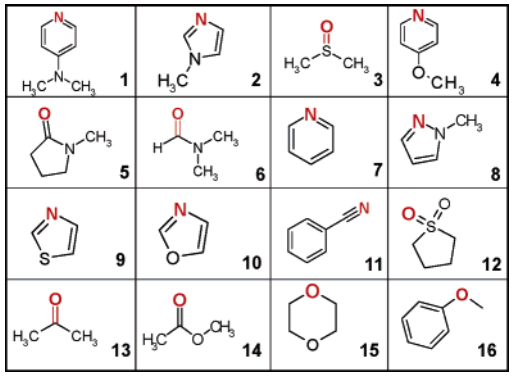


Figure 9. The structures of the hydrogen bond acceptors studied in this work. The atom in red color is the hydrogen bond acceptor.

variety of hydrogen bond acceptors, comprising sp^2 and sp^3 oxygen atoms and sp^1, sp^2, and some sp^3 nitrogen atoms from a large number of functional groups and heterocycles. It has been reported that theoretical modeling of hydrogen bond acceptors is more difficult than modeling hydrogen bond donors.[15] In this work, we selected 16 molecules that cover the major categories of acceptors in the list of Abraham et al.[1] In their experiment, the hydrogen-bonding constants for these acceptors were measured against a common hydrogen bond donor, 4-nitrophenol dissolved in 1,1,1-trichloroethane (TCE), which has a dielectric constant of 7.53. To avoid self-interference in the experiment, all hydrogen donors in an acceptor molecule have to be blocked. As a result, only a limited number of acceptors could be measured in such an experiment. The structures of our selected acceptors are shown in Figure 9. We calculated the interaction energies for these acceptors with the amide hydrogen bond donor (shown in Figure 2). In our calculations, the acceptors and donor were arranged in a fixed orientation based on the procedure described in the previous subsection. Energy calculations were carried out on each structure at 5 different donor−acceptor separations, from 2.8 to 3.2 Å at 0.1 Å intervals. Geometric optimization was done for each structure with constrained energy minimization previously described in the theory section. The lowest energy found from the 5 structures was used as the hydrogen-bonding energy for a given acceptor.

The computational results are collected in Table 1. Column 1 of the table is the label of the acceptor; column 2 contains the calculated hydrogen-bonding energy from this work; and column 3 shows the logarithm of experimental hydrogen-bonding constants, log K_β. The spreading scope of the calculated hydrogen bond energy and log K_β are not the same; and, physically, the first value is energy, while the second number corresponds to free energy. However, an excellent linear correlation exists between the two sets of values. Figure 10 shows the plot of the calculated ΔE_{HB} against experimental log K_β. The following relationship and statistics were found:

$$\log K_\beta = -0.892\Delta E_{HB} - 2.630 \qquad (2)$$

$$N = 16,\ R^2 = 0.94,\ F = 229$$

It is observed that the correlation obtained from this method

Table 1. Hydrogen-Bonding Energies of Different Acceptors with Respect to an Amide Donor

acceptor[a]	ΔE_{HB}[b]	log K_β [c]	acceptor[a]	ΔE_{HB}[b]	log K_β [c]
1	−6.78	3.54	9	−4.76	1.90
2	−6.73	3.68	10	−4.72	1.67
3	−6.20	3.06	11	−4.60	1.06
4	−5.95	2.87	12	−4.39	1.61
5	−5.84	3.12	13	−4.27	1.50
6	−5.37	2.81	14	−4.06	1.43
7	−5.36	2.52	15	−4.03	1.28
8	−5.10	2.22	16	−2.96	0.30

[a] The numerical labels of the acceptors correspond to those defined in Figure 9. [b] ΔE_{HB}, in units of kcal/mol, was defined by eq 1 and calculated with the DFT/B3LYP (6-31++G**) method including BSSE correction. [c] log K_β values adopted from Table 3 of ref 1, also listed in Figures 3 and 4 of ref 15.

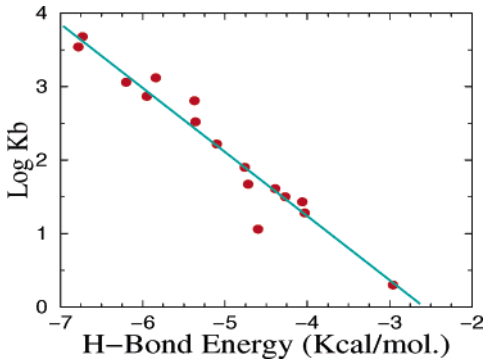


Figure 10. Correlation plot between theoretically calculated hydrogen-bonding energies (H-bond energy) and the logarithm of experimentally measured hydrogen-bonding constants (log K_β). The two sets of values fit the following relationship: log $K_\beta = -0.892E - 2.36$ with $N = 16$, $R^2 = 0.94$, $F = 229$.

is significantly better than that obtained by Gancia et al.[15] Those authors modeled hydrogen bond strength based on a parametrization of log K_β with respect to quantum-chemistry calculated properties of atomic electrophilic superdelocalizability (SE). The regression model between SE and log K_β has larger errors with R^2 in the range of 0.59−0.75. Moreover, in the parametrization approach, different types of acceptors, such as nitrogen and oxygen acceptors, have to be modeled with separated QSAR equations. In our present approach, all acceptors follow the same relationship between the calculated interaction energies and the logarithm of experimental hydrogen-bonding constants. This is consistent with a common physical process for the hydrogen bonding between the pairs of donors and acceptors.

The ab initio calculation results provide insights into the dominant factor for hydrogen-binding strength. While the hydrogen-bonding constant is determined by the binding free energy, the fact that there is an excellent linear correlation between the hydrogen-binding interaction energy and the binding constant suggests that the logarithm of binding constants for different acceptors with respect to given donors are largely determined by an enthalpic factor arising from the interaction energy between the donor and the acceptor. The entropic contribution to the hydrogen-bonding constant would then amount to a constant for all the acceptors. A constant entropic component is not a universal phenomenon;

very different entropic contributions have been observed in ITC experiments on ligands bound to a given receptor.[36] An alternative plausible explanation to the linear correlation between the hydrogen-binding energy and the logarithm of the hydrogen-binding constant could be that the entropy varies linearly with enthalpy. Further investigation will be needed to resolve these two possibilities.

There are interesting individual cases among the set of 16 acceptors. The oxazole molecule, #10 in Figure 9, has two hydrogen bond acceptors: a nitrogen and an oxygen atom. Because only one hydrogen-binding constant was reported for this molecule,[1] it was not clear whether the nitrogen or the oxygen is the actual acceptor in oxazole. Our calculation indicated that the hydrogen-bonding energy for the nitrogen and oxygen atoms in oxazole are −4.72 and −1.23 kcal/mol, respectively. Based on the strong statistic from eq 2, we can confidently predict that the nitrogen atom of the oxazole is the acceptor, whereas the oxygen atom has virtually no acceptor capacity in this particular molecule. A simple explanation of the difference between the hydrogen bond strength of nitrogen and oxygen atoms in oxazole is that the nitrogen atom is more basic than oxygen hence a better hydrogen bond acceptor. Another interesting case is sulfur containing acceptors. As shown in Table 1, sulfoxide is a stronger hydrogen bond acceptor than sulfone, and our calculation indicates that sulfonamide, with a calculated hydrogen-bonding energy of −3.87 kcal/mol, is an even weaker hydrogen bond acceptor than sulfone.

The hydrogen-bonding strength of an acceptor can be easily modified by chemical substitutions on the acceptor's structure. Attesting to this point are the three pyridine derivatives, #1, 4, and 7, shown in Figure 9 and Table 1. The calculated hydrogen-bonding energy of simple pyridine is −5.36 kcal/mol. Substitution on the para position with electron-donating groups such as methoxyl or dimethylamine gives molecules #1 and 4, which have larger hydrogen-bonding energies, −5.95 and −6.78 kcal/mol, respectively. Because dimethylamine is a stronger electron donating group than methoxyl, the former induces a larger increase in the hydrogen-bonding strength for pyridine. Following this line of reasoning, if electron-withdrawing groups are substituted on the pyridine, the hydrogen-bonding strength of the resulting molecules should be smaller than that of pyridine. Indeed, our calculations indicated that the hydrogen-bonding energy of pyridine para-substituted with nitro ($-NO_2$), nitrile ($-C{\equiv}N$), and −CF3 groups are −3.17, −4.27, and −4.26 kcal/mol, respectively. In general, substitution with an electron-withdrawing group will decrease the hydrogen bond strength of an acceptor but will increase the strength of a hydrogen-bonding donor. Electron-donating groups have opposite effects. The electronic effects can be transmitted effectively through conjugated systems. Modulating the hydrogen-bonding strength with proper substituents on a hydrogen donor and/or acceptor can be a very useful strategy in drug optimization and has been utilized in the drug optimization process.[2]

4. Effects of Hydrogen Bond Donors. One assumption in the previous subsection was that the *order* of the hydrogen-bonding energy for different acceptors should be generally the same regardless which donor was used in the calculation. The same assumption also holds for the experimentally measured hydrogen-bonding constants for acceptors,[1] which are transferable across different donors. We verified this assumption by calculating the interaction energy of the four acceptors shown in Figure 2 with the five donors shown in the same figure. Table 2 shows the hydrogen-bonding energies for these molecules: the columns correspond to the donors, while the rows correspond to acceptors. Crossing the columns, it is observed that the hydrogen-bonding strength of the donor is in the order of imidazole > indole > amide > hydroxyl. Crossing the rows, it is observed that the order of acceptor strength is as imidazole > amide > ketone > furan. As expected, the same order of hydrogen-bonding strength for the acceptor series holds across the different donors. In the other direction, the same order of hydrogen-bonding strength for the donor series holds across different acceptors.

Table 2. Hydrogen-Bonding Energy between Acceptors and Donors[a]

	amide-D	hydroxyl-D	indole-D	imidazole-D	alanine dipeptide-D
imidazole-A	−6.55	−6.24	−7.11	−7.96	−8.03
amide-A	−5.44	−4.51	−6.27	−6.92	−6.94
ketone-A	−4.06	−3.54	−5.12	−5.78	−5.52
furan-A	−2.13	−2.22	−2.47	−2.75	−3.01

[a] The hydrogen-bonding energy is in units of kcal/mol. The chemical structures of the donors (top row) and acceptors (first column) are shown in Figure 2. The energy values were calculated by the DFT/B3LYP(6-31++G**) method with BSSE correction.

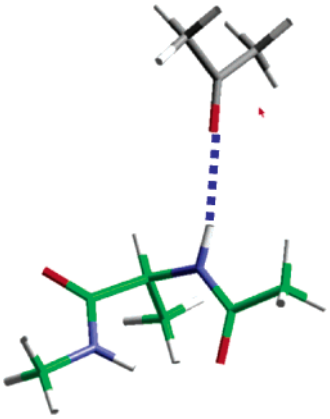


Figure 11. The conformation used in calculating the hydrogen-bonding energy for a ketone donor and an alanine dipeptide donor.

The case of alanine dipeptide as donor warrants further discussion. While we have been largely concerned so far with small donors such as amide and hydroxyl, the hydrogen-bonding groups of a receptor interacting with drug molecules are parts of the protein. It is of interest to know how a hydrogen bond donor from a larger structure interacts with a small acceptor or vice versa. Figure 11 shows the conformation of hydrogen-bonding interaction between a ketone acceptor and an alanine dipeptide donor. Compared to the amide donor shown in Figures 5 and 7, the alanine dipeptide clearly has more structural mass that is not directly involved in hydrogen bonding with the acceptor. Our

computational result has indicated that the amide donor is not the strongest in the set of five donors. Nonetheless, as shown in Table 2, the calculated interaction energy for all the acceptors with the alanine dipeptide donor is larger than with other donors. Clearly, there are secondary interactions other than hydrogen bonding that contributed to the calculated total interaction energy between the alanine dipeptide and the acceptors. This result suggests that, to obtain an accurate estimate for the hydrogen-bonding interaction in a given binding site of the receptor, secondary interactions have to be taken into account in the total energy calculation. It is likely that secondary interactions will vary in different binding sites and have to be dealt with on an individual basis. The constrained energy minimization procedure designed here could be used to calculate more complicated systems with a specified geometry. One difficulty in extending the present calculation to the protein binding site was that there is no experimental hydrogen bond parameter that can be compared with, thereby validate, theoretically calculated values.

Discussion

A key result to this work is the observation that ab initio calculated hydrogen-bonding energy linearly correlates with the logarithm of experimentally measured hydrogen-bonding constants. This provides a basis for theoretical prediction of hydrogen-bonding strength for drug design by ab initio calculations. The results presented here were obtained from DFT/B3LYP calculations with the 6-31++G** basis set. We had found that calculations at the MP2/6-31G** level gave quantitatively similar results. The MP2 calculations seemed to produce larger interaction energies for weaker hydrogen-bonding acceptors or donors, but the DFT/B3LYP calculation is faster and can provide consistent and satisfactory information.

It should be emphasized again that the calculated property here is the hydrogen-bonding energy. Since it is the free energy change from solvent to a receptor binding site that ultimately determines the hydrogen-bonding constant, it was not obvious that the interaction energy alone can explain the variations in hydrogen-binding strength from one acceptor to another. However, the excellent linear relationship between the calculated interaction energies at ab initio quantum level and logarithm of hydrogen-bonding constants for a variety of acceptors gives support to the theory that the enthalpic contribution from the hydrogen-bonding energy determines whether one hydrogen bond is stronger than another. Plausibly, entropic factors make only a constant contribution to the free energy change when different acceptors bind to a given donor. It is likely that a similar relationship also holds in the case of ligands binding to a receptor, at least for a series of compounds adopting the same binding mode. Another plausible explanation is that the entropic component varies proportionally with enthalpy in these systems.

Predicting the hydrogen-bonding strength by ab initio quantum chemical calculations has considerable advantages over semiempirical parametrization approaches.[8−15] First, the correlation between the ab initio calculated energies and experimental hydrogen-binding constants is much better than that from a parametrization approach.[15] Second, all different acceptors or donors can be evaluated by the same computational procedure in a unified correlation equation with the ab initio method. In contrast, with the parametrization approach, multiple equations would be needed for different types of donors or acceptors. For traditional QSAR models, the application range of a given model is even narrower. Third, the ab initio method can handle steric factors in a hydrogen bond system in a straightforward manner since ab initio calculation takes all electronic and steric factors of the system into account. Steric effects have been a major problem for the transferability of parametrization-based hydrogen bond models because, in such models, only one subunit of the donor−acceptor pair in a hydrogen bond is considered. A major drawback of the ab initio method is the high demand on computational resources. When fast Linux clusters become readily available, multiple ab initio calculations involved in distance scans can be carried out in parallel such that results for a handful of molecules typically encountered in an SAR study can be obtained overnight. Therefore, the computational demand of the ab initio method is not presently a major obstacle.

The directional properties of hydrogen bonds are important to both computational studies and practical applications. A survey of X-ray and neutron diffraction structures of organic and biological molecules[3] had found that the directional dependency of hydrogen bonds is weak: the distribution of hydrogen-bonding orientation almost evenly spreads over a broad band of ±50° from the normal direction for donor and/or acceptor. Our calculations also indicated that, at the normal hydrogen-bond distance, the interaction energy is not very sensitive to angular variations up to ±45° in certain directions. This property of the hydrogen bonds simplifies the task of calculating the minimum energy of hydrogen-bonding interaction. The orientation of the system needs only to be sparsely checked to ensure that there is no major steric problem. However, because of the sensitivity of the interaction energy to donor−acceptor separation, a scan over the distance variable would be required to identify the accurate energy minimum. Ireta et al.[26] reported that the accuracy of the DFT calculated hydrogen-bonding energy depends on the bond directionality. When the hydrogen bond deviates from linearity, the discrepancy between the DFT energy and the values calculated by the MP2 method[37,38] increases. The reason for the somewhat different observations between Ireta et al. and the present work is not clear.

Another interesting result of the present work is that the directionality of the hydrogen bond is not isotropic. While the statistics for a large number of hydrogen-bonding structures from X-ray data shows a nearly uniform distribution of hydrogen bond orientation over a broad range, individual molecules have different structures and steric characteristics. At the individual molecule level, hydrogen-bonding energy can deteriorate quickly in some directions. Using analogous X-ray structures as templates is a reasonable approach for setting up the geometry of a hydrogen bond system for energy calculation. When steric problems occur, the orientation of the hydrogen bond has to be adjusted to obtain a correct hydrogen-bonding energy.

Secondary interactions will be a complicating factor for an accurate estimate of the hydrogen-bonding interaction for ligand binding to a protein receptor. In an earlier study by Scheiner et al.,[35] it was reported that when the donor molecule increases in size from one to two amino acid units, the hydrogen-bonding energy of the water with the peptides changes only by ~10%. It is possible that a water molecule is a very small and simple system which would experience less secondary interactions with the donor molecules. For larger acceptors such as drug molecules interacting with protein receptors, secondary interactions accompanying the hydrogen-bonding interaction could be considerably large. To accurately estimate the strength for larger hydrogen bond donors or acceptors interacting with specific protein receptors, secondary interactions have to be dealt with adequately.

