JCTC Journal of Chemical Theory and Computation

Introducing the *Journal of Chemical Theory and Computation*

Welcome to the *Journal of Chemical Theory and Computation (JCTC).* This new American Chemical Society journal is being introduced to provide a focal point for publication of articles, letters, and reviews reporting new theories based on physical laws, advances in computational methods, and important applications to problems in chemistry. Though the contents will evolve, some specific topics that are appropriate today for *JCTC* include advances in or applications of ab initio quantum mechanics, density functional theory, statistical mechanics, design and properties of new materials, surface science, Monte Carlo simulations, solvation models, QM/MM calculations, biomolecular structure prediction, and molecular dynamics in the broadest sense including gas-phase dynamics, ab initio dynamics, biomolecular dynamics, and protein folding.

Many theoretical and computational chemists have believed for a long time that their rapidly expanding community should be represented in a dedicated ACS journal. Their work is currently dispersed over a wide variety of journals in which they often represent a minority interest. The ACS has now responded, and you are invited to participate in this undertaking to establish *JCTC* as the central source for the best work in the field. The ACS connection will provide reassurance that all aspects of the processing of manuscripts will be handled professionally and efficiently and that the resultant publications will benefit from state-of-the-art appearance and worldwide accessibility. To facilitate this, manuscript submission and review for *JCTC* will only occur through the online ACS Paragon System. The Journal is off to a great start as documented in this first issue. It is hoped that the birth of *JCTC* will be viewed in time as a notable event comparable to the introduction of predecessors such as the *Journal of Physical Chemistry* in 1896 or the *Journal of Chemical Physics* in 1933. In the latter's first issue, its editor, Harold Urey, observed: "The life and interests of the individual are short and of little consequence as compared to the lives of our sciences." This will always be true, and our only intention is for *JCTC* to help enrich the science through improved publication and sense of community.

There are many people who deserve thanks for bringing *JCTC* to fruition. The support, encouragement, and handling of innumerable details by the ACS Board of Directors, Executive Director & CEO Madeleine Jacobs, President of the ACS Publications Division Robert Bovenschulte, Vice President for the Journal Publishing Group Mary Scanlan, their endlessly helpful staffs, the Editorial Advisory Board, and Editorial Assistant Patricia Morales have been essential.

William L. Jorgensen
Editor

CT040003S

10.1021/ct040003s CCC: $30.25
Published on Web 10/21/2004

JCTC Journal of Chemical Theory and Computation

An Efficient Linear-Scaling Ewald Method for Long-Range Electrostatic Interactions in Combined QM/MM Calculations

Kwangho Nam, Jiali Gao, and Darrin M. York*

Department of Chemistry and Supercomputing Institute, University of Minnesota, Minneapolis, Minnesota 55455-0431

Received September 8, 2004

Abstract: A method is presented for the efficient evaluation of long-range electrostatic forces in combined quantum mechanical and molecular mechanical (QM/MM) calculations of periodic systems. The QM/MM-Ewald method is a linear-scaling electrostatic method that utilizes the particle mesh Ewald algorithm for calculation of point charge interactions of molecular mechanical atoms and a real-space multipolar expansion for the quantum mechanical electrostatic terms plus a pairwise periodic correction factor for the QM and QM/MM interactions that does not need to be re-evaluated during the self-consistent field procedure. The method is tested in a series of molecular dynamics simulations of the ion−ion association of ammonium chloride and ammonium metaphosphate and the dissociative phosphoryl transfer of methyl phosphate and acetyl phosphate. Results from periodic boundary molecular dynamics (PBMD) simulations employing the QM/MM-Ewald method are compared with corresponding PBMD simulations using electrostatic cutoffs and with results from nonperiodic stochastic boundary molecular dynamics (SBMD) simulations, with cutoffs and with full electrostatics (no cutoff). The present method allows extension of linear-scaling Ewald methods to molecular simulations of enzyme and ribozyme reactions that use combined QM/MM potentials.

1. Introduction

The profound effects of solvation on chemical reactions have been recognized for over a century and continue to attract intensive experimental and theoretical research effort.[1] For reactions catalyzed by enzymes or ribozymes, the environment is even more complicated. The challenge in theoretical studies of the mechanism and reactivity of chemical processes is to move accurate quantum electronic structure calculations from the gas phase into the condensed phase realm. However, for large systems such as proteins and nucleic acids, the complexity and system size preclude the use of even the most efficient linear-scaling electronic structure methods[2] to simulate the reaction dynamics explicitly. This is further exacerbated by the need for an adequate treatment of long-range electrostatic interactions in polar solvents and in the presence of mobile counterions. Fortunately, it is often the case that the vast majority of the system does not require a high-level and computationally intensive quantum mechanical model. This situation is ideally suited for application of a combined quantum mechanical and molecular mechanical (QM/MM) approach, in which the solute is treated quantum mechanically and the environment by classical force fields.[3−6]

Electrostatic interactions are generally perceived to be the dominant forces that stabilize transition states in biochemical reactions[7−10] and provide essential stability in long-time dynamic simulations of proteins and nucleic acids.[11−16] The development of efficient linear-scaling electrostatic methods[14,17,18] (methods for which the computational effort scales linearly with system size — or nearly so, see below) have greatly improved the reliability of molecular dynamics simulations of large biological systems. In combined QM/

* Corresponding author e-mail: york@chem.umn.edu. Corresponding address: Department of Chemistry, University of Minnesota, Minneapolis, MN 55455-0431.

10.1021/ct049941i CCC: $30.25
Published on Web 12/14/2004

MM potential models, the electrostatic environment affects the quantum electronic polarization of the solute[19] that plays a significant role in the stabilization of macromolecules in solution[20] and the rate enhancement of some enzymes.[21] Consequently, it is critical to compute long-range electrostatic interactions accurately in QM/MM simulations of biochemical reactions. Nonetheless, due to the lack of availability of algorithms that extend linear-scaling electrostatic methods to combined QM/MM potentials, a large percentage of QM/MM applications routinely employ electrostatic cutoffs.[10]

The present paper presents a linear-scaling Ewald method for efficient calculation of long-range electrostatic interactions in combined QM/MM simulations using semiempirical quantum models. The method can be easily extended to QM/MM ab initio molecular orbital and density functional theory. In the following, the essential theory and computational details are first outlined. Next, results obtained using the QM/MM-Ewald method are compared with those from other simulations by computing interionic potentials of mean force for ion association and for phosphoryl transfer reactions. Specifically, periodic boundary molecular dynamics (PBMD) simulations calculated with the QM/MM-Ewald method are compared with corresponding PBMD simulations using electrostatic cutoffs and with results from nonperiodic stochastic boundary molecular dynamics (SBMD) simulations, with cutoffs and with full electrostatics (no cutoff). Finally, the paper concludes with a summary of the key results and identifies directions of future research.

2. Theory

2.1. Electrostatic Energy of a Periodic System of Point Charges. Consider a periodic system of N point charges $\{q_i\}$ located at position $\{\mathbf{r}_i\}$, $i = 1, \cdots, N$, in a periodic unit cell U characterized by the set of real-space lattice vectors $\{\mathbf{a}_k\}$, $k = 1, 2, 3$. The classical electrostatic energy of this system, excluding the infinite self-energy of the point charges, is given by

$$E_{\text{elec}} = \frac{1}{2}\sum_{i}^{N}\sum_{j}^{N}\sum_{\mathbf{n}}^{\infty}{}' \frac{q_i q_j}{|\mathbf{r}_{ij} + \mathbf{n}|} \tag{1}$$

where $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$, and the summation over $\mathbf{n}$ is over all integer translations of the real space lattice vectors $\mathbf{n} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3$ for integers n_k ($k = 1, 2, 3$), and the prime symbol indicates that the terms where $|\mathbf{r}_{ij} + \mathbf{n}| = 0$ are neglected. The summation in eq 1 is not convergent unless the total charge of the system sums to zero (i.e., the monopole moment of the unit cell vanishes). If the unit cell has vanishing monopole *and* dipole moments, the sum converges absolutely; however, if the unit cell has a net dipole moment, the sum is only conditionally convergent and has different converged values depending on the order and limiting manner whereby the sum is affected.[22] In any case, the expression in eq 1, under the conditions where it does converge, does so very slowly, and is not a practical means of computing electrostatic energies for periodic systems.

The Ewald summation convention[23,24] uses an elegant mechanism of transforming the slowly convergent sum in eq 1 into two rapidly convergent sums over real-space and reciprocal space lattice vectors

$$E_{\text{elec}}(q^N,\mathbf{r}^N,U) = E_{\text{real}}(q^N,\mathbf{r}^N,U;\kappa) + E_{\text{recip}}(q^N,\mathbf{r}^N,U;\kappa) + E_{\text{surf}}(q^N,\mathbf{r}^N,U;P,\epsilon) \tag{2}$$

where

$$E_{\text{real}}(q^N,\mathbf{r}^N,U;\kappa) = \frac{1}{2}\sum_{i}^{N}\sum_{j}^{N} q_i q_j \left(\sum_{\mathbf{n}}^{\infty}{}' \frac{\text{erfc}(\kappa|\mathbf{r}_{ij} + \mathbf{n}|)}{|\mathbf{r}_{ij} + \mathbf{n}|} - \frac{2\kappa}{\sqrt{\pi}}\delta_{ij}\right) \tag{3}$$

$$E_{\text{recip}}(q^N,\mathbf{r}^N,U;\kappa) = \frac{1}{2}\cdot\frac{4\pi}{V}\sum_{|\mathbf{k}|\neq 0}^{\infty}\frac{\exp(-k^2/4\kappa^2)}{k^2}|S(\mathbf{k})|^2 \tag{4}$$

where $\text{erfc}(x)$ is the complementary error function, defined as $\text{erfc}(x) = 1 - \text{erf}(x)$, and $\text{erf}(x)$ is the error function.[25] The summation in eq 4 is over vectors $\mathbf{k} = 2\pi\mathbf{m}$, and $\mathbf{m}$ sums over all integer translations of the reciprocal lattice $\mathbf{m} = m_1\mathbf{a}_1^* + m_2\mathbf{a}_2^* + m_3\mathbf{a}_3^*$ for integers m_k ($k = 1, 2, 3$), where the set of reciprocal lattice vectors $\{\mathbf{a}^*_i\}$ are related to the real-space lattice vectors $\{\mathbf{a}_i\}$ by $\mathbf{a}_i^*\cdot\mathbf{a}_j = \delta_{ij}$. In eq 4, V is the volume of the unit cell U ($V = |\mathbf{a}_1\cdot\mathbf{a}_2 \times \mathbf{a}_3|$), and $S(\mathbf{k})$ is the structure factor[26,27] and is given by

$$S(\mathbf{k}) = \sum_{j}^{N} q_j \exp(i\mathbf{k}\cdot\mathbf{r}_j) \tag{5}$$

The two summations contain a parameter κ that adjusts the relative rates of convergence. The total energy is independent of the κ parameter, so long as the real-space and reciprocal space sums are both sufficiently converged. In practice, these sums are truncated at some point so as to fall below a fixed tolerance level in accuracy ϵ_{tol}. If the parameter κ is chosen such that only the $|\mathbf{n}| = 0$ term is required in eq 3 to obtain the desired level of accuracy (i.e., the *minimum image convention*[24,28] can be used to perform the summations over particles in the unit cell), then the number of reciprocal-space lattice vectors $\mathbf{k}$ required to obtain the same level of accuracy becomes constant with respect to scaling of the unit cell, and an order N^2 algorithm results.[26,29] If the parameter κ is optimized for scaling efficiency, then an order $N^{3/2}$ algorithm can be obtained.[30] However, to extend the method to very large systems, a so-called "linear-scaling" algorithm is required whereby the scaling is *better than* order N^λ, $\forall\lambda > 1$. Such algorithms have been developed previously,[18] perhaps the most commonly employed algorithm in molecular dynamics simulations is the particle mesh Ewald method[27,31] that has recently been extended to higher order multipole moments.[32]

Before proceeding further, it is worthwhile to briefly clarify further eq 4. It was mentioned previously that the original expression of eq 1 was subject to several convergence restrictions. These restrictions manifest themselves through the $|\mathbf{k}| = 0$ term, that in eq 4 has been neglected. Clearly care must be taken with the $|\mathbf{k}| = 0$ term since the sum involves a $1/k^2$ factor, that must be resolved via a limiting procedure involving the ratio $|S(\mathbf{k})|/|k|$. From eq 5

it is clear that $|S(0)| \neq 0$ for a non-neutral system, consistent with the statement earlier that eq 1 is nonconvergent under this condition. If the system is neutral, but the unit cell has a net dipole moment ($\mathbf{D} = \sum_j q_j \mathbf{r}_j$), then the $|\mathbf{k}| = 0$ term gives rise to the surface term, $E_{\text{surf}}(q^N,\mathbf{r}^N,U;P,\epsilon)$ in eq 2, that depends quadratically on the dipole moment $\mathbf{D}$

$$E_{\text{surf}}(q^N,\mathbf{r}^N,U;P,\epsilon) = \frac{1}{2}\alpha(P,\epsilon)|\sum_j q_j \mathbf{r}_j|^2 = \frac{1}{2}\alpha(P,\epsilon)\cdot|\mathbf{D}|^2 \quad (6)$$

where the proportionality constant $\alpha(P,\epsilon)$ depends on the macroscopic shape of the crystal, P, and the dielectric constant, ϵ, of the surrounding medium.[22] The physical interpretation of the surface term is that of an energy associated with a dipole layer on the surface of the crystal, embedded in a polarizable dielectric medium. In the limit that the surrounding dielectric constant becomes infinite (i.e., is a conductor), the energy of the surface dipole layer vanishes, i.e.,

$$\lim_{\epsilon\to\infty} \alpha(P,\epsilon) = 0 \quad (7)$$

In the literature, this is often referred to as employing "tinfoil" boundary conditions.[33] A number of studies have investigated whether the inclusion of the surface term is a physically reasonable model for macroscopic systems[22,34] (since in a real crystals instantaneous microscopic fluctuations of the unit cell dipole moment are not propagated synchronously to the macroscopic limit). For most simulations, this terms has little overall effect and is generally neglected. Consequently, this term will not be further discussed, and henceforth the assumption will be made that the sum in reciprocal space can be made neglecting the $|\mathbf{k}| = 0$ term (and hence the surface energy correction).

It is useful to note that the Ewald sum energy, for a neutral system, can be written in terms of a pair potential $\psi_E(\mathbf{r}_{ij})$ as[11,31]

$$E_{\text{elec}} = \frac{1}{2}\sum_i^N \sum_j^N q_i q_j \psi_E(\mathbf{r}_{ij}) \quad (8)$$

where the Ewald pair potential, $\psi_E(\mathbf{r}_{ij})$, (assuming the Ewald parameter κ is chosen such that the summation over real-space lattice vectors includes only the $|\mathbf{n}| = 0$ term) is given by

$$\psi_E(\mathbf{r}_{ij}) = \left(\frac{(1-\delta_{ij})\text{erfc}(\kappa|\mathbf{r}_{ij}|)}{|\mathbf{r}_{ij}|} - \frac{2\kappa}{\sqrt{\pi}}\delta_{ij} + \frac{4\pi}{V}\sum_{|\mathbf{k}|\neq 0}\frac{\exp(-k^2/4\kappa^2)}{k^2}\cos(\mathbf{k}\cdot\mathbf{r}_{ij})\right) \quad (9)$$

In fact, efficient Ewald sum algorithms have been designed that precompute the Ewald pair potential on a 3-dimensional grid and use multidimensional interpolation procedures to allow rapid evaluation in molecular dynamics simulations.[18,28,35] This procedure, although fast, still scales as order N^2 and hence becomes limiting for large systems. However, as will be seen shortly, for hybrid QM/MM calculations where the QM part of the system is small, and update of the QM contribution to the Ewald energy is required at each step of an SCF procedure, the use of a correction to the Ewald pair potential becomes computationally efficient.

2.2. Electrostatic Energy of a Periodic System with a Smooth Charge Density. The focus of the present paper is to develop a linear-scaling method for efficient calculation of electrostatic interactions specifically for hybrid QM/MM calculations. The case of QM/MM calculations is somewhat specialized in that the quantum mechanical region is typically fairly small in relation to the much larger surrounding molecular mechanical environment. It would be considerably costly to Fourier transform directly the localized QM density that would require many reciprocal space lattice vectors (or alternatively, a very fine fast Fourier transform grid) in a typically very large QM/MM unit cell. On the other hand, for semiempirical QM models, there are very efficient methods for solution of the Poisson equation for the QM charge distribution in real space.[36,37] As will be discussed in more detail below, the object of the present work is to develop a method that takes advantage of the specialized features of QM/MM calculations and capitalizes simultaneously on the most efficient methods for calculating electrostatics of point charge and smooth density distributions.

To facilitate development of the method, the following general notation is introduced for the electrostatic interaction energy between two generalized charge distributions, Q_A and Q_B, under real-space nonperiodic boundary conditions (RS) and periodic boundary conditions (PB) as

$$E^X[A,B] = \frac{2-\delta_{A,B}}{2}\int\int Q_A(\mathbf{r})G^X(\mathbf{r},\mathbf{r}')Q_B(\mathbf{r}')d^3\mathrm{r}d^3\mathrm{r}' \quad (10)$$

where $G^X(\mathbf{r},\mathbf{r}')$ is the generalized Green's function for the Poisson equation that is a solution of

$$\nabla_{\mathbf{r}}^2 G(\mathbf{r},\mathbf{r}') = -4\pi\delta(\mathbf{r}-\mathbf{r}') \quad (11)$$

and the superscript "X" in eq 10 specifies the boundary conditions, which can be either "RS" or "PB" for real-space or periodic boundary conditions, respectively. The term $\delta_{A,B}$ is equal to 1 when the charge distributions are the same, and zero when the charge distributions are different. In the case that Q_A and Q_B are the same charge distribution and that charge distribution contains point charges, it is further assumed that the infinite self-energy of the point charges are neglected. In short, $E^X[A,B]$ in eq 10 represents the normal classical electrostatic energy of the charge distribution Q_A interacting with the charge distribution Q_B, including the possibility that Q_A and Q_B are identical. Note also that, by this definition, $E^X[A,B] = E^X[B,A]$ and $E^X[A+B, A+B] = E^X[A,A] + E^X[A,B] + E^X[B,B]$.

For a QM/MM calculation, the charge distribution is partitioned into a QM charge distribution that consists of the quantum mechanical electron density and nuclear core charges and an MM charge distribution that consists of the partial atomic charges of the MM environment. The total

energy of the system, under periodic boundary conditions, is thus given by

$$E^{\mathrm{PB}}[\boldsymbol{\rho}+\mathbf{q},\boldsymbol{\rho}+\mathbf{q}] = E^{\mathrm{PB}}[\boldsymbol{\rho},\boldsymbol{\rho}] + E^{\mathrm{PB}}[\boldsymbol{\rho},\mathbf{q}] + E^{\mathrm{PB}}[\mathbf{q},\mathbf{q}] \quad (12)$$

where ρ represents the distribution of electron density ρ of the QM atoms (plus the core nuclear charges) and $\mathbf{q}$ represents the distribution of classical MM point charges. Recall that the number of atoms associated with the QM charge distribution (N_{QM}) is typically much smaller than the number of atoms associated with the MM charge distribution (N_{MM}) and that the latter distribution spans a much greater spatial extent. The main problem to overcome involves the calculation of the electrostatics in the $E^{\mathrm{PB}}[\boldsymbol{\rho},\boldsymbol{\rho}]$ and $E^{\mathrm{PB}}[\boldsymbol{\rho},\mathbf{q}]$ terms, since, for semiempirical methods, these terms involve a smooth charge density with high atomic multipolar character. However, the same electrostatic interactions for this term are straightforward to calculate in real-space and, in fact, are part of the computational machinery of any stand-alone or integrated semiempirical quantum method. This observation motivates rewriting the expression of eq 12 as

$$E^{\mathrm{PB}}[\boldsymbol{\rho}+\mathbf{q},\boldsymbol{\rho}+\mathbf{q}] = (E^{\mathrm{PB}}[\boldsymbol{\rho},\boldsymbol{\rho}] - E^{\mathrm{RS}}[\boldsymbol{\rho},\boldsymbol{\rho}]) + E^{\mathrm{RS}}[\boldsymbol{\rho},\boldsymbol{\rho}] + (E^{\mathrm{PB}}[\boldsymbol{\rho},\mathbf{q}] - E^{\mathrm{RS}}[\boldsymbol{\rho},\mathbf{q}]) + E^{\mathrm{RS}}[\boldsymbol{\rho},\mathbf{q}] + E^{\mathrm{PB}}[\mathbf{q},\mathbf{q}] \quad (13)$$

Consider now an approximate quantum mechanical charge distribution, $\mathbf{Q}$, that is modeled as a set of auxiliary point charges such that the electrostatic potential closely represents that of the full QM charge distribution at distances on the order of the distance between crystal images. In the present work, simple Mulliken charges[38] are used for this purpose. However, alternate charge partitioning,[39−41] charge mapping,[42−45] or charge fitting[46,47] procedures could also be used as well. The Mulliken charge is particularly convenient to incorporate into the Fock operator, as is discussed below, owing to the simple linear relation with the single-particle density matrix. At short range, the potential due to these charges will deviate significantly from the exact quantum mechanical potential, but at distances on the order of a full unit cell translation away, the differences are very small. This motivates introduction of the following approximation for eq 13 as

$$E^{\mathrm{PB}}[\boldsymbol{\rho}+\mathbf{q},\boldsymbol{\rho}+\boldsymbol{q}] \approx (E^{\mathrm{PB}}[\mathbf{Q},\mathbf{Q}] - E^{\mathrm{RS}}[\mathbf{Q},\mathbf{Q}]) + E^{\mathrm{RS}}[\boldsymbol{\rho},\boldsymbol{\rho}] + (E^{\mathrm{PB}}[\mathbf{Q},\mathbf{q}] - E^{\mathrm{RS}}[\mathbf{Q},\mathbf{q}]) + E^{\mathrm{RS}}[\boldsymbol{\rho},\mathbf{q}] + E^{\mathrm{PB}}[\mathbf{q},\mathbf{q}] \quad (14)$$

The above equation for the Ewald energy is useful for practical implementation into semiempirical QM/MM methods. It is clear, for example, that eq 14 requires evaluation of the periodic boundary energy only for the point charge distributions, $\mathbf{Q}$ and $\mathbf{q}$, whereas the more complicated exact quantum mechanical charge distribution, ρ, is required to be evaluated in real space.

2.3. Combined QM/MM Potential in Real Space. Combined QM/MM potential methods have been reviewed extensively elsewhere[5,6,48] and are only briefly outlined in this subsection. The effective Hamiltonian for the combined QM/MM potentials treated in the present work take the form

$$\hat{H}_{\mathrm{eff}} = \hat{H}^{0}_{\mathrm{QM}} + \hat{H}^{\mathrm{el}}_{\mathrm{QM/MM}} + \hat{H}^{\mathrm{vdW}}_{\mathrm{QM/MM}} + \hat{H}_{\mathrm{MM}} \quad (15)$$

where $\hat{H}^{0}_{\mathrm{QM}}$ is the Hamiltonian for the QM charge distribution represented as nuclei and electrons within the Born−Oppenheimer approximation and $\hat{H}_{\mathrm{MM}}$ is the molecular mechanical potential of MM atoms. Two coupling terms, $\hat{H}^{\mathrm{el}}_{\mathrm{QM/MM}}$ and $\hat{H}^{vdW}_{QM/MM}$, represent interactions between QM and MM sites: $\hat{H}^{\mathrm{el}}_{\mathrm{QM/MM}}$ accounts for the electrostatic interactions of electrons and nuclei on QM atoms with point charges on MM sites, and $\hat{H}^{\mathrm{vdW}}_{\mathrm{QM/MM}}$ represents the short-range Pauli exchange repulsion and the long-range dispersion interactions and is modeled by a Lennard-Jones form. In real-space calculation, which is modified below to include long-range electrostatic interactions, the electrostatic interaction Hamiltonian, $\hat{H}^{\mathrm{el}}_{\mathrm{QM/MM}}$, given in eq 15 is written, in atomic units, as an exact interaction Hamiltonian of QM nuclei and electrons with MM atoms represented by partial point charges

$$\hat{H}^{\mathrm{el-RS}}_{\mathrm{QM/MM}} = -\sum_{i}^{N_{\mathrm{MM}}}\sum_{a}^{N_{\mathrm{e}}}\frac{q_i}{r_{ia}} + \sum_{i}^{N_{\mathrm{MM}}}\sum_{\alpha}^{N_{\mathrm{QM}}}\frac{q_i Z_\alpha}{R_{i\alpha}} \quad (16)$$

where q_i and Z_α are charges on MM and QM nuclei, N_{e} is the total number of electrons in QM region, and r_{ia} and $R_{i\alpha}$ are the distances of the quantum electrons and nuclei from the classical charge sites, respectively.

Then, the real-space potential energy in the combined QM/MM potential is computed using eq 17

$$E^{\mathrm{RS}}[\boldsymbol{\rho}] = \langle\Phi|\hat{H}_{\mathrm{eff}}|\Phi\rangle = E_{\mathrm{QM}}[\boldsymbol{\rho}] + E^{\mathrm{el-RS}}_{\mathrm{QM/MM}}[\boldsymbol{\rho},\mathbf{q}] + E^{\mathrm{vdW-RS}}_{\mathrm{QM/MM}} + E_{\mathrm{MM}} \quad (17)$$

where Φ is the wave function of the solute in the field of MM environment, in which $E_{\mathrm{QM}} + E^{\mathrm{el-RS}}_{\mathrm{QM/MM}}$ is determined through Hartree−Fock self-consistent-field (SCF) MO calculation by solving the Roothaan-Hall equation[49]

$$\mathbf{F}^{\mathrm{RS}}\mathbf{C}^{\mathrm{RS}} = \mathbf{S}\mathbf{C}^{\mathrm{RS}}\mathbf{E}^{\mathrm{RS}} \quad (18)$$

where $\mathbf{F}^{\mathrm{RS}}$, $\mathbf{C}^{\mathrm{RS}}$, and $\mathbf{S}$ denote the Fock, eigenvectors, and overlap matrices in real space, respectively, and $\mathbf{E}^{\mathrm{RS}}$ is the diagonal matrix of orbital energies for molecular orbitals. The combined QM/MM potential constructed in this way is expected to be valid in the range that the QM model is adequately large so as to capture the essential chemical reaction process, and the MM model provides a sufficiently accurate representation of the electrostatic environment.