The usefulness of theoretically calculated hydrogen bond energies to drug design has been demonstrated by Bingham et al.[2] In their work, the SAR of a series of analogous molecules bound to the protein kinase IKK can be accounted for by the hydrogen bond strength of these compounds. That result was one of the motivations of the present study. With a more accurate and comprehensive calculation scheme, the present method can be used for problems involving hydrogen bond bioisostere in drug design; a comparison of the hydrogen-bonding energies of the bioisostere provides a more rational basis for the bioisostere approach. The calculated hydrogen-bonding energy by the present method can also be used in the refinement of force field parameters for molecular modeling simulation. Based on the author's experience, currently used force fields were not parameterized to reproduce the difference of the experimentally measured hydrogen bond strengths among different donors and acceptors.

Conclusion

We demonstrated in this work that ab initio calculated hydrogen-binding energy has an excellent linear correlation with logarithm of experimental hydrogen-bonding constants. This provides a basis for the theoretical prediction of hydrogen-bonding strength for a series of acceptors (or donors) with respect to a given donor (or acceptor). The ab initio approach advances the level of sophistication of theoretical modeling of hydrogen bonds and may provide deeper insights to the mechanism of hydrogen-bonding interactions. The method also provides a tool for handling steric effects, conformational properties, and secondary interactions of specific systems in predicting hydrogen bond strengths. These effects were traditionally not treated in empirical and parametrization approaches for hydrogen bonds. We believe that the method developed here is helpful in dealing with hydrogen bond related SAR problems in drug design.

Acknowledgment. The author thanks Dr. Ingo Mugge for his valuable comments and suggestions about this work and Drs. Sandy Farmer and James Stevenson for their thorough reading and editing of the paper.

References

(1) Abraham, M. H.; Duce, P. P.; Ptior, D. V.; Barratt, D. G.; Morris, J. J.; Taylor, P. J. *J. Chem. Soc., Perkin Trans*. **1989**, *11*, 1355−1374, 1989.

(2) Bingham, A. H.; Davenport, R. J.; Gowers, L.; Knight, R. L.; Lowe, C.; Owen, D. A.; Parry, D. M.; Pitt, W. R. *Bioorg. Med. Chem. Lett*. **2004**, *14*, 409−412.

(3) Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer-Verlag: Berlin, 1989.

(4) Umeyama, H.; Morokuma, K. *J. Am. Chem. Soc*. **1977**, *99*, 1316−1332.

(5) Morokuma, K.; Kitaura, K. In *Chemical applications of atomic and molecular electrostatic* potentials; Politzer, P., Truhlar, D. G., Eds.; Plenum: New York, 1981; pp 215−242.

(6) Jeffrey, G. A.; Maluszynska, H. *Int. J. Biol. Macromol*. **1982**, *4*, 173−185.

(7) Taylor, R.; Kennard, O.; Versichel, W. *J. Am. Chem. Soc*. **1983**, *105*, 5761−5766.

(8) Fujita, T.; Nishioka, T.; Nakajima, M. *J. Med. Chem*. **1977**, *20*, 1071−1081.

(9) Charton, M.; Charton, B. I. *J. Theor. Biol*. **1982**, *99*, 629−644.

(10) Wilson, L. Y.; Famini, G. R. *J. Med. Chem*. **1991**, *34*, 1668−1674.

(11) Murray, J. S.; Politzer, P. *J. Org. Chem*. **1991**, *56*, 6715−6717.

(12) Dearden, J. C.; Cronin, M. T. D.; Wee, D. In *QSAR and molecular modeling: concepts, computational tools and biological applications*; Sanz, F., Giraldo, J., Manaut, F., Eds.; Proust Science Publishers: Barcelona, Spain, 1995; pp 117−119.

(13) Dearden, J. C.; Cronin, M. T. D.; Wee, D. *J. Pharm. Pharmacol*. **1997**, 49 Suppl. 4, 110.

(14) Dearden, J. C.; Ghafourian, T. *J. Chem. Inf. Comput. Sci*. **1999**, *39*, 231−235.

(15) Gancia, E.; Montana, J. G.; Manallack, D. T. *J. Mol. Graphics Modell*. **2001**, *19*, 349−362.

(16) Kollman, P. A.; Allen, L. C. *Chem. Rev*. **1972**, *72*, 283−303.

(17) *Modeling the hydrogen bond;* Smith, D. A., Ed.; American Chemistry Society: Washington, DC, 1994; Vol. 569.

(18) Scheiner, S.; Kar, T. *J. Am. Chem. Soc*. **1995**, *117*, 6970−6973.

(19) Pan, Y.; McAllister, M. A. *J. Org. Chem*. **1997**, *62*, 8171−8176.

(20) Scheiner, S. *Hydrogen Bonding. A Theoretical Perspective*; Oxford University Press: New York, 1997.

(21) Rablen, P. R.; Lockman, J. W.; Jorgensen, W. L. *J. Phys. Chem. A* **1998**, *102*, 3782−3797.

(22) Gu, J.; Leszczynski, J. *J. Phys. Chem. A* **2000**, *104*, 7353−7358

(23) Lukin, O.; Leszczynski, J. *J. Phys. Chem. A* **2002**, *106*, 6775−6782.

(24) Sukhanov, O. S.; Shishkin, O. V.; Gorb, L.; Podolyan, Y.; Leszczynski, J. *J. Phys. Chem. B* **2003**, *107*, 2846−2852.

(25) Domingo, L. R.; Andres, J. *J. Org. Chem.* **2003**, *68*, 8662−8668.

(26) Ireta, J.; Neugebauer, J.; Scheffler, M. *J. Phys. Chem. A* **2004**, *108*, 5692−5698.

(27) Guo, H.; Gresh, N.; Roques, B. P.; Salahub, D. R. *J. Phys. Chem. B* **2000**, *104*, 9746−9754.

(28) Alia, J. M.; Edwards, H. G. M. *J. Phys. Chem. A* **2005**, *109*, 7977−7987.

(29) Quinonero, D.; Garau, C.; Frontera, A.; Ballester, P.; Costa, A.; Deya, P. M. *J. Phys. Chem. A* **2005**, *109*, 4632−4637.

(30) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 4209−4212.

(31) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553−566.

(32) Jencks, W. P. *Adv. Enzymol*. **1975**, *43*, 219−410;

(33) Page, M. I. *Angew. Chem., Int. Ed. Engl*. **1977**, *16*, 449−459.

(34) *Jaguar v.5.5*; Schrodinger, Inc.: 2005.

(35) Scheiner, S.; Kar, T.; Gu, Y. *J. Biol. Chem*. **2001**, *276*, 9832−9837.

(36) Kroe, R. R.; Regan, J.; Proto, A.; Peet, G. W.; Roy, T.; Landro, L. D.; Fuschetto, N. G.; Pargellis, C. A.; Ingraham, R. H. *J. Med. Chem.* **2003**, *46*, 4669−4675.

(37) Vargas, R.; Garza, J.; Dixon, D. A.; Hay, B. P. *J. Am. Chem. Soc*. **2000**, *122*, 4750−4755.

(38) Vargas, R.; Garza, J.; Friesner, R. A.; Stern, H.; Hay, B. P.; Dixon, D. A. *J. Phys. Chem. A* **2001**, *105*, 4963−4968

CT0600262

JCTC Journal of Chemical Theory and Computation

Computing the Amino Acid Specificity of Fluctuations in Biomolecular Systems

K. Hamacher*,†,‡ and J. A. McCammon†,‡,§

Center for Theoretical Biological Physics, Department of Chemistry and Biochemistry, Department of Pharmacology, and Howard Hughes Medical Institute, University of California at San Diego, La Jolla, California 92093-0365

Received October 7, 2005

Abstract: We developed a new amino acid specific method for the computation of spatial fluctuations of proteins around their native structures. We show the consistency with experimental values and the increased performance in comparison to an established model, based on statistical estimates for a set of test proteins. We apply the new method to HIV-1 protease in its wild-type form and to a V82F−I84V mutant that shows resistance to protease inhibitors. We further show how the method can be successfully used to explain the molecular biophysics of drug resistance of the mutant.

1. Introduction

Fluctuations of biomolecular complexes around their native states are of great importance in functional studies in molecular biophysics. One can deduce several features such as entropy changes upon binding,[1] probable binding sites for drugs,[2] or the overall stability and function of complexes[3] from the detailed analysis of those fluctuations.

Traditional normal-mode analysis (NMA)[4−6] utilizes the carefully developed force fields of molecular dynamics (MD) packages. Here, one minimizes the energy of the structure and diagonalizes the Hessian of the energy function in the minimum to obtain normal modes. One has, however, to bear in mind two shortcomings: first, NMA requires a lot of computational time in the minimization step, and second, the MD force fields are optimized to mimic several static and dynamical properties at the same time and, therefore, are not necessarily optimal for low-energy fluctuations around the native state.[7]

To tackle the first problem, one-parameter potentials were introduced: the Tirion potential[8] and the related Gaussian network model (GNM).[9] The sole parameter is here adjusted to reproduce experimental B factors, and one starts from an experimental protein structure without any minimization. The anisotropic network model (ANM)[10] and the GNM were thoroughly investigated with respect to molecular dynamics simulation[11] and proved to be in good agreement despite their simplicity.

NMA, GNM, and ANM were successfully applied to a variety of molecules such as tRNAs,[12] tryptophan synthase,[13] cowpea chlorotic mottle virus,[14] and tubulin.[3]

The success of the Tirion potential and the GNM stems from the observation that the large number of atomic contacts distributed in the protein makes the central limit theorem valid—leading to the mean strength γ in, for example, the Tirion potential $V = \sum_{(ij)\in\mathcal{C}}(\gamma/2)(s_{ij} - s_{ij}^{o})^2$ with $s_{ij} = |\vec{r}_{ij}|$, where $\vec{r}_{ij}$ is the distance vector between the C_α atoms of contacting residues i and j. Members of the set of contacting residues are defined by $\mathcal{C} = \{(i, j)|s_{ij}^{o} < R_c\}$ with some given R_c, and the superscript o indicates the native conformation of the N residues throughout this paper. These methods were extended in several aspects: larger systems,[15,16] anisotropy,[10] or side-chain geometries (βGM).[17] Additionally, investigations, for example, ref 18, included anharmonic effects.

These methods, however, still lack—because of the intrinsic averaging procedure—the ability to incorporate the chemical nature of the residues constituting the protein. Therefore, there is a gap between NMA based on empirical MD force fields, which provide for specificity, and the one-

* Corresponding author phone: 1-(858)-534-7356; e-mail: hamacher@ctbp.ucsd.edu; URL: http://www.kay-hamacher.de.

† Center for Theoretical Biological Physics.

‡ Department of Chemistry and Biochemistry.

§ Department of Pharmacology and Howard Hughes Medical Institute.

10.1021/ct050247s CCC: $33.50
Published on Web 04/15/2006

parameter models that provide for fast computations and, therefore, large-scale screening capabilities important for studies of stability and function as well as drug design. We fill this gap with the proposed method.

2. The Model

Let the potential consist of two terms: a Hookean pairwise potential for covalent bonds $(K/2)\sum_i(s_{i,i+1} - s^{o}_{i,i+1})^2$ and a contact potential that assigns to contacting residues i and j an amino acid type-specific energy κ_{ij} in the native structure, which is at the same time assumed to be the global minimum. Fluctuations can break contacts if they get larger than some value a. These constraints can be modeled (see ref 19 for a similar approach) by a harmonic term inside a well with width a: $(\kappa_{ij}/a^2)[((s_{ij} - s^{o}_{ij})^2 - a^2)\Theta(a^2 - (s_{ij} - s^{o}_{ij})^2)]$, where $\Theta(x)$ is the step function [$\Theta(x) = 1$ for $x \geq 0$ and is 0 otherwise].

In the native configuration $\vec{s}^{o}_{ij}$, this results in the value κ_{ij}. We will use this particular property later on.

If the typical displacement is smaller than the well's width, we can omit the step function and obtain

$$V = \alpha a^{-2}\left[\frac{a^2K}{2}\sum_i(s_{i,i+1} - s^{o}_{i,i+1})^2 + \sum_{(i,j)\in\mathcal{C}'}\kappa_{ij}(s_{ij} - s^{o}_{ij})^2\right] \quad (1)$$

where we dropped an unimportant constant and introduced the scaling factor α to fit experimental B factors later on. Here, $\mathcal{C}' = \mathcal{C}\backslash\{(i,j)||i-j| = 1\}$ is the set of noncovalent-bonded contacts.

We note in passing that starting with the bond-breaking model and deriving eq 1 makes the whole approach consistent with the ANM and the knowledge-based potentials we will use later on to parametrize κ_{ij}: the prefactor αa^{-2} is an overall fit parameter that the ANM chooses so that the average B factor equals the one of the experiment and the various κ_{ij}'s *weight* the different contacts against each other, while we do not have to demand that the κ_{ij}'s are exact energy values (an overall scaling factor might be included).

Because we are interested in fluctuations, we develop V to the second order around the native state, which is the global minimum of eq 1 generically. The constant and the linear term vanish, and we obtain $V \approx {}^1/_2\Delta\vec{R}^{T}\mathbf{H}\Delta\vec{R}$, where $\Delta\vec{R}$ is the $3N$-dimensional displacement vector of all residues. The matrix $\mathbf{H}$ is the Hessian of eq 1 with respect to all $3N$ coordinates. We omit the somewhat clumsy expressions[20] for brevity here.

By means of singular value decomposition,[21] we obtain the eigenvalues λ_k and the eigendirections $\vec{u}_k$ of the movements. We effectively extend the anisotropic network model[10] to describe amino acid specific interactions. We refer to this method as extended ANM (EANM) in this paper because we would regain the ANM potential by setting uniformly $(2\alpha\kappa_{ij}/a^2) = \alpha K = \gamma$. Spatial fluctuations can afterward be computed from the inverse of $\mathbf{H}$ by

$$\langle(R^{\alpha}_i)^2\rangle = 3k_{B}T\sum_k\frac{[\vec{u}_k\vec{u}^{T}_k]^{\alpha\alpha}_{ii}}{\lambda_k} \quad (2)$$

for $\alpha = x, y$, and z, and the B factor of residue i can be computed from $B_i = (8\pi^2/3)\sum_{\alpha=x,y,z}\langle(R^{\alpha}_i)^2\rangle$.

Here, we weight the interactions κ_{ij} inside one protein (hereafter called "intrachain") according to the absolute value of the established contact potential of Miyazawa and Jernigan,[22] while the κ_{ij}'s for interchain contact between residues in multiple-chain proteins (see below for the HIV-1 protease) are taken as the absolute values as suggested by Keskin et al.[23] The rationale behind this is the property of our model (see above) that, in the native state, the potential is just the sum of the contact energies κ_{ij} of all the contacts of the native state. The intra- and interchain contacts are weighted differently because there are subtle differences in the contact potentials for these two classes of contacts. The major physical difference between the two types of contacts is illustrated: (a) The packing density inside a single protein is higher; therefore, the statistics of contacts is different. (b) The "hydrophobic force" acts differently on residues of one chain during folding and residues in protein−protein-binding events.

The values of these knowledge-based potentials take various effects into account that are responsible for the observed frequencies of contacts: electrostatics, solvent effects, hydrogen bonds, and so on. Although, this choice of κ_{ij} is strictly valid only for ensembles in the Bethe approximation in which the formation of pair contacts is thought to be equivalent by a chemical reaction and thus governed by a reaction law for the association/dissociation of pairs of residues individually. The underlying principles of minimal frustration and self-averaging were first discussed in the works of Wolynes and others and Onuchic and others.[24,25] It turns out that this approximation is precise enough for our purposes (see section "Performance" below for a justification).

The hard-core repulsion of Miyazawa and Jernigan's contact potential[22] is not taken into account as we are dealing with the native fold of the protein that cannot show any clashes. This proved to be a reasonable choice in other applications of the contact potential values,[22] too.[26] The low-frequency movements we are interested in are also the ones with large-scale changes in the structure, thus tending to involve solvent effects that are included in κ_{ij}. The strength along the peptide bond K is a fitting parameter, while we adjust αa^{-2} to agree with experimental B factors in the spirit of the GNM. The fitting allows one to fix the energy scale, avoiding the setting of a "folding" temperature a priori as was discussed to be crucial for contact potentials.[22,27] Additionally, we compensate with the fitted K for the various temperatures used in experiments and the unknown well width a.

Because we start with an experimental structure and the κ_{ij}'s were derived also from those structures, we are consistent. Because of the experimental structure, we also do not have to worry about finding a "problematic" minimum as might happen in NMA. The necessary structure minimization step in NMA might lead to different configurations for different mutants—making the comparison of eigenvectors and the implied principal directions of motion difficult or impossible.

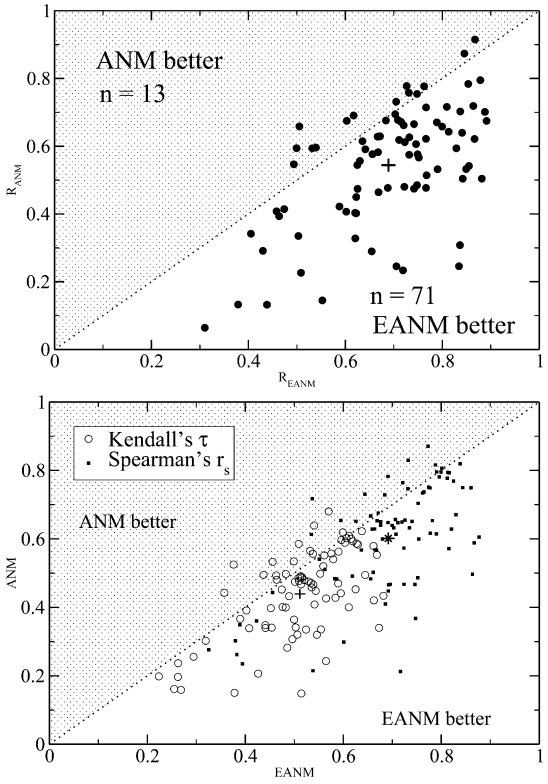


Figure 1. Graphs comparing the performance of EANM and ANM with respect to the set of test proteins. For every point, we computed the B factors and measured the agreement with three different data analysis methods. The upper graph shows the linear regression coefficient of the predicted vs the experimental values, while the lower uses two ranking measures. The crosses and the asterisk indicate the average over the set. The dashed lines are guides to the eye to distinguish the "winning method". The data points, for which ANM was better than EANM, are close to this line—indicating that EANM is even then not much off.

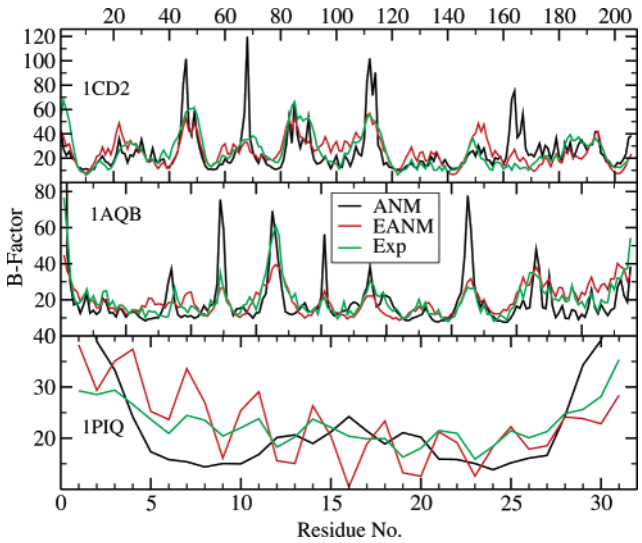


Figure 2. B factors for three proteins. Upper part: Note the failure of ANM for residues 162−168 and 68; both the experimental and the EANM values are much smaller. This is not just a common factor between those values and ANM, as can be seen from the areas around 20−35, 120−140, and 175−200. Middle: the ANM magnitudes are, at certain locations, good (e.g., 1−45 and 66). EANM is, however, capable of giving more-detailed insight over the whole range of the backbone. Lower part: EANM is not just lowering excessively high B factors but also reflects more of the underlying topology.

3. Performance

We applied EANM to 84 proteins[28] taken from the PDB database and set the contact defining distance to $R_c = 13$ Å. To reproduce the experimental B factors, we scanned for K of eq 1 in steps of five for every protein individually. We optimized the linear correlation between the experimental values and the computed ones.

In a recent coarse-grained study, a value of 83.333 RT/Å^2 was derived for the peptide bond.[29] In another study[30] on allostery of trypsiogen, Wall and Ming found $K = 53.4$ kcal/(mole Å^2) $= 89$ RT/Å^2 to be optimal. The obtained best K values here were found to be 82 ± 2.86 RT/a^2.

Because the all-atom average B factor of the proteins is 18.8 Å^2 (17.5 Å^2 for the C_α alone), an average displacement/resolution of 0.85 Å results. The K value and the displacement are both consistent with the above-mentioned value of Trylska et al. and Wall and Ming for $a \approx 1$ Å. The small standard deviation of K is in accordance with the small temperature range of experimental B factors.

The quality of every computation is measured by the linear correlation R, Spearman's r_s, and Kendel's τ[21] between the predicted and the experimental B factors. In Figure 1, we plotted for every protein the performance of the best K EANM computation in the x direction and the respective ANM result in the y direction.[31] Note that EANM is variational in the sense that increasing the accuracy of the K scan can only decrease the error with respect to the experimental values. Therefore, the results become, in general, better for smaller ΔK values, and EANM will perform even better than ANM.

Figure 2 shows the B factors for three arbitrarily chosen proteins from the test set. Even under this crude K scan, not only is the agreement better but, additionally, structural features are reproduced in more detail with EANM.[20]

On the basis of the top two proteins of Figure 2, one might be tempted to attribute the increased performance of EANM most of all to lowering excessively high B values. To account for this, we first note that the ranking measure by Kendel and Spearman does not take into the account the overall magnitude of the B values, only their respective ranking—therefore, neglecting the order of magnitude of the values. Nevertheless, the performance of EANM is much better, giving evidence of the fact that EANM is indeed capable of giving a higher level of detail of the *shape* of the B values of the molecular system. A potential lowering of the B values can only result from the strengthening of the covalent bonds by larger K values (an order of magnitude larger than the κ_{ij} values). To test whether the increased performance of the EANM can be attributed to the larger K values alone, we repeated the computation in the test set with the best $K_{best} = 82RT$ and a uniform contact interaction $\kappa_{ij}^{uniform} = 3.563$ RT, which is the average of the contact potential by Miyazawa

and Jernigan.[22] We found that, under all three approaches to measure the performance (linear correlation coefficient, Spearman's r_s, and Kendel's τ), the results were worse than those with the amino acid specific κ_{ij}. We thus conclude that the increased performance cannot be attributed to the lowering of excessively high B values or the stronger covalent bond alone.

4. Application to HIV-1 Protease

HIV-1 protease[32] is a crucial protein during HIV infection. Protease inhibitors bind to a "pocket" of this dimer and prevent further activity, thus reducing the spread of HIV viruses in an organism.

GNM was used by Kurt et al.[33] to investigate the effect of a ligand bound to the protein complex. Because the effects are only due to the change of the contact map for the complex upon binding the ligands, no amino acid specificity could be investigated.

Drug resistance due to mutation of the protease has, however, become a major problem[34] and prompts for sequence-specific studies. The dissociation rate of protease inhibitors is greatly influenced by the dynamics of the flaps of the pocket. A recent and comprehensive investigation[35] on several mutants of HIV-1 protease used normal-mode analysis. This, however, involves time-consuming minimization, energy function evaluation, and so forth, as discussed in the Introduction.

We want to show here how our model closes the gap between traditional NMA and GNM/ANM/βGM for this molecular complex. We used the experimentally determined wild-type structure[32] and applied the mutations V82F and I84V to be able to focus on those residues solely. We set the distance for which contacts are obtained to $R_c = 20$ Å because we found a more pronounced effect for this value.

Our setting is justified by the fact that, in the pocket, the small number of water molecules eventually bridges the spatial gap of size ~22 Å, extending the range of interaction partially throughout the pocket of HIV-1 protease as a water-molecule network, as recently discussed by Papoian et al.[36] Our choice of $R_c = 20$ Å, therefore, does not introduce contacts between the mutation sites and the flaps. All of the influence of the mutations is therefore carried by a mechanical mechanism encoded in the structure of the HIV-1 protease homodimer (see below for results on the mechanism). To confirm our choice further, we note that this R_c leads to the correct shape of the frequency distribution. The frequency distribution of a general protein motif was first obtained by Elber and Karplus[38] with the effective medium approximation. For the particular case of the HIV-1 protease, Zoete et al.[35] showed the contribution to the RMSD, which is, by means of eq 2, proportional to the square of the frequency—our results are in agreement with these RMSD contributions.

Our main results are depicted in Figures 3 and 4. The mutation of just two residues in each molecule of the protease dimer results in an increased movement of the flaps.

One can now use the eigenvectors to investigate the change in the motion of the residues upon mutation. One approach—the one we pursue in the following—is to look for orthogonality between modes in the different mutants for particular residues of interest. This allows an investigation of whether some residues are subject to larger effects than others.

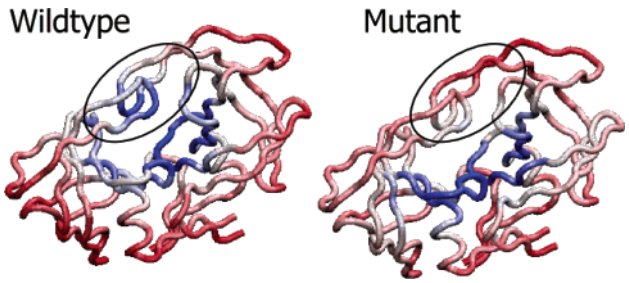


Figure 3. Structures of HIV-1 protease, its mutant, and the respective B factors. Overlaying the predicted B factors onto the structure of HIV-1 protease shows the increased flexibility of the "flaps". While the base of the pocket shows small fluctuations (blue), the mutant provides for the possibility of larger movements (red) of the "flaps". Note that the mutation occurs *on the base* of the pocket.

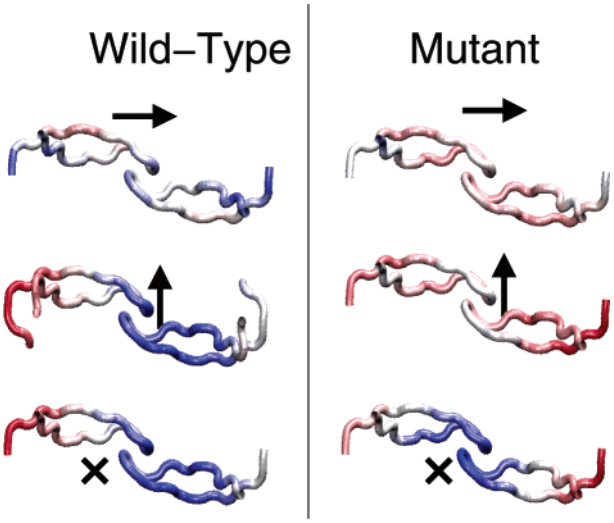
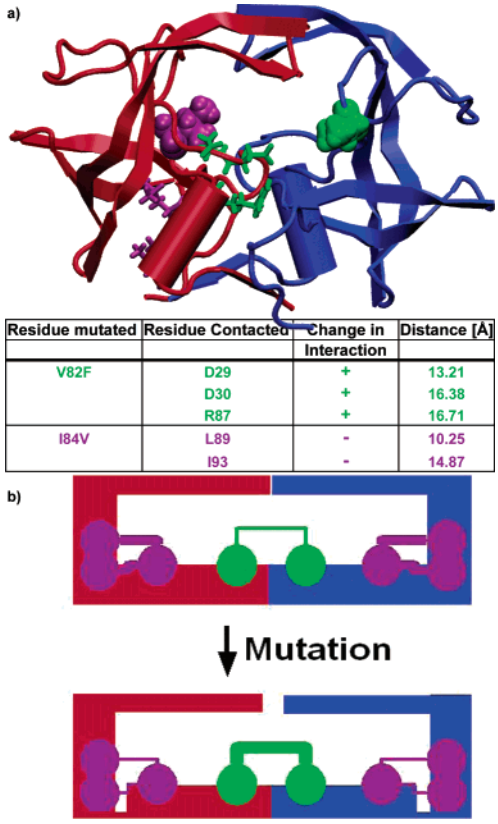


Figure 4. Anisotropic fluctuations for the HIV-1 protease in the directions indicated by the arrows. The mutation causes larger movements to open/curl the flaps, while the other two spatial directions are not affected on a larger scale (red, higher fluctuations; blue, lower fluctuations).

The residues 43−58 constitute the flaps in the homodimer of HIV-1 protease. Analyzing the overlap of all 594 obtained modes in the subspace of those residues between the wild-type and the mutant, we found that the 21st modes are the slowest mutual orthogonal ones.

These modes are $\sim\sqrt{4.1}$ faster than the overall slowest mode. The small ratio guarantees that this mode is substantially realized.[39] From eq 2, we can immediately conclude that the contribution of this mode to the overall B factor is $\approx 1/4.1 \approx 24\%$ of the first mode. The mode provides for the curling of the tips as a precursor of opening.[20] The curling was observed to be a prerequisite to opening the flaps.[37]

To elucidate the mechanism that results in increased floppiness of the flaps, we performed some thousand mutations studies: we changed the interaction strength between the contacts of residues 82 and 84 pairwise while leaving the rest alone. We therefore computed a "hybrid"



Residue mutated	Residue Contacted	Change in Interaction	Distance [Å]
V82F	D29	+	13.21
	D30	+	16.38
	R87	+	16.71
I84V	L89	-	10.25
	I93	-	14.87

Figure 5. (a) Residues that are mutated (spheres) and the respective residues they contact (sticks). These provided for most of the change in the spectrum. The table shows the residue numbers and the respective change in the interaction strengths upon mutation. In addition, we show the distances between residues which supports our choice of $R_c = 20$ Å here. (b) The emerging picture: upon mutation, the "base" becomes more stiff because of the increased strength of the bonds between the two monomers; at the same time, the flaps and their "anchors" are weaker bonded and, therefore, can move more freely in comparison to the base.

system, which contained the original contacts of residues 82 and 84 in the wild type to most of their contacting neighbors while changing κ_{ij} to the mutant values for pairs of contacting residues. This allows for a systematic investigation of the origin of the effect of the mutation. A *global* measure of the influence of such a hybrid mutation is the difference in the spectrum of the eigenvalues induced by the mutation. We extracted up to a noticeable gap in the above-mentioned difference the top of these pairs, which consisted remarkably of the same residues in different combinations. The results are shown in Figure 5.

We clearly see that the mutation stiffens the base while at the same time loosening its grip on the "anchors" of the flaps through which the loosened stiffness is then transferred to the flaps through the connections of contacts between residues in the anchor and the flaps. Note that the stiffening of the base occurs because of interaction between the two chains in the homodimer, while the weakening occurs in each monomer alone. Note also that this finding explains why our choice of $R_c = 20$ Å is sound here: the most important residues, D30 and R87, are farther apart than the generic distance of $R_c = 13$ Å.

In patients, the V82F−I84V mutant shows some additional mutations that seem, on the basis of our investigation, not necessary for the adaptation/evolution of HIV toward protease inhibitor resistance. They might, however, be necessary for protein stability or folding dynamics, issues that are beyond the scope of the present study.

5. Conclusions

In this study, we extended the widely used ANM method to allow for amino acid specificity. We justified the extended model, discussed its parametrization, and showed its superior performance on a set of proteins in comparison to the ANM. After fitting appropriate couplings for the bonded interactions, we applied the method to a problem most relevant to pharmacology and molecular biology: the drug resistance of a V82F−I84V mutant of HIV-1 protease. We showed the effect of this mutation on the movements of the "flaps" of this dimer. Time-consuming molecular dynamics were employed to reveal this mechanism in other studies,[37] while our approach for the protease took merely an hour of CPU time, including the refinement of *K*. We hope that our work facilitates further investigations into the molecular biophysics of HIV and other pharmacological problems.

Acknowledgment. K.H. is supported through a Liebig Fellowship of the Fonds der Chemischen Industrie. Other support has been provided by NSF, NIH, HHMI, CTBP, NBCR, W.M. Keck Foundation, and Accelrys, Inc. Stimulating discussions with A. Perryman and T. Hwa are gratefully acknowledged. Figures 3 and 4 were generated using VMD.[40] We are grateful to anonymous reviewers' comments and suggestions, which helped us to improve the presentation of our results and steered us to elucidate more on the mechanism of the mutation in the HIV-1 protease mutant.

References

(1) Lazaridis, T. *Curr. Org. Chem.* **2002**, *6*, 1319.

(2) Micheletti, C.; Cecconi, F.; Flammini, A.; Maritan, A. *Protein Sci.* **2002**, *11*, 1878.

(3) Keskin, O.; Durell, S.; Bahar, I.; Jernigan, R.; Covell, D. *Biophys. J.* **2002**, *83*, 663.

(4) Brooks, B.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 6571.

(5) Brooks, B.; Janezic, D.; Karplus, M. *J. Comput. Chem.* **1995**, *16*, 1522.

(6) Kitao, A.; Go, N. *Curr. Opin. Struct. Biol.* **1999**, *9*, 164.

(7) Novotny, J.; Bruccoleri, R.; Karplus, M. *J. Mol. Biol.* **1984**, *177*, 787.

(8) Tirion, M. *Phys. Rev. Lett.* **1996**, *77*, 1905.

(9) Bahar, I.; Atilgan, A. R.; Erman, B. *Folding Des.* **1997**, *2*, 173.

(10) Atilgan, A.; Durrell, S.; Jernigan, R.; Demirel, M.; Keskin, O.; Bahar, I. *Biophys. J.* **2001**, *80*, 505.

(11) Doruker, P.; Atilgan, A. R.; Bahar, I. *Proteins: Struct., Funct., Genet.* **2000**, *40*, 512.

(12) Bahar, I.; Jernigan, R. L. *J. Mol. Biol.* **1998**, *281*, 871.

(13) Bahar, I.; Jernigan, R. *Biochemistry* **1999**, *38*, 3478.

(14) Tama, F.; Brooks, I. *J. Mol. Biol.* **2002**, *318*, 733.

(15) Tama, F.; Gadea, F.; Marques, O. *Proteins: Struct., Funct., Genet.* **2000**, *41*, 1.

(16) Doruker, P.; Jernigan, R.; Bahar, I. *J. Comput. Chem.* **2002**, *23*, 119.

(17) Micheletti, C.; Carloni, P.; Maritan, A. *Proteins* **2004**, *55*, 635.

(18) Miyashita, O.; Onuchic, J. N.; Wolynes, P. G. *PNAS* **2003**, *100*, 12570.

(19) This approximation is similar to Micheletti's self-consistent pair contact probability approximation.[41]

(20) Details to be published elsewhere.

(21) Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical Recipies in C*; Cambridge University Press: Cambridge, England, 1995.