2.4. Ewald Modifications to the Effective Hamiltonian (Fock) Matrix Elements. The elements of the effective Hamiltonian (Fock) matrix, $F_{\mu\nu}$, in a periodic boundary system are defined as

$$F_{\mu\nu} = \frac{\delta E[\rho]}{\delta \rho_{\mu\nu}} \quad (19)$$

where $E[\rho]$ is the total energy that depends on the single-particle density matrix, ρ, with elements $\rho_{\mu\nu}$ and is related to the terms in eq 14 by

$$E[\rho] = (E^{\mathrm{PB}}[\mathbf{Q},\mathbf{Q}] - E^{\mathrm{RS}}[\mathbf{Q},\mathbf{Q}]) + E^{\mathrm{RS}}[\boldsymbol{\rho},\boldsymbol{\rho}] + (E^{\mathrm{PB}}[\mathbf{Q},\mathbf{q}] - E^{\mathrm{RS}}[\mathbf{Q},\mathbf{q}]) + E^{\mathrm{RS}}[\boldsymbol{\rho},\mathbf{q}] \quad (20)$$

This energy can be decomposed into QM and QM/MM

components, each of which consists of a real-space term plus a *periodic boundary correction* (PBC) as

$$E[\rho] = E^{\mathrm{RS}}_{\mathrm{QM}}[\boldsymbol{\rho}] + \Delta E^{\mathrm{PBC}}_{\mathrm{QM}}[\mathbf{Q}] + E^{\mathrm{RS}}_{\mathrm{QM/MM}}[\boldsymbol{\rho}] + \Delta E^{\mathrm{PBC}}_{\mathrm{QM/MM}}[\mathbf{Q}]$$
$$= E^{\mathrm{RS}}[\boldsymbol{\rho}] + \Delta E^{\mathrm{PBC}}[\mathbf{Q}] \quad (21)$$

where

$$E^{\mathrm{RS}}_{\mathrm{QM}}[\boldsymbol{\rho}] = E^{\mathrm{RS}}[\boldsymbol{\rho},\boldsymbol{\rho}] \quad (22)$$

$$\Delta E^{\mathrm{PBC}}_{\mathrm{QM}}[\mathbf{Q}] = (E^{\mathrm{PB}}[\mathbf{Q},\mathbf{Q}] - E^{\mathrm{RS}}[\mathbf{Q},\mathbf{Q}]) \quad (23)$$

$$E^{\mathrm{RS}}_{\mathrm{QM/MM}}[\boldsymbol{\rho}] = E^{\mathrm{RS}}[\boldsymbol{\rho},\mathbf{q}] \quad (24)$$

$$\Delta E^{\mathrm{PBC}}_{\mathrm{QM/MM}}[\mathbf{Q}] = (E^{\mathrm{PB}}[\mathbf{Q},\mathbf{q}] - E^{\mathrm{RS}}[\mathbf{Q},\mathbf{q}]) \quad (25)$$

$$E^{\mathrm{RS}}[\boldsymbol{\rho}] = E^{\mathrm{RS}}_{\mathrm{QM}}[\boldsymbol{\rho}] + E^{\mathrm{RS}}_{\mathrm{QM/MM}}[\boldsymbol{\rho}] \quad (26)$$

$$\Delta E^{\mathrm{PBC}}[\mathbf{Q}] = \Delta E^{\mathrm{PBC}}_{\mathrm{QM}}[\mathbf{Q}] + \Delta E^{\mathrm{PBC}}_{\mathrm{QM/MM}}[\mathbf{Q}] \quad (27)$$

Note that although the energy $E[\rho]$ depends on the density matrix ρ in a fairly complicated way, the periodic boundary correction terms (eqs 23 and 25) depend on the density matrix only through the atomic charge vector $\mathbf{Q}$. The Fock matrix can similarly be decomposed into a real-space term plus a periodic boundary correction as

$$F^{\mathrm{PB}}_{\mu\nu} = F^{\mathrm{RS}}_{\mu\nu} + \Delta F^{\mathrm{PBC}}_{\mu\nu} \quad (28)$$

where

$$F^{\mathrm{RS}}_{\mu\nu} = \frac{\delta}{\delta \boldsymbol{\rho}_{\mu\nu}}\{E^{\mathrm{RS}}_{\mathrm{QM}}[\boldsymbol{\rho}] + E^{\mathrm{RS}}_{\mathrm{QM/MM}}[\boldsymbol{\rho}]\} = \frac{\delta E^{\mathrm{RS}}[\boldsymbol{\rho}]}{\delta \boldsymbol{\rho}_{\mu\nu}} \quad (29)$$

$$\Delta F^{\mathrm{PBC}}_{\mu\nu} = \frac{\delta}{\delta \rho_{\mu\nu}}\{\Delta E^{\mathrm{PBC}}_{\mathrm{QM}}[\mathbf{Q}] + \Delta E^{\mathrm{PBC}}_{\mathrm{QM/MM}}[\mathbf{Q}]\} = \frac{\delta \Delta E^{\mathrm{PBC}}[\mathbf{Q}]}{\delta \rho_{\mu\nu}} \quad (30)$$

For the purposes of the present work, only the *correction* to the Fock matrix ($\Delta F^{\mathrm{PBC}}_{\mu\nu}$) that arises from introduction of the periodicity is described, because it is presumed that the machinery for construction of the complete Fock matrix in real space ($F^{\mathrm{RS}}_{\mu\nu}$) is already available.[36,50] For the purposes of implementation, the present work formulates extension to periodic systems as an additional term that can be included by an auxiliary computer subroutine or module.

The task that remains is to write the periodic boundary correction to the energy in eq 27 in terms of the set of Mulliken charges[38] $\{Q_\alpha\}$, defined in NDDO-based semiempirical methods as

$$Q_\alpha = Z_\alpha - \sum_{\mu\in\alpha} \rho_{\mu\mu} \quad (31)$$

where α is an atom index, in a manner that is efficient to calculate and update during the self-consistent field (SCF) procedure. As mentioned previously, other charge partitioning,[39−41] charge mapping,[42−45] or charge fitting[46,47] methods may also be employed so long as a rigorous mapping to the single-particle density matrix can be affected such that the Fock matrix may be modified accordingly. From the periodic correction to the energy ($\Delta E^{\mathrm{PBC}}[\mathbf{Q}]$ of eq 27), the chain relation is used to obtain the periodic correction to the Fock matrix elements as

$$\frac{\delta \Delta E^{\mathrm{PBC}}[\mathbf{Q}]}{\delta \rho_{\mu\nu}} = \sum_{\alpha\in\mu,\nu} \frac{\delta Q_\alpha}{\delta \rho_{\mu\nu}} \cdot \frac{\partial \Delta E^{\mathrm{PBC}}[\mathbf{Q}]}{\partial Q_\alpha} = - \sum_{\alpha\in\mu,\nu} \delta_{\mu\nu} \cdot \frac{\partial \Delta E^{\mathrm{PBC}}[\mathbf{Q}]}{\partial Q_\alpha} \quad (32)$$

To maximize efficiency of the method, one must bear in mind that the number of MM atoms, N_{MM}, usually greatly exceeds the number of quantum atoms, N_{QM}. Moreover, at each molecular dynamics integration step, the calculation of the QM/MM total energy and gradient requires an SCF procedure to be performed, and consequently, the periodic potential due to the QM charge distribution must be updated (recalculated) at each SCF iteration. Note that the atomic positions of all the atoms in the system remain fixed during the SCF procedure.

Consider the first periodic energy correction term, $\Delta E^{\mathrm{PBC}}_{\mathrm{QM}}[\mathbf{Q}]$, of eq 23. If the convention is used that the Ewald parameter κ is chosen such that the summation over real-space lattice vectors includes only the $|\mathbf{n}| = 0$ term, the first periodic energy correction term can be written concisely as

$$\Delta E^{\mathrm{PBC}}_{\mathrm{QM}}[\mathbf{Q}] = \frac{1}{2}\sum_{\alpha}^{N_{\mathrm{QM}}}\sum_{\beta}^{N_{\mathrm{QM}}} Q_\alpha Q_\beta \Delta\psi_E(\mathbf{R}_{\alpha\beta}) \quad (33)$$

where $\Delta\psi_E(\mathbf{R}_{\alpha\beta})$ is the periodic correction to the Ewald pair potential (eq 9), and with the choice of κ above, is given by

$$\Delta\psi_E(\mathbf{R}_{\alpha\beta}) = \left(\frac{4\pi}{V}\sum_{|\mathbf{k}|\neq 0}\frac{\exp(-k^2/4\kappa^2)}{k^2}\cos(\mathbf{k}\cdot\mathbf{R}_{\alpha\beta}) - \frac{\mathrm{erf}(\kappa|\mathbf{R}_{\alpha\beta}|)}{|\mathbf{R}_{\alpha\beta}|}\right) \quad (34)$$

Note that in the derivation of eq 34 the limiting relation was used

$$\lim_{r\to 0}\frac{\mathrm{erf}(\kappa r)}{r} = \frac{\kappa}{\sqrt{\pi}} \quad (35)$$

For the QM periodic correction term (eq 33), the corresponding correction to the Fock matrix must be recalculated at each step of the SCF procedure. Since the number of quantum atoms is small, the correction to the Ewald pair potential (eq 34) needed for the Fock matrix can easily be calculated once as an $N_{\mathrm{QM}} \times N_{\mathrm{QM}}$ matrix and stored and hence not be recalculated during the SCF. This makes calculation of the periodic correction to the Fock matrix consist of a simple matrix multiplication of the Ewald pair potential correction with the Mulliken charge vector, only the latter of which changes at each iteration.

Consider now the second periodic energy correction term, $\Delta E^{\mathrm{PBC}}_{\mathrm{QM/MM}}[\mathbf{Q}]$, of eq 30 (see also eq 25) that can be written as

$$\Delta E^{\mathrm{PBC}}_{\mathrm{QM/MM}}[\mathbf{Q}] = \sum_{\alpha}^{N_{\mathrm{QM}}} Q_{\alpha} \sum_{j}^{N_{\mathrm{MM}}} q_j \Delta\psi_E(\mathbf{R}_{\alpha} - \mathbf{r}_j) \qquad (36)$$

In this case, the corresponding correction to the Fock matrix does *not* need to be updated during the SCF procedure, since the MM charge distribution is not changing. Consequently, the periodic correction to the static potential of the MM charges at the QM charge positions can be calculated once as a $N_{\mathrm{QM}} \times 1$ vector and simply added to the 1-electron terms of the Fock matrix (sometimes referred to as the *core* Hamiltonian matrix). The above procedure leads to an efficient method for calculation of long-range electrostatic interactions in combined QM/MM calculations.

3. Computational Details

The combined QM/MM-Ewald sum method has been implemented into a modified version of CHARMM[51] (version c30a1) interfaced with the MNDO97 program[52] and MOPAC.[53] To test the method introduced in the present work, simulations of ion association processes and dissociative mechanisms of phosphoryl transfer were performed. The interionic potential of mean force was calculated for each simulation and compared with calculated values from nonperiodic (full electrostatic or cutoff) and periodic (cutoff) simulations. In particular, the following systems were examined: the ionic association of (1) ammonium chloride and (2) ammonium metaphosphate and the dissociative phosphoryl transfer mechanism of (3) methyl phosphate and (4) acetyl phosphate. The semiempirical AM1 model[50] was used for the ammonium chloride system, and MNDO/d[54] for the phosphorus-containing systems for which *d*-orbitals have been shown to be important.[55] The solutes were treated fully quantum mechanically at the semiempirical level and were solvated in a 40.0 Å cubic box of TIP3P water molecules,[56] resulting in a total of 2042 waters for the ammonium chloride system, 2040 waters for the ammonium metaphosphate system, 2038 waters for the methyl phosphate system, and 2035 waters for the acetyl phosphate system. Internal water geometries were constrained using the SHAKE algorithm in all simulations.[57]

A spherical cutoff scheme was used to evaluate the van der Waals and the real-space electrostatic interactions in the QM/MM-Ewald method and in the non-Ewald approaches. In all cases, water molecules were included in the cutoff list if the geometrical center of water was less than the cutoff distance from any group center of the solute. The solute ion pair was divided into two groups, one for each formal ion. It should be emphasized that in evaluating QM/MM interaction energies, solute−solvent (or QM-MM) interactions were determined for the entire QM system, whenever a solvent molecule was within the cutoff distance from any solute (QM) group. In simulations using the QM/MM-Ewald sum technique, a 10.0 Å group-based cutoff was used. The nonbonded list and crystal images were updated every 25 steps during molecular dynamics simulations. For Ewald summation, the Ewald κ value (see above) was chosen to be 0.340 $\mathrm{Å}^{-1}$, and the smooth particle mesh Ewald (PME) method was employed for reciprocal space summations between MM sites with an approximate grid size of 0.8 Å ($50 \times 50 \times 50$ FFT grid)[27,31] and with net charge correction to the Ewald potential suggested by Bogusz et al.[33] All simulations were propagated using the leapfrog Verlet algorithm with 1 fs integration time step.[24] Periodic boundary conditions were used along with the isothermal−isobaric ensemble (NPT) at 1 atm and 298 K using extended system pressure algorithm of Andersen[58] with effective mass of 500.0 amu and Hoover thermostat[59] with effective mass of 1000.0 kcal/mol-ps^2, respectively. In the QM/MM simulations under PB without Ewald summation, electrostatic interactions were determined using a spherical cutoff scheme based on group separation with switching between 10.5 and 11.5 Å.

The potential of mean force (PMF) profiles have been determined using umbrella sampling,[60] in which PMF is represented as a function of internuclear distance defined as the N···Cl distance ($R_{\mathrm{N-Cl}}$) in ammonium chloride, the N···P distance ($R_{\mathrm{N-P}}$) in ammonium metaphosphate, and the O···P distance ($R_{\mathrm{O-P}}$) in methyl phosphate and acetyl phosphate. After initial 200 ps of equilibration, 25 separate umbrella sampling windows (28 windows in the simulations of dissociative phosphoryl transfer reactions) were executed to span the internuclear separation up to 12.0 Å by applying a harmonic restraining potential centered at the center of the particular umbrella window. The spacing between neighboring windows was a function of the interionic separation distance: 2.0, 2.5, and 5.0 Å spacings were used for interionic separation distances R in the range of $R \leq 3.0$ Å, 3.0 Å $\leq R \leq$ 4.0 Å, and $R \geq 4.0$ Å, respectively. The force constants used were chosen and adjusted based on the shape of PMF profile for each system tested to guarantee sufficient overlap of the probability distribution with neighboring windows (force constant values ranged between 80.0−150.0 kcal/mol-Å in the region of steep repulsive wall at small R, 40.0−80.0 kcal/mol-Å in the intermediate separation, and 10.0−40.0 kcal/mol-Å in the region of large interionic separation). Each umbrella sampling window was equilibrated for 35 ps followed by 50 ps of production with data collected every step. The weighted histogram analysis method (WHAM)[61] was employed to compute the potential of mean force as a function of internuclear separation.

For additional comparison of the ion association simulations, umbrella sampling simulations were also performed to compute interionic PMF profiles with stochastic boundary molecular dynamics (SBMD)[62,63] by using spherical water box for ammonium chloride with 1034 TIP3P water molecules and ammonium metaphosphate with 1031 TIP3P waters. Simulations were performed without cutoff as well as with a 11.5 Å nonbonded cutoff as in the periodic simulations. The radius of water sphere is 20.0 Å to keep the size of the simulation fairly close to that of the PB simulations.

Table 1. Relative Fluctuations in Potential Energy (σ_E), Volume (σ_V), Temperature (σ_T), and Average Total Force (F_{ave}^{tot}) from MD Simulations Using QM/MM-Ewald with PME for MM Electrostatics[a]

	σ_E	σ_V	σ_T	F_{ave}^{tot} [b]
$NH_4^+\cdots CL^-$	3.60×10^{-3}	5.52×10^{-3}	1.01×10^{-2}	2.55×10^{-2}
$NH_4^+\cdots PO_3^-$	3.49×10^{-3}	4.77×10^{-3}	0.99×10^{-2}	2.51×10^{-2}
$CH_3O^-\cdots PO_3^-$	3.75×10^{-3}	4.04×10^{-3}	1.05×10^{-2}	2.51×10^{-2}
$CH_3CO_2^-\cdots PO_3^-$	3.47×10^{-3}	5.42×10^{-3}	1.03×10^{-2}	2.54×10^{-2}

[a] $\sigma_E = \text{rms}_E/E_{ave}$, $\sigma_V = \text{rms}_V/V_{ave}$, and $\sigma_T = \text{rms}_T/T_{ave}$ where rms is the root mean square deviation of those properties. [b] The average total force goes to zero when regular Ewald sum method is used for MM electrostatics.

4. Results and Discussion

The focus of the current paper is to describe the development and implementation of an efficient QM/MM-Ewald method and provide benchmark simulation tests and to characterize the effects of treatment of electrostatic interactions on the QM/MM free energy profiles. Emphasis will not be placed on detailed analysis of the simulations and comparison with experiment that would first require a more quantitative assessment of the accuracy of the quantum and solvation models and QM/MM parameters. Development of new semiempirical quantum models for chemical reactions is an area of intense effort,[64−70] and consequently the extension of methods for efficient treatment of long-range electrostatic interactions in these calculations is of prime importance.

This section presents results of simulations of ion association and dissociative phosphoryl transfer using the QM/MM-Ewald method developed in the present work. The effects of periodicity and electrostatic cutoff are compared from PMF profiles of separate simulations. For the ion association tests (ammonium chloride and ammonium metaphosphate), PBMD and SBMD simulations are compared both with and without electrostatic cutoffs (the simulations without cutoffs are referred to as the QM/MM-Ewald and full-electrostatic SBMD for periodic and nonperiodic systems, respectively). For the phosphoryl transfer reactions (methyl phosphate and acetyl phosphate), PBMD simulations with and without electrostatic cutoff are compared (i.e., PBMD with cutoff and QM/MM-Ewald simulations). A summary of the observed fluctuations in potential energy, volume, temperature, and total force over a 10 ps interval for each of the QM/MM-Ewald simulations are shown in Table 1. The relative force errors for cutoff and alternative Ewald methods with respect to the QM/MM-Ewald method are compared in Table 2. A noteable feature is that inclusion of the Ewald term as a post-SCF correction still leads to considerable force error. The comparison underscores the importance of inclusion of long-range electrostatics in the simulation and updates of the periodic contribution to the quantum polarization during the SCF procedure since, although the post-SCF molecular mechanical Ewald correction of QM/MM calculation using Mulliken charge representation on quantum atoms improves the force error significantly, it still leads to considerable error.

4.1. Association of Oppositely Charged Ions. 4.1.1. Ammonium Chloride. Figure 1 compares the PMF profiles for ion association of ammonium chloride as a function of the N···Cl distance from simulations with and without electrostatic cutoffs under nonperiodic and periodic boundary conditions. The zeros of the PMFs were set to the minimum value for the ionic complex.

Table 2. Comparison of the Force Errors on the QM Atoms for Several Electrostatic Methods: Root Mean Square Error (RMSE), Mean Signed Error (MSE), Mean Unsigned Error (MUE) of Force, and Maximum Force Error (MAXE)[a]

	RMSE	MSE	MUE	MAXE
	Methyl Phosphate[b]			
10.0 Å cutoff	1.618	−0.010	1.221	4.612
11.5 Å cutoff	2.221	0.052	1.726	6.073
Post-SCF Ewald[c]	1.155	0.006	0.808	3.864
	Acetyl Phosphate[d]			
10.0 Å cutoff	2.836	−0.064	1.961	8.404
11.5 Å cutoff	1.830	−0.015	1.195	5.271
Post-SCF Ewald[c]	1.810	−0.009	1.138	6.642

[a] All units are kcal·mol^{-1}·Å^{-1}. [b] The forces are computed at the transition state of $R_{O-P} = 3.2$ Å. [c] The forces from Ewald potential are added to the forces computed from 10.0 Å cutoff, in which the Ewald potential has been computed from the Mulliken charges of the QM atoms. [d] The forces are computed at the transition state of $R_{O-P} = 2.9$ Å.

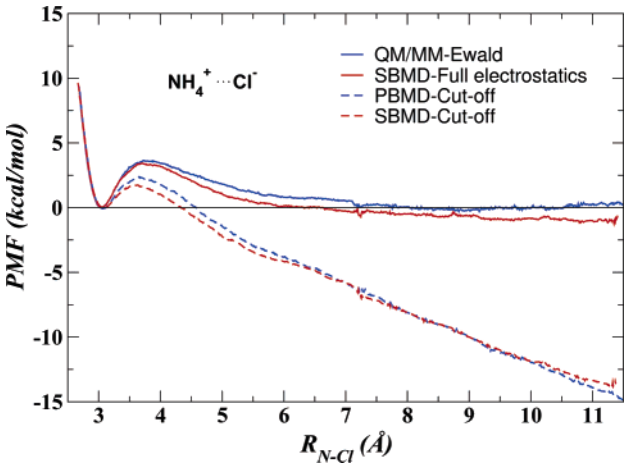


Figure 1. Comparison of potential of mean force (PMF) profiles for the ionic separation (R_{N-Cl}) of ammonium chloride ($NH_4^+\cdots Cl^-$) in water. Profiles were constructed from MD simulations with periodic boundary conditions using the combined QM/MM-Ewald sum (solid blue line) and PBMD method with 11.5 Å cutoff (dashed blue line) and with full-electrostatic SBMD (solid red line) and SMBD with 11.5 Å (dashed red line).

The electrostatic cutoff causes an artificial decrease in the PMF for oppositely charged ions at a large separation irrespective of whether the system is treated with spherical stochastic boundary (SBMD with cutoff) or with periodic boundary (PBMD with cutoff) conditions. Alternately, the full-electrostatic simulations (QM/MM-Ewald and full-

electrostatic SBMD) show the expected flattening of the PMF at large distances. These profiles are similar between the QM/MM-Ewald and the full-electrostatic SBMD simulations. The behavior of all the PMF profiles are similar up to an interion separation of around 3.0 Å, after which, the results of the cutoff simulations begin to diverge significantly from those of the QM/MM-Ewald and full-electrostatic SBMD simulations. The barrier for ionic dissociation is about 3.6 kcal/mol (QM/MM-Ewald), 3.4 kcal/mol (full-electrostatic SBMD), 2.3 kcal/mol (PBMD with cutoff), and 1.7 kcal/mol (SBMD with cutoff), respectively. The cutoff methods affect the energy barrier by over 1 kcal/mol at the transition state (a distance of less than 0.8 Å from the minimum) and have an even more profound effect at larger distances in the PMF profile.

At large separation, the PMF profiles of the simulations with long-range electrostatics become relatively flat after about 7.0 Å, indicating the ions are effectively shielded by the nonlocal solvent response. Alternately, the PMF profiles of the cutoff simulations show a steady linear drift from around 6.0 Å out past 10 Å. This linear drift of the PMF profiles for the cutoff simulations is due to an imbalance in the electrostatic interactions. The dipole moments of the waters solvating the individual oppositely charged ions are favorably aligned in the region between the ions and unfavorably aligned at opposite ends. The unfavorable interactions at the ends fall outside the cutoff first as the ions separate, while the interactions of the favorably aligned waters in the center are retained and result in the artificial drift in the PMF profiles. It is likely, therefore, that QM/MM simulations of biochemical reactions that involve association or dissociation of oppositely charged species, such as seen in many biochemical S_N1 reactions and photodissociation processes, may be subject to artificial overstabilization of the separated ionic species if electrostatic cutoffs are used.

4.1.2. Ammonium Metaphosphate. Figure 2 compares the PMF profiles for ion association of ammonium metaphosphate as a function of the N···P distance from simulations with and without electrostatic cutoffs under nonperiodic and periodic boundary conditions. The zeros of the PMFs were set to the limiting long-range value for the QM/MM-Ewald and full-electrostatic SBMD simulations and adjusted such that the short-ranged repulsive wall were coincident for the cutoff simulations. In these simulations, the MNDO/d Hamiltonian was employed since it has been demonstrated to provide a reliable description of biological phosphorus compounds.[54,55,70] Unlike the ammonium chloride PMF (Figure 1) that exhibits a stable free energy minimum for the ion−ion complex, the PMF profile for ammonium metaphosphate decreases monotonically. In the short range (N···P distances less than 4.0 Å), all of the PMF profiles are similar; however, after 4.0 Å, the PMF values for the cutoff simulations diverge from those of the QM/MM-Ewald and full-electrostatic SBMD simulations. The long ranged behavior of the PMF for the QM/MM-Ewald and full-electrostatic SBMD are quite similar, exhibiting a flat asymptotic limit after around 6.5 Å, indicating the oppositely charged ions are effectively screened. The PMF profiles for

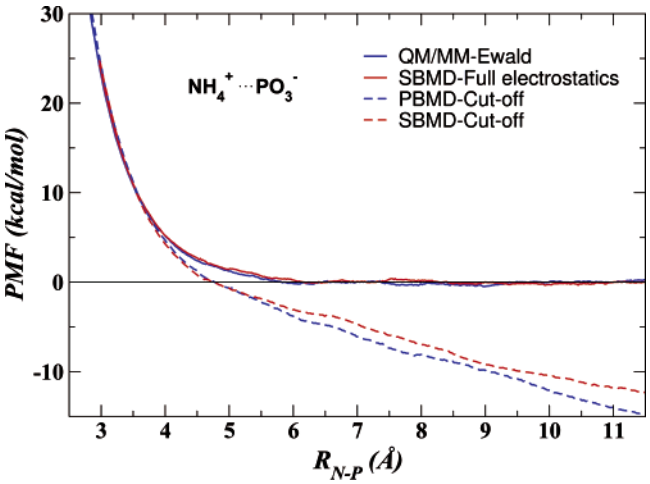


Figure 2. Comparison of potential of mean force (PMF) profiles for the ionic separation (R_{N-P}) of ammonium metaphosphate ($NH_4^+\cdots PO_3^-$) in water. Profiles were constructed from MD simulations with periodic boundary conditions using the combined QM/MM-Ewald sum (solid blue line) and PBMD simulations with 11.5 Å cutoff (dashed blue line) and with full-electrostatic SBMD (solid red line) and SMBD with 11.5 Å (dashed red line).

the PBMD and SBMD cutoff simulations show a linear drift after 6.0 Å, as in the ammonium chloride case (Figure 1). It is likely that the metaphosphate plane has a random orientation relative to the ammonium ion at large separation, but, as the ions approach each other, the plane is aligned perpendicular to the N···P vector. At this short separation, the partial positive nature of phosphorus atom cancels the favorable interactions between ammonium and oxygens in metaphosphate and results in no stable ion−ion complex. The results of these simulations echo those for the ammonium chloride system: QM/MM simulations of reactions that involve the dissociation of oppositely charged ions may lead to separated ionic species that are significantly overstabilized.