(22) Miyazawa, S.; Jernigan, R. L. *J. Mol. Biol.* **1996**, *256*, 623.

(23) Keskin, O.; Bahar, I.; Badretdinov, A.; Ptitsyn, O.; Jernigan, R. *Protein Sci.* **1998**, *7*, 2578.

(24) Bryngelson, J.; Wolynes, P. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *16*, 7524.

(25) Bryngelson, J.; Onuchic, J.; Socci, N.; Wolynes, P. *Proteins: Struct., Funct., Genet.* **1995**, *21*, 167.

(26) Shen, T.; Canino, L.; McCammon, J. *Phys. Rev. Lett.* **2002**, *89*.

(27) Miyazawa, S.; Jernigan, J. *Macromolecules* **1985**, *18*, 534. Miyazawa, S.; Jernigan, J. *Protein Eng.* **1994**, *7*, 1209. Gutin, A. M.; Badretdinov, A. Y.; Finkelstein, A. V. *Mol. Biol. (USSR)* **1992**, *26*, 94. Sali, A.; Shakhnovich, E.; Karplus, M. *J. Mol. Biol.* **1994**, *235*, 1614.

(28) We have chosen subsets of the well-known WHAT-IF and PDBSelect lists, namely, these proteins: 1a55, 1a7g (2 models), 1aba, 1abe, 1amp, 1aqb, 1aru, 1byb, 1c53, 1cd2, 1chd, 1cnv, 1coo, 1cot, 1ctf, 1ctt, 1d3z (10 models), 1dhr, 1eca, 1ede, 1enh, 1es6, 1esc, 1ezm, 1fkj, 1flp, 1fxd, 1gca, 1gof, 1hfc, 1hyp, 1iae, 1lid, 1mba, 1mjc, 1mml, 1npk, 1nxb, 1oyc, 1paz, 1pea, 1php, 1pht, 1piq, 1poa, 1ptq, 1r69, 1rcb, 1ris, 1sbp, 1tca, 1tde, 1ten, 1tig, 1ubq, 1xnb, 2cmd, 2cpl, 2ctc, 2ebn, 2end, 2i1b, 2liv, 2mcm, 2mnr, 2rn2, 2sas, 2sn3, 3dfr, 3il8, 3lzm, 451c, 4fgf, and 5p21. The sizes of these homomers range from 32 to 639 residues. We set $R_c = 13$ Å, as was determined to be optimal in general ANM applications.[10]

(29) Trylska, J.; Tozzini, V.; McCammon, J. Exploring Global Motions and Correlations in the Ribosome. *Biophys. J.* **2005**, *89*, 1455−1463.

(30) Preprint: qbio.BM/0506031. Ming, D.; Wall, M. E. *Phys. Rev. Lett.* **2005**, *95*, 198103.

(31) Note that by focusing on the linear correlation coefficient and the two ranking measures we avoid the adjustment of the overall magnitude completely.

(32) Reiling, K.; Endres, N.; Dauber, D.; Craik, C.; Stroud, R. *Biochemistry* **2002**, *41*, 4582.

(33) Kurt, N.; Scott, W.; Schiffer, C.; Haliloglu, T. *Proteins: Struct., Funct., Genet.* **2003**, *51*, 409.

(34) Little, S. J.; Holte, S.; Routy, J. P.; Daar, E. S.; Makrowitz, M.; Collier, A. C.; Koup, R. A.; Mellors, J. W.; Connick, E.; Conway, B.; Kilby, M.; Want, L.; Whitcomb, J. M.; Hellmann, N. S.; Richman, D. D. *N. Engl. J. Med.* **2002**, *347*, 385.

(35) Zoete, V.; Michielin, O.; Karplus, M. *J. Mol. Biol.* **2002**, *315*, 21.

(36) Papoian, G.; Ulander, J.; Wolynes, P. *J. Am. Chem. Soc.* **2003**, *125*, 9170.

(37) Perryman, A.; Lin, J.-H.; McCammon, J. *Protein Sci.* **2004**, *13*, 1108.

(38) Elber, R.; Karplus, M. *Phys. Rev. Lett.* **1986**, *56*, 394.

(39) A movie of this mode is available from the following URL: http://McCammon.ucsd.edu/khamache/EANM.html.

(40) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33.

(41) Micheletti, C.; Banavar, J. R.; Maritan, A. *Phys. Rev. Lett.* **2001**, *87*, 088102.

CT050247S

JCTC Journal of Chemical Theory and Computation

The Stabilities of N−Cl Bonds in Biocidal Materials

Akin Akdag, Serife Okur, Michael L. McKee,* and S. D. Worley*

Department of Chemistry and Biochemistry, Auburn University, Auburn, Alabama 36849

Received January 5, 2006

Abstract: N-halamine chemistry has been a research topic of considerable importance in these laboratories for over two decades. N-halamine compounds are useful in preparing biocidal materials. There are three N−Cl moieties available in cyclic N-halamine compounds: imide, amide, and amine. The stabilities toward the release of free halogen have been experimentally shown to decrease in the order amine > amide > imide. In this work, this generalization has been tested theoretically at the level of B3LYP/6-31+G(d) and using the conductor-like polarizable continuum aqueous solvation model with UAKS cavities. Excellent accord was observed between theory and experiment. It was also found that the imide and amide N-halamine stabilities on hydantoin rings could be reversed with substitution patterns at the 5 position.

Introduction

Compounds containing nitrogen−halogen (N-halamine) bonds have been shown to be excellent biocides.[1] N-halamines act as sources of free halogens, which are potent oxidizing agents.[2] The oxidations in a cellular system lead to the destruction of microorganisms in a direct contact mechanism.[1b]

Numerous N-halamine compounds have been synthesized in these laboratories over the past 25 years.[3] These compounds have been incorporated into macromolecules (e.g., polymers) and surfaces (e.g., textiles).[4] The goal of the work has been to develop stable biocidal compounds which do not lose their oxidative chlorine contents readily in the presence of water and organics.[3b]

In principle, there are three types of N-halamine structures possible: imide, amide, and amine. The stabilities of these to hydrolysis and reactions with organic receptors vary.[5] The reaction of a general N-halamine with water to give the amine, amide, or imide and hypochlorous acid is given in eq 1.

$$R_2N{-}Cl + H_2O \rightarrow R_2NH + HOCl \qquad (1)$$

The dissociation constants for different N-halamine functionalities are given in Table 1.[6] The table shows that the stabilities toward dissociation of the N−Cl moieties are in the order amine > amide > imide halamine.

Table 1. N-Halamine Dissociation Constants in Aqueous Solution

nature of N-halamine	dissociation constant
imide	$<10^{-4}$
amide	$<10^{-9}$
amine	$<10^{-12}$

As previously mentioned, Worley and co-workers have demonstrated that many examples of N-halamine compounds are efficient biocides. The kinetics of inactivation of the bacteria depend on the nature of the N−Cl moiety. The higher the dissociation constant, the more rapid is the biocidal action. For example, an N-halamine containing an imide N−Cl functionality kills bacteria more rapidly than the one which has an amide N−Cl functionality.[7] However, the stability (i.e., the persistence) of the N-halamine is important because the presence of the N-halamine is necessary for preventing the re-establishment of the microbes.

Some examples of the potential precursor N-halamine compounds synthesized and studied in these laboratories are 5,5-dimethylhydantoin (DMH) derivatives,[8] 2,2,5,5-tetramethylimidazolidinone (TMIO),[9] and 7,7,9,9-tetramethyl-1,3,8-triazaspiro[4.5]decane-2,4-dione (TTDD).[10] As seen from the structures in Figure 1, DMH has an imide and an

* Corresponding authors. Phone: (334) 844-6953; e-mail: mckee@chem.auburn.edu (M.L.M.). Phone: (334) 844-6944; e-mail: worlesd@auburn.edu (S.D.W.).

10.1021/ct060007s CCC: $33.50
Published on Web 03/16/2006

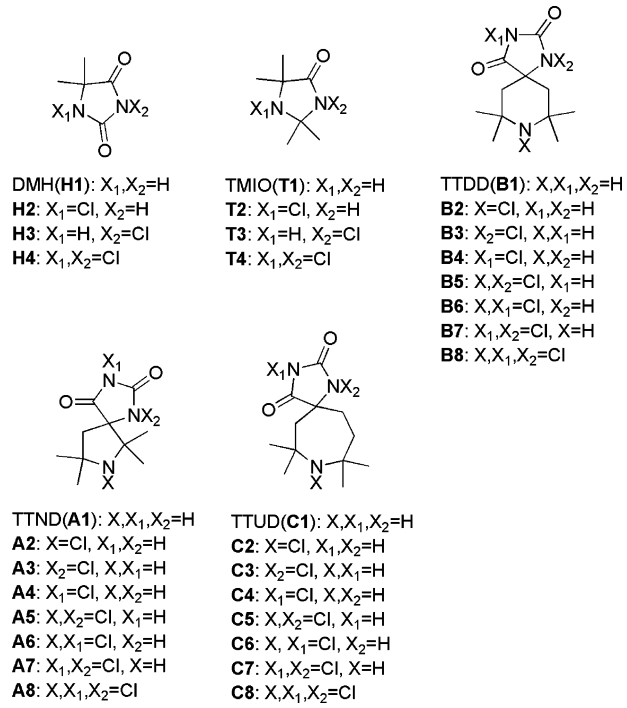


Figure 1. Structures of the compounds under study.

amide functionality, TMIO has an amide and an amine functionality, and TTDD has an imide, an amide, and an amine functionality. All of the above compounds, when chlorinated, gave N-halamine compounds which acted as excellent biocides.

In this study, we seek to determine natures of the effects which are playing an important role in N-halamine stabilities. The chlorinated forms of DMH and TMIO were first subjected to the high-level calculations in order to see if the theoretical results were in accord with the experimental ones. Then, TTDD (**B**) and its five-membered (**A**) and seven-membered (**C**) ring analogues were subjected to the high-level calculations to examine the effect of ring size on chlorine stability. This study may lead to new classes of more effective N-halamine compounds.

Computational Details

All calculations were performed with Gaussian 03.[11] The structures were optimized at the B3LYP/3-21G level initially, and then, optimized structures were subjected to reoptimization at the B3LYP/6-31+G(d) level. The zero-point and thermal corrections were calculated at the B3LYP/6-31+G(d) level. Natural bond order (NBO) analyses were performed on the geometry obtained at the B3LYP/6-31+G(d) level for **H4** and **T4**. Solvation effects were included in the geometry obtained at the B3LYP/6-31+G(d) level with the conductor-like polarizable continuum model (CPCM) using UAKS cavities.[12] The free energy of the reaction was computed using eq 2:

$$\Delta G(\text{solution phase}) = \Delta G(\text{gas phase}) - \Delta G(\text{solvation of products}) + \Delta G(\text{solvation of reactants}) \quad (2)$$

A 1.9 kcal/mol correction was included in the calculation due to the fact that the molecules are changing in state from an ideal gas (1 mol/22.4 L at 298 K) to an ideal solution (1 mol/L).[13] A correction factor of 2.4 kcal/mol was also applied due to the fact that the water molarity is 55.56.[13] All of the results are tabulated in Table 2. In ordered to check the reliability of the DFT relative energies, calculations were also carried out at the MP4SDQT/6-31+G(d)//B3LYP/6-31+G(d) level for **H1**−**H4**. Free energies of reactions using MP4SDQT/6-31+G(d) energies are reported in Table 3 for reactions involving **H1**−**H4** with HOCl and H_2O.

Table 2. Total Energies, Enthalpies, and Free Energies for **H**, **T**, **A**, **B**, and **C** Series[a]

compound	*H*(298 K) (Hartree)	*G*(g) (Hartree)	Δ*G*(sol) (kcal/mol)
H1	−455.223 389	−455.264 681	−16.75
H2	−914.776 951	−914.823 284	−9.88
H3	−914.775 843	−914.820 889	−11.04
H4	−1374.328 763	−1374.378 716	−2.30
T1	−459.720 398	−459.767 615	−10.71
T2	−919.282 827	−919.333 251	−7.56
T3	−919.279 877	−919.330 134	−4.94
T4	−1378.838 750	−1378.893 926	−0.33
A1	−705.785 633	−705.841 811	−16.60
A2	−1165.351 218	−1165.409 885	−13.72
A3	−1165.332 984	−1165.392 420	−12.11
A4	−1165.338 243	−1165.397 807	−10.43
A5	−1624.898 869	−1624.961 326	−8.36
A6	−1624.903 655	−1624.966 109	−7.11
A7	−1624.884 749	−1624.948 239	−4.45
A8	−2084.450 284	−2084.516 704	−0.41
A3-i	−1165.332 954	−1165.392 008	−11.97
B1	−745.068 831	−745.125 335	−18.53
B2	−1204.631 788	−1204.691 205	−16.92
B3	−1204.611 948	−1204.673 655	−13.31
B4	−1204.621 193	−1204.681 226	−12.47
B5	−1664.173 875	−1664.238 753	−11.20
B6	−1664.183 867	−1664.247 180	−10.62
B7	−1664.163 577	−1664.228 836	−5.96
B8	−2123.725 767	−2123.792 423	−4.40
B3-i	−1204.618 544	−1204.677 658	−11.22
C1	−784.345 708	−784.406 077	−16.71
C2	−1243.898 694	−1243.962 430	−15.68
C3	−1243.889 753	−1243.953 061	−11.45
C4	−1243.898 666	−1243.962 906	−10.48
C5	−1703.444 025	−1703.510 128	−10.28
C6	−1703.451 343	−1703.519 069	−9.31
C7	−1703.441 668	−1703.508 536	−4.03
C8	−2162.995 817	−2163.065 856	−2.67
C3-i	−1243.891 450	−1243.953 361	−11.13

[a] Total energies are at the B3LYP/6-31+G(d)//B3LYP/6-31+G(d) level; thermal corrections are made at the B3LYP/6-31+G(d) level, and solvation corrections are made at the CPCM/B3LYP/6-31+G(d) level.

Results and Discussion

5,5-Dimethylhydantoin (DMH). 5,5-Dimethylhydantoin contains two nitrogen atoms, which can be halogenated to give two different N-halamine moieties. Experimentally, it was shown that an imide halamine is more labile than an amide one.[14] Theoretically, the same pattern was observed (see Table 2). The calculations showed that the bond length for the imide N-halamine is shorter than that for the amide

Table 3. Free Energies (kcal/mol) in the Gas Phase and in Solution for Selected Reactions at the B3LYP/6-31+G(d) Level

		B3LYP/6-31+G(d)//B3LYP/6-31+G(d)[a]		MP4SDQT/6-31+G(d)//B3LYP/6-31+G(d)[a]	
eq	reaction	ΔG(gas)	ΔG(aq)	ΔG(gas)	ΔG(aq)
3	**H1** + HOCl → **H2** + H_2O	−9.71	−2.16	−14.39	−6.84
4	**H1** + HOCl → **H3** + H_2O	−8.21	−1.82	−11.67	−5.28
5	**H4** + H_2O → **H2** + HOCl	7.72	−0.54	11.56	3.30
6	**H4** + H_2O → **H3** + HOCl	9.23	−0.19	14.28	4.86
7	**T1** + HOCl → **T2** + H_2O	−14.35	−10.52		
8	**T1** + HOCl → **T3** + H_2O	−12.17	−5.72		
9	**T4** + H_2O → **T2** + HOCl	11.01	3.10		
10	**T4** + H_2O → **T3** + HOCl	12.97	7.68		
11	**A1** + HOCl → **A3** + H_2O	−4.70	0.47		
12	**A1** + HOCl → **A4** + H_2O	−8.08	−1.23		
13	**A2** + HOCl → **A5** + H_2O	−5.22	0.82		
14	**A2** + HOCl → **A6** + H_2O	−8.22	−0.93		
15	**B1** + HOCl → **B3** + H_2O	−3.26	2.64		
16	**B1** + HOCl → **B4** + H_2O	−8.01	−1.27		
17	**B2** + HOCl → **B5** + H_2O	−2.78	3.62		
18	**B2** + HOCl → **B6** + H_2O	−8.06	−1.08		
19	**C1** + HOCl → **C3** + H_2O	−2.42	3.52		
20	**C1** + HOCl → **C4** + H_2O	−8.60	−1.69		
21	**C2** + HOCl → **C5** + H_2O	−2.87	3.21		
22	**C2** + HOCl → **C6** + H_2O	−8.48	−1.43		

[a] Zero-point energy, heat capacity, and entropy corrections to 298 K were made at the B3LYP/6-31+G(d) level. Solvation effects in water were determined at the CPCM/B3LYP/6-31+G(d) level with UAKS cavities at B3LYP/6-31+G(d) optimized geometries.

N-halamine (1.70 and 1.72 Å, respectively) in both isolated systems (**H2** and **H3**) and the combined system (**H4**). The experimental value for the N−Cl bond length in N-chlorosuccinimide is 1.69 Å, and an amide N-halamine N−Cl bond length is 1.71 Å in N-chloroazasteroids,[15] in excellent accord with the computations. Although the N−Cl bond length for the imide is, in fact, shorter than that for the amide, the amide N−Cl bond is clearly stronger, as evidenced by both experimental and computational data. This could be due to the increased polarity of the imide N−Cl bond over the amide one, leading to more-ionic character in the former and more-rapid dissociation by hydrolysis.

When **H1** is subjected to a reaction with hypochlorous acid in water, the formation of **H2** is more spontaneous and exothermic than is the formation of **H3**. To test the idea that the imide N-halamine chlorine is first to dissociate, the reaction of **H4** with water was treated computationally. It was found that forming the amide N-halamine **H2** is more favorable than forming imide N-halamine **H3** (eqs 3−6). The order of N-halamine stability in the gas phase and in solution (water as solvent) did not change. However, the Gibbs free energy differences between reaction 3 and reaction 4 decreased by 1.16 kcal/mol when solvation was included. This decrease arises from the solvation differences of **H1** and **H2**. Thus, the computational methods were in accord with experimental observations for the hydantoin derivatives in that the formation of **H2** was predicted to occur more spontaneously than that of **H3**.

2,2,5,5-Tetramethylimidazolidin-4-one (TMIO). TMIO has a hindered secondary amine and an amide functional group, and when chlorinated, both functional groups have been shown experimentally to bind chlorine more tightly than does DMH.[16] It was demonstrated that the amine N-halamine group bound chlorine more strongly than did the amide N-halamine group (Table 2). Moreover, experiments showed that the monohalogenation of the title compound produced the amide N-halamine as the kinetically controlled product, which rearranges to the thermodynamically controlled product of the amine N-halamine over time.[17] Surprisingly, the N−Cl bond lengths calculated for the amide and the amine were 1.73 and 1.77 Å, respectively.

As seen in Table 2, the amide N-chloramine **T3** was predicted to be higher in energy than the amine **T2**. The reactions examined in eqs 7−10 show that the reaction forming **T2** is more favored than the one forming **T3** thermodynamically and that the order did not change with the addition of solvation.

The free energy of solvation for **T2** and **T3** shows a 2.62 kcal/mol difference. This difference is due to the fact that the amide is more polarized than the amine.

The Spiro Systems (A, B, C). 6,6,8,8-Tetramethyl-1,3,7-triazaspiro[4.4]nonane-2,4-dione. On the basis of the experimental and the theoretical results for DMH and TMIO, similar results were expected for the order of stabilities for N-halamines of the **A** structure. However, Table 2 shows that the order is amine > imide > amide N-halamine in terms of stability. The reversal between amide and imide has been a focal point of the analysis. As seen in Figure 2, steric interaction between methyl groups and chlorine seems to affect the stability of an amide N-halamine group. The distances between the chlorine and hydrogens on the methyl group are the reason for this steric hindrance. It was expected that changing **A3** to **A3-i** (the conformation has been changed, see Figure 2) would alter the N-halamine stability order. But, it was observed that the stability order did not change. Furthermore, inverting the pyrolidine ring with respect to the hydantoin ring caused a negligible increase in energy, enthalpy, and free energy (**A3** → **A3-i**).

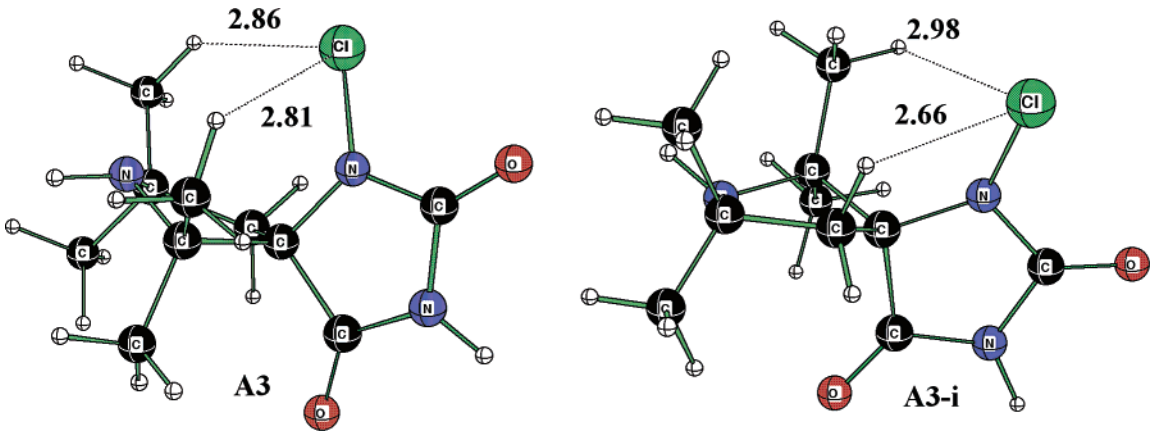


Figure 2. Minimized structures at the B3LYP/6-31+G(d) level of **A3** and **A3-i**.

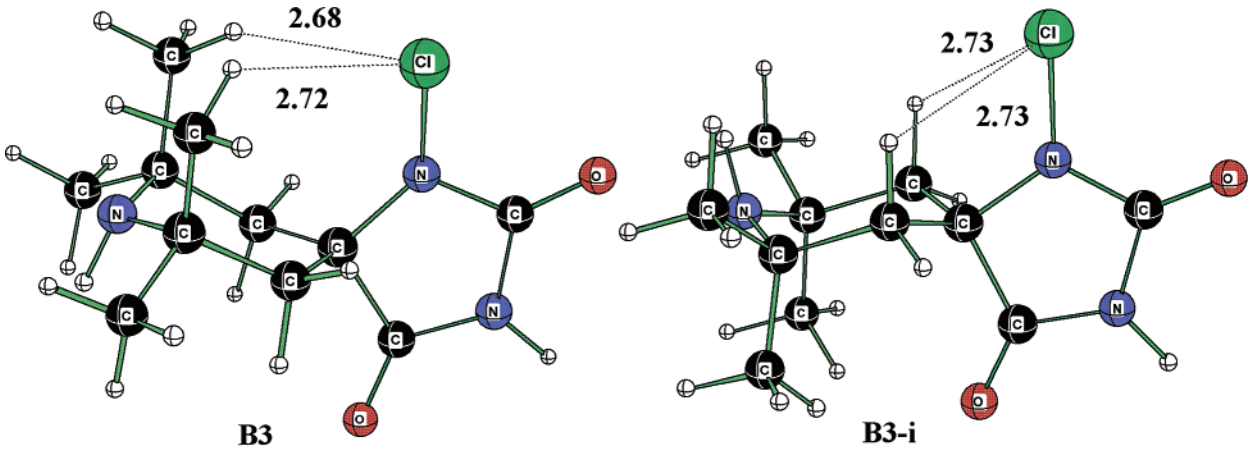


Figure 3. Minimized structures at the B3LYP/6-31+G(d) level of **B3** and **B3-i**.

The N−Cl bond lengths are similar to those for DMH and TMIO. The amine has the longest bond length at 1.77 Å, then the amide at 1.71 Å, and then the imide at 1.70 Å. Thus, as discussed earlier, the N−Cl bond length becomes shorter as the bond polarity increases.

The energy difference between **A3** and **A4** is 3.34 kcal/mol and that for **A5** and **A6** is 2.94 kcal/mol. The chlorination reaction thermodynamics showed the same pattern in gas and aqueous solution (eqs 11−14).

7,7,9,9-Tetramethyl-1,3,8-triaza-spiro[4.5]decane-2,4-dione. From an energetic point of view, the order of N-halamine stability is the same as that for the **A** series: amine > imide > amide. This order was largely affected by the steric interactions between chlorine connected to the amide nitrogen and the methyl groups (see Figure 3). The distance between chlorine and the methyl groups for the **B** series is even smaller than that for the **A** and is within the range of 2.62−2.72 Å. These shorter distances can be rationalized from 1 to 3 diaxial interactions (six-membered ring in the chair conformation).

The **B3** was inverted to **B3-i** (see Figure 3); that is, the amide functionality was brought to the equatorial position. It was seen that 1−3 diaxial interaction between methyl groups and chlorine has diminished because of the chair conformation of the piperidine moiety. Although the amide N-halamine stability has been affected by steric interaction with the methyl group, the energy was not low enough to be more stable than the imide N-halamine. Because the amide N-halamine chlorine is above the ring, inevitable steric interactions with hydrogens on the axial position occur.

The analysis of data for the **B** series (see Table 2) showed the same pattern as that for the **A** series. **B4** is thermodynamically more stable than **B3**, and similarly, **B6** is more stable than **B5**. Furthermore, the reactions of **B1** and **B2** with hypochlorous acid showed that the imide N-halamine is formed more spontaneously than is the amide N-halamine, both in the gas phase and in water (eqs 15−18).

7,7,9,9-Tetramethyl-1,3,8-triaza-spiro[4.6]undecane-2,4-dione. The same pattern was observed for the **C** series as in the cases of the **A** and **B** series (see Table 2). The reason for the amide N-halamine being less stable than the imide N-halamine appears to be due to steric hindrance in the case of the **C** series overall. The distance between the amine chlorine and one methyl group was observed to be within the range of 2.67−2.69 Å. Even though the methyl group for one side (methyl on the 7 position) is far enough away to not have significant steric interactions, it was seen that the chlorine is too close to a hydrogen on the ring (distance is changing from 2.70 to 2.73 Å), as in Figure 4. Therefore, the same procedure was followed as above; the ring was inverted. The inversion process caused the energy to become lower, but not enough to be lower than that for **C4**.

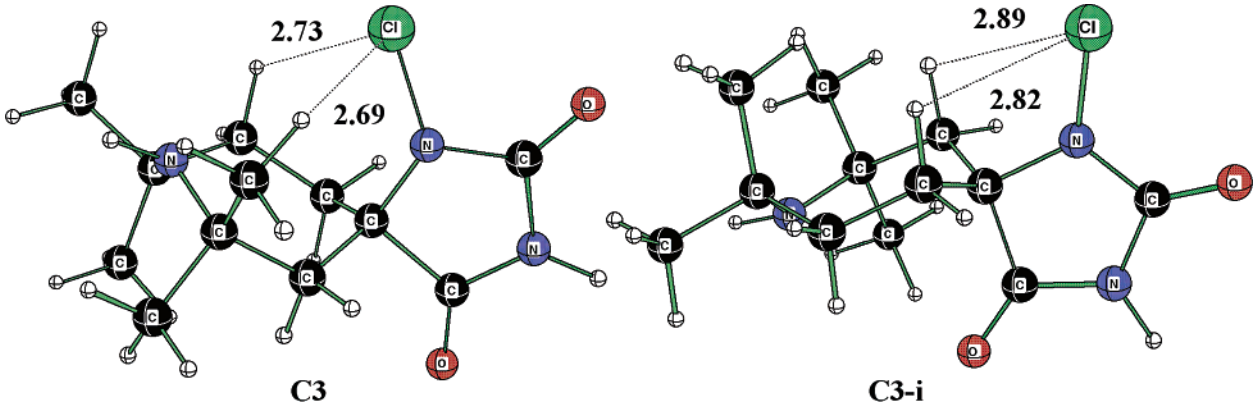


Figure 4. Minimized structures at the B3LYP/6-31+G(d) level of **C3** and **C3-i**.

It was expected that, by relaxing the ring (larger ring systems), the steric effect would decrease. Thus, the amide N-halamine could become more stable than the imide N-halamine. To test this hypothesis, a seven-membered ring has been attached to the 5 position of the hydantoin. Despite the initial expectations, it was shown that imide N-halamine is still more stable than the amide N-halamine (Table 2). The reactions for the **C** series showed the same order as that in cases of the **A** and **B** series in terms of spontaneity, both in the gas and solution phases (eqs 19−22).

As seen for the cases of **A**, **B**, and **C**, the N-halamine stability is not only determined by the electronic effects but is also determined by the substituents at the 5 position of the hydantoin ring. The effect of the substituents seems to be steric, with electronic effects being less important. The conformation of the cyclic systems at the 5 position affects the stability of the amide N-halamine from a steric point of view.

As alluded to earlier, the N−Cl bond lengths for the N-halamines showed an interesting pattern: amine(N−Cl) > amide(N−Cl) > imide(N−Cl). The NBO analysis for **H4** and **T4** showed that the hybridization of the nitrogen atom remains the same in all nitrogen cases. However, the natural population charges differ. For **H4**, the imide nitrogen has a charge of −0.54 (the Cl bonded to this nitrogen has a charge of 0.23) and the amide nitrogen has a charge of −0.52 (the Cl has a charge of 0.18). For **T4**, the amide nitrogen has a charge of −0.51 (for Cl, it is 0.18) and the amine nitrogen has a charge of −0.50 (Cl has 0.08). This polarization agrees with a literature report.[18] Moreover, the Wiberg bond orders are consistent with the bond lengths. As a result, the imide bond is more polarized than the amide, which is more polarized than the amine. The higher bond polarization causes the shorter bond length because of Coulombic attraction.

Conclusions

In this study, N-halamines were investigated theoretically. It was shown that the experimental stability order could be reproduced with high-level calculations both in the gas phase and in solution. However, the polar N−Cl bonds have the higher solvation energy.

The calculated N−Cl bond length was surprising in that the strongest bond as evidenced by ease of hydrolytic dissociations, amine(N−Cl) > amide(N−Cl) > imide(N−Cl), was the longest bond. It could be concluded that the greater the ionic character of the bond, the shorter was the bond length.

The steric effect of cyclic moieties at the 5 position of the hydantoin ring causes the amide N-halamine moiety to be more labile than the imide N-halamine one for the **A**, **B**, and **C** series. Varying the ring sizes does not affect the amine N-halamine stability in the ranking.

Acknowledgment. This work has been supported by the U.S. Air Force and the Vanson HaloSource Company. The computer time was provided by Auburn University and the Alabama supercomputer. S.O. thanks Ahmet Okur for helpful discussions.

Note Added after ASAP Publication. This article was released ASAP on March 16, 2006, with the incorrect Received Date. The correct version was posted on April 12, 2006.

Supporting Information Available: The optimized gas-phase structures and Cartesian coordinates of **H**, **T**, **A**, **B**, and **C** series at the B3LYP/6-31+G(d) level are given. The absolute energies at the B3LYP/3-21G and B3LYP/6-31+(G) levels, zero-point energies, enthalpy corrections, entropies, and solvation energies were tabulated for **H**, **T**, **A**, **B**, **C**, water, and hypochlorous acid. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) Worley, S. D.; Sun, G. *Trends Polym. Sci.* **1996**, *4*, 364−370. (b) Worley, S. D.; Williams, D. E. *Crit. Rev. Environ. Control* **1988**, *18*, 133−175.

(2) The term “free halogen” refers to a halogen with an oxidation state of +1. In water, this forms hypohalous acid or the hypohalite anion depending upon pH.

(3) For example: (a) Tsao, T.-C.; Williams, D. E.; Worley, C. G.; Worley, S. D. *Biotechnol. Prog.* **1991**, *7*, 60−66. (b) Worley, S. D.; Williams, D. E.; Barnela, S. B. *Water Res.* **1987**, *21*, 983−988.

(4) For example: (a) Chen, Y.; Worley, S. D.; Kim, J.; Wei, C.-I.; Chen, T.-Y.; Suess, J.; Kawai, H.; Williams, J. F. *Ind. Eng. Chem. Res.* **2003**, *42*, 5715−5720. (b) Worley, S. D.; Chen, Y.; Wang, J.-W.; Wu, R.; Cho, U.; Broughton, R. M.; Kim, J.; Wei, C.-I.; Williams, J. F.; Chen, J.; Li, Y. *Surf. Coat. Int., Part B* **2005**, *88*, 93−99. (c) Sun, Y.; Sun, G. *J. Appl. Polym. Sci.* **2001**, *81*, 617−624.

(5) Chen, Y.; Worley, S. D.; Kim, J.; Wei, C.-I.; Chen, T. Y.; Suess, J.; Kawai, H.; Williams, J. F. *Ind. Eng. Chem. Res*. **2003**, *42*, 5715−5720.

(6) Table 1 was reproduced from the given reference. Qian, L.; Sun, G. *J. Appl. Polym. Sci.* **2003**, *89*, 2418−2425.

(7) (a) Lauten, S. D.; Sarvis, H.; Wheatly, W. B.; Williams, D. E.; Mora, E. C.; Worley, S. D. *Appl. Environ. Microbiol*. **1992**, *58*, 1240−1243. (b) Qian, L.; Sun, G. *J. Appl. Polym. Sci*. **2004**, *91*, 2588−2593.

(8) 1,3-Dichloro-5,5-dimethylhydantoin was incorporated into polystyrene in the given reference. Sun, G.; Wheatley, W. B.; Worley, S. D. *Ind. Eng. Chem. Res*. **1994**, *33*, 168−170.

(9) (a) Elrod, D. B.; Worley, S. D. *Ind. Eng. Chem. Res*. **1999**, *38*, 4144−4149. (b) Tsao, T.-C.; Williams, D. E.; Worley, S. D. *Ind. Eng. Chem. Res.* **1990**, *29*, 2161−2163.

(10) (a) Mailey, E. A.; Day, A. R. *J. Org. Chem*. **1957**, *22*, 1061−1065. (b) Worley, S. D.; Chen, Y.; Liang, J.; Wu, R.; Barnes, K.; Broughton, R. M.; Cho, U.; Lee, J. U.S. Patent App. US 2005186173; *Chem. Abs.* 143:99011.

(11) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.

(12) Takano, Y.; Houk, K. N. *J. Chem. Theory Comput*. **2005**, *1*, 70−77.

(13) McKee, M. L. *J. Phys. Chem. A* **2003**, *107*, 6819−6827.

(14) Corral, R.; Orazi, O. O. *J. Org. Chem*. **1963**, *28*, 1100−1104.