4.1.3. Effect of Cutoff on the Association of Oppositely Charged Ions. The effect of treatment of electrostatic interactions for the association of oppositely charged ions has been studied previously, although to our knowledge, not with the same QM/MM-Ewald model as presented in the present work. At long range, both ion association PMF profiles exhibit an artificial linear drift using either PBMD or SBMD with cutoff. The slope of the linear drift from 6 to 10.0 Å is similar between the PBMD and SBMD cutoff simulations with values of −2.25 and −1.66 kcal/mol-Å, respectively, for ammonium chloride and −2.01 and −1.86 kcal/mol-Å, respectively, for ammonium metaphosphate. For opposite-charged ionic systems, the results from current simulations indicate an unphysical roughly linear downward drift in the PMF profiles as the ions separate. Rozanska and Chipot also observed a similar artifact for the PMF profile of guanidinium-acetate association from molecular dynamics simulations using a smoothed electrostatic cutoff, whereas the simulations using the Ewald sum showed the expected solvent shielded behavior of solvent-separated ion pair, and the generalized reaction field correction also significantly

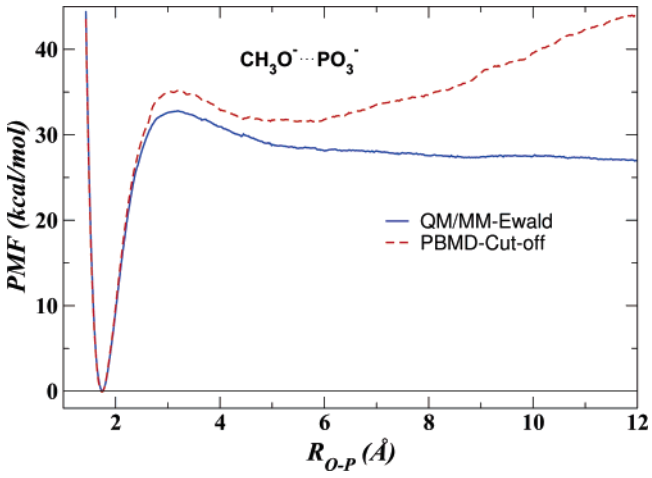


Figure 3. The computed PMF from PBMD simulations using combined QM/MM-Ewald sum potential (solid blue line) and with 11.5 Å cutoff (dashed red line) for the dissociation of methyl phosphate ($CH_3O^-\cdots PO_3^-$) in water.

improved the description relative to the cutoff method except toward the edge of the cutoff sphere.[71]

4.2. Dissociative Phosphoryl Transfer Mechanisms. The nonenzymatic and enzymatic chemical mechanism of phosphate hydrolysis reactions remains a topic of discussion and considerable debate.[72] Phosphate hydrolysis reactions are often discussed in terms of their associative or dissociative character[73] that can sometimes be distinguished kinetically. Kinetic measurements provide crucial data for these reactions, although they do not always provide a unique mechanistic interpretation.[74,75] The associative versus dissociative character is governed by many factors including the degree of esterification of the phosphate, the protonation state, the nature of the leaving group, and interactions with solvent, ions, and macromolecular environment.

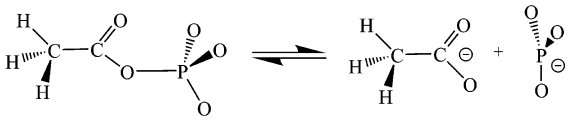
To test the QM/MM-Ewald method, the dianionic dissociative phosphoryl transfer pathways for methyl phosphate and acetyl phosphate were examined using PBMD simulations both with electrostatic cutoff and with the QM/MM-Ewald method. Methyl phosphate is a commonly employed model for phosphoryl transfer reactions in kinases and phosphatases, and acetyl phosphate represents a model for a high-energy intermediate in the metabolism of many bacteria.[76] The first step of the dissociative pathway (Scheme 1) involves a dephosphorylation step characterized by the departure of a solvated metaphosphate (PO_3^-) group. The second step of the reaction involves the nucleophilic substitution to the metaphosphate by a nucleophile (usually a hydroxide ion or a water molecule in solution). This type of reaction is referred to as a $D_N + A_N$ type mechanism in IUPAC nomenclature.[77]

4.2.1. Methyl Phosphate. Figure 3 shows the computed PMF profile for methyl phosphate. The free energy of dissociation from current PMF profiles with QM/MM-Ewald method is 27.1 kcal/mol for methyl phosphate, and the activation free energy barrier is 32.8 kcal/mol. The effect of cutoff in the PBMD simulations is to raise the activation free energy barrier to 35.4 kcal/mol (an increase of 2.6 kcal/mol, or 8%). The effect of cutoff is even more pronounced

Scheme 1. Dissociative Phosphoryl Transfer Mechanism

Methyl phosphate

Acetyl phosphate



on the free energy of dissociation due to cutoff artifacts of the like-charged ions at fairly large separation (see below).

There have been several studies from experiment and theory for the dissociative reaction of methyl phosphate. The reaction free energy estimated from experiment by Guthrie is 37 ± 3 kcal/mol.[75,78] This indicates the reaction free energy calculated with the present work may be as much as 10 kcal/mol in error. Although it is not the purpose here to present free energy profiles with the greatest accuracy, it is worthwhile to point out the likely sources of error in order to assist in the development of improved QM/MM models. The main sources of error involve the semiempirical quantum model itself, the simplistic molecular mechanical model for water, and the QM/MM van der Waals interactions. The latter has a tremendous effect on the reaction free energies and barrier heights for processes that involve ion association/dissociation. For example, the heat of formation of methoxide (CH_3O^-) ion computed from MNDO/d Hamiltonian gives −39.7 kcal/mol in the gas phase,[36,54] while the experimentally determined value is −32.2 ± 1.1 kcal/mol.[79] The error in the methoxy ion from semiempirical MNDO/d model alone is 7.5 kcal/mol, which is close to the 10 kcal/mol difference between computed reaction free energy and experiment. A promising approach toward improvement of semiempirical quantum models is to develop reaction-specific parameters[70] that closely reproduce high-level quantum results.[80,81] Additionally, the model for water that was employed lacks explicit electronic polarizability, which is expected to be important for the stabilization of highly ionic systems such as those studied here. Finally, the optimization of the QM/MM van der Waals radii[82−84] to reproduce correct relative solvation free energies is critical. All of these areas will be addressed in future work.

4.2.2. Acetyl Phosphate. Figure 4 shows the computed PMF profiles for acetyl phosphate. The free energy of dissociation from current PMF profiles with QM/MM-Ewald method is 6.8 kcal/mol for acetyl phosphate, and the activation free energy barrier is 12.2 kcal/mol. The effect of cutoff in the PBMD simulations raised the activation free energy barrier to 13.8 kcal/mol (an increase of 1.6 kcal/mol or 13%). As for the dissociative phosphoryl transfer of methyl phosphate, the effect of cutoff has an even more profound effect on the free energy of dissociation (see below).

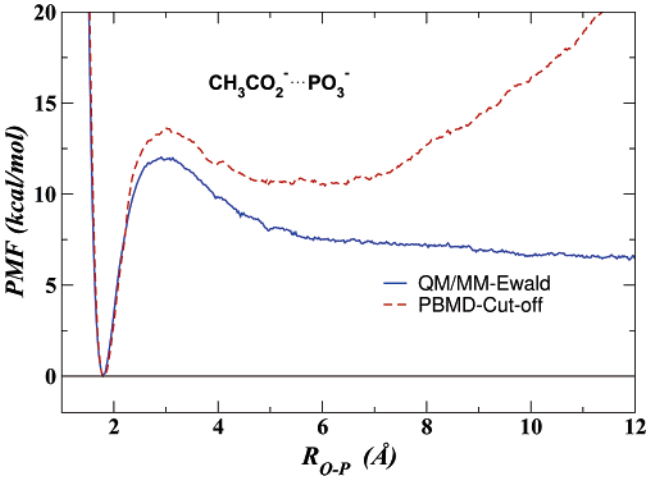


Figure 4. The computed PMF from PBMD simulations using combined QM/MM-Ewald sum potential (solid blue line) and with 11.5 Å cutoff (dashed red line) for the dissociation of acetyl phosphate ($CH_3CO_2^-\cdots PO_3^-$) in water.

The dissociative free energy activation barrier for acetyl phosphate is predicted to be 20.6 kcal/mol less than that for methyl phosphate. This is largely due to the increased stability of the acetate anion in solution, which has a much lower pK_a value (4.8) than methanol (15.5),[85] and is a considerably better leaving group. The reaction free energy for acetyl phosphate is similarly predicted to be lower than that of methyl phosphate by 20.3 kcal/mol. Thus, the lowering of the activation barrier can be explained by the added stabilization of the acetate anion relative to the methoxide anion, in accord with the Hammond postulate.[86,87]

4.2.3. Effect of Cutoff on the Dissociation of Like Charged Ions. The free energy of dissociation from current PMF profiles with Ewald sum is 27.1 kcal/mol for methyl phosphate and 6.8 kcal/mol for acetyl phosphate, respectively. The barrier height is 32.8 kcal/mol for methyl phosphate and 12.2 kcal/mol for acetyl phosphate from simulations with Ewald sum, while it is 35.4 kcal/mol and 13.8 kcal/mol with the cutoff method, respectively. The radial distribution function of water molecule around the solute at the transition state of both reactions has been checked, but no significant differences between QM/MM-Ewald and cutoff simulations were observed (data not shown). Thus, the difference in the barrier heights of dissociation reactions can most likely be attributed to long-range electrostatic effects that involve the ions and solvent.

The PMF profiles of the like-charged ionic systems of the present work are nonmonotonic and exhibit a broad minimum between 5 and 7 Å. Nonmonotonic PMF profiles arising from the use of electrostatic cutoffs have been previously observed by Bader and Chandler[88] in the dissociation of ferric and ferrous ion models in aqueous solution, whereas with the use of Ewald sums, correct monotonic PMF profiles were obtained. Furthermore, a comparison of the spatial distribution functions of like ion pairs have been investigated by Dang and Pettitt using molecular dynamics simulation and integral equation theory.[89] The results suggest the existence of a minimum for $Cl^-\text{—}Cl^-$ pair at close distances and also a slight minimum at larger distance (6—7 Å) using a cutoff method with switching.[89] Del Buono et al. studied $Cl^-\text{—}Cl^-$ pair and ferric and ferrous ion pair by computing solvent dielectric response[90] and also found an artificial minimum at large separation by using electrostatic cutoff methods as Dang and Pettitt,[89] in which the smooth truncation of electrostatics even amplifies this artifact. Alternately, the PMF profiles from simulations using Ewald sums produced the correct high dielectric shielding of the ions by water and removed the artificial minimum at large separation.

5. Conclusion

The current paper presents an extension of Ewald summation method to combined QM/MM calculations with semiempirical quantum models. The method is tailored to systems where the number of quantum atoms is small compared to the number of molecular mechanical atoms such that the Ewald contribution to the Fock matrix elements can be evaluated efficiently during the self-consistent field procedure required at each step of a molecular dynamics calculation. The method is based on a partition of the total Ewald potential into a short-ranged real-space interaction and a long-range periodic correction. The periodic correction term requires only a Mulliken charge representation of the charge density and hence can be used with any efficient linear-scaling Ewald method for point charge (or multipolar) systems, such as the particle mesh Ewald method. If the number of quantum atoms is sufficiently small, a considerable reduction in computational cost can be achieved through direct computation of the Ewald pair potential correction for only the quantum atoms such that the periodic correction to the electrostatic energy can be efficiently affected at each SCF iteration by a simple matrix multiplication with the Mulliken charge vector. Although the method is applied with semiempirical quantum methods in the present work, the methodology can be extended to other quantum models such as density-functional methods without significant code modifications. The implementation and performance of the method is tested in simulations of ion—ion association and on dissociative phosphoryl transfer reactions. The PMF profiles from these simulations are compared with those of full-electrostatic SBMD simulations and PBMD and SBMD simulations with electrostatic cutoff. Significant artifacts arise in the reaction free energies and activation barriers when a cutoff is used. These artifacts vanish when the QM/MM-Ewald method is employed. Despite the known problems associated with the use of electrostatic cutoffs, the majority of present day applications of QM/MM methods routinely employ cutoffs in simulations of biological reaction dynamics. Consequently, the present method offers an important extension of linear-scaling Ewald techniques to combined QM/MM calculations of large biological systems.

Acknowledgment. This work was partially supported by grants from the National Institutes of Health (GM62248 and GM46736) and from the Army High Performance Computing Research Center (AHPCRC) under the auspices of the Department of the Army, Army Research Laboratory (ARL) under Cooperative Agreement number DAAD19-01-2-0014. Computational resources were provided by the

Minnesota Supercomputing Institute and the Army High Performance Computing Research Center.

References

(1) Reichardt, C. *Solvents and solvent effects in organic chemistry*; VCH: Weinheim, 1990.

(2) Goedecker, S.; Scuseria, G. E. *IEEE Comput. Sci. Eng.* **2003**, *5*, 14−21.

(3) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227−249.

(4) Åqvist, J.; Warshel, A. *Chem. Rev.* **1993**, *93*, 2523−2544.

(5) Gao, J. *Rev. Comput. Chem.* **1995**, *7*, 119−185.

(6) Gao, J.; Truhlar, D. G. *Annu. Rev. Phys. Chem.* **2002**, *53*, 467−505.

(7) Warshel, A.; Åqvist, J. *Annu. Rev. Biophys. Biophys. Chem.* **1991**, *20*, 267−298.

(8) Villa, J.; Warshel, A. *J. Phys. Chem. B* **2001**, *105*, 7887−7907.

(9) Warshel, A. *Acc. Chem. Res.* **2002**, *35*, 385−395.

(10) Garcia-Viloca, M.; Gao, J.; Karplus, M.; Truhlar, D. G. *Science* **2004**, *303*, 186−195.

(11) York, D. M.; Darden, T.; Pedersen, L. G. *J. Chem. Phys.* **1993**, *99*, 8345−8348.

(12) York, D. M.; Wlodawer, A.; Pedersen, L. G.; Darden, T. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 8715−8718.

(13) York, D. M.; Yang, W.; Lee, H.; Darden, T.; Pedersen, L. G. *J. Am. Chem. Soc.* **1995**, *117*, 5001−5002.

(14) Sagui, C.; Darden, T. A. *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 155−179.

(15) Cheatham, T. E., III; Young, M. A. *Biopolymers* **2001**, *56*, 232−256.

(16) Makarov, V.; Pettitt, B. M.; Feig, M. *Acc. Chem. Res.* **2002**, *35*, 376−384.

(17) Grengard, L. *Science* **1994**, *265*, 909−904.

(18) Hockney, R. W.; Eastwood, J. W. *Computer Simulation Using Particles;* Institute of Physics Publishing: 1998.

(19) Gao, J.; Xia, X. *Science* **1992**, *258*, 631.

(20) York, D. M.; Lee, T.-S.; Yang, W. *J. Am. Chem. Soc.* **1996**, *118*, 10940−10941.

(21) Garcia-Viloca, M.; Truhlar, D. G.; Gao, J. *J. Mol. Biol.* **2003**, *327*, 549−560.

(22) de Leeuw, S. W.; Perram, J. W.; Smith, E. R. *Proc. R. Soc. London, Ser. A* **1980**, *373*, 27−56.

(23) Ewald, P. P. *Ann. Phys.* **1921**, *64*, 253−268.

(24) Allen, M.; Tildesley, D. *Computer Simulation of Liquids*; Oxford University Press: 1987.

(25) Arfken, G. B.; Weber, H. J. *Mathematical methods for physicists*, 5th ed.; Academic Press: San Diego, 2001.

(26) Karasawa, N.; Goddard, W. A., III. *J. Phys. Chem.* **1989**, *93*, 7320−7327.

(27) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Hsing, L.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577−8593.

(28) Adams, D. J.; Dubey, G. S. *J. Comput. Phys.* **1987**, *72*, 156−176.

(29) Heyes, D. M. *J. Chem. Phys.* **1981**, *74*, 1924−1929.

(30) Perram, J. W.; Petersen, H. G.; Leeuw, S. W. D. *Mol. Phys.* **1988**, *65*, 875−893.

(31) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.

(32) Sagui, C.; Pedersen, L. G.; Darden, T. A. *J. Chem. Phys.* **2004**, *120*, 73−87.

(33) Bogusz, S.; Cheatham, T. E., III; Brooks, B. R. *J. Chem. Phys.* **1998**, *108*, 7070−7084.

(34) Boresch, S.; Steinhauser, O. *Ber. Bunsen-Ges. Phys. Chem.* **1997**, *101*, 1019−1029.

(35) Smith, P. E.; Pettitt, B. M. *J. Am. Chem. Soc.* **1991**, *113*, 6029−6037.

(36) Dewar, M. J.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4899−4907.

(37) Dewar, M. J. S.; Thiel, W. *Theor. Chim. Acta* **1977**, *46*, 89−104.

(38) Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 1833−1840.

(39) Löwdin, P. O. *J. Chem. Phys.* **1950**, *18*, 365−375.

(40) Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 735−746.

(41) Bader, R. F. W. *Acc. Chem. Res.* **1985**, *18*, 9−15.

(42) Storer, J. W.; Giesen, D. J.; Cramer, C. J.; Truhlar, D. G. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 87−95.

(43) Li, J.; Zhu, T.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. A* **1998**, *102*, 1820−1831.

(44) Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. *J. Comput. Chem.* **2000**, *21*, 132−146.

(45) Thompson, J. D.; Cramer, C. J.; Truhlar, D. G. *J. Comput. Chem.* **2003**, *24*, 1291−1304.

(46) Francl, M. M.; Carey, C.; Chirlian, L. E.; Gange, D. M. *J. Comput. Chem.* **1996**, *17*, 367−383.

(47) Sigfridsson, E.; Ryde, U. *J. Comput. Chem.* **1998**, *19*, 377−395.

(48) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700−733.

(49) Roothaan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69−89.

(50) Dewar, M. J. S.; Zoebisch, E.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902−3909.

(51) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.

(52) Thiel, W. *MNDO97*, version 5.0; 1998.

(53) Stewart, J. J. P. *J. Comput.-Aided Mol. Des.* **1990**, *4*, 1−105.

(54) Thiel, W.; Voityuk, A. A. *J. Phys. Chem.* **1996**, *100*, 616−626.

(55) Gregersen, B. A.; Lopez, X.; York, D. M. *J. Am. Chem. Soc.* **2003**, *125*, 7178−7179.

(56) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.

(57) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.

(58) Andersen, H. C. *J. Chem. Phys.* **1980**, *72*, 2384−2393.

(59) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695−1697.

(60) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187−199.

(61) Kumar, S.; Bouzida, D.; Swendsen, R.; Kollman, P.; Rosenberg, J. *J. Comput. Chem.* **1992**, *13*, 1011−1021.

(62) Brooks, C. L., III; Karplus, M. *J. Chem. Phys.* **1983**, *79*, 6312−6325.

(63) Brooks, C. L., III; Brunger, A.; Karplus, M. *Biopolymers* **1985**, *24*, 843−865.

(64) Hutter, M. C.; Reimers, J. R.; Hush, N. S. *J. Phys. Chem. B* **1998**, *102*, 8080−8090.

(65) Schürer, G.; Gedeck, P.; Gottschalk, M.; Clark, T. *Int. J. Quantum Chem.* **1999**, *75*, 17−31.

(66) Bernal-Uruchurtu, M.; Ruiz-López, M. *Chem. Phys. Lett.* **2000**, *330*, 118−124.

(67) Bredow, T.; Geudtner, G.; Jug, K. *J. Comput. Chem.* **2001**, *22*, 861−887.

(68) Repasky, M. P.; Chandrasekhar, J.; Jorgensen, W. L. *J. Comput. Chem.* **2002**, *23*, 498−510.

(69) Sherer, E. C.; York, D. M.; Cramer, C. J. *J. Comput. Chem.* **2003**, *24*, 57−67.

(70) Lopez, X.; York, D. M. *Theor. Chem. Acc.* **2003**, *109*, 149−159.

(71) Rozanska, X.; Chipot, C. *J. Chem. Phys.* **2000**, *112*, 9691−9694.

(72) Perreault, D. M.; Anslyn, E. V. *Angew. Chem., Int. Ed.* **1997**, *36*, 432−450.

(73) Florián, J.; Warshel, A. *J. Phys. Chem. B* **1998**, *102*, 719−734.

(74) Florián, J.; Warshel, A. *J. Am. Chem. Soc.* **1997**, *119*, 5473−5474.

(75) Florián, J.; Åqvist, J.; Warshel, A. *J. Am. Chem. Soc.* **1998**, *120*, 11524−11525.

(76) Ahn, N. *Chem. Rev.* **2001**, *101*, 2207−2208.

(77) Guthrie, R. D.; Jencks, W. P. *Acc. Chem. Res.* **1989**, *22*, 343−349.

(78) Guthrie, J. P. *J. Am. Chem. Soc.* **1977**, *99*, 3991−4001.

(79) Ramond, T. M.; Davico, G. E.; Schwartz, R. L.; Lineberger, W. C. *J. Chem. Phys.* **2002**, *112*, 1158−1169.

(80) Range, K.; McGrath, M. J.; Lopez, X.; York, D. M. *J. Am. Chem. Soc.* **2004**, *126*, 1654−1665.

(81) López, C. S.; Faza, O. N.; Gregersen, B. A.; Lopez, X.; de Lera, A. R.; York, D. M. *Chem. Phys. Chem.* **2004**, *5*, 1045−1049.

(82) Gao, J. Computation of intermolecular interactions with a combined quantum mechanical and classical approach. In *Modeling the hydrogen bond*; Smith, D. A., Ed.; ACS Symposium Series 569; Oxford University Press: New York, 1994.

(83) Freindorf, M.; Gao, J. *J. Comput. Chem.* **1996**, *17*, 386-395.

(84) Riccardi, D.; Li, G.; Cui, Q. *J. Phys. Chem. B* **2004**, *108*, 6467−6478.

(85) *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*; Linstrom, P., Mallard, W., Eds.; National Institute of Standards and Technology: Gaithersburg MD, 20899, 2003 (http://webbook.nist.gov).

(86) Hammond, G. S. *J. Am. Chem. Soc.* **1955**, *77*, 334−338.

(87) Jencks, W. P. *Chem. Rev.* **1985**, *85*, 511−527.

(88) Bader, J. S.; Chandler, D. *J. Phys. Chem.* **1992**, *96*, 6423−6427.

(89) Dang, L. X.; Pettitt, B. M. *J. Phys. Chem.* **1990**, *94*, 4303−4308.

(90) Del Buono, G. S.; Figueirido, F. E.; Levy, R. M. *Chem. Phys. Lett.* **1996**, *263*, 521−529.

CT049941I

JCTC Journal of Chemical Theory and Computation

Neural Network Models of Potential Energy Surfaces: Prototypical Examples

James B. Witkoskie[†] and Douglas J. Doren*

Department of Chemistry and Biochemistry, University of Delaware, Newark, Delaware 19716

Received June 30, 2004

Abstract: Neural networks can be used generate potential energy hypersurfaces by fitting to a data set of energies at discrete geometries, as might be obtained from ab initio calculations. Prior work has shown that this method can generate accurate fits in complex systems of several dimensions. The present paper explores fundamental properties of neural network potential representations in some simple prototypes of one, two, and three dimensions. Optimal fits to the data are achieved by adjusting the network parameters using an extended Kalman filtering algorithm, which is described in detail. The examples provide insight into the relationships between the form of the function being fit, the amount of data needed for an adequate fit, and the optimal network configuration and number of neurons needed. The quality of the network interpolation is substantially improved if gradients as well as the energy are available for fitting. The fitting algorithm is effective in providing an accurate interpolation of the underlying potential function even when random noise is added to the data used in the fit.

I. Introduction

The accuracy of a Monte Carlo or molecular dynamics simulation is limited by the accuracy of the potential energy surface used. Empirical models using explicit analytic potential functions are fast to evaluate but limited in accuracy. On the other hand, first-principles quantum mechanics can provide a very accurate potential energy surface, but the computational expense of these calculations limits the size and length of the simulation. An ideal potential energy surface should have the accuracy of the quantum mechanical calculations and yet be as fast to evaluate as empirical models. One approach to this problem is to construct potential energy surfaces by interpolating among the results of first-principles energy calculations at specific configurations. However, standard interpolation methods (such as splines or orthogonal polynomial methods) are difficult to apply to systems with more than a few degrees of freedom. In recent years, a number of techniques have been developed to address this problem, and substantial progress has been made.[1−17] Nevertheless, more work is needed to develop an accurate, efficient method that can be routinely applied to reacting systems with many degrees of freedom.

One method that has shown promise in a number of applications is the artificial neural network. Several efforts to use neural networks to describe potential functions have now been reported, in most cases with good results.[18−29] These efforts include some rather complex problems, for example: dissociation of H_2 on silicon[18] and palladium[24] surfaces; tetra-atomic van der Waals complexes;[20] water dimer[21] and a polarizable model of liquid water;[27] $H_2O-Al^{3+}-H_2O$ three-body interactions;[23] covalent C−C and C−H interactions in hydrocarbon molecules and pure carbon phases;[26] and a number of problems in vibrational and electronic spectroscopy.[22,25,28,29] The neural network is a very general form that does not require adaptation or special coordinates for each application. As such, it may offer a simple, general approach to potential fits.

A particular strength of the neural network is that it can effectively model data with noise. While nonsystematic errors in quantum chemistry calculations are expected to be small, there may be variation in the energy due to degrees of

* Corresponding author e-mail: doren@UDel.edu.

[†] Current address: Department of Chemistry, MIT, Cambridge MA 02139.

10.1021/ct049976i CCC: $30.25
Published on Web 10/21/2004

freedom that are not included in the model. For example, it is common to model only a relatively small, chemically active region of a larger system. The energy of the smaller subsystem still depends on the coordinates of the surrounding system, but if the smaller system and its surroundings are not strongly correlated, the larger system may be modeled as a thermal bath (as in Langevin models). This approach is likely to be useful for small molecules on metal surfaces. Even for systems where there are strong correlations between the smaller system and its surroundings, the uncorrelated variation in the larger system may be treated as noise. Indeed, this has proven to be effective in the H_2/Si(100) model discussed in ref 18, where the ab initio calculations allowed optimization of an entire Si_9H_{12} cluster, while the network modeled only the coordinates of two H and two Si atoms.

Given the effectiveness of neural networks in representing a number of realistic multidimensional potential functions, it is somewhat surprising that there has been little effort to explore applications of neural networks in simple model systems. The insight available from simple problems may be a helpful guide to understanding the limitations of neural networks and improving their performance.

This paper has two aims: to describe a method for optimizing neural network parameters that is efficient and can tolerate noise in the available data and to develop a qualitative picture of neural network representations of potential functions through applications to several prototype problems. Section II describes extensions of earlier work[18,30] on parameter optimization to allow use of information about the derivatives of the potential and to allow network architectures with multiple hidden layers. In section III, we report applications of this approach to a number of simple functions of one, two, and three dimensions. The density of data and number of parameters needed to achieve a good fit, the effect of using gradient information in addition to energy information, effects of noise in the data, and convergence properties of the Kalman filter are discussed in section IV. The network's ability to extrapolate and the relationship between the form of the potential and the properties of the optimal network were also examined. These divisions among topics are artificial, since the network structure, quality of fit, optimization routine, and data used to achieve the fit are all intricately related. Full information about the fit achieved for each function is summarized in the figure captions.

II. Theory of Neural Networks and Kalman Filtering

We begin by reviewing the terminology of neural networks and the basic concepts of the parameter optimization routine.