(15) (a) For N−Cl bond in *N*-chlorosuccinimide, see: Brown, R. N. *Acta Crystallogr*. **1961**, *14*, 711−715. (b) For N−Cl bond in amide, see: Roszak, A. W.; Brunner, K.; Back, T. G.; Codding, P. W. *Acta Crystallogr., Sect. B* **1991**, *47*, 383−389.

(16) Qian, L.; Sun, G. *Ind. Eng. Chem. Res*. **2005**, *44*, 852−856.

(17) (a) Naquib, I.; Tsao, T.-C.; Sarathy, P. K.; Worley, S. D. *Ind. Eng. Chem. Res.* **1991**, *30*, 1669−1671. (b) Naquib, I.; Tsao, T.-C.; Sarathy, P. K.; Worley, S. D. *Ind. Eng. Chem. Res*. **1992**, *31*, 2046−2050.

(18) Poleshchuk, O. K.; Makiej, K.; Ostafin, M.; Nogaj, B. *Magn. Reson. Chem.* **2001**, *39*, 329−333.

CT060007S

JCTC Journal of Chemical Theory and Computation

Small Carbon Clusters Doped with Vanadium Metal: A Density Functional Study of VC_n ($n = 1-8$)

Pilar Redondo,* Carmen Barrientos, and Antonio Largo

Departamento de Química Física, Facultad de Ciencias, Universidad de Valladolid, 47005 Valladolid, Spain

Received January 31, 2006

Abstract: A theoretical study of the different isomers of neutral VC_n ($n = 1-8$) clusters has been carried out. Predictions for their electronic energies, rotational constants, dipole moments, and vibrational frequencies have been made using the B3LYP method with different basis sets. For linear and cyclic VC_n clusters, the lowest-lying states correspond to quartet states, with the only exception being VC, which has a $^2\Delta$ ground state. In the fan-type structures, the electronic ground state is found to be a quartet state for even *n* values, whereas for odd *n* values, the ground state is a doublet for VC_3 and VC_5 and a quartet for VC_7. From the incremental binding energies, we can deduce an even−odd parity effect, with *n*-even clusters being more stable than *n*-odd ones in the linear and fan clusters. It was also found that neutral VC_n clusters with $n \leq 6$ prefer fan structures over linear and cyclic isomers, whereas cyclic ground states are predicted for the clusters with $n > 6$.

Introduction

The inclusion of metal atoms in pure carbon clusters provides a way to modify their properties and generate new materials. A wide variety of materials can be formed from the interaction of carbon with transition metals. *Networked* metallofullerenes[1] can be originated when transition metals are incorporated into carbon cages, whereas some rare-earth elements can be trapped inside fullerene cages to form *endohedral* metallofullerenes.[2] On the other hand, late transition metals have interesting properties as catalysts for carbon nanotube formation.[3] Finally, early transition metals have been found to form stable gas-phase metal−carbon clusters, known as metallocarbohedrenes, "*met-car*",[4−7] with a M_8C_{12} stoichiometry; observed examples are Ti_8C_{12}[8] and V_8C_{12}.[9]

Because of the presence of transition-metal atoms, the interpretation of experimental information is rather difficult. In particular, the prediction of the ground-state structures is complicated because of the large number of local minima on their potential energy surfaces, related to the presence of *nd* shells in the transition-metal atom. However, the determination of geometry is one of the fundamental problems in clusters research because it may affect chemical properties, and therefore, accurate theoretical information on the structures of transition metal carbides is desirable. In addition, theoretical studies of small carbon clusters containing transition metals provide a way to obtain a detailed description of the metal−carbon interactions and are very useful to understand the growth mechanism of the various metal−carbon nanomaterials. In fact, the type of structure formed depends essentially on the nature of these interactions. Obviously, the study of the structures and reactivity of metal carbides are also important in other areas, such as surface chemistry, combustion chemistry, or astrochemistry.

The competition between different isomers on small first-row transition-metal carbides has recently been the subject of some theoretical studies. In the case of dicarbide compounds (MC_2), linear and cyclic structures have been characterized on ScC_2,[10−12] VC_2,[13] FeC_2,[14,15] and CoC_2.[16] In all cases, cyclic ground states are predicted. Whereas systematic studies as a function of the size of the clusters on heteroatom-doped carbon clusters containing first- and second-row elements have been extensively carried out by theoretical methods, only a few works are devoted to the study of first-row transition-metal-doped carbon clusters. In

* Corresponding author fax: 34-983-423013; e-mail: predondo@qf.uva.es.

10.1021/ct060033q CCC: $33.50
Published on Web 03/31/2006

the past few years, different theoretical approaches have been applied to the study of systems such as CrC_n ($n = 1-8$),[17] TiC_n ($n = 2-4$),[18,19] FeC_n ($n = 1-4$),[20] NiC_n,[21−23] and ScC_n ($n = 1-8$).[24,25] In all cases, it is shown that different geometrical configurations can be reached for this type of system.

Vanadium is one of the transition metals that has been shown to form stable "met-car".[9] However, to the best of our knowledge, theoretical studies of vanadium carbide compounds have only been performed, at different levels of theory, for VC,[26−29] VC_2,[13,30] and V_2C_2[30] systems. In addition, there are some experimental studies reporting the photoelectron spectra of first-row transition-metal carbides (MC_2 and MC_3).[31,32] In the present work, we have carried out a theoretical study of the VC_n (with n ranging from 1 to 8) systems taking into account different geometrical conformations. We analyze the relative stability of the different structures as a function of n, and for the most stable isomers, we report their equilibrium structures and some spectroscopic data that could be helpful for their eventual experimental detection. In addition, the knowledge about the behavior of vanadium-doped carbon clusters as a function of the size of the cluster will allow the identification of possible systematic trends, which could help to understand the structure of these systems and could be useful to make extrapolation for some properties and, therefore, predictions for larger clusters.

Computational Methods

We have employed the same theoretical approach as that in our previous studies on ScC_n systems.[24,25] Therefore, all of our calculations have been made using density functional theory (DFT), in particular, employing the B3LYP exchange-correlation functional.[33,34] This consists of the Lee−Yang−Parr[35] correlation functional in conjunction with a hybrid exchange functional first proposed by Becke.[36] The latter is a linear combination of local density approximation, Becke's gradient correction,[37] and the Hartree−Fock exchange energy based on Kohn−Sham orbitals.[38] As is known, The DFT/B3LYP method has been widely applied to the study of many medium-sized heteroatom-doped carbon clusters, providing structures in good agreement with the experimental results. In addition, previous studies on vanadium carbide systems, such as the works on VC[26,29] and VC_2,[13] have shown that the B3LYP method is in reasonable agreement with multiconfigurational approaches predicting similar ground states and lowest-lying excited states. These facts suggest that the B3LYP method can be applied in the study of these compounds.

Computations have been performed using the 6-311+G(d) basis set that includes diffuse functions and is constructed employing the triple split-valence 6-311G[39] for carbon atoms and the Wachters[40] and Hay[41] basis set with the scaling factor of Ragavachari and Trucks[42] for vanadium atoms. The use of effective core potential methods for the description of clusters reduces significantly the computational time. In particular, we have also employed a mixed basis set formed by Los Alamos ECP plus DZ, LanL2DZ,[43−45] for vanadium atoms in conjunction with the 6-311+G(d) basis set for carbon atoms, denoted as LAN-6+(d). In our previous study on the ScC_n system,[24] we analyzed the behavior of different basis sets, showing that the 6-311+G(d) and LAN-6+(d) basis sets give the most reliable results.

Harmonic vibrational frequencies were computed from analytical gradient techniques. This allows an estimate of the zero-point vibrational energy (ZPVE) correction for each structure, as well as an assessment of the nature of the stationary points and, therefore, a characterization of whether they are true minima on the respective potential surface.

All calculations reported in this work were carried out with the Gaussian 98 program package.[46]

Results and Discussion

We have carried out a study of the neutral VC_n ($n = 1-8$) system, on the doublet, quartet, and sextuplet potential energy surfaces. In this work, we will only present the three most stable conformations on each potential energy surface: a linear structure with the vanadium atom sited at the end position, a cyclic isomer where the vanadium atom is bonded to the two terminal carbon atoms of the chain, and finally a fan-type structure where the vanadium atom interacts with the whole carbon chain. Other isomers, such as a linear isomer with the vanadium atom in an intermediate position or a cyclic isomer with an exocyclic vanadium atom, have been studied, but they are not presented here because they lie higher in energy.

We will present the results for each type of structure separately in order to analyze systematic trends in their properties with the size of the clusters. After this, the relative stability of the three structures along the series was discussed. It should be pointed out that, for some structures reported in this work, for example, those with $^4\Phi$ or $^6\Phi$ electronic states, the B3LYP method, as well as single-reference-based methods such as HF, MP2, QCISD, and so forth, provides nondegenerate π frequencies. On the other hand, for the states of Σ and Δ symmetry studied in this work, it is possible to obtain degenerate frequencies, and they are denoted in Table S1 (Supporting Information) as (2).

VC_n Linear Clusters. We present in Table 1 some properties, such as absolute and relative electronic energies, $\langle S^2 \rangle$ expectation values, dipole moments, or binding energies, at the B3LYP/6-311+G(d) and B3LYP/LAN-6+(d) levels of theory, for the lowest-lying open-chain VC_n species on the doublet, quartet, and sextuplet potential surfaces. Their corresponding harmonic vibrational frequencies and rotational constants are given as Supporting Information in Table S1. For n-even clusters, we have included results for two quartet states ($^4\Sigma$ and $^4\Phi$) because they are very close in energy and, in all cases, are more stable than the corresponding lowest-lying doublet and sextuplet states. These molecular properties might be helpful in an experimental search for these species. In addition, the optimized geometries at the B3LYP/6-311+G(d) and B3LYP/LAN-6+(d) levels of theory for the most stable structures are shown in Figure 1.

In general, it can be seen that the two basis sets considered here lead to very close results for equilibrium geometries and other properties such as dipole moments, vibrational frequencies, or relative energies. Therefore, it can be

Table 1. Electronic Energies, $\langle S^2 \rangle$ Values, Binding Energies (for the Ground States), Dipole Moments, and Relative Energies for Linear VC_n Clusters with the B3LYP/6-311+G(d) (First Line) and B3LYP/LAN-6+(d) (Second Line) Methods

isomer	state	$-E$ (au)	$\langle S^2 \rangle$	BE (eV)	μ (D)	ΔE (kcal mol^{-1})
VC	$^2\Delta$	981.893063	1.1532	3.84	6.02	0.00
		109.224179	1.2432	3.84	6.04	0.00
	$^4\Delta$	981.885398	3.8785		2.84	4.73
		109.215534	3.9237		2.86	5.35
	$^6\Sigma$	981.861752	8.7990		3.93	19.28
		109.191867	8.8047		3.93	19.95
VC_2	$^2\Delta$	1019.960382	0.9400		6.46	30.33
		147.289287	0.9520		6.55	31.43
	$^4\Sigma$	1020.008738	3.9190	10.78	8.72	0.00
		147.339374	3.9339	10.76	8.79	0.00
	$^4\Phi$	1020.002094	3.9137		9.42	4.31
		147.333478	3.9328		9.55	3.77
	$^6\Delta$	1019.983546	8.7947		4.51	16.19
		147.316392	8.7972		4.76	14.82
VC_3	$^2\Delta$	1058.052393	1.5419		13.65	19.16
		185.385069	1.5021		13.42	18.22
	$^4\Phi$	1058.081416	4.4493		8.99	0.64
		185.414436	4.4665		9.10	−0.40
	$^6\Sigma$	1058.082820	8.7844	16.54	8.63	0.00
		185.414098	8.7853	16.54	8.79	0.00
VC_4	$^2\Delta$	1096.141157	1.3277		8.02	28.68
		223.471035	1.3233		8.15	29.57
	$^4\Sigma$	1096.186952	4.0768	23.11	12.22	0.00
		223.518281	4.0900	23.10	12.38	0.00
	$^4\Phi$	1096.183059	4.3260		12.00	2.20
		223.515642	4.3793		11.96	1.44
	$^6\Phi$	1096.167774	8.7654		10.03	11.54
		223.501194	8.7659		10.35	10.31
VC_5	$^2\Delta$	1134.233757	1.7721		18.03	21.39
		261.566333	1.7326		18.13	21.64
	$^4\Phi$	1134.267174	4.5837	29.06	11.70	0.00
		261.600513	4.5971	29.11	11.87	0.00
	$^6\Sigma$	1134.264878	8.8132		11.53	1.82
		261.596343	8.8145		11.76	2.92
VC_6	$^2\Delta$	1172.318571	1.5493		9.30	28.75
		299.648651	1.5519		9.37	30.57
	$^4\Sigma$	1172.363472	4.1751		15.13	0.55
		299.694901	4.1870		15.36	1.48
	$^4\Phi$	1172.363720	4.5184	35.40	14.02	0.00
		299.696815	4.5544	35.45	14.07	0.00
	$^6\Phi$	1172.353181	8.7709		12.43	6.48
		299.686707	8.7714		12.78	6.22
VC_7	$^2\Delta$	1210.413532	1.9346		22.49	22.94
		337.745953	1.8984		22.64	22.39
	$^4\Phi$	1210.449339	4.6735	41.48	14.14	0.00
		337.782777	4.6856	41.54	14.35	0.00
	$^6\Sigma$	1210.444147	8.8435		14.35	3.69
		337.775620	8.8452		14.65	4.54
VC_8	$^2\Delta$	1248.496004	1.7095		10.55	30.66
		375.826217	1.7206		10.54	32.44
	$^4\Sigma$	1248.540325	4.2539		17.96	2.71
		375.871820	4.2641		18.26	3.47
	$^4\Phi$	1248.543919	4.6246	47.76	16.10	0.00
		375.877237	4.6493	47.82	16.23	0.00
	$^6\Phi$	1248.535968	8.7757		14.63	4.89
		375.869529	8.7763		15.00	4.62

concluded that, for linear VC_n clusters, the employment of effective core potentials (in particular, LanL2DZ) for the vanadium atom gives similar results to those obtained with 6-311+G(d).

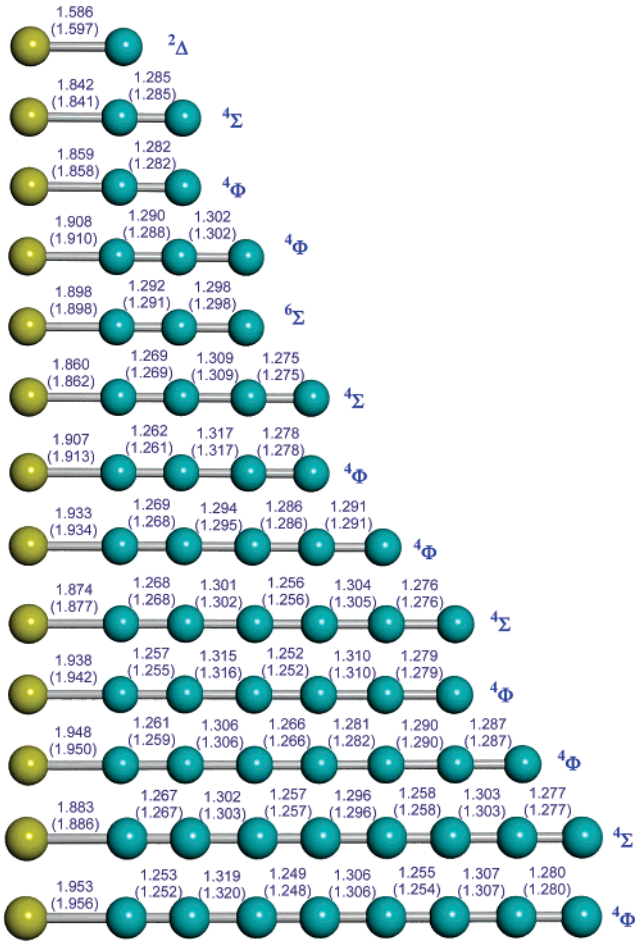


Figure 1. Equilibrium geometries of VC_n linear clusters at the B3LYP/6-311+G(d) and B3LYP/LAN-6+(d) (in parentheses) levels of theory. Distances are given in angstroms.

Before discussing our results, we are going to compare them with the previous results that have been reported for VC and VC_2 isomers. The first member of the series, vanadium carbide VC, has been studied both theoretically[26−29] and experimentally,[47−49] and it is well-established that the lowest-lying state corresponds to a $^2\Delta$ symmetry. The most complete study of this species was carried out recently by Kalemos et al.,[29] where they explored 29 states of VC using MRCI methods with large atomic natural orbital basis sets. From this study, they concluded that the ground state is of $^2\Delta$ symmetry, with the first two excited states, $^4\Delta$ and $^2\Sigma$, located 4.16 and 6.95 kcal/mol, respectively, above the fundamental state. The MRCI(+Q) binding energy of $^2\Delta$ is estimated to be 88.5 (89.3) kcal/mol, and after some corrections (BSSE, ZPVE, relativistic effects, and $3s^23p^6$ semicore), they obtained a value of 95.3 kcal/mol, in good agreement with the experimental one of 100.1 ± 5.7 kcal/mol.[48] As can be seen in Table 1, we also found $^2\Delta$ and $^4\Delta$ as the ground and first excited states, respectively, and the relative energy (4.73 and 5.35 kcal/mol, obtained with 6-311+G(d) and LAN-6+(d), respectively) is very similar to that obtained at the MRCI(+Q) level. We have also characterized a $^2\Sigma$ state lying 15.57 kcal/mol above the ground state. As in the study of Kalemos et al., this state is more stable than the lowest-lying sextuplet state, $^6\Sigma$. In

addition, our results for the binding energy of $^2\Delta$ are very close to those obtained at the MRCI(+Q) level by Kalemos et al.