A. Neural Network Architecture. The neural network is a nonlinear function of many parameters that maps particular inputs (in our case, coordinates) to an output (in our case, a potential energy). The network is composed of linear and nonlinear transfer functions, called neurons. For the networks we will consider, the neurons can be classified into different "layers" depending on where the input to the neuron comes from and how the output is used. The first ("input") layer takes an m-dimensional vector of coordinates, $\boldsymbol{x}$, at which the potential is to be evaluated and outputs an affine transformation of $\boldsymbol{x}$ to the nodes in the second ("hidden") layer. Each node in the hidden layer applies a nonlinear transfer function to its input. In the work described here, we have used the sigmoid transfer function, $(1 + e^{\boldsymbol{a}\bullet\boldsymbol{x}+b})^{-1}$. The outputs from this layer may become the inputs to another hidden layer of nonlinear functions. Whether there is one hidden layer or more, the outputs from the neurons in the last hidden layer become the inputs to the single neuron in the "output" layer, which performs an affine transformation to produce the network output. This paper will primarily concern a single-layer network architecture, which is essentially a superposition of sigmoids, i.e., the potential is approximated by

$$\tilde{V}(\boldsymbol{x};\boldsymbol{\theta}) = c_0 + \sum_{i=1}^{n} c_i (1 + e^{\boldsymbol{a}_i\bullet\boldsymbol{x}+b_i})^{-1} \qquad (1)$$

for a network with n hidden nodes. In this case, the adjustable parameters are the n vectors $\boldsymbol{a}_i$ (each with m components), the n values, b_i, and the $n+1$ values, c_i. These $n(m+2)+1$ parameters are collectively denoted by the vector $\boldsymbol{\theta}$. The goal is to find network parameters such that the network output is equal to the potential energy at the coordinates specified by the network input. Note that the network is defined at all values of the inputs and can be analytically differentiated with respect to both the coordinates and network parameters.

In the "feed-forward" architectures we have considered, neurons do not receive input from other neurons in the same layer, or from lower layers, so there is no feedback. This simplifies parameter fitting, while retaining adequate flexibility. We have considered only architectures in which all neurons in a given layer take inputs from all neurons in the layer above. Thus, for a given number of coordinates, the network architecture is completely defined by the number of hidden neurons (or, if there are multiple hidden layers, the number in each hidden layer).

Other transfer functions might be used, but the sigmoid has the advantages that it is bounded (so that the coefficients do not have to allow for cancellation between large positive and negative values), and it has a single variable region (so that fitting to local variation does not interfere with sigmoids that fit variation in other regions). Moreover, a fundamental theoretical result guarantees the efficacy of sigmoids: On a finite interval in $\mathbf{R}^n$, every differentiable function is the limit of a uniformly convergent sequence of finite sums of sigmoids.[31] As a result, a feed-forward network with a single, finite, hidden layer can uniformly fit any smooth, continuously differentiable function on the interval to within a specified tolerance.

Finding the optimal network for fitting a particular set of data involves both trying different numbers of hidden neurons and optimizing the parameters of each network configuration. In many cases, the quality of fit may be improved for a given network configuration by increasing the amount of data used in the parameter optimization. However, when the data are obtained from electronic structure calculations, it is much faster to optimize many different networks than generate significant amounts of new data. Thus, we generally do

Table 1. Equations of the Extended Kalman Filter[a]

$\lambda_k = \lambda_0 \cdot \lambda_{k-1} + (1 - \lambda_0)$	(T1)
$(\mathbf{D}_k)_{i,j} = (\boldsymbol{d}_k)_i^2 \cdot \delta_{i,j}$	(T2)
$\mathbf{H}_k = \partial \tilde{\boldsymbol{f}}(\mathbf{x}_k;\boldsymbol{\theta})/\partial \boldsymbol{\theta}\|_{\boldsymbol{\theta}=\boldsymbol{\theta}_k}$	(T3)
$\mathbf{R}_k, \mathbf{Q}_k = (1 - \alpha)\cdot(\mathbf{R}_{k-1}, \mathbf{Q}_{k-1}) + \alpha \cdot \mathbf{D}_k$	(T4)
$\mathbf{A}_k = [\mathbf{I} + \lambda_k^{-1} \cdot \mathbf{H}_k^T \mathbf{P}_{k-1} \mathbf{H}_k + \mathbf{R}_k]^{-1}$	(T5)
$\mathbf{K}_k = \lambda_k^{-1} \cdot \mathbf{P}_{k-1} \mathbf{H}_k \mathbf{A}_k$	(T6)
$\mathbf{P}_k = \lambda_k^{-1} \cdot [\mathbf{I} - \mathbf{K}_k \mathbf{H}_k^T] \mathbf{P}_{k-1} + \mathbf{Q}_k$	(T7)
$\boldsymbol{\theta}_k = \boldsymbol{\theta}_{k-1} + \mathbf{K}_k [\boldsymbol{f}(\boldsymbol{x}_k) - \tilde{\boldsymbol{f}}(\boldsymbol{x}_k;\boldsymbol{\theta}_k)]$	(T8)

[a] The filter is used to update the network parameters, $\boldsymbol{\theta}_k$, so that the network output, $\tilde{\boldsymbol{f}}(\boldsymbol{x};\boldsymbol{\theta}_k)$, approximates the desired output $\boldsymbol{f}(\boldsymbol{x})$, at the coordinates, $\boldsymbol{x}_k$. The significance of the other filter variables and suitable values are discussed in the text.

thorough searches of various network architectures in attempting to optimize the fit before adding more data.

B. The Extended Kalman Filter. For a given choice of network configuration, the network parameters must be optimized to fit the function. The optimization method must avoid local minima of the squared-error surface with respect to the parameters and should be stable in the presence of noise. The extended Kalman filter has this robustness and achieves a least-squares fit to data.[32] We provide here a qualitative guide to the filter, with enough details to allow others to reproduce our method. However, this section is not essential to the remainder of the paper.

A derivation of the extended Kalman filter and the role of each element in the filter can be found in ref 33, but the final equations are given in Table 1 (eqs T1−T8). For a finite data set with specified inputs (i.e., coordinate values), $\boldsymbol{x}$, the goal is to find a network model that produces the same result, $V(\boldsymbol{x})$, obtained from ab initio calculations at those points. Here we have denoted the desired output, $\boldsymbol{f}$, as a vector. In the simplest case, this will have a single element, the potential energy, $V(\boldsymbol{x})$. More generally, the network can be used to fit a vector of outputs. In particular, we will simultaneously fit the potential energy and its gradient, so that

$$\boldsymbol{f}(\boldsymbol{x}) = \left[V(\boldsymbol{x}), \frac{\partial V}{\partial x_1}\bigg|\boldsymbol{x}, \ldots, \frac{\partial V}{\partial x_m}\bigg|\boldsymbol{x}\right] \tag{2}$$

Having specified a network estimate for the potential, $\tilde{V}(\boldsymbol{x};\boldsymbol{\theta})$, it is a simple matter to differentiate this network with respect to the coordinates to obtain the corresponding estimate of the gradient. Thus, there is a known functional relationship, $\tilde{\boldsymbol{f}}(\boldsymbol{x};\boldsymbol{\theta})$, that gives the network approximation to $\boldsymbol{f}(\boldsymbol{x})$ for a given input, $\boldsymbol{x}$, and given network parameters, $\boldsymbol{\theta}$. Starting from an initial random guess at $\boldsymbol{\theta}$, the Kalman filter uses this functional relationship and comparisons between $\tilde{\boldsymbol{f}}(\boldsymbol{x};\boldsymbol{\theta})$ and $\boldsymbol{f}(\boldsymbol{x})$ to determine optimal parameters for the model network.

Unlike a traditional least-squares approach that examines the entire data set before determining a correction, the Kalman filter examines individual pieces of data to determine the next step. That is, the disparity between the "ideal" network output and the current network model is calculated for a single point in the data set using the current estimate of $\boldsymbol{\theta}$, denoted $\boldsymbol{\theta}_k$. The parameter set is updated to a new value, $\boldsymbol{\theta}_{k+1}$, to reduce the difference between $\tilde{\boldsymbol{f}}(\boldsymbol{x};\boldsymbol{\theta})$ and $\boldsymbol{f}(\boldsymbol{x})$. The filter is sequentially applied to each element of the data set, and then the data are randomly shuffled and examined again. This process continues until a converged value of $\boldsymbol{\theta}$ is reached. The sequential approach of the filter reflects the original application of the Kalman filter, which was developed for models of discrete time sequences.[34] In the present application, where the data have no intrinsic order, the sequence is randomized on each pass through the data, to eliminate any sequence dependence.

For linear models, the final parameters are the same as for a least-squares fit. For nonlinear problems, the Kalman filter corrections to θ_k are found by (locally) linearizing the problem and finding the zero, similar to Newton's method. This requires the matrix $\mathbf{H}_k$ of derivatives of $\tilde{\boldsymbol{f}}(\boldsymbol{x};\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$ (eq T3), which are explicit analytic functions.

The filter does not assume the values in the data set are exact, and the updates do not jump straight to the solution as each piece of data is added. Instead, the data are assumed to include Gaussian-distributed noise, which the filter tries to remove by developing three covariance matrices, $\mathbf{P}_k$, $\mathbf{Q}_k$, and $\mathbf{R}_k$. The latter two, $\mathbf{Q}_k$ and $\mathbf{R}_k$ (eq T4), measure uncertainties of the parameters and data, respectively.[35] Even if there is no noise in the data, there will still be uncertainty in the optimal parameters because the network cannot fit the data exactly. $\mathbf{Q}_k$ and $\mathbf{R}_k$ also serve to add numerical stability by allowing the program to converge to a solution without fitting the data perfectly. These matrices are not known in advance and must be approximated,[36] though the approximations can often be crude. For this study, these matrices are approximated as a weighted average between the old matrix and an update matrix (eq T4), with $\mathbf{Q}_k$ and $\mathbf{R}_k$ initialized as zero matrices. The update matrix consists of the diagonal elements of the latest error vector. This is indicated in eq T2, where the correction to the covariance matrix ($\mathbf{Q}_k$ or $\mathbf{R}_k$) is indicated by $\mathbf{D}_k$ and the corresponding error vector is represented by $\boldsymbol{d}_k$. The error vector $\boldsymbol{f}(\boldsymbol{x}) - \tilde{\boldsymbol{f}}(\boldsymbol{x};\boldsymbol{\theta})$ is used to calculate the update for $\mathbf{R}_k$, while $\mathbf{H}_k(\boldsymbol{f}(\boldsymbol{x}) - \tilde{\boldsymbol{f}}(\boldsymbol{x};\boldsymbol{\theta}))$ is used in the update for $\mathbf{Q}_k$. α is an empirical weighting parameter (eq T4);[36] it is usually much less than unity ($\alpha \leq 0.15$).

The heart of the Kalman filter is $\mathbf{P}_k$, the covariance matrix that measures the uncertainty in the coefficients during the execution of the filter (eq T7). The exact definition of $\mathbf{P}_k$ and the derivation of its updates are given in ref 33. $\mathbf{P}_k$ stores information on the direction of previous parameter updates. The correction term in the update expression, $\mathbf{K}_k \mathbf{H}_k^T \mathbf{P}_k$, describes the change in network parameter uncertainty resulting from the update. We initialized $\mathbf{P}_k$ as a multiple of the identity matrix, with trace equal to the mean squared error, i.e.,

$$[\mathbf{P}_0]_{i,j} = \sum_{l=1}^{n} ||\boldsymbol{f}(\boldsymbol{x}_l) - \tilde{\boldsymbol{f}}(\boldsymbol{x}_l;\boldsymbol{\theta})||^2 \cdot \frac{1}{n} \cdot \delta_{i,j} \tag{3}$$

with $\boldsymbol{\theta}$ set to the initial guess of the parameter values.

With the statistics from $\mathbf{P}_k$, $\mathbf{Q}_k$, and $\mathbf{R}_k$ and the derivative, $\mathbf{H}_k$, the filter determines an optimal step to reduce the error in the network prediction for the current piece of data, without increasing the error in the fit to previous elements of the data set. This is achieved by moving in a direction (in parameter space) that satisfies an orthogonality condition while reducing the error in the current measurement. The

covariance matrices allow larger changes in those parameters with the greatest uncertainties. The net direction of the step is given by the gain matrix, $\mathbf{K}_k$, which is multiplied by the magnitude of the error to determine the correcting step size (eq T8).

We have modified the traditional extended Kalman filter by including a forgetting function, λ_k, that multiplies some terms (eqs T1 and T5−T7).[30] This gives the most recent steps in the algorithm a higher weight. During the course of the filtering operation, the forgetting function approaches unity to develop better statistics by giving high weights to a longer sequence of the previous data points. λ_0 is the constant that determines the rate at which λ_k approaches unity and λ_1 is the initial value of λ_k. The values of λ_0 and λ_1 are close to unity ($\lambda_0 \geq 0.99$, $\lambda_1 \geq 0.90$), but specific choices for their values depend on how often λ_k is updated as well as the data and potential.[30]

III. Network Architecture and Quality of Fit

In principle, a finite number of hidden-layer neurons can uniformly fit any well-behaved function to a specified tolerance.[31] We have investigated the pragmatic issue of how many neurons are needed to interpolate some common functions to a reasonable degree of accuracy. Clearly, the number of neurons needed will depend on the function being fitted. The sigmoid transfer function has a single variable region and two constant regions. To fit a function, the network uses the variable region of each sigmoid to match the variation in the function in some region, with multiple neurons used to fit the function over an extended domain. Because the sigmoid function is monotonically increasing, a single sigmoid cannot fit a function where the derivative changes sign. An interval between points where the derivative changes sign will be referred to as a monotonic region. The number of monotonic regions being fit is a lower bound on the number of hidden-layer neurons needed to fit the function.

In this section, we show the fits achieved for some exemplar functions with features that are common to typical interatomic interaction potentials. We explore the actual number of neurons needed to achieve a qualitative fit and the effect of additional neurons on improving the quantitative fit. Unless stated otherwise, the domain and range for all functions in this study is the unit interval. Data for training the function were selected from this interval as well. Errors reported are the integrated root-mean-square difference between the network and the function being fitted, calculated by integration over the unit interval (the functions are known exactly, there is no need to base errors on a finite test set).

A. One-Dimensional Model Functions. *1. Harmonic Oscillator and Morse Potentials.* As a simple example we begin with the harmonic oscillator on the positive *x*-axis, which has only one monotonic region. A single sigmoid gives a good qualitative fit to this half-oscillator, though the best fit achieved with one sigmoid had an rms error of 10^{-2}. Turning to a full harmonic oscillator (with the minimum at $x = 0.5$; see Figure 1), there are two monotonic regions and at least two hidden-layer neurons are required. A network with two sigmoids fits the function to a surprisingly high accuracy, with a rms error on the order of 10^{-3} over the unit interval (unshaded region of Figure 1).

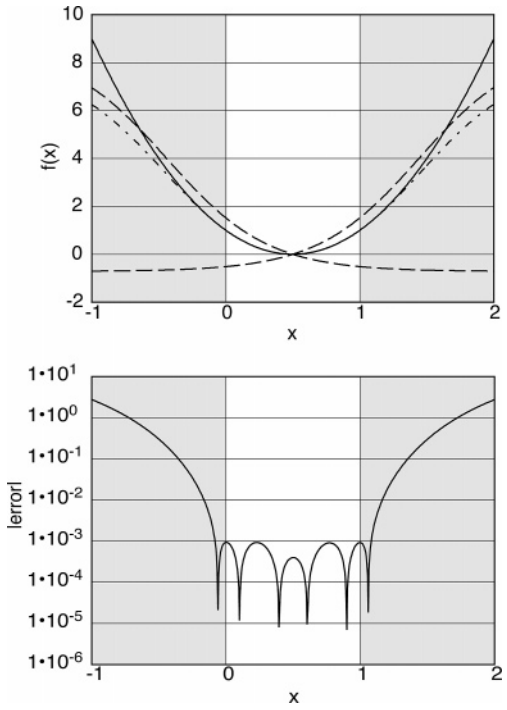


Figure 1. (Top) Harmonic oscillator potential, $V(x) = 4(x - 1/2)^2$ (solid curve), the individual outputs of the two neurons used in the network fit (dashed curves), and the total network fit (dot−dashed curve). The outputs due to the individual neurons are shifted vertically, as described in the text, to aid in visualizing the fit. (Bottom) Error in the neural network fit as a function of *x*; the integrated rms error is 0.8×10^{-3}. Energy and derivative data at six equally spaces points in the unshaded region ($0 \leq x \leq 1$) were used to optimize the network parameters. Each neuron predominantly fits one side of the oscillator. However, each neuron also makes a non-negligible contribution to the other side of the oscillator function, which prevents use of a single neuron and symmetry to determine the parameters. The network fit remains qualitatively correct as it extrapolates into the shaded region, where no data were used in parameter optimization.

Plotting the contributions of individual neurons (Figure 1) demonstrates why the fit for the full oscillator is more accurate than that for the half-oscillator. These individual-neuron contributions are determined by turning off the output from other neurons in the hidden layer. For clarity, a constant has been added to the resulting network output is equal to the fitted function at the point where the single-neuron network and the fitted function have equal derivatives. For the full harmonic oscillator, most of the variation in each monotonic region is matched by a single neuron, as in the fit of the half-oscillator, but the second neuron makes a nontrivial contribution that improves the fit. Although no symmetry constraints were imposed on the network, the two optimal neurons are related by the reflection symmetry of the oscillator. Note that the small region of positive curvature at the bottom of each sigmoid makes the dominant contribution over the range $0 \leq x \leq 1$. This is achieved by scaling the coordinates with fairly large coefficients.

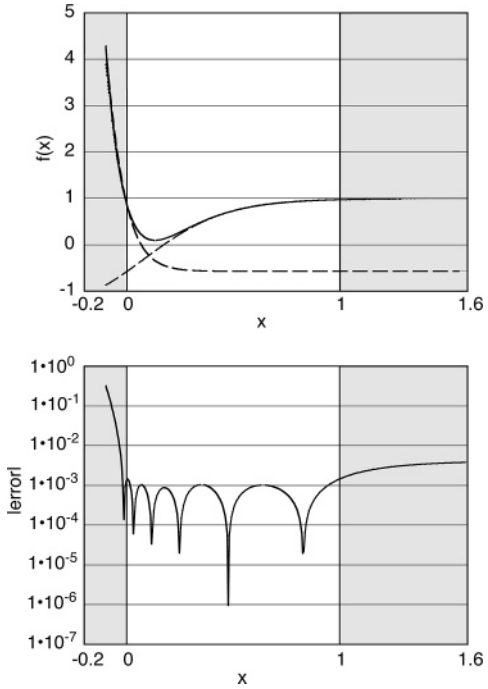


Figure 2. (Top) Morse potential, $V(x) = 0.1 + 0.9(1 - \exp(0.65-5x))^2$ (solid curve), the individual outputs due to the two neurons used in the fit (dashed curves), and the total network fit (dotted curve). (Bottom) Error in the neural network fit as a function of x; the integrated rms error is 2.0×10^{-3}. Energy and gradient data at 21 points in the unshaded region ($0 \le x \le 1$) were used for parameter optimization. Note the range over which extrapolation remains accurate. A fit with data at only 11 points had a similar rms error of 2.2×10^{-3} but does not extrapolate as well as the fit shown.

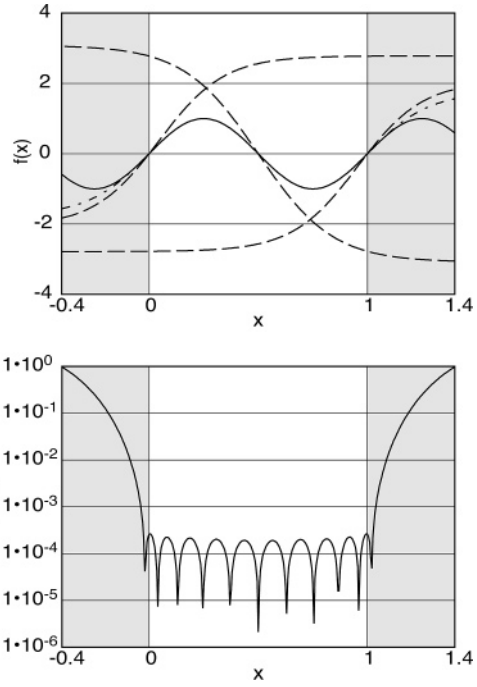


Figure 3. (Top) Single-period sine function, $V(x) = \sin(2\pi x)$ (solid curve), individual output from the three neurons used to fit it (dashed curves), and the total network fit (dot−dashed curve). (Bottom) Corresponding error plot; the integrated rms error is 1.0×10^{-3}. Energy and gradient data at 21 equally spaced points in the unshaded region ($0 \le x \le 1$) were used for parameter optimization. Each monotonic segment is primarily fit by a single neuron. The symmetry of the function being fit is reflected by symmetry relations among the neurons, with one antisymmetric neuron and the other neurons related by translation. These symmetry relations make the network fit symmetric, even when extended to the (shaded) extrapolation region.

Two neurons are also successful at fitting the Morse potential to 10^{-3} rms error (Figure 2). This illustrates the flexibility of networks with the minimum number of neurons. However, convergence was difficult to achieve for the Morse potential, and the search had to be repeated with several choices of initial network parameters. The large derivatives of the repulsive wall are the likely origin of this difficulty. Matching steep derivatives requires large network coefficients, so the filter must search a large region of coefficient space to achieve the optimal fit.

2. Sine Waves. Figure 3 illustrates the fit to a sine wave. Good accuracy (10^{-4} rms error) was achieved with three neurons, which is again the minimum number for this case. Again, most of the variation in each region of high slope is matched by a single neuron, but there is an appreciable contribution from the other neurons. The network configuration also reflects the symmetry of the function. One neuron has the antisymmetry of the sine wave and the other two neurons are related by a periodic translation.

Functions with several monotonic regions, like the double-period sine wave, are more difficult to fit (Figure 4). Over the unit interval, the double-period sine wave has 5 monotonic regions, requiring a minimum of 5 neurons. With so many neurons, several are varying at any given coordinate value, causing more interference among them. The best fit with 5 neurons had an rms error of 10^{-1} (Table 2). Four additional compensating neurons allowed a fit with 10^{-4} rms error. This is a rather subtle effect: Table 2 shows that there is a dramatic decrease in the error upon increasing the network from eight to nine neurons. This is not simply the gradual improvement expected with addition of more parameters; a full set of four compensating neurons appears to be an important addition to the model.

The symmetry relationships between individual neurons are less clear as the number of neurons increases, but the dominant neurons of the network for the double-period sine wave retain some symmetry properties. The plots of major contributing sigmoids in Figure 4 show one antisymmetric sigmoid and two pairs of sigmoids that are related by periodic translation and reflection. However, the additional compensating neurons needed for a more accurate fit show no clear symmetry. As a result, symmetry is not maintained on extrapolation beyond the unit interval. If the network fails to reflect some of the symmetry of the system, it may indicate that the network has too many (or too few) neurons. A single neuron on one side of a symmetry plane might be matching the same variation as multiple neurons on the other side of the plane. For example, initial searches for a network with nine neurons had problems converging for the double-period sine wave, while a ten-neuron network converged quickly. However, the ten-neuron network did not reflect any of the

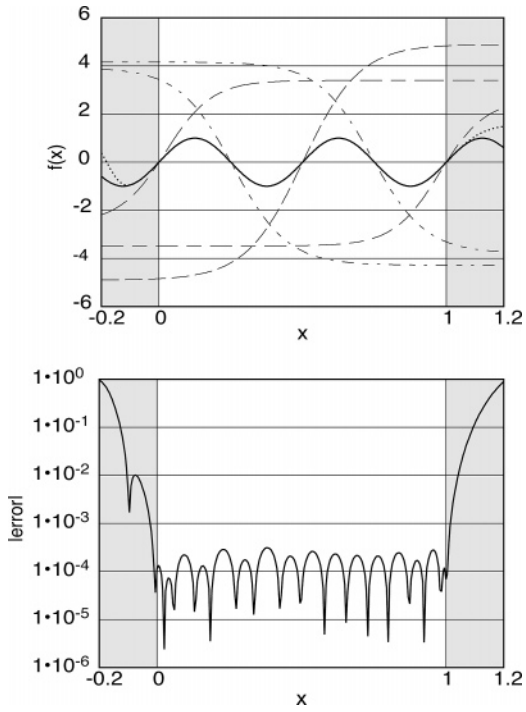


Figure 4. (Top) Double-period sine function, $V(x) = \sin(4\pi x)$ (solid curve), individual outputs from the five neurons that fit the majority of the variation (various dashed curves), and the total nine-neuron network fit (dotted curve). (Bottom) Corresponding error plot for a fit with nine neurons; the integrated rms error is 0.12×10^{-3}. One neuron is antisymmetric, and the other four neurons are related by reflection and translation operations. Even with energy and gradient data at 101 points in the unshaded region ($0 \le x \le 1$), these five neurons alone cannot achieve a quantitative fit. Four other neurons (not shown) make small contributions to the fit (magnitudes less than 0.06 in the fitted region). The additional neurons quickly grow outside the fitted region and cause the breakdown in symmetry.

symmetry of the system. Further searches for a nine-neuron network were eventually successful in achieving the fit described above, in which the dominant contributions reflect the symmetry of the function.

B. Higher Dimensional Models. As the number of dimensions increases, more neurons are generally needed, though there is no clear lower bound on their number. A single sigmoid transfer function varies only along one generalized coordinate. This direction is a linear transformation of the input coordinates for the single hidden-layer networks considered here, though a nonlinear dependence on the input coordinates could be achieved with a multiple hidden-layer network.

As in one dimension, a few neurons usually account for most of the variation in the examples we have examined. This can be seen for a two-dimensional model made by coupling a Morse potential in one direction to a sinusoidal wave in the other (Figure 5). Two neurons account for most of the repulsive wall of the Morse potential, while two other neurons account for the asymptotic approach to zero at large bond lengths. Six other neurons account for small variations caused by the more subtle sinusoidal element. Similar trends

Table 2. Sensitivity of Network Error to Number of Neurons for the Double-Period One-Dimensional Sine Function and the Two-Dimensional Anharmonic Oscillator[a]

no. of neurons	RMS error (% of scale) (%)
1-D Double-Period Sine Wave (Energy and Derivative at 101 Points; No Noise)	
5	8.38
8	0.71
9	0.012
10	0.019
2-D Anharmonic Oscillator (Energy and Gradient at 36 Points; No Noise)	
3	0.17
4	0.13
5	0.16
2-D Anharmonic Oscillator (Energy and Gradient at 100 Points; 6.5% rms Added Noise)	
3	3.0
4	0.9
5	1.6

[a] In each case, over 30 converged optimizations were performed to find the best fit.

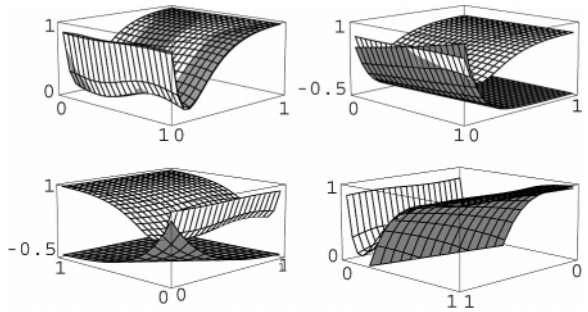


Figure 5. Two-dimensional surface model, $V(x, y) = 1 + 0.1\text{-}(\sin(2\pi x) + 9)\,((1-e^{0.65-5y})^2 - 1)$ (white surface in all figures). This function is compared to the best network fit (gray surface, upper left figure) and the individual contributions of three neurons (gray surfaces) that account for different aspects of the variation in the function. The network fit used 10 hidden-layer neurons; energy and gradient data at 121 points were used to achieve an rms error of 0.9×10^{-3}.

were observed in a three-dimensional surface model (not shown) in which the two-dimensional model of Figure 5 was coupled to a sinusoidal variation in the third direction.

The two-dimensional anharmonic oscillator (Figure 6) illustrates some of the subtleties of finding an optimal network. The quality of the fit is quite good with three, four, or five neurons, though the best fit was achieved with four neurons (Table 2). It was difficult to find a converged solution for a three-neuron model, though in the end the rms error was low. The five-neuron model has one neuron that is virtually inactive. The fifth neuron actually increases the error in the fit by allowing the network to over-fit the data. With added noise in the data, the quality of the fit is even more sensitive to the number of neurons (Table 2). If there is any noise in the data, a network with more parameters can fit the noise, so the importance of minimizing the number of neurons is magnified.

A two-dimensional sine function (Figure 7) has more sign changes than the two-dimensional oscillator or surface model.

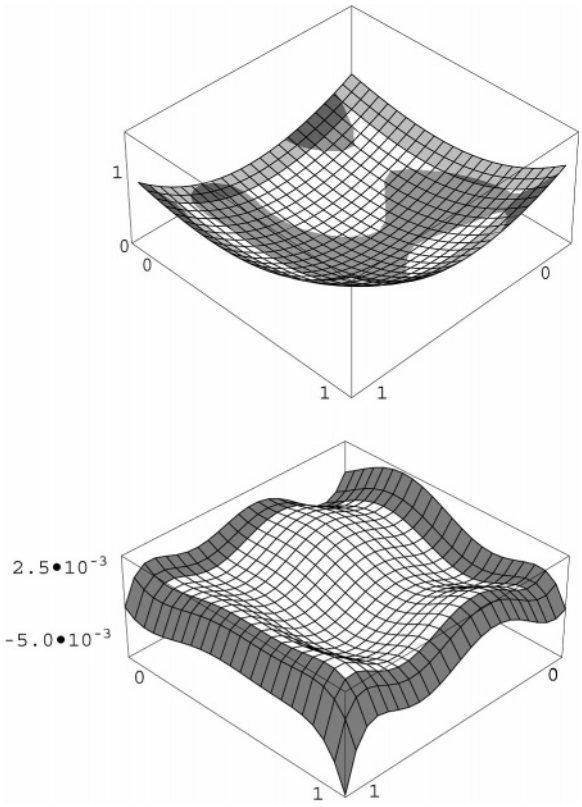


Figure 6. (Top) Two-dimensional anharmonic oscillator (with coefficients chosen as rational approximations to random numbers, with the condition that there is a single minimum and no maxima in the region being fit), $V(x, y) = (4/9)x^3 + (2/27)x^2y + (2/45)xy^2 + (16/27)y^3 + (124/135)x^2 + (4/75)xy + (358/675)y^2 - (1042/1125)x - (76/125)y + (1684/5625)$ (white surface) and the four-neuron network fit to the function (gray surface). (Bottom) Corresponding error plot. Energy and gradient data at 36 equally spaced data points were taken from the unshaded region (note in this case, $0 \leq x, y \leq 0.9$) to fit the network parameters. The network extrapolates this surface into the shaded region and achieves an rms error of 1.3×10^{-3} in the unshaded region.

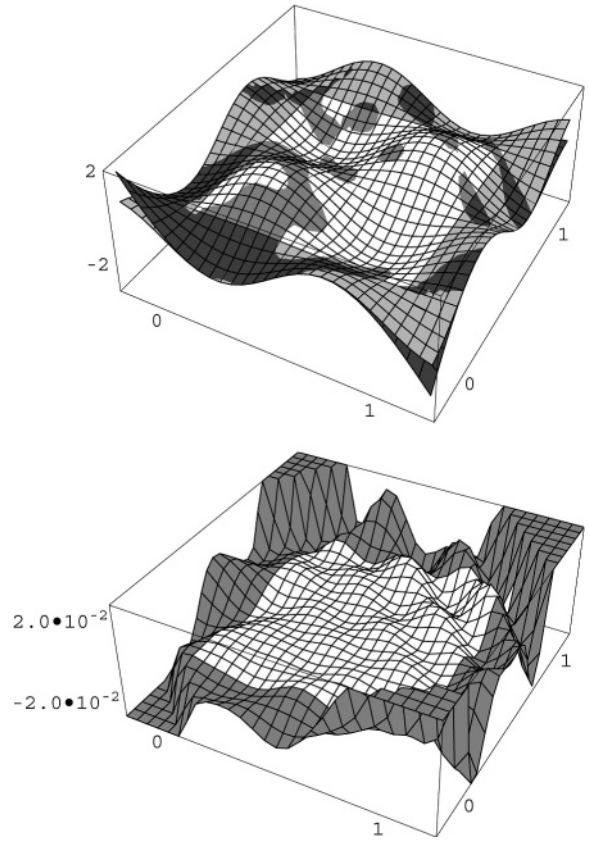


Figure 7. (Top) Two-dimensional sinusoidal surface model, $V(x, y) = \sin(2\pi x)\sin(2\pi y)$ (white surface). (Bottom) Corresponding error plot. The function was fit with 12 neurons. Energy and gradients at 121 points (for a total of 363 pieces of data) were required to achieve the fit in the unshaded region ($0 \leq x, y \leq 1$) to an rms error of 3×10^{-3}. The network extrapolation breaks down at the corners, the points farthest from the data used for parameter optimization.

In this case, an error of 3×10^{-3} is achieved with 12 sigmoids. This can be compared to the three sigmoids needed for comparable accuracy in the one-dimensional case.

C. Extrapolation. As with any interpolation method, neural networks are not expected to make accurate predictions beyond the region where data are available for parameter optimization. However, extrapolation is accurate over a surprising range for many of the functions (Figures 1−7). Typically, the network fit is qualitatively reasonable to about 10% beyond the domain where the function was fitted, though in some cases it can be more. For the two-dimensional sinusoid (Figure 7), this extrapolation breaks down around the corners of the extrapolation, i.e., the points farthest from the data used for parameter optimization. The network extrapolation of the higher curvature functions, like the double period sine wave (Figure 4), is accurate over a much shorter range than lower curvature functions, like the single-period sine wave and the large-x side of the Morse potential (Figures 2 and 3). Since the sigmoid eventually becomes constant, it extrapolates best for functions that approach a constant (e.g., the Morse oscillator at large x, Figure 2).

D. Degeneracy. We have found many examples where different optimizations of the same network with the same data yield different coefficients, while the rms errors of the different optimizations are almost identical. These network models are distinct in the sense that no simple permutation of the neurons makes them identical. We refer to such sets of networks as degenerate. One might expect degeneracy in networks with too many parameters, since making the network too flexible would allow it to fit the data many different ways. However, we have observed degeneracy in networks that do not over-fit the data. For the two-dimensional anharmonic oscillator in Figure 6, three neurons fit the function very well (rms error of 1.7×10^{-3}) and do not over-fit the function, since the optimal fit was achieved with four neurons (rms error of 1.3×10^{-3}). Nevertheless, we have found 10 distinct sets of coefficients for the three-neuron model. Plots of the error surface show at least two distinct topologies, though several sets of network parameters produce similar fits to the function. Some cases of degen-

eracy may result from the symmetry of the sigmoid function

$$1-(1+e^{-ax+b})^{-1} = (1+e^{ax-b})^{-1}$$

but this cannot explain the different error topologies.

E. Multiple Hidden Layers. In principle, using multiple hidden layers can provide an alternate model that may require fewer parameters than a single-layer network. Additional layers act as nonlinear coordinate transformations that become the input to the final layer. The shape of the network fit is still a superposition of the sigmoids from the final layer, though the nonlinear coordinate transformation can add curvature to the sigmoids in multiple directions. However, our results for multiple layers in one- and two-dimensional examples are not encouraging. We find that more parameters are required to fit functions with multiple layers than a single layer. On the other hand, other groups studying higher-dimensional systems have found multiple layers to be useful,[23,24] so this issue deserves further exploration.

IV. Parameter Optimization and Data Density

The examples above show the type of fits that can be achieved with neural networks, though we have not addressed the kind of data needed for successful parameter optimization. The number and distribution of elements in the data set, the type of data available (energies or energies and derivatives), and conditioning of the data will all have a significant impact on the ability of the optimization algorithm to find network parameters to fit the function. In this section we describe the quality of fit that can be achieved when different types of data are used to determine network parameters. To isolate the effects of the data sets, we have used only the network architecture that yields the best fit for each function.

A. Functional Form and Data Density. Since functions requiring more neurons for an accurate fit will have more parameters in the network, we expect that more data will be needed to accurately determine all the parameters. This is observed in the two sine waves described above. For a single-period sine wave (Figure 3), which is fit with three neurons and a total of 10 parameters, only 42 pieces of information (the energy and derivative at 21 different points) are required to achieve a rms error of less than 10^{-3}. For the shorter wavelength sine wave (Figure 4), 202 pieces of data (energies and derivatives at 101 points) are required to get a comparable fit with nine neurons and 28 parameters. In every case reported, including those with noise added to the data, fewer than 10 pieces of data per parameter (whether energy or derivative data) were needed for accurate fits. The greatest amount of data is needed when there are large network coefficients associated with high gradients (as for the Morse potential) or curvature, because the output is sensitive to small changes in the large parameters.

B. Gradient Data. In the prior section, we described fits using a combination of energy and derivative data. Most prior work with neural networks has used energy data alone for parameter optimization. Here we explore the effectiveness of using gradient data in addition to the energy. Intuitively, one expects that a fit to both energy and gradient data will lead to smaller fluctuations between points in the data set than a fit to energy values alone. This expectation is borne out in fits to the Morse potential where energy and derivative data at 11 points results in a better fit (2×10^{-3} rms error) than energy data alone at 41 points (1×10^{-2} rms error). Thus, gradient data can dramatically reduce the density of data needed, an important consideration if these data are to obtained from ab initio calculations. Moreover, many ab initio methods can calculate the gradient at a computational expense comparable to that of the energy, so m components of the gradient vector can be obtained for the effort of one energy (m denotes the number of coordinate inputs to the network). In many-dimensional systems, this may partly compensate for the increased size of the space to be sampled. Other partial derivative information (such as the curvature) could be used as well, though such data are less likely to be available.

Table 3. Effects of Gradient Information and Data Density on the Fit to the Two-Dimensional Anharmonic Oscillator (Figure 6) with 4 Neurons

no. of data points	RMS error (% of range)	
	no gradients (%)	with gradients (%)
	Date with No Noise	
36	0.4	0.1
100	0.2	0.1
	Data with 6.5% rms Added Noise	
36	3.2	1.0
49	2.4	1.0
100	1.4	0.9

Several other examples confirm that use of gradient data allows sampling at many fewer points than needed with energy data alone. Table 3 shows the results of several fits to a two-dimensional anharmonic oscillator using uniformly spaced data. A fit with four neurons using energy data at 36 points gives an rms error of 0.4%. This error is cut in half if energy data are used at 100 points. However, using both energy and gradient data at only 36 points gives a fit with even smaller error (0.1%). Since the gradient is a two-dimensional vector, there are three pieces of data at each of the 36 points or 108 pieces of information in total. Thus, in this example, a lower-density data set with energy and gradient information is more effective in allowing a good fit than energy data alone at higher-density (with a similar total amount of information).

The two-dimensional surface model (Figures 5) also shows how gradient information improves the fit. Using energy data alone at 361 points, the filter was unable to match the sinusoidal variation of the function and a best fit of 2.4×10^{-2} rms error was achieved. Using gradient information (in addition to energies) at 121 points allowed a fit to less than 10^{-2} rms error. Again, while the total amount of information is the same in these two data sets, the combination of energy and gradient information is more effective in achieving a good fit, and (if analytic derivatives are available in an ab initio calculation) the combined data will be more economical to calculate.

C. Noise. The Kalman filter is designed to allow for noise in the data, as might be caused by variation in unmodeled

coordinates. To test this capability, we have added white noise with a root-mean-square value of 6.5% of the total range to both the energy and gradient data from the two-dimensional anharmonic oscillator. With energy and derivative information at 36 points, the network removed most of the added noise, fitting the oscillator function to 1% error (Table 3). Increasing the data density to 100 points helps very little, with a remaining error of 0.9%. Even with noise added to the gradients, they improve the fit: without gradients, energies at 100 points are needed to achieve 1.4% rms error. Clearly, this method is effective even when there is significant variation due to unmodeled coordinates.

D. Filter Convergence. The initial guess at the network parameters is random, so there is a risk that the optimizations will converge on a local minimum in parameter space. The worst case in our experience was for the 3-dimensional surface model (Morse potential coupled to 2D sinusoidal variation) discussed above. Local minima were found in 90% of the runs. A simple solution is to do a series of optimizations and take the best results, though it is never clear how many examples might be needed. Nevertheless, the computing requirements for a network optimization are relatively small. Most of the calculations described here take only a few minutes on a typical desktop computer. The number of iterations through the whole data set depends on the function being fit and the size of the data set. Fewer iterations are needed when more data are used, presumably because of redundancy in the data.

We have found examples where the network parameters do not converge at all for some initial parameter sets. For example, if there are not enough neurons to match the variation of the function, the coefficients of the network will oscillate. This behavior has been observed even in networks that can fit almost all of the variation, like the three-neuron fit to the two-dimensional anharmonic oscillator. On the other hand, if the network has too many neurons, there are more likely to be local minima.

V. Summary and Final Comments

The neural network using sigmoidal neurons is very flexible and can fit a range of functions typical of interatomic potentials in low dimensions, with a small number of neurons and a modest amount of data. The unbiased nature of the network fit is illustrated by the remarkable fact that a two-neuron network can model both the harmonic and Morse oscillators with comparable accuracy over a wide range. No assumptions about the functional form of the potential are required to achieve a good representation. The resulting function is infinitely differentiable and globally defined. The fit is not only merely local, like a spline, but can also reproduce global features, such as symmetry, without special choices of coordinates. Even in the presence of random noise, the network can recover the underlying relationship between coordinates and potential. Thus, neural networks remain promising candidates for fitting models of interatomic interactions.

The extended Kalman filter is a robust method of achieving a least-squares fit of network parameters to the data. However, the number of data points needed to sample the variation in a function will increase when there is significant variation in more dimensions. Using gradient data is an economical way to obtain information about variation in higher-dimensional functions, which partly compensates for the need to sample more coordinates. For many systems, it may also be possible to isolate the dependence on a few degrees of freedom, treating most degrees of freedom as a statistically defined bath. However, more efficient methods for data sampling will ultimately have to replace the uniform sampling used here. Collins[9] has developed an approach that samples configurations that are likely to occur in a thermal reaction process. By running classical molecular dynamics (MD) trajectories on an estimate of the potential, important regions of configuration space can be found, and new ab initio data can be calculated in those regions. Another example of trajectory-based sampling has been reported by Wood et al.[37] in an application to solvation thermodynamics. Classical molecular dynamics (MD) simulations are run with an approximate potential, followed by ab initio calculations of solute−solvent interactions at a random sample of configurations. Thermodynamic perturbation theory is used to calculate the difference in free energy between the approximate and ab initio potentials. However, if the initial approximate potential is not reasonably close to the ab initio potential, there will be large sampling errors. In such a case, the approximate potential can be improved using information from the ab initio calculations (followed by new MD simulations with the improved potential).[37] Since these ab initio calculations are done at configurations sampled from a thermal ensemble (albeit with an approximate potential), they are inherently more dense at the low-energy configurations that dominate the thermal ensemble. Neural networks may be a convenient tool for modeling these improvements to the approximate potential.

Acknowledgment. This material is based upon work supported by the National Science Foundation under Grants No. 9971241 and 0316223.

References

(1) Ho, T.-S.; Rabitz, H. *J. Chem. Phys.* **1996**, *104*, 2584. Hollebeek, T.; Ho, T.-S.; Rabitz, H. *Annu. Rev. Phys. Chem.* **1999**, *50*, 537. Ho, T.-S.; Rabitz, H. *J. Chem. Phys.* **2000**, *113*, 3960.

(2) Frishman, M.; Hoffman, D. K.; Rakauskas, R. J.; Kouri, D. J. *Chem. Phys. Lett.* **1996**, *252*, 62. Frishman, D. K.; Frishman A.; Kouri, D. J. *Chem. Phys. Lett.* **1996**, *262*, 393.

(3) Boothroyd, A. I.; Keogh, W. J.; Martin, P. G.; Peterson, M. R. *J. Chem. Phys.* **1991**, *95*, 4343; **1996**, *104*, 7139.

(4) Mielke, S. L.; Garrett, B. C.; Peterson, K. A. *J. Chem. Phys.* **2002**, *116*, 4142.

(5) Salazar, M.; Simons, J. *J. Chem. Phys.* **1996**, *105*, 10919.

(6) Hack, M. D.; Truhlar, D. G. *J. Chem. Phys.* **1999**, *110*, 4135.

(7) Pu, J. Z.; Truhlar, D. G. *J. Chem. Phys.* **2002**, *116*, 1468.

(8) Szalay, V. *J. Chem. Phys.* **1999**, *111*, 8804.

(9) Bettens, R. P. A.; Collins, M. A. *J. Chem. Phys.* **1999**, *111*, 816. Duncan, A. H.; Collins, M. A. *J. Chem. Phys.* **1999**, *111*, 1346. Zhang, D. H.; Collins, M. A.; Lee, S.-Y. *Science* **2000**, *290*, 961. Moyano, G. E.; Collins, M. A. *J. Chem. Phys.* **2003**, *119*, 5510. Crespos, C.; Collins, M. A.; Pijper, E.; Kroes, G. J. *J. Chem. Phys.* **2004**, *120*, 2392.

(10) Lopez, M. J.; Jellinek, J. *J. Chem. Phys.* **1999**, *110*, 8899.

(11) Berweger, C. D.; van Gunsteren, W. F.; Müller-Plathe, F. *J. Chem. Phys.* **1998**, *108*, 8773.

(12) Mattson, T. R.; Wahnström, G.; Bengtsson, L.; Hammer, B. *Phys. Rev. B* **1997**, *56*, 2258.

(13) Gross, A.; Scheffler, M.; Mehl, M. J.; Papaconstantopoulos, D. A. *Phys. Rev. Lett.* **1999**, *82*, 1209.

(14) Olsen, R. A.; Busnengo, H. F.; Salin, A.; Somers, M. F.; Kroes, G. J.; Baerends, E. J. *J. Chem. Phys.* **2002**, *116*, 3841.

(15) Liu, Y.-P.; Kim, K.; Berne, B. J.; Friesner, R. A.; Rick, S. W. *J. Chem. Phys.* **1998**, *108*, 4739.

(16) Burnham, C. J.; Xantheas, S. S. *J. Chem. Phys.* **2002**, *116*, 1479; **2002**, *116*, 1493; **2002**, *116*, 1500; **2002**, *116*, 5115.

(17) Groenenboom, G. C.; Mas, E. M.; Bukowski, R.; Szalewicz, K.; Wormer, P. E. S.; van der Avoird, A. *Phys. Rev. Lett.* **2000**, *84*, 4072. Mas, E. M.; Bukowski, R.; Szalewicz, K.; Groenenboom, G. C.; Wormer, P. E. S.; van der Avoird, A. *J. Chem. Phys.* **2000**, *113*, 6687. Mas, E. M.; Bukowski, R.; Szalewicz, K. *J. Chem. Phys.* **2003**, *118*, 4386.

(18) Blank, T. B.; Brown, S. D.; Calhoun, A. W.; Doren, D. J. *J. Chem. Phys.* **1995**, *103*, 4129.

(19) Skinner, A. J.; Broughton, J. Q. *Modell. Simul. Mater. Sci. Eng.* **1995**, *3*, 371.

(20) Brown, D. F. R.; Gibbs, M. N.; Clary, D. C. *J. Chem. Phys.* **1996**, *105*, 7597.

(21) No, K. T.; Chang, B. H.; Kim, S. Y.; Jhon, M. S.; Scheraga, H. A. *Chem. Phys. Lett.* **1997**, *271*, 153.

(22) Prudente, F. V.; Soares Neto, J. J. *Chem. Phys. Lett.* **1998**, *287*, 585. Prudente, F. V.; Acioli, P. H.; Soares Neto, J. J. *J. Chem. Phys.* **1998**, *109*, 8801.

(23) Gassner, H.; Probst, M.; Lauenstein, A.; Hermansson, K. *J. Phys. Chem. A* **1998**, *102*, 4596.

(24) Lorenz, S.; Gross, A.; Scheffler, M. *Chem. Phys. Lett.* **2004**, *395*, 210.

(25) Muñoz, C.; Niño, A. *Comput. Chem.* **1998**, *22*, 355.

(26) Hobday, S.; Smith, R.; Belbruno, J. *Modell. Simul. Mater. Sci. Eng.* **1999**, *7*, 397.

(27) Cho, K.-W.; No, K. T.; Scheraga, H. A. *J. Mol. Struct.* **2002**, *641*, 77.

(28) Rocha Filho, T. M.; Olivera, Z. T., Jr.; Malbouisson, L. A. C.; Gargano, R.; Soares Neto, J. J. *Int. J. Quantum Chem.* **2003**, *95*, 281.

(29) Bittencourt, A. C. P.; Prudente, F. V.; Vianna, J. D. M. *Chem. Phys.* **2004**, *297*, 153.

(30) Blank, T. B.; Brown, S. D. *J. Chemometrics* **1994**, *8*, 391.

(31) Cybenko, G. *Mathematical Control Signals Systems* **1989**, *2*, 303.

(32) Puskorius, G. V.; Feldkamp, L. A. *IEEE Trans. Neural Networks* **1994**, *5*, 279.

(33) Gelb, A. *Applied Optimal Estimation;* MIT Press: Cambridge, 1974.

(34) Kalman, R. E. *J. Basic Eng. Ser. D* **1960**, *5*, 35.

(35) Grewal, M. S.; Andrews, A. P. *Kalman Filtering: Theory and Practice*; Prentice Hall: Englewood Cliffs, 1993.

(36) Tsoi, A. C. Gradient Based Learning Methods. In *Adaptive Processing of Sequences and Data Structures*. (Lecture Notes in Computer Science; Vol. 1387: Lecture notes in artificial intelligence). Giles, C. L., Gori, M., Eds.; Springer-Verlag: New York, 1998.

(37) Wood, R. H.; Yezdimer, E. M.; Sakane, S.; Barriocanal, J. A.; Doren, D. J. *J. Chem. Phys.* **1999**, *110*, 1329. Sakane, S.; Yezdimer, E. M.; Liu, W.; Barriocanal, J. A.; Doren, D. J.; Wood, R. H. J. *Chem. Phys.* **2000**, *113*, 2583.

CT049976I

JCTC Journal of Chemical Theory and Computation

Parametrization of Reversible Digitally Filtered Molecular Dynamics Simulations

Adrian P. Wiley, Martin T. Swain, Stephen C. Phillips, and Jonathan W. Essex*

School of Chemistry, University of Southampton, Highfield, Southampton SO17 1BJ, U.K.

Colin M. Edge

GlaxoSmithKline, New Frontiers Science Park (North), Coldharbour Road, Harlow CM19 5AD, U.K.

Received July 16, 2004

Abstract: Reversible Digitally Filtered Molecular Dynamics (RDFMD) is a method of amplifying or suppressing motions in a molecular dynamics simulation, through the application of a digital filter to the simulation velocities. RDFMD and its derivatives have been previously used to promote conformational motions in liquid-phase butane, the Syrian hamster prion protein, alanine dipeptide, and the pentapeptide, YPGDV. The RDFMD method has associated with it a number of parameters that require specification to optimize the desired response. In this paper methods for the systematic analysis of these parameters are presented and applied to YPGDV with the specific emphasis of ensuring a gentle and progressive method that produces maximum conformation change from the energy put into the system. The portability of the new parameter set is then shown with an application to the M20 loop of E-coli dihydrofolate reductase. A conformational change is induced from a closed to an open structure similar to that seen in the DHFR−$NADP^+$ complex.

1. Introduction

Molecular dynamics (MD) simulations have been frequently applied to protein systems to provide atomistic detail that is unavailable to experimental methods.[1−3] The time scale that traditional MD simulations can simulate is typically in the order of tens of nanoseconds and is severely limited by computational resources. However, even with this length of simulation, the large-scale conformational motions of proteins are rare events, due to the significant energy barriers that can lie on the potential energy surface. Methods that enhance the rate of these events so that they occur within the length of typical simulations are therefore of great interest.[4]

To address the sampling issues of molecular dynamics a number of equilibrium and nonequilibrium methods exist. These include generalized-ensemble methods that induce a system's random walk in potential energy space, thereby overcoming conformational energy barriers.[5] Within this collection of algorithms are simulated tempering,[6] the multicanonical algorithm,[7] and replica-exchange molecular dynamics.[8] Other well-studied methods range from self-guided molecular dynamics[9] (and its most recent refinement, self-guided Langevin dynamics[10]), that applies an evolving guiding force to a simulation, to CONCOORD[11] ('from constraints to coordinates'), which predicts protein conformations based on distance constraints.

Reversible Digitally Filtered Molecular Dynamics (RDFMD) has previously been shown to amplify or suppress motions of a specific frequency, and the amplification of low-frequency motions has produced increased conformational sampling in a range of systems.[1,2] RDFMD can be tailored through interdependent parameter selection, and methods to choose appropriate parameters for different applications are therefore required. In this paper, a range of analysis methods are presented to develop an RDFMD

* Corresponding author e-mail: J.W.Essex@soton.ac.uk.

10.1021/ct049970t CCC: $30.25
Published on Web 12/17/2004

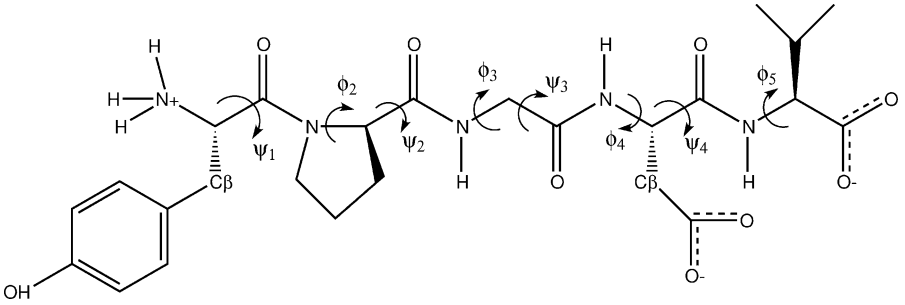


Figure 1. YPGDV with the backbone dihedral angles of interest labeled.

protocol that amplifies the conformational motions of a short peptide chain. The generated parameter set is shown to be suitable for flexible regions of larger protein systems, inducing a conformational change between known forms of the E-coli dihydrofolate reductase M20 loop.