Theoretical information for the VC_2 system has been reported by Majumdar et al.[13] using DFT and MRSDCI levels with relativistic effective core potentials (RECPs) for vanadium and carbon atoms. In contrast with our results shown in Table 1, they report $^6\Sigma$ and $^4\Sigma$ (lying about 4.61 kcal/mol above the $^6\Sigma$ state at the MRSDCI+Q level) states as the most stable linear isomers. We have found a $^4\Sigma$ state as the lowest-lying state. This isomer has two imaginary π-type frequencies (as can be seen in Table S1) and corresponds to the transition state for the rearrangement of the cyclic isomer ($^4A_1$). In addition, we have located another $^4\Sigma$ state that corresponds to a true minimum and is located above the $^6\Sigma$ isomer. However, we have characterized $^4\Phi$ and $^6\Delta$ states (included in Table 1), both lower in energy than the $^6\Sigma$ isomer (about 23.92 and 12.03 kcal/mol below the $^6\Sigma$ isomer, respectively), reported by Majumdar et al. as the lowest-lying linear one.

It is readily seen in Table 1 that all linear VC_n clusters present quartet ground states, with the only exception being VC, where the scheme is different because the carbon is only bonded to vanadium and presents a $^2\Delta$ ground state. Linear clusters with *n*-odd values have a $^4\Phi$ ground state, and the first excited state corresponds to a $^6\Sigma$ symmetry. In the case of VC_3, the B3LYP results show that both states are virtually isoenergetic (ΔE is less than 1.00 kcal/mol). The quartet−sextuplet energy differences slightly increase along the series for *n*-odd clusters, being 3.69 kcal/mol for the VC_7 isomer, so both states are interesting from an experimental point of view. The stability of the different states can be explained in terms of their electronic configurations. Vanadium-doped carbon clusters, VC_n, present $4n + 5$ valence electrons, and the corresponding electronic configuration for the $^4\Phi$ state is

$$\{\text{core}\}1\sigma^2...1\pi^4....(n+2)\sigma^1\left(\frac{n+1}{2}\right)\pi^3 1\delta^1 \qquad n\text{-odd clusters}$$

The lowest-lying sextuplet state is obtained upon a $[(n + 1)/2]\pi^3 \rightarrow 1\delta$ promotion. Both orbitals are mainly located at the vanadium atom, and the energy difference between the $\sigma^1\pi^3\delta^1$ and $\sigma^1\pi^2\delta^2$ configurations is very small. The lowest-lying doublet state corresponds to a $\pi^4\delta^1$ configuration and is located higher in energy (about 20.00 kcal/mol).

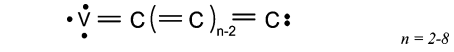
On the other hand, *n*-even linear VC_n clusters present two quartet states that are very close in energy, $^4\Sigma$ and $^4\Phi$, corresponding to the following electronic configurations:

$$\{\text{core}\}1\sigma^2...1\pi^4....(n+2)\sigma^1\left(\frac{n}{2}\right)\pi^4 1\delta^2 \qquad {}^4\Sigma\ n\text{-even clusters}$$

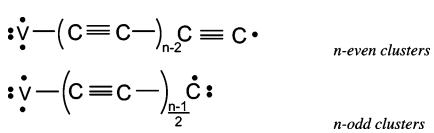
$$\{\text{core}\}1\sigma^2...1\pi^4....(n+2)\sigma^1\left(\frac{n+2}{2}\right)\pi^1 1\delta^1 \qquad {}^4\Phi\ n\text{-even clusters}$$

For the smaller members of the series ($n = 2$, 4), the ground state is found to be $^4\Sigma$, whereas for clusters with $n = 6$ and 8, the $^4\Phi$ state is more favorable. The similar energy of the two quartet states can be related to the fact that the $[(n + 2)/2]\pi$ and the 1δ molecular orbitals are mainly located at the vanadium atom. The lowest-lying sextuplet state, $^6\Phi$, is obtained from the $^4\Phi$ state upon $(n/2)\pi \rightarrow [(n + 2)/2]\pi$ promotion. As can be seen in Table 1, the quartet−sextuplet energy difference decreases as *n* increases for *n*-even linear clusters. As in the case of *n*-odd isomers, the lowest-lying doublet state, $^2\Delta$, corresponding to a $...(n + 2)\sigma^2(n/2)\pi^4 1\delta^1$ electronic configuration, is located higher in energy than quartet and sextuplet ones.

By looking at Figure 1, it can be seen that the C−C bond distances are all in the range 1.24−1.32 Å and, therefore, can be assimilated to moderately strong and typical double bonds, characteristic of cumulenic structures:



On the other hand, a clear alternation in the C−C bond distances can be observed, $C_{odd}-C_{even}$ being shorter than $C_{even}-C_{odd}$. This fact suggests that there is some polyacetylenic character with alternate triple C−C bonds starting in the C_1-C_2 bond:



It can also be observed that $^4\Phi$ states present bigger V−C bond distances (ranging from 1.859 Å for VC_2 to 1.953 Å for VC_8) than those of $^4\Sigma$ states (1.842 Å for VC_2 to 1.883 Å for VC_8). Consequently, the contribution of polyacetylenic-type structures is higher for $^4\Phi$ states than for the $^4\Sigma$ ones, where the cumulenic-type structure is dominant.

Dipole moments are, in all cases, quite high, reflecting a significant charge transfer from the vanadium atom to the C_n unit. Furthermore, the dipole moment increases regularly with the number of carbon atoms for both *n*-odd and *n*-even clusters, being higher for *n*-even members.

As in our previous studies on carbon-doped clusters,[24,25] the relative stability of linear VC_n compounds will be discussed following the suggestion by Pascoli and Lavendy,[50] in terms of the incremental binding energies.[51] The incremental binding energy can be defined as the change in energy accompanying the process

$$VC_n \rightarrow VC_{n-1} + C \qquad (1)$$

and can be computed as the consecutive binding energy differences between adjacent VC_n and VC_{n-1} clusters.

The incremental binding energy for the different linear VC_n clusters as a function of the number of carbon atoms is shown in Figure 2. From this figure, it can be observed that a clear even−odd alternation in stability exists for linear clusters, with their *n*-even members being more stable than the corresponding $n - 1$ and $n + 1$ ones. The difference in stability between *n*-odd and *n*-even clusters is attenuated for the last members of the series. This parity effect can be attributed to the electron number in π-type and δ-type highest occupied molecular orbitals. In the case of the smaller clusters of the series, a ground state corresponding to a δ^2 electronic configuration (*n*-even) is energetically more favor-

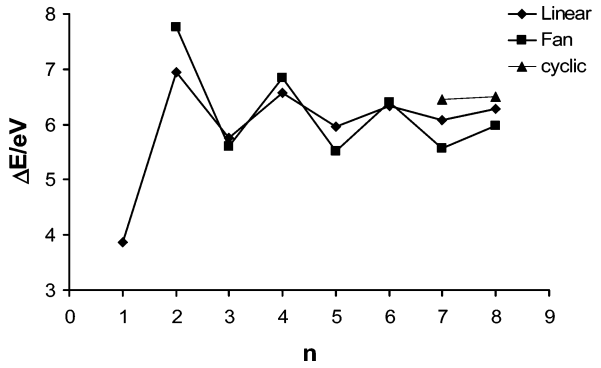


Figure 2. Incremental binding energies (eV) for the VC_n linear, fan, and cyclic clusters vs the number of carbon atoms at the B3LYP/6-311+G(d) level.

able than a $\pi^3\delta^1$ one (n-odd). On the other hand, linear VC_n clusters with $n = 6$ and 8 have ground states resulting from an electronic configuration of $\pi^1\delta^1$, which is only slightly more stable than the corresponding $\pi^3\delta^1$ for n-odd clusters.

VC_n "Fan" Clusters. The molecular properties (absolute and relative energies, S^2 expectation values, and dipole moments) for the lowest-lying doublet, quartet, and sextuplet states for fan-type VC_n clusters are given in Table 2, whereas the corresponding vibrational frequencies and rotational constants are included as Supporting Information in Table S2. In addition, in Figure 3, the geometrical parameters for the ground states are shown. In general, all of the structures reported in Figure 3 are true minima on their respective potential energy surfaces, with the only exception being the $^4B_2$ state of VC_7 at the B3LYP/LAN-6+(d) level. The doublet state of VC_6 also presented an imaginary frequency when the LAN-6+(d) basis set was employed (as can be seen in Table S2). In both cases, all of our attempts to find a true minimum failed. For the ^4B state of VC_8, we only reached the fan structure with the 6-311+G(d) basis set; all of our attempts to obtain this structure with the LAN-6+(d) basis set led to the cyclic conformation, which is lower in energy. As in the case of linear isomers, the inclusion of effective core potentials in the basis set of the vanadium atom does not seem relevant for the description of fan VC_n isomers, because not only are the geometrical parameters obtained with both basis sets, 6-311+G(d) and LAN-6+(d), very coincident (see Figure 3) but also all other properties are very close (see Tables 2 and S2).

There are only available theoretical results for the first member of the series, VC_2.[13,30] Majumdar el al.[13] have studied the different states of the VC_2 system in C_{2v} symmetry. As in the case of linear structures, they employed MRSDCI and DFT/B3LYP methods with a RECPs basis set for vanadium and carbon atoms. They found that, in general, the geometries and the energy separations of the reported electronic states at the DFT level are similar to the MRSDCI data for most of the states. An $^2A_2$ symmetry and an $^6A_1$ state are reported as the lowest-lying doublet and sextuplet states, which are located about 27.90 and 28.59 kcal/mol, respectively, above the $^4B_1$ ground state at the MRCISD+Q level. We also found $^2A''$ (corresponding to an $^2A_2$) and $^6A_1$ states with energy differences respective to the $^4B_1$ state very similar to that estimated by Majumdar et al. (27.08 and 34.14 kcal/mol, respectively; see Table 2).

Table 2. Electronic Energies, $\langle S^2\rangle$ Values, Binding Energies (for the Ground States), Dipole Moments, and Relative Energies for Fan VC_n Clusters with the B3LYP/6-311+G(d) (First Line) and B3LYP/LAN-6+(d) (Second Line) Methods

isomer	state	$-E$ (au)	$\langle S^2\rangle$	BE (eV)	μ (D)	ΔE (kcal mol^{-1})
VC_2	$^2A''$	1019.995993	0.7753		6.41	27.08
		147.326054	0.7789		6.48	27.01
	$^4B_1$	1020.039824	3.7897	11.78	6.77	0.00
		147.370027	3.7906	11.75	6.80	0.00
	$^6A_1$	1019.985001	8.7680		4.28	34.14
		147.314893	8.7691		4.66	34.29
VC_3	$^2A_1$	1058.109391	0.8748	17.56	5.34	0.00
		185.436671	0.8909	17.44	5.28	0.00
	$^4B_2$	1058.100062	3.8157		5.32	5.23
		185.427137	3.8256		5.47	5.31
	$^6A_1$	1058.081370	8.7977		5.02	16.57
		185.409567	8.7994		5.24	15.94
VC_4	$^2B_1$	1096.209770	1.6403		3.67	8.35
		223.534658	1.6368		3.68	9.11
	$^4B_1$	1096.223033	3.7891	24.53	3.96	0.00
		223.549048	3.7910	24.39	4.01	0.00
	$^6B_1$	1096.150580	8.7721		3.61	44.25
		223.479310	8.7707		4.07	42.52
VC_5	$^2A'$	1134.287688	1.6061	30.18	2.32	0.00
		261.609891	1.6533	29.92	2.29	0.00
	$^4A'$	1134.281289	3.8096		1.86	3.71
		261.603569	3.8195		1.83	3.31
	$^6B_2$	1134.250284	8.8448		1.98	22.78
		261.575483	8.8449		2.14	20.81
VC_6	$^2A'$	1172.361429	0.8137		3.77	14.81
		299.681196	0.8324		3.82	16.23
	$^4A_2$	1172.385474	3.7883	36.72	0.84	0.00
		299.708020	3.7892	36.48	0.74	0.00
	$^6B_1$	1172.316781	8.7860		0.58	41.62
		299.642708	8.7864		0.23	39.66
VC_7	^2A	1210.445070	1.6656		3.12	5.29
	$^2A_1$	337.764383	1.7390		3.21	5.16
	$^4B_2$	1210.453569	3.8122	42.46	3.28	0.00
		337.772336	3.8278	42.11	3.35	0.00
	$^6A'$	1210.438130	8.8444		1.43	9.52
		337.763581	8.8460		1.34	5.61
VC_8	$^2A''$	1248.531749	1.7721		3.30	2.53
		375.841624	1.7449		3.42	
	^4B	1248.535327	3.7743	48.57	4.24	0.00
	$^6B_1$	1248.496690	8.8059		1.38	23.70
		375.824748	8.8095		1.26	

As can be seen in Table 2, all n-even fan VC_n clusters have quartet ground states, and the quartet−doublet energy difference has a tendency to decrease as the number of carbon atoms increases, obtaining an energy difference of only 2.53 kcal/mol for $n = 8$. For n-odd clusters, doublet and quartet states are close in energy, and we have found that for VC_3 and VC_5 isomers the ground state corresponds to $^2A_1$ ($^2A'$), whereas in the case of VC_7, a $^4B_2$ state is predicted to lie lower in energy than the corresponding doublet one. Along all of the series, the sextuplet states are the most unstable

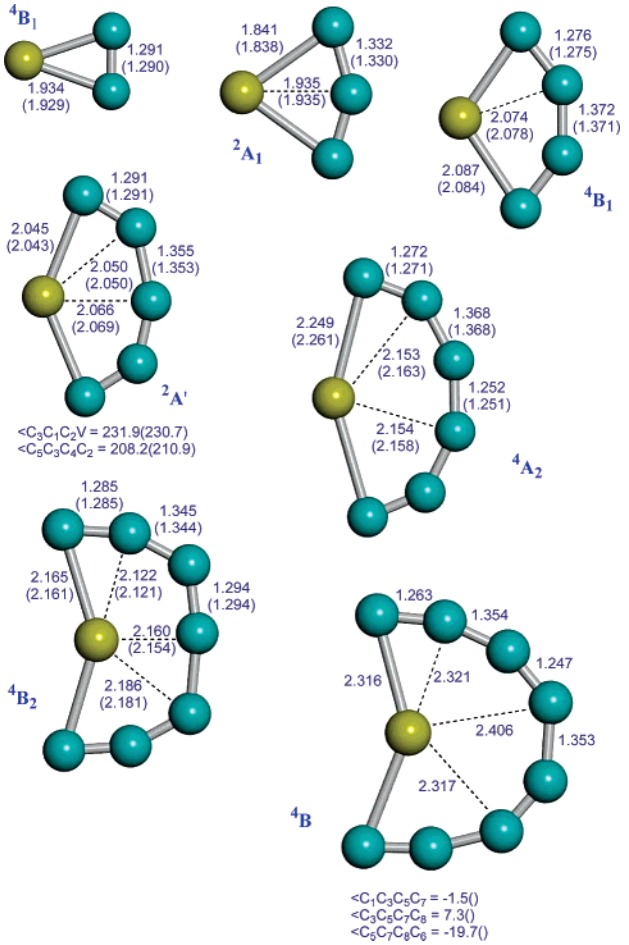


Figure 3. Equilibrium geometries of VC_n fan clusters at the B3LYP/6-311+G(d) and B3LYP/LAN-6+(d) (in parentheses) levels of theory. Distances are given in angstroms.

ones, with an energy difference respective to the quartet states that is lower for *n*-odd clusters. In general, we can see that in fan type structures there is a stabilization of doublet states with respect to the linear ones, whereas sextuplet states increase their energy relative to quartet ones.

There are some interesting features of the geometrical parameters for fan VC_n clusters. It can be seen, in Figure 3, that the V−C distances for fan structures are, in general, longer than in the case of linear isomers. An analysis of the V−C bond distances shows that there is a π-type interaction between the vanadium atom and the entire carbon unit, and consequently, all of the V−C distances are very close. As an example, V−C_1, V−C_2, V−C_3, and V−C_4 are respectively 2.165, 2.122, 2.186, and 2.160 Å for VC_7, suggesting that, in fact, there is a similar interaction between the vanadium and each of the carbon atoms. On the other hand, V−C bond lengths are higher for *n*-even clusters, and in both cases (*n*-even and *n*-odd), these distances are systematically longer as *n* increases, showing a lower degree of participation of the vanadium atom in each bonding when the number of carbon atoms in the C_n unit is higher. The C−C bond distances are within the range 1.25−1.37 Å, following a pattern that resembles the behavior found for linear isomers, with a clear alternation of C−C distances (C_{odd}−C_{even} distances shorter than C_{even}−C_{odd} ones).