2. RDFMD

A digital filter is a list of coefficients, c_i, that can be used to weight a discrete vector input, $\mathbf{x}_i$, and summed to give a vector output, $\mathbf{y}$ (eq 1). The filter response is the filter's effect on the phase and amplitude of the input signal. The greater the number of coefficients, the closer the filter's response will be to that desired.

$$\mathbf{y} = \sum_{i=-m}^{m} c_i \mathbf{x}_i \tag{1}$$

The RDFMD filter sequence begins by filling a buffer of velocities, $\mathbf{v}$, using a microcanonical (NVE) ensemble MD simulation. This buffer has the same number of steps as the number of filter coefficients ($2m+1$) in eq 2. The external rotational and translational motion of the target system, $\mathbf{v}_{\mathrm{Ext}}$, is removed (eq 3), and a digital filter is then applied to the Cartesian components of the internal velocities, $\mathbf{v}_{\mathrm{Int}}$. Filters used for this work are designed using the *fircls* function in MATLAB[12] and have the property of yielding a frequency response of unity over the frequency range to be amplified and zero elsewhere.[2]

$$\mathbf{v}_{\mathrm{Filt}} = \sum_{i=-m}^{m} c_i \mathbf{v}_{\mathrm{Int},i} \tag{2}$$

$$\mathbf{v}_{\mathrm{Int}} = \mathbf{v} - \mathbf{v}_{\mathrm{Ext}} \tag{3}$$

The filters used are typically designed to extract all components of velocities corresponding to motions most likely to induce conformational change. These filtered velocities, $\mathbf{v}_{\mathrm{Filt}}$, are multiplied by an amplification factor, A, that can be changed to adjust the energy put into the system. The amplified velocities are then summed with the original velocity set for the central buffer point, $\mathbf{v}_0$, producing a new set of velocities, $\mathbf{v}'$, that are in phase with the coordinates at the center of the buffer (eq 4). An amplification factor of 2 would therefore produce a final set of velocities for which the targeted frequency components have been extracted, multiplied by 2, and then summed with the original velocities to give a 3-fold increase in the kinetic energy of targeted motions (as they will also be included in the full set), with all other frequencies left unchanged, in contrast to the self-guided Langevin dynamics method.[10]

$$\mathbf{v}' = \mathbf{v}_0 + A\mathbf{v}_{\mathrm{Filt}} \tag{4}$$

From the new velocities, $\mathbf{v}'$, and the central coordinates of the buffer, conventional simulation continues both backward and forward in time so that a new buffer is filled. Another filter can then be applied, separated from the last by a specified time delay (the filter delay). The purpose of this delay is to allow the system to relax from the effects of the previous filter and to allow some progression over the potential energy surface. Filters are repeatedly applied in this manner until the kinetic energy in the system rises beyond some defined limit (the internal temperature cap) or until a certain number of filters have been applied (the filter cap).

RDFMD is typically run as the combination of periods during which the system velocities are modified by repeated applications of a digital filter and of traditional MD in the canonical (NVT) or the isothermal−isobaric (NPT) ensemble. During the MD for which a thermostat is applied, the temperature can be returned to that desired and new equilibrated velocities and coordinates are generated. From these another set of filter applications can be performed.

In this paper, the parameters associated with the repeated applications of the filter will be examined, their effects on the simulation determined, and methods for their optimization described.

3. RDFMD Parameters

To analyze the response of a system to different RDFMD parameters the YPGDV pentapeptide (tyrosine-proline-glycine-aspartic acid-valine) is used (Figure 1). YPGDV has been used in previous studies as a test case for RDFMD,[2] conformational analysis,[13] the ART-2′ clustering algorithm,[14] and self-guided molecular dynamics.[15] NMR data from *trans*-proline YPGDV in water is known to show an approximately equal proportion of reverse turn and extended conformations at 273 K.[16]

The system was set up from an all-trans Z-matrix of YPGDV, generated and solvated within the MCPRO package[17] using 805 water molecules and 1 sodium ion. Cubic periodic boundary conditions were used throughout. Simulation was performed using the NAMD package[18] with a switching function applied to the Lennard-Jones interactions between 8 and 12 Å, a PME treatment of electrostatics,[19] and SHAKE[20] was applied to all bonds involving a hydrogen atom, with a tolerance of 10^{-8} Å. Explicit water was modeled

by the TIPS3P water model as implemented by CHARMM, and the protein was described by the CHARMM22 force field.[21] NAMD was developed by the Theoretical Biophysics Group in the Beckman Institute at Urbana-Champaign.

Initially, minimization was performed with the conjugate gradient line-search algorithm[18] for a total of 22 000 steps. The system was then gradually heated with a 20 000 step canonical simulation at each temperature between 50 and 300 K at 50 K intervals using a 2 fs time step. A Langevin Thermostat[22] was used with a damping parameter of 10 ps^{-1}. 80 000 steps were then performed at the target 300 K followed by 500 000 steps (1 ns) in the isothermal−isobaric ensemble. A Nosé-Hoover Langevin piston barostat[23] with a pressure target of 1 atm, a piston temperature of 300 K, a damping decay parameter of 200 fs, and an oscillation period of 400 fs was used. A further 200 000 steps (100 ps) of isothermal−isobaric MD were then performed with a 1 ps^{-1} thermostat damping parameter, 300 fs barostat damping decay, and 500 fs piston oscillation period. The simulation state at this point is used when randomising velocities to create a large number of starting points for RDFMD trials. Parameters not specified for later simulations are as above.

A 20 ns isothermal−isobaric MD simulation of YPGDV has been performed from the equilibrated YPGDV system, and coordinate, velocity, and box dimensions were extracted every 2 ns to give 10 largely independent starting points for filter applications.

To investigate the parametrization of RDFMD, parameters are systematically varied from a previously reported protocol[2] and their interdependence discussed. Discussion of the internal temperature cap parameter will be published elsewhere through comparison with parallel tempering[8] and MD simulations. Unless otherwise stated, the RDFMD parameters are as follows: a filter buffer length of 1001 steps, a target frequency of 0−25 cm^{-1}, an amplification factor of 4, a filter delay of 20 steps, an internal temperature cap of 2000 K, and a filter cap of 10 filters. All atoms in the YPGDV peptide are targeted by the digital filter.

3.1. Frequency Target. The frequency target is the range of frequencies that are desired for amplification or suppression by RDFMD. This parameter is the most important and is required before all others can be optimized. It is suggested that some form of frequency analysis on MD simulations is performed to validate the frequency target, or alternatively the response of the system to different filters can be tested. If the range chosen is too broad, energy will be put into or removed from motions that are not intended for manipulation. If the range is too narrow however, the frequencies at which desirable motions occur may be missed.

A 16 678 step microcanonical (NVE) ensemble simulation was performed from the equilibrated YPGDV system for a Fourier analysis, the results of which are reported elsewhere.[2] A 16 678 step simulation with a time step of 2 fs yields approximately a 1 cm^{-1} resolution. A previous suggestion for an RDFMD frequency target of 0−25 cm^{-1} was based upon an amplitude spectra argument that showed the majority of motion in the backbone ψ and ϕ angles (labeling as in Figure 1) occurring at very low frequencies.[2] A further frequency analysis using the Hilbert-Huang Transform[24,25]

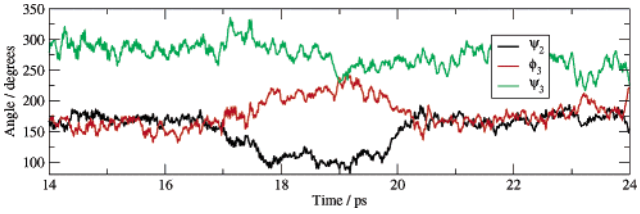


Figure 2. Relevant dihedral angles during the YPGDV conformational change event.

showed the presence of significant energy in the backbone dihedral motions of YPGDV, during conformational changes, with frequencies in the 0−25 cm^{-1} region.[26]

To examine the frequency target further, the Empirical Mode Decomposition[24] Method (EMD) has been applied to the simulation trajectory. EMD is a method of decomposing a signal into a set of intrinsic mode functions (IMFs). IMFs are iteratively determined by following the highest frequency motion and then removing the created IMF from the signal. Thus a set of IMFs are produced, each describing the evolution of a frequency present in the input signal. The last IMFs describe the lowest frequencies present in the system and leave a trend component with no wavelike properties, similar to the dc component of a Fourier analysis. Full details of the algorithm are described elsewhere.[24]

Using EMD, each backbone dihedral angle trajectory from the NVE simulation was split into three components: high frequency 'noise' from interactions with higher frequency degrees of freedom (angle and bond stretching motions), the dihedral angle motion that persists throughout the simulation, and the low-frequency conformational motions resulting from intermittent large scale conformational changes in the system. Fortunately, the YPGDV system is inherently flexible, and a conformational event was captured within the short NVE simulation. This was the brief formation of a β 3-turn between 17.3 and 17.7 ps, as determined by the DSSP (Dictionary of Protein Secondary Structure[27]) algorithm using default options. This secondary structure change was accompanied, and followed, by significant rearrangement in ψ_2, ϕ_3, and ψ_3, as shown in Figure 2.

EMD performed on the ψ_2 signal produces 11 IMFs, the first two (those with the highest frequencies) describe motions over 200 cm^{-1}, as determined by Fourier transform (FT). The last five IMFs describe motions below 25 cm^{-1}. The physical relevance of the IMFs is best described when summed in frequency groups as shown in Figure 3. The Fourier transform of the summed signals are shown in Figure 4, showing the frequency ranges that incorporate the different signals. Although there is some frequency overlap between the signals, it cannot be determined whether this is due to the deficits of the EMD or the FT methods. Limitations of the application of EMD to MD analysis have been discussed elsewhere.[28]

Results show that the vibrations below 25 cm^{-1} are associated with the conformational change event itself. Therefore, the previous frequency target of 0−25 cm^{-1} would only be able to amplify physically relevant signals (those of significant amplitude present in the system) if a large scale conformational event such as that shown is occurring. To

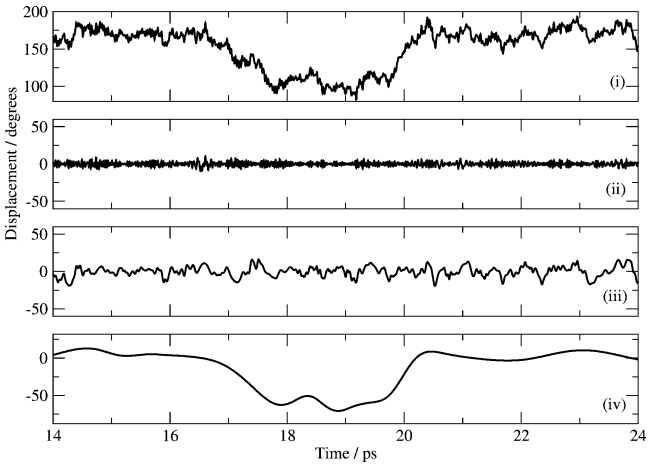


Figure 3. (i) ψ_2 signal during the conformational change event, (ii) sum of IMFs 1 and 2 showing motions above 200 cm^{-1}, (iii) sum of IMFs 3 to 6 showing motions predominantly between 25 and 250 cm^{-1}, and (iv) sum of IMFs 7 to 11 showing motions below 25 cm^{-1}.

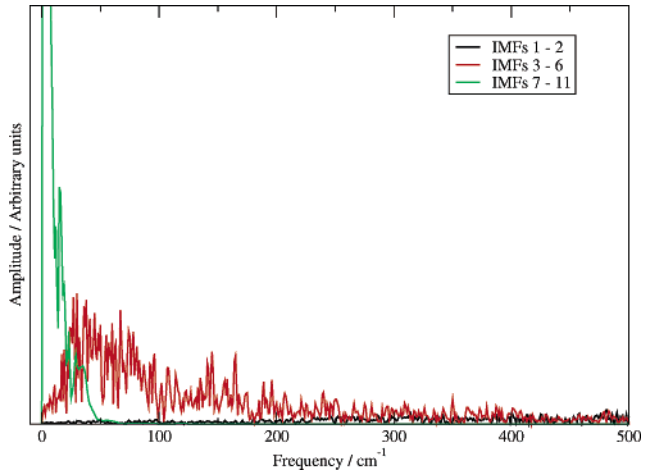


Figure 4. Fourier transforms of summed IMFs derived from the ψ_2 trajectory. A Hanning window was used.

target the dihedral angle motions present in the entire simulation, the frequency range 25−100 cm^{-1} is clearly significant. Amplifying frequencies from 0 cm^{-1} should be considered desirable in case motions are occurring at very low frequencies. It is clear that the inherent dihedral motions, i.e., vibrations within a potential energy well, occur predominately between 25 and 100 cm^{-1} and that these frequencies should be targeted if large-scale conformational change involving escape from local minima is to be induced.

To measure the response of the system to different frequency targets, filters designed to amplify frequencies below a specified value have been used. For each filter tested, 50 simulations of a single filter application to the YPGDV system were performed. The simulation start points were obtained from the equilibrated YPGDV system after a random reassignment of velocities at 300 K and 1000 steps of NPT MD simulation. Filters designed to target broader frequency ranges will see more degrees of freedom, and thus the amount of energy put into the system cannot be controlled with a constant amplification factor as previously used (eq 4). Instead the amplification factor is recalculated for each filter application so that the kinetic energy of the system is always increased by 25 kcal mol^{-1}. This is a comparable energy increase to that produced by a single filter application using the original parameter set.

For analysis, the root-mean-squared deviation (RMSD) is calculated between the trajectories of the eight relevant dihedral angles before and after a filter application. This is done for 100 steps (200 fs) before and after each filter application so that only the effects of the filter application are measured. The sum of the RMSDs for each filter application is considered to be a measure of induced conformational change, and from the 50 simulations an average and standard error can be calculated. Averaged results are shown in Figure 5 (a), showing a drop off in the amount of induced conformational change after 100 cm^{-1}. For higher frequencies, less conformational change is induced, and energy is being less efficiently placed in the dihedral angle motions. To demonstrate the importance of frequency specificity, a control simulation has been performed in which the filter used has a central coefficient of 1, and all other coefficients are zero. 25 kcal mol^{-1} of kinetic energy is therefore put directly into the internal velocities, across all frequencies. The upper and lower error limits of the induced conformational change for the control simulation are shown in Figure 5(a) by dashed lines. The amount of conformational change induced by this simple heating procedure is much less than that obtained using the targeted filters.

Repeated applications of filters have been shown to be more effective than use of a single filter and a greater amplification factor.[2] The different filter targets have been therefore been tested for longer filter sequences, using a delay parameter of 20 steps and a filter cap of ten. Kinetic energy is increased by 25 kcal mol^{-1} by each filter application, and the 10 equilibrated YPGDV states are used as simulation starting points. The induced conformational change is calculated as before for each filter buffer and summed across all 10 buffers for each run. The results are shown in Figure 5(b). A nonfrequency specific energy input has again been applied, and the error limits are shown with dashed lines. As the upper limit on the frequency target is increased, the induced conformational change is reduced. The nonfrequency specific energy input is once again far less efficient than any of the tested frequency targets.

There is a noticeable reduction in the conformational change induced by the 0−25 cm^{-1} filter when repeated filter applications are used. It is believed that this is due to amplifying a frequency range that includes only low amplitude motions (such as 0−25 cm^{-1} when a rare event is not occurring), rather than a range in which significant motions are apparent (such as 25−100 cm^{-1} where dihedral motions persist throughout the simulation). This effect is best investigated by analyzing an individual dihedral trajectory. Figure 6 (a) shows a breakdown of the ψ_2 dihedral angle from the first buffer of one of the simulations. Once again, the signal is separated by EMD into IMFs that can be grouped together as a high-frequency component (> 250 cm^{-1}), a dominant motion of intermediate frequencies (30 cm^{-1}−150 cm^{-1}), and a low-frequency motion (<60 cm^{-1}).

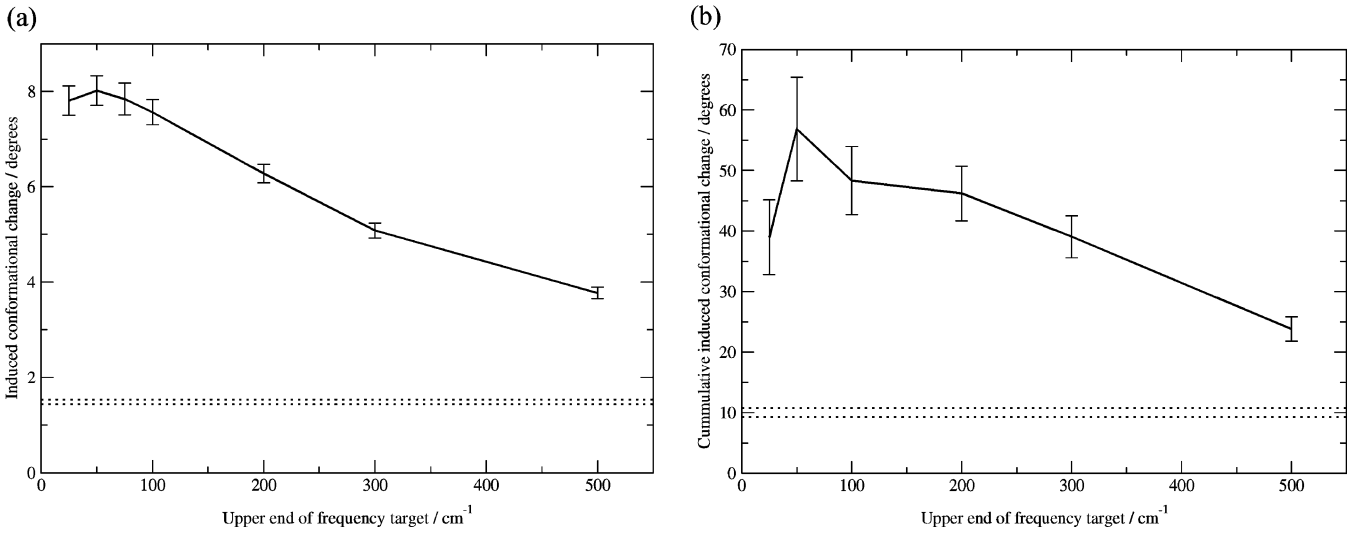


Figure 5. Measure of induced conformational change with different frequency targets. Dashed lines indicate error bounds of conformational change induced by an input of nonfrequency specific energy: (a) 50 applications of a single filter and (b) 10 filter sequences of 10 filters.

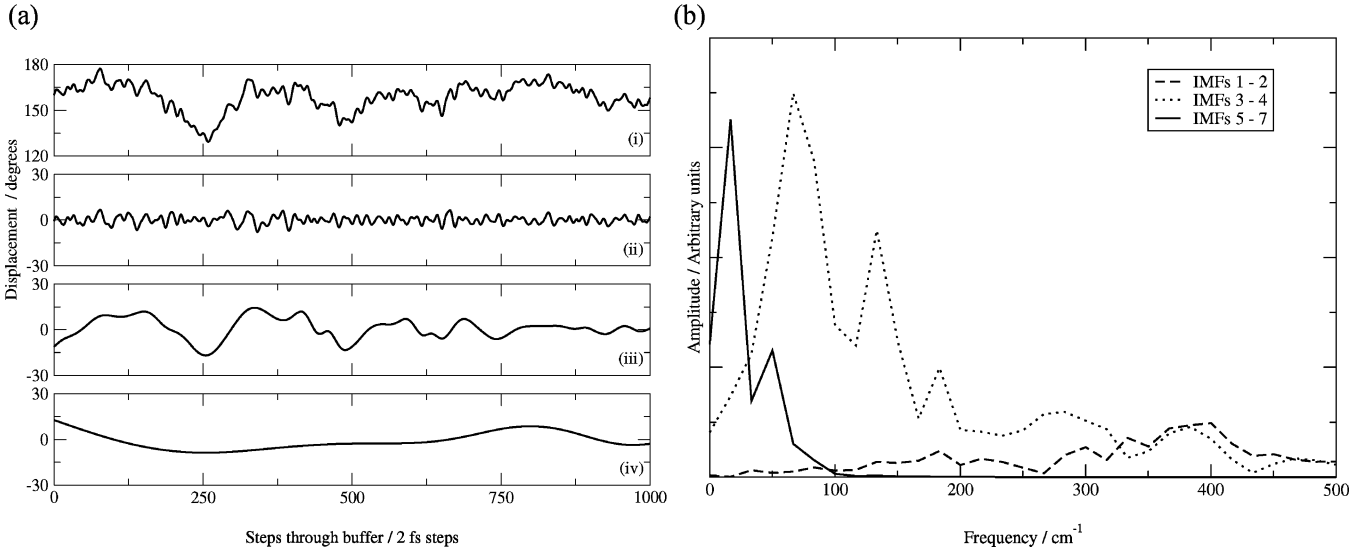


Figure 6. Decomposition of the ψ_2 trajectory from the first buffer of an RDFMD simulation: (a) (i) signal, (ii) sum of IMFs 1 and 2, (iii) sum of IMFs 3 and 4, (iv) sum of IMFs 5 to 7 and (b) Fourier transforms of split signal using a Hanning window.

The Fourier spectra of the grouped IMFs are reported in Figure 6(b). The intermediate frequency motion has significant energy around 70 cm^{-1}, outside the previously proposed frequency target for YPGDV. Figure 7 shows the effect of targeting either the low ($0-25$ cm^{-1}) or the low- and intermediate- ($0-100$ cm^{-1}) frequency regions identified here. Significant conformational motions are induced in both cases. However, amplifying the intermediate frequencies progressively targets the highest amplitude motion, yielding greater conformational change. In this analysis the frequency resolution is limited by the length of the buffer, and a reduced resolution can be expected compared with the 16 678 step NVE simulation analyzed previously.

The filter target can therefore be selected by either measuring the system's response to a range of filters or by predicting the desired frequency range from analysis of a sample trajectory. As has been shown, it is important to target frequencies relevant to the system, and the YPGDV results suggest no benefit in targeting above 100 cm^{-1}. The $0-100$ cm^{-1} range is suggested as optimal from both measurements of the system's response and from analysis of frequencies related to inherent dihedral motions.

3.2. Buffer Length. To apply a digital filter to a buffer of velocities, the buffer must contain at least the number of steps as there are coefficients in the filter. The larger the number of coefficients, the closer the frequency response of the filter will be to that for which it is designed. A shorter buffer requires reduced computational expense, but a filter of insufficient length will not produce a precise response and may result in undesired amplification or suppression of motions.

The previous filter target of $0-25$ cm^{-1} requires 1001 coefficients to achieve a reasonable filter response.[2] A higher cap on the frequency response suggested by the analysis presented here does not require as many coefficients to produce a sufficiently precise response. The responses of a

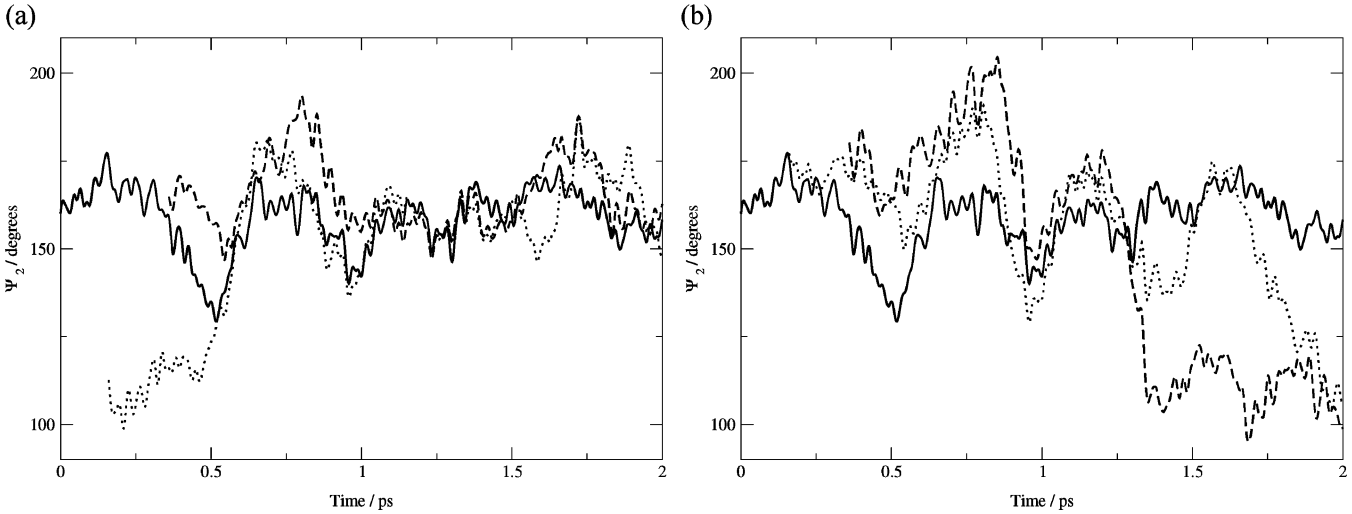


Figure 7. ψ_2 trajectory after 0 (solid), 5 (dotted), and 10 (dashed) filter applications: (a) 0−25 cm^{-1} filter applied and (b) 0−100 cm^{-1} filter applied.

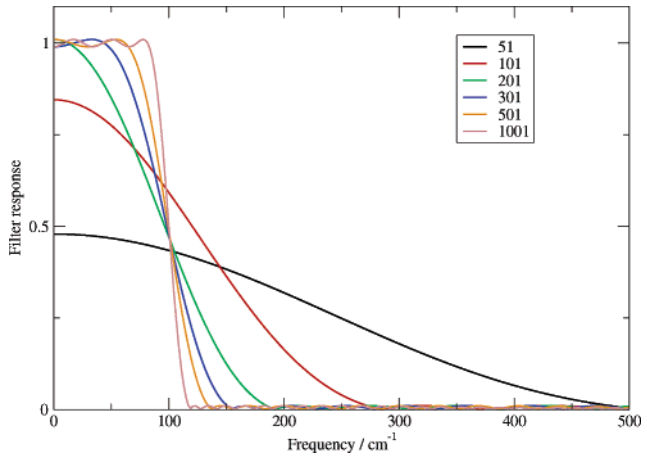


Figure 8. Filter responses of 0−100 cm^{-1} filters using a different number of coefficients (shown in legend).

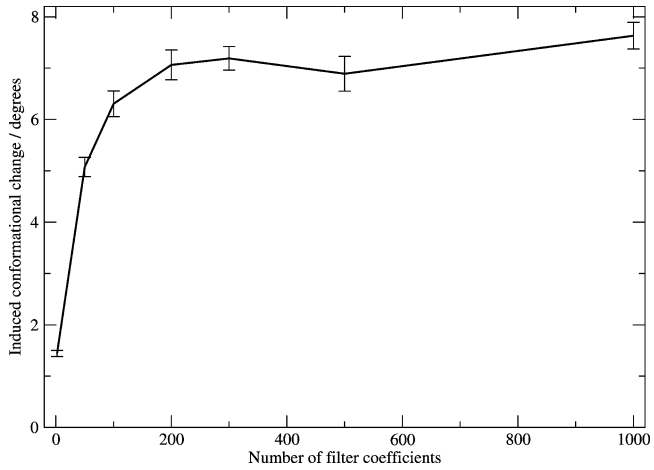


Figure 9. Results averaged from 50 RDFMD simulations using filters with different numbers of coefficients.

number of 0−100 cm^{-1} filters using different numbers of coefficients are shown in Figure 8.

RDFMD simulations using 0−100 cm^{-1} filters with different numbers of coefficients have been performed. 50 starting points for single filter applications were produced from the equilibrated YPGDV system with randomized velocities and 1000 steps (2 ps) of NPT MD. The amplification of each filter is adjusted to increase the system's kinetic energy by 25 kcal mol^{-1}. The induced conformational change is calculated as before, and the results are shown in Figure 9.

No benefit of using a greater number of coefficients than that required to produce an accurate filter response curve (200−300 coefficients) is seen. Fewer coefficients than this amplify higher than desired frequencies, for example, 100 coefficients targets 0−200 cm^{-1} and induces a similar level of conformational change to a filter targeting this region with 1001 coefficients in Figure 5(a).

3.3. Filter Delay. A delay between filters allows energy put into the system by the previous filter to dissipate and thus prevents overheating, keeping the system away from the internal temperature cap. Ideally the delay should be as long as possible so that the energy build up is slow and the simulation advances over the potential energy surface. However the effects of each filter application can quickly dissipate and too long a filter delay will result in a series of essentially independent filter applications.

RDFMD simulations starting from each of the 10 equilibrated YPGDV states were run using different filter delays. Other parameters are as specified in the original protocol, including an amplification factor of 4 and a 0−25 cm^{-1} filter. The induced conformational change produced by each filter application is shown in Figure 10 (a). The shorter the delay, the higher the initial induced conformational change, and the quicker the temperature cap is reached. Longer delays (> 40 steps) do not quickly reach the temperature cap, and more buffers are completed, each yielding progressive amplification of targeted motions. Very long delays (> 150 steps) yield constantly low conformational changes as energy is dissipating between buffers and the filter applications become more independent and less progressive. The energy build up is more clearly seen in Figure 10(b) in which the internal protein temperature after a filter application is plotted against the buffer number. Short delays quickly reach the 2000 K temperature cap, and long delays see no progressive increase in internal temperature.

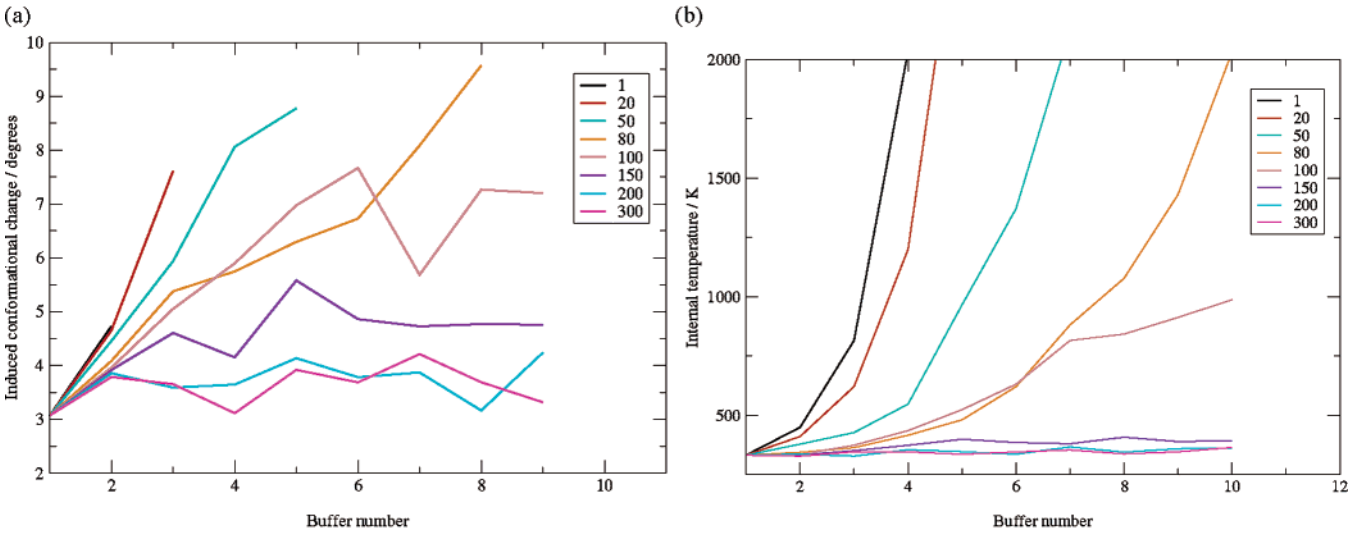


Figure 10. Effects of changing the filter delay parameter. Results averaged over 10 simulations: (a) conformational change induced in filter buffer and (b) internal protein temperature after filter application.

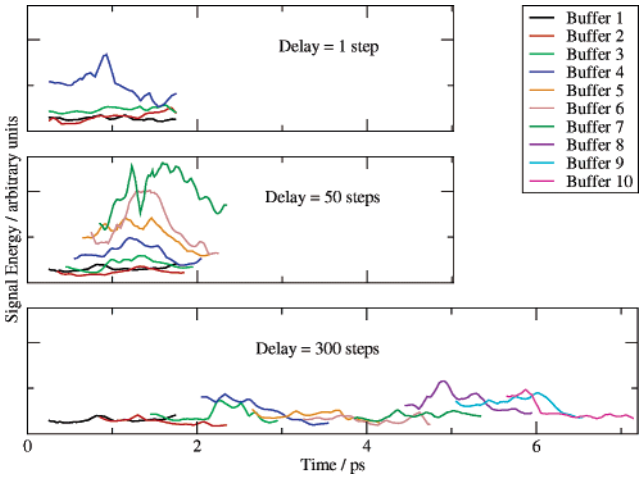


Figure 11. HHT of dihedral angle trajectories for different filter delay parameters. Buffers completed before reaching the internal temperature cap shown. *y*-axes are shown to the same scale for comparison.

Using the Hilbert-Huang Transform, it is possible to see the energy build up in dihedral angle signals (the signal energy) for the targeted frequency range. Figure 11 shows the results for one of the RDFMD simulation starting points showing energy in motions occurring in the 0−50 cm^{-1} region. A running average over 50 steps has been performed on the data and the first and last 100 steps of each buffer removed as frequency information at the edges of HHT data can be unreliable.[26] A delay of 1 step shows the greatest increase in energy for each amplification performed, but only four buffers are completed before reaching the temperature cap. A delay of 50 steps completes seven buffers as the system is able to relax between filter applications. The signal energy in the targeted frequency region reaches a higher level, before breaching the temperature cap, than the simulation with a 1 step delay, an improvement that shows the importance of the filter delay. A long delay of 300 steps shows little amplification, and each buffer does not necessarily reach higher levels of low-frequency energy than the last, suggesting that a long delay produces a nonprogressive protocol.

For the original RDFMD parameter set, a filter delay between 50 and 100 steps is clearly most suitable, showing progressive amplification without overheating the system.

3.4. Amplification Factor. A specified input of kinetic energy (as used to analyze the frequency target) or increase of temperature can be used to adjust the amplification of the velocities in RDFMD. A fixed amplification factor was originally implemented that adjusts the level of kinetic energy put into the system according to the amount that was already there. If there is a small amplitude, low-frequency motion present in the system, the filter will increase the kinetic energy by less than if the motion is of larger amplitude.

Regardless of the method of energy insertion, for a progressive protocol, sufficient energy must be put into the system so the effect of one filter application has not dissipated before the next. Equally, too great an amplification of velocities could overheat the system, a protocol that would risk denaturing larger protein systems.

A range of RDFMD simulations using different amplification factors from each of the 10 equilibrated YPGDV states has been performed. The induced conformational change is measured as previously described. The amount of kinetic energy in the low-frequency range is calculated from the velocities extracted by the applied 0−25 cm^{-1} filter ($\mathbf{v}_{Filt}$ in eq 2). A linear correlation is found between low-frequency energy and induced conformational change, even when results are averaged over several simulations (typically with a correlation coefficient > 0.8). Lines of best fit are plotted in Figure 12 (a) to the edges of the data set. The amount of conformational change induced when a specific quantity of energy is measured in the filter region increases as the amplification factor is raised. An amplification factor of 1 is inefficient, and later buffers do not explore regions of significantly increased low-frequency energy. RDFMD is stopped by reaching the maximum filter cap of ten. An

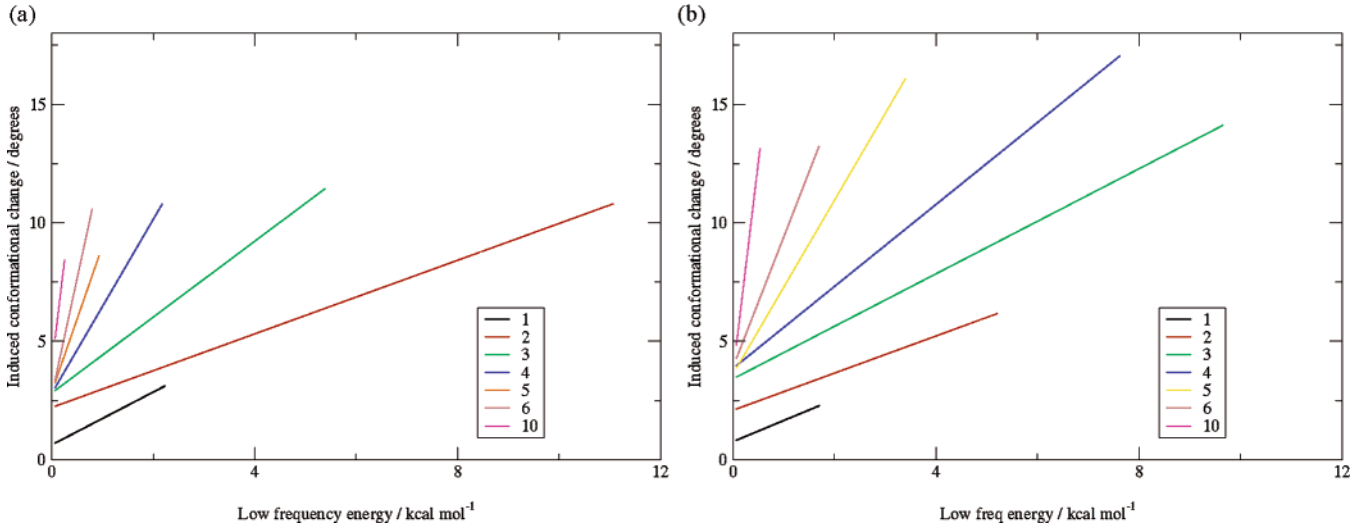


Figure 12. Effects of differing amplification factor (shown in legend): (a) filter delay of 20 steps and (b) filter delay of 50 steps.

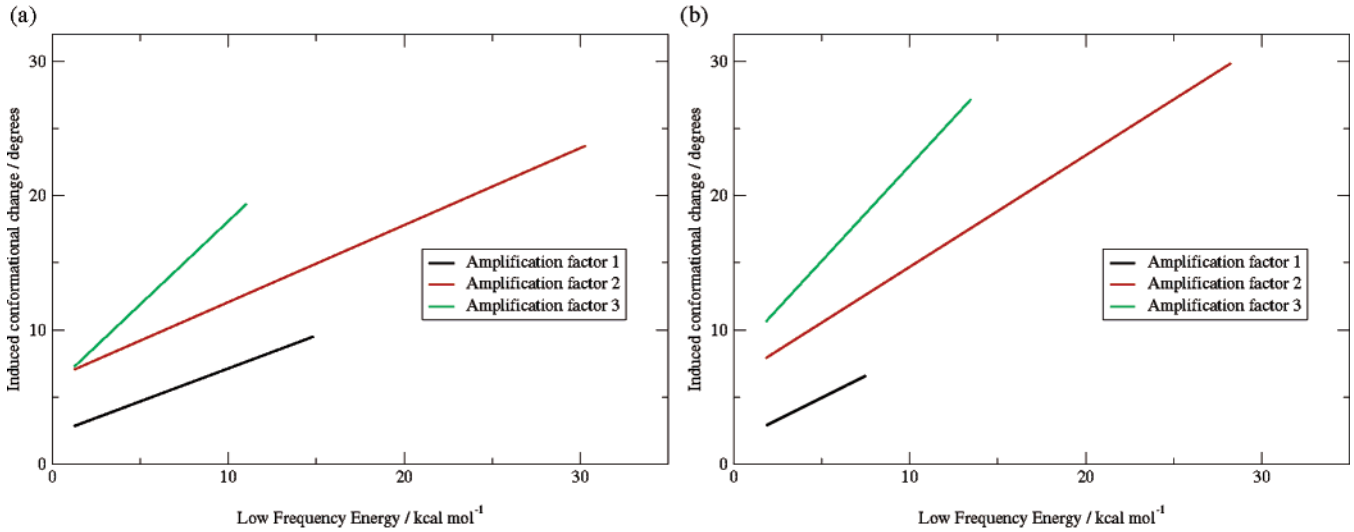


Figure 13. Results for a range of filter delay parameters and amplification factors. Data are averaged over 10 simulations and presented as previously described: (a) filter delay of 50 steps and (b) filter delay of 100 steps.

amplification factor of 2 yields greatly increased energy in the filter region and is clearly the most suitable compliment to the rest of the parameter set. At, and above, an amplification factor of 3, the internal temperatures reach the cap of 2000 K, and simulations are stopped. Although significant conformational motion is induced by high amplification factors, there are minimal increases in low-frequency energy over a small number of buffers before the temperature cap is reached. The protocol is therefore neither gentle nor progressive, as is desired.

The energy input required for gentle but progressive low-frequency energy amplification is heavily dependent on the rest of the parameter set. For example if the delay parameter is increased to 50 steps (a suitable choice as seen from previous analysis), an amplification factor of 2 shows slow energy build up and factors of 3 to 4 are more suitable, as shown in Figure 12(b). The improvement using the longer delay is clear, with 50% more conformational motions induced when comparing the best result to that of the simulations using a 20 step filter delay.

Once a frequency target and filter delay have been chosen for application to a system, analysis similar to that shown here will assist in tailoring the method of amplification.

3.5. Interdependence of Parameters. As previously discussed, the parameters used for RDFMD are heavily interdependent. Once a frequency target has been chosen however, optimizing the other parameters discussed here can be done systematically with a small number of trial simulations.

Results presented so far suggest use of a 0−100 cm^{-1} filter, a delay between 50 and 100 steps (0.1−0.2 ps), and amplification factors less of than 4. An accurate frequency response for a 0−100 cm^{-1} digital filter is produced using 301 coefficients (Figure 8). Averaged results for delay parameters and amplification factors in the suggested regions are presented in Figure 13(a),(b) as previously described. An amplification factor of 2 is shown to be the best choice for a filter delay of 50 or 100 steps. It is worth noting that for both parameter sets the levels of induced conformational

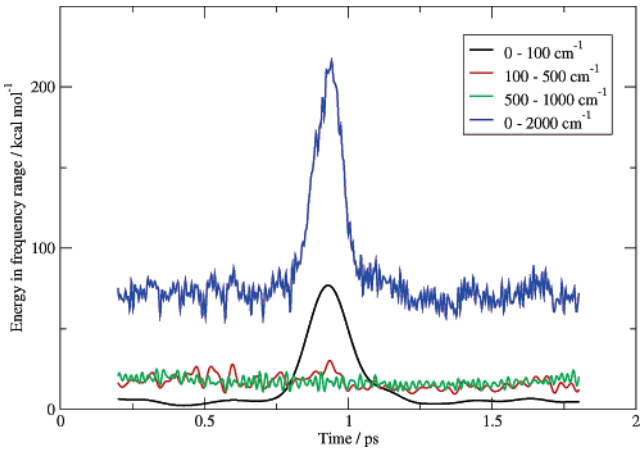


Figure 14. Kinetic energy in different frequency bands (shown in legend) of an RDFMD buffer.

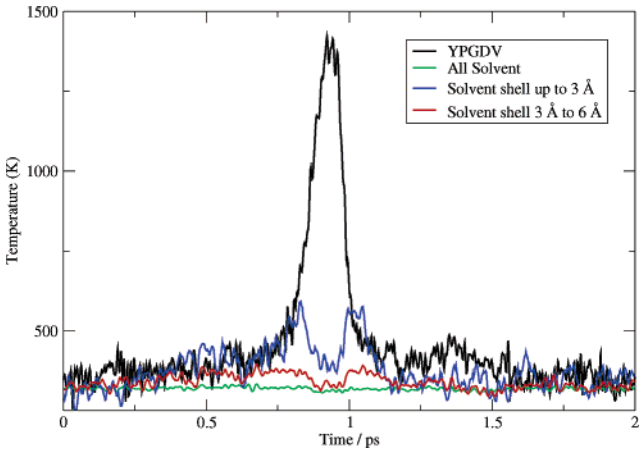


Figure 15. Temperature fluctuations in solvent and solute regions (shown in legend) across an RDFMD buffer.

change and of low-frequency energy are much higher than with previous protocols.

The protocols developed here have been optimized with the goal of promoting maximum dihedral motion while limiting the energy put into the system. The parameters should be applicable to any similar protein, or part of a protein, for which this is desired.

Discussion

It has been shown that energy dissipates between filter applications, and this energy can dissipate into either solute degrees of freedom with higher frequencies of motion than those targeted (i.e. nonselective heating) or into surrounding solvent molecules. To analyze this effect, the kinetic energy in motions of a particular frequency can be calculated across a buffer of velocities using a sliding digital filter. For example, a filter that multiplies all motions with frequencies of 0−100 cm^{-1} by a factor of 1 and all else by a factor of 0 can be produced with reasonable accuracy using 201 coefficients (Figure 8). By applying this filter to the first 201 velocities in an RDFMD buffer, an estimation of the velocities in low-frequency motions can be attained for step 101. The filter can be applied to steps 2−202 to estimate velocities at step 102 and so on across the buffer. Figure 14 shows the kinetic energy seen by a range of filters applied in this manner. The buffer analyzed is the last before reaching the temperature cap from an RDFMD simulation of YPGDV using the previously reported protocol (0−25 cm^{-1} frequency target, filter delay of 50 steps, 1001 coefficient filter, amplification factor of 4, and an internal temperature cap of 2000 K). The energy that has been added is clearly localized under 100 cm^{-1}, as desired.

To analyze the dissipation of energy into surrounding solvent molecules, the average temperature of solvent shells around the YPGDV protein has been calculated across a filter buffer. Figure 15 shows the results of a YPGDV RDFMD simulation using the previously reported protocol. Again the last buffer before reaching the temperature cap is used. The first solvent shell includes all water molecules with the oxygen atom within 3 Å of a protein atom. The number of water molecules included in the first shell fluctuates between 84 and 141 across this buffer. The second shell includes waters for which the oxygen atom lies within 3−6 Å of the closest protein atom and includes between 453 and 549 water molecules. The temperatures of the two shells increase on either side of the filter application, showing energy dissipating into the solvent. This is considered to be desirable, as the conformational flexibility of a protein will be linked to the mobility of surrounding solvent molecules.

Thus it has been shown that the energy dissipation observed after the application of a digital filter occurs predominantly into the solvent and not to higher frequency vibrations in the protein.

5. Application to DHFR

Escherichia coli dihydrofolate reductase (EcDHFR) catalyzes the reduction of 7,8-dihydrofolate (H_2F) to 5,6,7,8-tetrahydrofolate (H_4F) using nicotinamide adenine dinucleotide phosphate hydride (NADPH) as a reducing agent. During this reaction, the M20 loop (residues 15 to 20) adopts various conformations termed “closed”, “open”, and “occluded”.[29] NMR experiments show evidence of motions occurring in this region for apo-EcDHFR.[30,31] X-ray structures of various DHFR complexes can be found in the protein data bank[32] including entry 1RX2[33] for which DHFR exhibits a closed conformation complexed with folate and $NADP^+$, and entry 1RA9[33] with DHFR in an open form complexed with $NADP^+$. Here we show the application of RDFMD to the M20 loop, inducing a conformational change from the closed to open state.

Simulations of the thermal unfolding of EcDHFR have been published[34] revealing that the adenosine-binding domain (ABD, residues 38 to 106) partially unfolds at 400 K, and this region has been used to monitor the affect of RDFMD on protein stability.

An equilibrated solution structure of apo-EcDHFR has been produced from the closed 1RX2 structure. A cubic system with TIPS3P solvent extending at least 10 Å from the protein surface was prepared and neutralized by the addition of sodium ions, using the XLEAP module provided with AMBER 7.0.[35] Polar hydrogens positions and protonation states were determined by WHATIF.[36] The NAMD[18] package was used for further simulation with force field parameters taken from the CHARMM27 force field.[21]

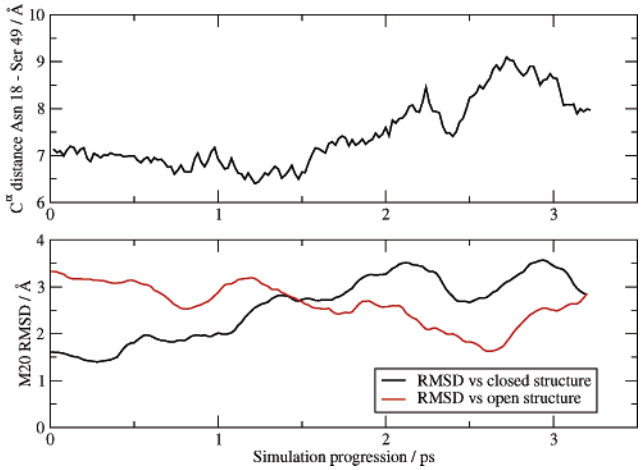


Figure 16. DHFR opening event during RDFMD simulation.

Mimimization was performed with the conjugate gradient line-search algorithm,[18] applying 25 000 steps to the solvent, 200 steps to the ions, 150 000 to solvent and ions, 3000 to the protein, and a further 30 000 steps to the entire system. Annealing was performed using a Langevin thermostat[22] with a damping parameter of 10 ps^{-1}, for 5000 steps at 50 K, 75 000 steps at 100 K, 10 000 steps at 150 K, 125 000 steps at 200 K, and 30 000 steps at 298 K. A Nosé-Hoover Langevin piston barostat[23] with a pressure target of 1 atm, a piston period of 400 fs, and a piston decay of 300 fs was applied for 125 000 steps of NPT simulation, followed by 50 000 steps with a thermostat damping parameter of 1 ps^{-1}. A 2 fs time step is used throughout the protocol with a switching function applied to Lennard-Jones interactions between 9.0 and 10.5 Å, SHAKE[20] applied to all bonds containing a hydrogen, and a PME treatment of electrostatics.[19]

A 4 ns NPT MD simulation was performed using the final equilibration parameters. Root-mean-squared deviations (RMSDs) have been calculated against the known closed (1RX2) and open (1RA9) structures with superposition across all secondary structure units. The RMSD of the protein secondary structure against the starting structure does not rise significantly, with an average of 0.82 Å and standard deviation of 0.10 Å. Motion of the M20 loop fragment is limited with a RMSD reaching 3 Å against the closed structure on three brief occasions, with no accompanying decreases in RMSD against the open structure. The closed conformer is characterized by a short distance between residues 18 (Asn) and 49 (Ser) where favorable polar interactions hold the M20 loop in place. The α-carbon distance between these two residues averages at 6.80 Å with a standard deviation of 0.64 Å for this simulation. The average RMSD of the ABD domain against the starting structure is 0.90 Å for the first 1.5 ns, with a standard deviation of 0.10 Å. It rises after 1.7 ns and remains stable for the remainder of the simulation, with an average of 1.29 Å and standard deviation of 0.17 Å for the final 2 ns. Full analysis of the simulation will be presented elsewhere; however, it does not suggest that DHFR leaves the closed conformation.

A suitable RDFMD protocol as suggested by YPGDV analysis has been chosen for application to DHFR. A frequency analysis of the dihedral angles in the M20 loop of DHFR yields similar results to those observed for YPGDV. A paper reporting this analysis in more detail is in preparation. The protocol includes use of a 301 coefficient filter targeting 0−100 cm^{-1} and a filter delay of 50 steps. The filter is applied to all atoms in DHFR residues 15 to 20. An amplification factor of 2 has been chosen based upon the results shown in Figure 13 and the internal temperature capped at 1500 K. 30 filter buffers have been performed for each simulation, and should an amplification factor of 2 bring the internal temperature above 1500 K, it is lowered so that the temperature cap is not breached and all buffers are completed.

Once again 10 simulations are performed using randomly assigned velocities at 300 K. Each simulation showed significantly increased loop mobility, with several opening events occurring. One of the clearest results measured the M20 loop RMSD against that of the 1RX2 closed structure rising from 1.38 Å to a maximum of 3.54 Å. As this occurs the loop RMSD against the 1RA9 open structure falls from 3.31 Å to 1.60 Å. The interresidue (carbon-α) distance between residues 18 and 49 rises from 7.13 Å to a maximum of 9.10 Å. Figure 16 shows the progression of these factors

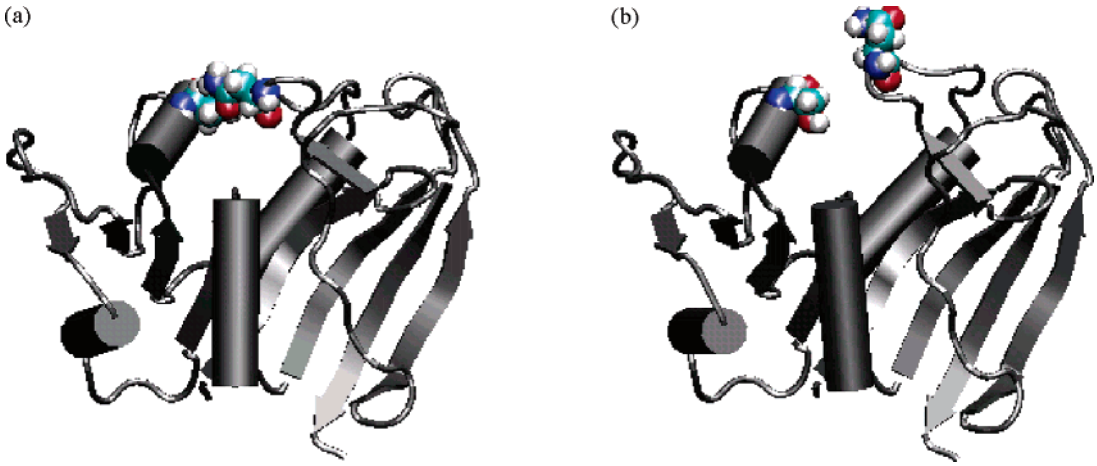


Figure 17. Conformers sampled during RDFMD DHFR simulation. van der Waals radii shown for residues 18 and 49: (a) conformation most similar to known closed structure and (b) conformation most similar to known open structure.

extracted from the forward components of the RDFMD buffers into a single trajectory. During the simulation the RMSD of the ABD domain does not rise above 1 Å. The most closed and open structures from this trajectory are shown in Figure 17.

The parameter set successfully used here has not undergone any further optimization, confirming the suitability of the derived protocol for application to similar systems. Positive results have also been produced using T4 lysozyme and HIV-1 protease (to be presented elsewhere).

Conclusions

In this paper we have discussed methods of assigning RDFMD parameters using the YPGDV pentapeptide as a test case. The goal has been to maximize the induced conformational change and promote the gentle and progressive amplification of low-frequency motions. The analysis methods developed here are intended as a guide for the application of RDFMD to any system and for any purpose. We anticipate, and are investigating, uses in protein folding, ligand binding, and inducing large-scale conformational motion in proteins of significant size.