Table 3. Electronic Energies, $\langle S^2 \rangle$ Values, Binding Energies (for the Ground States), Dipole Moments, and Relative Energies for Cyclic VC_n Clusters with the B3LYP/6-311+G(d) (First Line) and B3LYP/LAN-6+(d) (Second Line) Methods

isomer	state	−*E* (au)	$\langle S^2 \rangle$	BE (eV)	μ (D)	ΔE (kcal mol^{-1})
VC_6	$^2A_2$	1172.333646	1.7228		3.99	21.11
		299.661755	1.7121		4.69	20.49
	$^4A''$	1172.365666	3.8083	36.18	5.14	0.00
		299.694943	3.8144	36.12	5.21	0.00
	$^6A_1$	1172.332800	8.7985		3.07	20.28
		299.655621	8.8017		3.29	24.26
VC_7	$^2B_1$	1210.465646	1.1611		4.48	0.24
		337.793869	1.2476		4.27	−0.34
	$^2A_1$	1210.463828	1.1736		4.44	1.32
		337.792043	1.2732		4.19	0.76
	$^4B_2$	1210.465486	3.8578	42.78	3.44	0.00
		337.792814	3.8880	42.67	3.29	0.00
	$^6A_1$	1210.448359	8.8787		0.20	9.68
		337.776627	8.8811		0.47	9.12
VC_8	$^2B_1$	1248.543512	1.5622		0.26	14.51
		375.870105	1.6257		0.24	15.26
	$^4A''$	1248.566679	3.7820	49.42	0.12	0.00
	$^4B_1$	375.894644	3.7880	49.33	0.05	0.00
	$^6A_2$	1248.510595	8.8032		0.25	34.25
		375.838716	8.8108		0.42	34.10

It is also worth mentioning that the dipole moment decreases when *n* increases for small fan clusters, but an inversion in the trend is observed from $n = 6$. This change can be related with the fact that the vanadium atom tends to be included into the C_n unit for larger fan structures. As expected, the dipole moments of fan clusters are lower than the corresponding ones of linear species.

In Figure 2, the incremental binding energies for the fan VC_n clusters are also represented. It can be readily seen that incremental binding energies for fan clusters follow the same general pattern as their linear analogues. Again, *n*-even clusters are more stable than *n*-odd ones, and the difference in stability between *n*-odd and *n*-even clusters is attenuated when *n* increases. But now, the incremental binding energies vary more drastically for consecutive members in the series than for linear clusters.

VC_n "Cyclic" Clusters. Cyclic structure where the vanadium atom is bonded to the two terminal carbon atoms of the C_n unit are only reached for $n = 6$, 7, and 8 (in the case of $n = 5$, a cyclic structure was located, but we have not included it because it is lying very high in energy). All of our attempts to obtain this rearrangement for $n < 5$ collapsed to the fan-type structure with the two basis sets employed in this work. The main properties for the lowest-lying doublet, quartet, and sextuplet states of cyclic VC_n clusters are given in Table 3, whereas geometries for the ground states are shown in Figure 4. As in previous sections, the vibrational frequencies and rotational constants for the lowest-lying doublet, quartet, and sextuplet states of cyclic clusters are given as Supporting Information in Table S3. As in the case of linear and fan clusters, the effect of employing effective core potentials to describe the vanadium

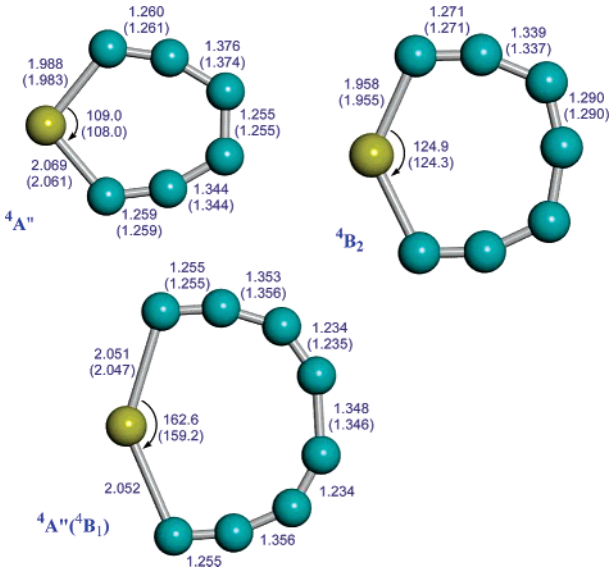


Figure 4. Equilibrium geometries of VC_n cyclic clusters at the B3LYP/6-311+G(d) and B3LYP/LAN-6+(d) (in parentheses) levels of theory. Distances are given in angstroms.

atom on the geometrical parameters, harmonic frequencies, dipole moments, and energies is not particularly important. All of the isomers included in Table 3 are true minima on their respective potential energy surfaces, except the $^2A_2$ state of VC_6, which has an imaginary frequency at the B3LYP/LAN-6+(d) level. All of our attempts to find a true minimum at this level in C_s symmetry failed.

From Table 3, it is readily seen that cyclic clusters have quartet ground states. However, in the case of the VC_7 cluster, there are two doublet states, $^2B_2$ and $^2A_1$, that are very close in energy to the $^4B_2$ ground state (energy differences between the three states are about 1.00 kcal/mol), and therefore, the three states could be accessible to experimental detection. The lowest-lying sextuplet state of VC_7, $^6A_1$, is only located about 9.68 kcal/mol above the ground state. The doublet−quartet, and sextuplet−quartet energy differences are larger for *n*-even clusters than for VC_7. In the case of *n*-even, it seems that the doublet states tend to be more stable when *n* increases, whereas the energy differences for sextuplet ones increase when going from VC_6 to VC_8 clusters.

Concerning the geometrical parameters, it can be seen in Figure 4 that V−C bond distances in cyclic systems are between those found for linear and fan isomers, according to the fact that the vanadium atom is bonded to the two terminal carbon atoms. On the other hand, the C−C bond distances exhibit a behavior quite similar to that found in the linear and fan isomers, with an alternation of C−C bond distances (C_{odd}−C_{even} distances being shorter than C_{even}−C_{odd} ones). In cyclic isomers, the dipole moments are relatively high but decrease significantly for VC_8.

Incremental binding energies for cyclic VC_n clusters can only be calculated for $n = 7$ and 8. As can be seen in Figure 2, the incremental binding energies for cyclic VC_7 and VC_8 are very similar, the VC_8 cluster being only slightly more stable. Therefore, the parity effect for cyclic systems seems to be lower than that in linear and fan structures.

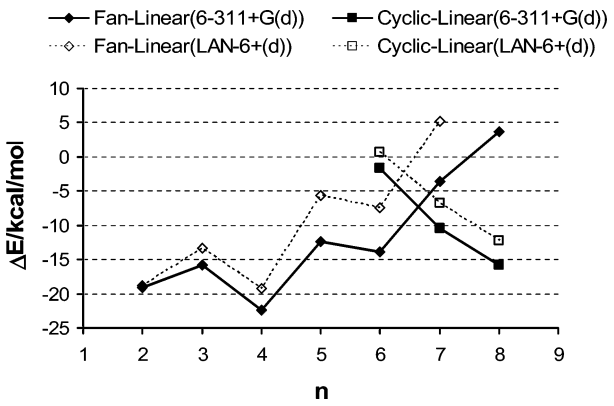


Figure 5. Relative energies (kcal/mol) of linear, fan, and cyclic VC_n clusters vs the number of carbon atoms.

Linear−Fan−Cyclic Stability. The energy differences between linear, fan, and cyclic structures for VC_n clusters are shown in Figure 5 as a function of the number of carbon atoms. We have represented the relative energies of fan and cyclic structures with respect to the linear ones, and a negative value means that the fan or the cyclic structure is more stable than the corresponding linear one. In general, we can see that the results obtained with the LAN-6+(d) basis set show similar trends to those obtained with the 6-311+G(d) basis set. Nevertheless, it can be seen that the LAN-6+(d) basis set slightly favors the linear isomer compared to the 6-311+G(d) basis set. This effect is larger for the last members of the series.

From Figure 5, we can seen that for $n \leq 6$, VC_n clusters are predicted to prefer fan structures, and it seems that fan isomers could be the most favorable to be characterized in experimental studies. There is also a clear even−odd alternation, *n*-even fan clusters being comparatively more stable than *n*-odd ones. This alternation can be related to the fact that fan structures have higher incremental binding energies for *n*-even clusters than do linear ones, whereas for *n*-odd the incremental binding energy is smaller than that of the corresponding linear isomers. On the other hand, when $n > 6$, the trend is to favor linear structures, and for $n = 8$, the linear isomer is even more stable than the fan structure. However, the behavior of the cyclic−linear energy differences is the opposite. For $n = 6$, linear and cyclic isomers are nearly isoenergetic, and the differences decrease almost linearly when going from VC_6 to VC_8. This fact results in the prediction that cyclic structures are the most stable VC_n isomers with $n > 6$.

Finally, as it can be seen in Figure 5, although the LAN-6+(d) basis set favors linear structures over fan and cyclic ones, both basis sets predict that, for $n \leq 6$, the most stable isomer is the fan structure and, for $n > 6$, the cyclic one is.

Conclusions

A theoretical study using the B3LYP method with the 6-311+G(d) and LAN-6+(d) basis sets has been carried out for the most stable isomers (linear, fan, and cyclic) of VC_n ($n = 1-8$) neutral clusters. Predictions for their geometrical parameters and other properties that could be useful for an eventual experimental characterization have been provided.

According to our calculations, linear VC_n clusters have quartet lowest-lying states, with the only exception being VC, which has a $^2\Delta$ ground state. *n*-odd clusters have a $^4\Phi$ ground state, and the lowest-lying sextuplet, $^6\Sigma$, is very close in energy, whereas for *n*-even, the $^4\Sigma$ and $^4\Phi$ states are very close in energy. The $^4\Sigma$ state is the most stable for clusters with $n = 2$ and 4, whereas VC_6 and VC_8 have a $^4\Phi$ ground state. In the case of fan VC_n clusters, *n*-even species have quartet ground states. For *n*-odd clusters, doublet states are favored for the first members of the series ($n = 3, 5$), whereas quartet ground states are found for the VC_7 isomer. The three cyclic VC_n clusters characterized have quartet lowest-lying states.

The stability of linear and fan clusters as a function of the size has been discussed in terms of the incremental binding energies. For both linear and fan structures, a clear even−odd alternation in stability is found, *n*-even clusters being more stable than *n*-odd ones. It is also found that this parity effect decreases along the series.

Concerning a possible experimental work on this system, one of the most interesting results of the present work refers to the competition between linear, fan, and cyclic structures. Our results suggest that for $n \leq 6$ there is a preference for fan structures over linear ones, especially for low *n* values. A parity effect is also observed, fan isomers being more stable for *n*-even clusters. For larger clusters ($n = 7$ and 8), the cyclic structure is the most stable. The linear structure is the least stable structure considered in this work for neutral VC_n clusters, only for VC_8, the linear structure is slightly more stable than the fan isomer.

Finally, the results obtained with both basis sets for the descriptions of geometrical parameters, energies, and other properties are very close. Therefore, the employ of effective core potentials for the description of the vanadium atom can be useful in the study of large vanadium−carbon clusters.

Acknowledgment. This research has been supported by the Ministerio de Ciencia y Tecnología of Spain (Grant BQU2004-07405-C02-01) and by the Junta de Castilla y León (Grant VA 085/03).

Supporting Information Available: Vibrational frequencies and rotational constants for linear VC_n, *fan* VC_n, and *cyclic* VC_n clusters are given in Tables S1−S3. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Clemmer, D. E.; Hunter, J. M.; Shelomov, K. B.; Jarrold, M. F. *Nature* **1994**, *372*, 248−250.

(2) Chai, Y.; Guo, T.; Jin, C.; Haufler, R. E.; Chibante, L. P. F.; Fure, J.; Wang, L.; Alford, J. M.; Smalley, R. E. *J. Phys. Chem.* **1991**, *95*, 7564−7568.

(3) Iijima, S.; Ichihashi, T. *Nature* **1993**, *363*, 603−605.

(4) Wei, S.; Guo, B. C.; Purnell, J.; Buzza, S.; Castleman, A. W. *Science* **1992**, *256*, 818−820.

(5) Wei, S.; Guo, B. C.; Purnell, J.; Buzza, S.; Castleman, A. W. *J. Phys. Chem.* **1992**, *96*, 4166−4168.

(6) Pilgrim, J. S.; Duncan, M. A. *J. Am. Chem. Soc.* **1993**, *115*, 6958−6961.

(7) Rohmer, M. M.; Benard, M.; Poblet, J. M. *Chem. Rev.* **2000**, *100*, 495−542.

(8) Guo, B. C.; Kearns, K. P.; Castleman, A. W. *Science* **1992**, *225*, 1411−1413.

(9) Guo, B. C.; Wei, S.; Purnell, J.; Buzza, S.; Castleman, A. W. *Science* **1992**, *256*, 515−516.

(10) Roszak, S.; Balasubramanian, K. *J. Phys. Chem. A* **1997**, *101*, 2666−2669.

(11) Jackson, P.; Gadd, G. E.; Mackey, D. W.; van der Wall, H.; Willett, G. D. *J. Phys. Chem. A* **1998**, *102*, 8941−8945.

(12) Hendrickx, M.; Clima, S. *Chem. Phys. Lett.* **2004**, *388*, 284−289.

(13) Majumdar, D.; Roszak, S.; Balasubramanian, K. *J. Chem. Phys.* **2003**, *118*, 130−141.

(14) Arbuznikov, A. V.; Hendrickx, M.; Vanquickenborne, L. G. *Chem. Phys. Lett.* **1999**, *310*, 515−522.

(15) Hendrickx, M.; Clima, S. *Chem. Phys. Lett.* **2004**, *388*, 290−296.

(16) Arbuznikov, A. V.; Hendrickx, M. *Chem. Phys. Lett.* **2000**, *320*, 575−581.

(17) Zhai, H. J.; Wang, L. S.; Jena, P.; Gutsev, G. L.; Bauschlicher, C. W., Jr. *J. Chem. Phys.* **2004**, *120*, 8996−9008.

(18) Sumathi, R.; Hendrickx, M. *Chem. Phys. Lett.* **1998**, *287*, 496−502.

(19) Sumathi, R.; Hendrickx, M. *J. Phys. Chem. A* **1998**, *102*, 4883−4889.

(20) Noya, E. G.; Longo, R. C.; Gallego, L. J. *J. Chem. Phys.* **2003**, *119*, 11130−11134.

(21) Froudakis, G. E.; Muhlhauser, M.; Andriotis, A. N.; Menon, M. *Phys. Rev. B* **2001**, *64*, 241401-1−241401-4.

(22) Grey, G.; Alemany, M. M. G.; Dieguez, O.; Gallego, L. G. *Phys. Rev. B* **2000**, *62*, 12640−12643.

(23) Longo, R. C.; Alemany, M. M. G.; Fernandez, B.; Gallego, L. J. *Phys. Rev. B* **2003**, *68*, 167401-1−167401-3.

(24) Redondo, P.; Barrientos, C.; Largo, A. *J. Phys. Chem. A* **2005**, *109*, 8594−8603.

(25) Redondo, P.; Barrientos, C.; Largo, A. *J. Phys. Chem. A* **2006**, *110*, 4057−4064.

(26) Maclagan, R. G. A. R.; Scuseria, G. E. *Chem. Phys. Lett.* **1996**, *262*, 87−90.

(27) Gutsev, G. L.; Andrews, L.; Bauschlicher, C. W., Jr. *Theor. Chem. Acc.* **2003**, *109*, 298−308.

(28) Majumdar, D.; Balasubramanian, K. *Mol. Phys.* **2003**, *101*, 1369−1376.

(29) Kalemos, A.; Dunning, T. H.; Mavridis, A. *J. Chem. Phys.* **2005**, *123*, 14301−14309.

(30) Tono, K.; Torasaki, A.; Ohta, T.; Kondow, T. *Chem. Phys. Lett.* **2002**, *351*, 135−141.

(31) Li, X.; Wang, L.-S. *J. Chem. Phys.* **1999**, *111*, 8389−8395.

(32) Wang, L.-S.; Li, X. *J. Chem. Phys.* **2000**, *112*, 3602−3608.

(33) Becke, A. D. *J. Chem. Phys.* **1986**, *84*, 4524−4529.

(34) Becke, A. D. *J. Chem. Phys.* **1988**, *88*, 2547−2553.

(35) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.

(36) Becke, A. D. *J. Chem. Phys.* **1988**, *88*, 1053−1062.

(37) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.

(38) Kohn, W.; Sham, L. J. *Phys. Rev. A* **1965**, *140*, 1133−1139.

(39) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650−654.

(40) Wachters, A. J. H. *J. Chem. Phys.* **1970**, *52*, 1033−1036.

(41) Hay, P. J. *J. Chem. Phys.* **1977**, *66*, 4377−4384.

(42) Ragavachari, K.; Trucks, G. W. *J. Chem. Phys.* **1989**, *91*, 1062−1065.

(43) Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 270−283.

(44) Wadt, W. R.; Hay, P. J. *J. Chem. Phys.* **1985**, *82*, 284−298.

(45) Hay, P. J.; Wadt, W. R. *J. Chem. Phys*. **1985**, *82*, 299−310.

(46) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millan, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelly, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*; Gaussian Inc.: Pittsburgh, PA, 1998.

(47) Gingerich, K. A. *J. Chem. Phys.* **1969**, *50*, 2255−2256.

(48) Gupta, S. K.; Gingerich, K. A. *J. Chem. Phys.* **1981**, *74*, 3584−3590.

(49) Hamrick, Y. M.; Weltner, W., Jr. *J. Chem. Phys.* **1991**, *94*, 3371−3380.

(50) Pascoli, G.; Lavendy, H. *J. Phys. Chem. A* **1999**, *103*, 3518−3524.

(51) Raghavachari, K.; Binkley, J. S. *J. Chem. Phys.* **1987**, *87*, 2191−2197.

CT060033Q