The first parameter to be chosen is the frequency target. For this, frequency analysis can be performed on signals extracted from a sample trajectory, using Fourier or Hilbert based methods. Alternatively, the response of the system to filters amplifying different frequency regions can be used. Both methods have been presented here, suggesting filters targeting 0−100 cm^{-1} for promotion of dihedral motions. Empirical Mode Decomposition has been used to separate dihedral signals into high-frequency noise from coupled degrees of freedom, a signal in the region 25−100 cm^{-1} that persists throughout simulation, and a low-frequency component that only has significant amplitude during rare conformational events (targeted in previous studies). Use of a higher upper limit on the frequency target allows the length of the buffer to be dramatically lowered, thus reducing the computational expense of each filter application. This is particularly important when applying RDFMD to larger systems.

The number of steps between filter applications (the filter delay parameter) has been analyzed in detail, including the use of the Hilbert-Huang Transform, which has only recently been introduced to molecular dynamics simulations. Short delays show rapid energy increases that induce dihedral angle motions but quickly overheat the system, and so few filter applications are performed. Long delays show no progressive increase in low-frequency energy or induced conformational change, the effect of each filter application having dissipated before the next. Intermediate delays of 50−100 steps show both slow energy increases and significant induced conformational change.

There are several methods of tailoring the amount of energy put into the system. A fixed amplification factor inputs energy according to how much is already present in the target region; this prevents energy being placed into motions of low amplitude that have little relevance to the system. Dynamically adjusting the amplification factor so that a set amount of energy is put into the system is also of use when comparing the system's response to filters that target different frequency ranges. An increase in either kinetic energy (as presented in frequency target section) or an adjustment dependent on the system temperature (used for application to DHFR) can be specified.

The resulting parameter set has been successfully applied to the M20 loop of DHFR: a frequency target of 0−100 cm^{-1}, a filter delay of 50 steps, and an amplification factor of 2, which is reduced should the internal temperatures reach 1500 K. A conformational change is induced from the closed to open form, without affecting the rest of the molecule.

Further work includes the development of parameter sets for alternative applications, and the use of derived protocols with other systems of significant size, including T4 lysozyme and HIV-1 protease. Long MD and parallel tempering simulations on these systems will be presented in subsequent publications to validate the conformational motions observed using RDFMD.

Acknowledgment. This work has been funded by GlaxoSmithKline, EPSRC, and BBSRC. Many thanks to Dr. R. Gledhill and Prof. N. Huang for their guidance with the EMD method and Hilbert transform and to Dr. C. Woods with his assistance with the preparation of this paper. We should also like to thank Prof. W. L. Jorgensen for his generous provision of the MCPRO program.

References

(1) Phillips, S. C.; Essex, J. W.; Edge, C. M. *J. Chem. Phys.* **2000**, *112*, 2586.

(2) Phillips, S. C.; Swain, M. T.; Wiley, A. P.; Essex, J. W.; Edge, C. M. *J. Phys. Chem. B* **2003**, *107*, 2098.

(3) Karplus, M.; McCammon, J. A. *Nat. Struct. Biol.* **2002**, *9*, 646.

(4) Tai, K. *Biophys. Chem.* **2004**, *107*, 213.

(5) Mitsutake, A.; Sugita, Y.; Okamoto, Y. *Biopolymers* **2001**, *60*, 96.

(6) Lyubartsev, A. P.; Martsinovski, A. A.; Shevkunov, S. V.; Vorontsovvelyaminov, P. N. *J. Chem. Phys.* **1992**, *96*, 1776.

(7) Berg, B. A.; Neuhaus, T. *Phys. Lett. B* **1991**, *267*, 249.

(8) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141.

(9) Wu, X. W.; Wang, S. M. *J. Phys. Chem. B* **1998**, *102*, 7238.

(10) Wu, X. W.; Brooks, B. R. *Chem. Phys. Lett.* **2003**, *381*, 512.

(11) de Groot, B. L.; van Aalten, D. M. F.; Scheek, R. M.; Amadei, A.; Vriend, G.; Berendsen, H. J. C. *Proteins: Struct., Funct., Genet.* **1997**, *29*, 240.

(12) *MATLAB 6.1.0*; The MathWorks Inc., Natick, MA, 2001.

(13) Tobias, D. J.; Mertz, J. E.; Brooks, C. L. *Biochemistry* **1991**, *30*, 6054.

(14) Karpen, M. E.; Tobias, D. J.; Brooks, C. L. *Biochemistry* **1993**, *32*, 412.

(15) Wu, X. W.; Wang, S. M. *J. Phys. Chem. B* **2000**, *104*, 8023.

(16) Dyson, H. J.; Rance, M.; Houghten, R. A.; Lerner, R. A.; Wright, P. E. *J. Mol. Biol.* **1988**, *201*, 161.

(17) Jorgensen, W. L. *MCPRO 1.4*, Yale University, New Haven, CT, 1996.

(18) 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.

(19) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.

(20) Ryckaert, J. P.; Ciccoti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.

(21) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; 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, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586.

(22) Paterlini, M. G.; Ferguson, D. M. *Chem. Phys.* **1998**, *236*, 243.

(23) Feller, S. E.; Zhang, Y. H.; Pastor, R. W.; Brooks, B. R. *J. Chem. Phys.* **1995**, *103*, 4613.

(24) Huang, N. E.; Shen, Z.; Long, S. R.; Wu, M. L. C.; Shih, H. H.; Zheng, Q. N.; Yen, N. C.; Tung, C. C.; Liu, H. H. *Proc. R. Soc. London, Ser. A* **1998**, *454*, 903.

(25) Bendat, J. S. *The Hilbert Transform;* Brüel & Kjær: Nærum, Denmark.

(26) Phillips, S. C.; Gledhill, R. J.; Essex, J. W.; Edge, C. M. *J. Phys. Chem. A* **2003**, *107*, 4869.

(27) Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577.

(28) Wiley, A. P.; Gledhill, R. J.; Phillips, S. C.; Swain, M. T.; Edge, C. M.; Essex, J. W. *Hilbert-Huang Transform Engineering: The analysis of molecular dynamics simulations by the Hilbert-Huang transform*; Marcel Dekker: New York, in press.

(29) Rod, T. H.; Brooks, C. L. *J. Am. Chem. Soc.* **2003**, *125*, 8718.

(30) Li, L. Y.; Falzone, C. J.; Wright, P. E.; Benkovic, S. J. *Biochemistry* **1992**, *31*, 7826.

(31) Falzone, C. J.; Wright, P. E.; Benkovic, S. J. *Biochemistry* **1994**, *33*, 439.

(32) 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.

(33) Sawaya, M. R.; Kraut, J. *Biochemistry* **1997**, *36*, 586.

(34) Sham, Y. Y.; Ma, B. Y.; Tsai, C. J.; Nussinov, R. *Proteins: Struct., Funct., Genet.* **2002**, *46*, 308.

(35) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E.; Debolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. *Comput. Phys. Commun.* **1995**, *91*, 1.

(36) Vriend, G. *J. Mol. Graphics* **1990**, *8*, 52.

CT049970T

JCTC Journal of Chemical Theory and Computation

Superposition State Molecular Dynamics

Arun Venkatnathan 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 October 14, 2004

Abstract: The ergodic sampling of rough energy landscapes is crucial for understanding phenomena like protein folding, peptide aggregation, polymer dynamics, and the glass transition. These rough energy landscapes are characterized by the presence of many local minima separated by high energy barriers, where Molecular Dynamics (MD) fails to satisfy ergodicity. To enhance ergodic behavior, we have developed the Superposition State Molecular Dynamics (SSMD) method, which uses a superposition of energy states to obtain an effective potential for the MD simulation. In turn, the dynamics on this effective potential can be used to sample the configurational free energy of the real potential. The effectiveness of the SSMD method for a one-dimensional rough potential energy landscape is presented as a test case.

1. Introduction

The Molecular Dynamics (MD) simulation method is an extremely important tool in chemistry, materials science, and biology. A number of MD methods developed so far have effectively simulated processes such as protein folding, peptide aggregation, and surface deposition. Needless to say, for a large set of molecules (e.g, large proteins), there are many competing interactions (intermolecular and intramolecular), leading to an extremely rough energy landscape and thus making the process of MD simulation a challenging task. The difficulty in sampling these rough energy landscapes is severe due to the presence of many local minima separated by high energy barriers. At normal temperatures, conventional MD or Monte Carlo (MC) simulations will be largely trapped in one of the various local minima. In turn, only some parts of the entire phase space will be sampled, and consequently structural, dynamical, and thermodynamic properties of the system cannot be reliabily calculated.

A number of MC algorithms have been developed to overcome the ergodicity problems posed by rough energy landscapes, such as multicanonical MC sampling,[1] parallel tempering,[2] multicanonical parallel tempering,[3] and simulated tempering.[4] Multicanonical sampling replaces the Boltzmann weight by a non-Boltzmann weight to effectively obtain a flat energy distribution between neighboring local minima. Parallel tempering improves sampling by using noninteracting configuration (or replicas) with different temperatures. This method can be further enhanced through a parallel implementation by distributing the noninteracting configurations across computational nodes. Furthermore, a multicanonical parallel tempering method, combining the advantages of parallel tempering and multicanonical sampling, has been developed by Faller et al.[3]

In the simulated tempering method, a random walk in temperature space is used to sample the free energy space by escaping the local energy minima. This method has been applied to study protein folding.[5] Stolovitzky and Berne[6] have also developed the catalytic tempering method to reduce the free energy barriers and to accelerate the sampling in phase space of complex systems without disturbing the actual potential minima. The basin-hopping MC algorithm of Wales et al.[7] also reduces the free energy barriers as seen in their implementation to perform global optimization on various atomic clusters. Here, the actual potential energy surface has been modified by a staircase-like effective potential without disturbing the positions of any minima.

The methods discussed in the previous paragraph are primarily based on MC. On the other hand, the Replica Exchange Molecular Dynamics (REMD) method of Sugita and Okamoto[8] is similar to the original parallel tempering approach but instead uses MD. Voter and co-workers[9] have also developed a class of methods called accelerated MD methods, which includes hyperdynamics, temperature accelerated dynamics, and parallel replica dynamics (PRD). In the hyperdynamics method of Voter,[10] the acceleration of the MD simulation requires a computation of the gradients

10.1021/ct0499175 CCC: $30.25
Published on Web 12/08/2004

and Hessian of the potential. The PRD method can offer a large boost in simulation time, and its success has been seen, e.g., in a study of Cu(100) surface vacancy diffusion.[11] The accelerated MD method of Miron and Fichthorn[12] is similiar to the work of Voter where the potential energy close to the local minima is modified by a boost potential to obtain an effective potential. The success of this method has been shown for a surface diffusion problem as well. Unfortunately, the choice of the boost potential can lead to certain instablities in the numerical derivatives, giving rise to artificial energy peaks near the local minima. The conformational flooding method of Grubmüller[13] has been employed to predict structural transitions in irregular or disordered systems where conventional MD fails. This method uses a "fictitious" potential derived by defining a conformational substate which samples all regions pertaining to low free energy thereby reducing the free energy barriers in the original energy landscape. Also, the metadynamics approach[14,15] has been developed to sample rough energy landscapes by using a time-dependent Gaussian distribution to fill in the minima of the original potential as they are visited, thus biasing the dynamics to explore the more inaccessible regions of the energy landscape. The puddle-skimming method[16] and accelerated MD method of Hamelberg et al.[17] have also explored mechanisms to obtain an effective potential energy surface by using an appropriate choice of bias potential and boost energy. The puddle-skimming method[16] does not perform well for high values of boost energy and also has certain discontinuities in the computation of derivatives of the real potential along the potential energy surface. This flaw has been removed by Hamelberg et al.[17] through a better choice of the bias potential. Although, both of these methods seem to be promising, they involve a single choice of an effective potential surface.

The purpose of this paper is to introduce a simple and flexible MD method to help sample rough energy landscapes. Here, we will study the canonical ensemble generated by the Nose-Hoover Chain (NHC)[18] thermostat to sample free energy profiles, although any MD ensemble should be possible. The NHC has proven to be a reliable thermostat for many systems. However, as might be expected, the NHC does not sample ergodically in cases where there are rough energy landscapes.

The new method presented here is called Superposition State Molecular Dynamics (SSMD), which uses an effective superposition potential to help sample the entire phase space. The motivating concept behind this method is the way in which nonstationary states in quantum mechanics can explore the Hilbert space by undergoing transitions between basis states. The SSMD method also has some similiarities to the hyperdynamics approach used by Voter[10] to overcome the high energy barriers. The recent work using the puddle-skimming method[16] and the accelerated MD method of Hamelberg et al.[17] can also be largely recovered as one possible limit of the SSMD approach. However, in SSMD one performs a MD simulation on the effective potential obtained by using a superposition of energy states, thus allowing for a very large number of possible choices to aid in the energy landscape sampling.

To demonstrate the effectiveness of the SSMD method, we have chosen a one-dimensional potential formed from a distribution of Gaussian barriers in which standard canonical MD with a NHC fails to sample ergodically.

2. Method

Let us denote $V(x)$ as the real potential on which MD sampling fails to be ergodic due to the presence of a rough energy landscape. In SSMD, the MD simulation is carried out on an effective potential using a superposition of "states". The SSMD effective potential is constructed by using a combination of the real energy state $V(x)$ and a set of N "fictitious" states. To illustrate this N-state superposition concept, the simplest 2-state implementation is shown here, where $V(x)$ and $V_0(x)$ are the potential energy functions in the first and second state, respectively. It should be noted that $V_0(x)$ is *any fictitious potential* which can be freely chosen to enhance the MD sampling. The effective potential for the MD forces is then obtained by solving the 2×2 matrix eigenvalue problem, where the highest and lowest eigen-energies are the upper 2-state $[V_+(x)]$ and lower 2-state $[V_-(x)]$ surface on which the MD sampling can then be performed. It should be noted that one has the flexibility to run the MD on either $V_+(x)$ or $V_-(x)$, depending on the choice of $V_0(x)$ and/or the physical problem at hand.

Mathematically, the effective 2-state SSMD potential can be written as

$$V_{\pm}(x) = \frac{V(x) + V_0(x)}{2} \pm \frac{1}{2}\sqrt{4V_{12}^2(x) + (V(x) - V_0(x))^2} \quad (1)$$

where one chooses either $V_+(x)$ or $V_-(x)$ for the dynamics as discussed earlier.

The governing sampling equations for SSMD simulation are as follows: Consider, the exact canonical distribution function, given by

$$P(x) = \frac{e^{-\beta V(x)}}{Q} \quad (2)$$

where β is equal to $1/k_BT$. Here, Q is the partition function related to $V(x)$ which can be written as

$$Q = Q_{\pm}\langle e^{-\beta \Delta V(x)}\rangle_{\pm} \quad (3)$$

where $Q_{\pm}$ is the partition function corresponding to the superposition potential $V_{\pm}(x)$ and $\Delta V(x) = V(x) - V_{\pm}(x)$. Using eqs 2 and 3 and the definition of $\Delta V(x)$, we can write the exact probability distribution function as

$$P(x) = \frac{e^{-\beta V_{\pm}(x)} e^{-\beta \Delta V(x)}}{Q_{\pm}\,\langle e^{-\beta \Delta V(x)}\rangle_{\pm}} \quad (4)$$

or

$$P(x) = P_{\pm}(x)\frac{e^{-\beta \Delta V(x)}}{\langle e^{-\beta \Delta V(x)}\rangle_{\pm}} \quad (5)$$

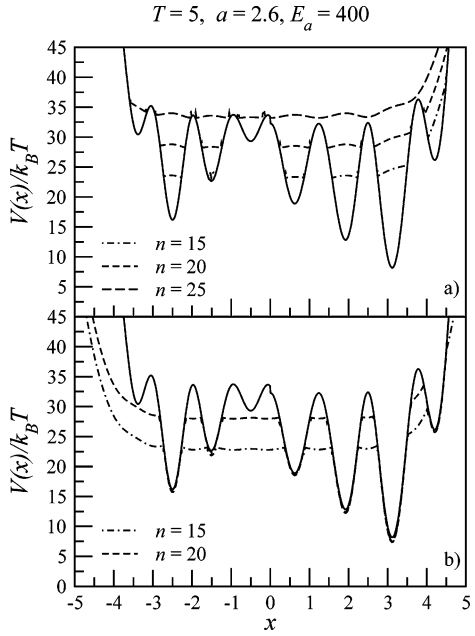


Figure 1. Real potential energies (solid lines) and effective SSMD potentials (dashed and dot−dashed lines) from an (a) upper 2-state SSMD surface and (b) lower 2-state SSMD surface.

where

$$P_{\pm}(x) = \frac{e^{-\beta V_{\pm}(x)}}{Q_{\pm}} \tag{6}$$

It should be noted that for the purposes of testing the SSMD sampling, the above equations can be rearranged to give

$$\frac{V(x)}{k_B T} = -\ln P_{\pm}(x) + \ln\langle e^{-\Delta V(x)/k_B T}\rangle_{\pm} - \ln Q + \frac{\Delta V(x)}{k_B T} \tag{7}$$

where the first two terms in the right-hand side of eq 7 are obtained from the SSMD simulation and the last two terms are easily calculated quantities for the simple potential used in this paper to test the SSMD method.

The choice of $V_0(x)$ is intended to enhance the SSMD sampling and also depends on the selection of effective potential surface, i.e., whether the SSMD is run on the upper surface, $V_+(x)$, or the lower surface, $V_-(x)$. If the MD sampling is performed on the upper 2-state surface, one simple choice is $V_0(x) = V_{\min} + nk_BT$, where $V_{\min}$ is an estimate of global minima of the real potential, while nk_BT corresponds to some user-defined height above V_{min}. For various choices of n, the different upper 2-state surfaces, $V_+(x)$, are shown in Figure 1a. As a result, when the real potential is lower than $V_0(x)$, the SSMD sampling is naturally performed on a "filled" potential energy surface with shallow wells. This approach is very similiar to that of the puddle-skimming method[16] and accelerated MD method of Hamelberg et al.,[17] although it is in principle much more flexible in the choice of $V_0(x)$.

Alternatively, the lower 2-state surface can be chosen as the SSMD potential, in which case a simple choice for $V_0(x)$ is for it to be equated to $V_{min} + nk_BT + V_{conf}$ to avoid being trapped in one of the local minima. Here V_{conf} is a confining potential to stop the trajectory from wandering away from the region of interest. For various values of n, the different lower 2-state surfaces obtained using the above criteria are shown in Figure 1b.

In both cases, the off-diagonal coupling elements (V_{12}) can be chosen as follows: When the difference between $V(x)$ and $V_0(x)$ is less than 1/2 k_BT, we chose V_{12} to be equal to k_BT. If the difference was more than 1/2 k_BT, then V_{12} was set as $d(V(x) - V_0(x))$, where $d = 1/k_BT$, though other choices of V_{12} are also possible.

The NHC equations of motion for a canonical ensemble are described in detail elsewhere.[18] However, for the sake of completeness, the NHC equations are written as follows:

$$\dot{r}_i = \frac{p_i}{m_i};\ \dot{p}_i = -\frac{\partial V(r)}{\partial r_i} - p_i\frac{p_{\eta_1}}{Q_1};\ \dot{\eta}_i = \frac{p_{\eta_i}}{Q_i}$$

$$\dot{p}_{\eta_1} = \left[\sum_{i=1}^{N_p}\frac{p_i^2}{m_i} - N_p k_B T\right] - p_{\eta_1}\frac{p_{\eta_2}}{Q_2}$$

$$\dot{p}_{\eta_j} = \left[\frac{p_{\eta_{j-1}}^2}{Q_{j-1}} - k_B T\right] - p_{\eta_j}\frac{p_{\eta_{j+1}}}{Q_{j+1}}$$

$$\dot{p}_{\eta_M} = \left[\frac{p_{\eta_{M-1}}^2}{Q_{M-1}} - k_B T\right] \tag{8}$$

where N_p is the number of particles, M is the number of thermostats, r's and p's are the particle positions and momenta, respectively, T is the kinetic temperature, and k_B is the Boltzmann constant. In eq 8, Q_i and m_i are masses of the ith thermostat and particle, respectively, while η and p_η are the thermostat positions and momenta, respectively. The conserved energy quantity corresponding to NHC dynamics is expressed as

$$H'(r, p, \eta, p_\eta) = V(r) + \sum_{i=1}^{N_p}\frac{p_i^2}{2m_i} + \sum_{i=1}^{M}\frac{p_{\eta_i}^2}{2Q_i} + N_p k_B T\eta_1 + \sum_{i=1}^{M} k_B T\eta_i \tag{9}$$

where $V(r)$ is the total potential energy of all particles in the system.

The Liouville operator for the equations of motion in eq 8 is given by

$$iL = \sum_{i=1}^{N_p} v_i\cdot\nabla_{r_i} + \sum_{i=1}^{N_p}\left[\frac{F_i(r)}{m_i}\right]\cdot\nabla_{v_i} - \sum_{i=1}^{N_p}\nu_{\eta_1} v_i\cdot\nabla_{v_i} + \sum_{i=1}^{M}\nu_{\eta_i}\frac{\partial}{\partial\eta_i} + \sum_{i=1}^{M-1}(\dot{p}_{\eta_i} - \nu_{\eta_i}\nu_{\eta_{i+1}})\frac{\partial}{\partial\nu_{\eta_i}} + \dot{p}_{\eta_M}\frac{\partial}{\partial\nu_{\eta_M}} \tag{10}$$

where v_i is the velocity of the ith particle and ν_{η_M} is the velocity associated with the Mth thermostat. The first two terms on the right-hand side of eq 10 denotes a shift in the particle positions and velocities and is performed by using

Table 1. Parameters Used To Compute the Analytic Potential in Eq 11

i	C_i	σ_i	μ_i ($x \le 0$)	μ_i($x > 0$)
1	0.8	0.4	−9.0	11.25
2	0.8	0.4	−8.0	10.00
3	0.8	0.3	−7.0	8.75
4	0.6	0.3	−6.0	7.50
5	0.6	0.4	−5.0	6.25
6	0.6	0.4	−4.0	5.00
7	0.4	0.3	−3.0	3.75
8	0.4	0.3	−2.0	2.50
9	0.4	0.4	−1.0	1.25
10	0.4	0.4	0.0	0.00

the velocity verlet algorithm.[19] The multistep time propagation algorithm of Martyna et al.[20] was used to solve eq 10.

In the present work, the real potential is computed as follows

$$V(x) = E_a \sum_{i=0}^{10} C_i \, exp\{-(x-\mu_i)^2/2\sigma_i^2\} + (x/a)^l \quad (11)$$

where a and E_a are chosen constants and l is the order of the polynomial. For example, values of $a = 1.3$, $E_a = 300$, and $l = 4$ were used as well as $a = 2.6$, $E_a = 400$, when $l = 8$, as described later. Also, C_i and σ_i were chosen to be a set of random numbers between 0 and 1 as given in Table 1, whereas μ_i was set on a equi-spaced grid with a spacing length of 1.0 when $x \le 0$, and 1.25 when $x > 0$, as also shown in Table 1.

3. Results and Discussion

Results from the SSMD simulation using the NHC are for the analytic (real) potentials computed from eq 11. The following parameters were used for all calculations: The mass (m) of the particle was equated to 1.0, the number of thermostats (M) was set to 10, and the mass of the thermostats [$Q(1{:}M)$] was equated to 1.0. The starting thermostat positions (η) and velocities (ν_η) were initialized to 1.0. All simulations were performed for 1×10^8 time steps.

In the first test, the use of the upper 2-state surface as an effective SSMD potential is demonstrated. The upper 2-state surface is computed as the highest eigenvalue from eq 1. The analytic potential is computed from eq 11. Figure 2a−c shows the potential energies predicted using SSMD from eq 7 in comparison to the analytic potential. Here, the different effective SSMD potentials are computed using the choice of fictitious "filling potential", $V_0(x) = V_{min} + nk_BT$, with various values for the parameter n. It is seen that the use of various effective SSMD potentials predicts potential energies which are in very close agreement to the analytic potential energy. Also, this example demonstrates that the upper 2-state surface acts as a very effective SSMD potential surface provided a reasonable choice of effective potential $V_0(x)$ is made.

The choice of the lower 2-state surface as the effective SSMD potential is given by the lowest eigenvalue from eq 1. For $V_0(x)$ in this case, we use $V_0(x) = V_{min} + nk_BT + (x/a)^l$. Figure 3 shows the predicted potential energies from

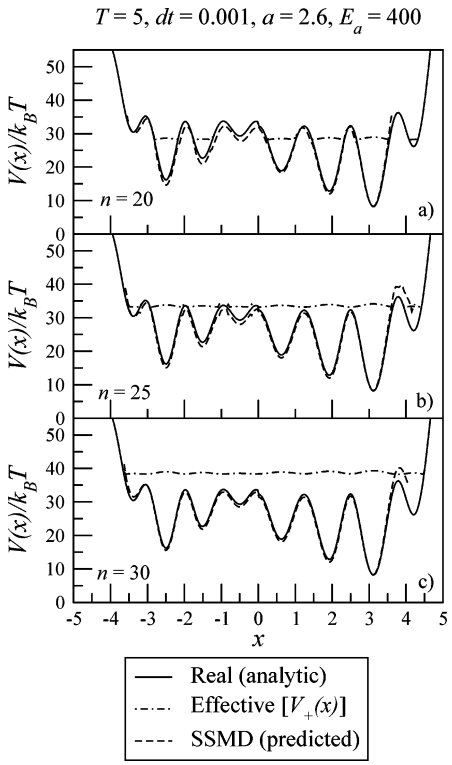


Figure 2. Predicted potential energies using the upper SSMD effective potential [$V_+(x)$], in comparison to the analytic potential, with the SSMD fictitious potential given by $V_0(x) = V_{min} + nk_BT$, at heights (a) $n = 20$, (b) $n = 25$, (c) $n = 30$, above the global minima [V_{min}].

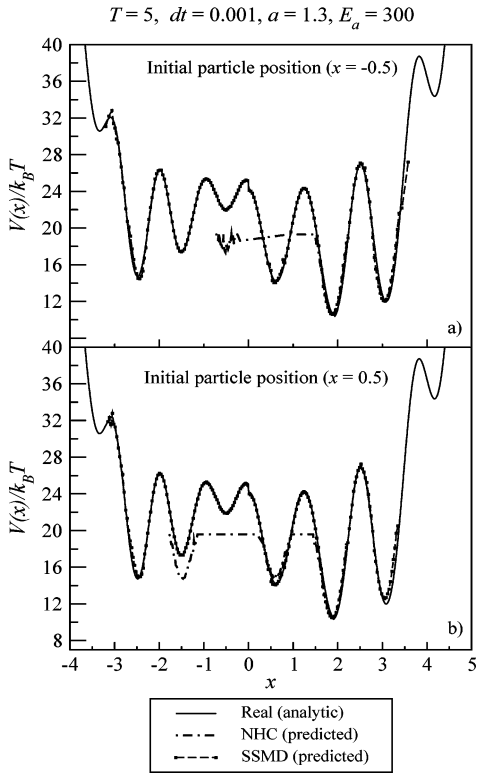


Figure 3. Predicted potential energies using NHC and the lower SSMD effective potential, [$V_-(x)$], in comparison to the analytic potential at $T = 5.0$ for different initial starting particle position at (a) $x = -0.5$ and (b) $x = 0.5$, as described in the text.

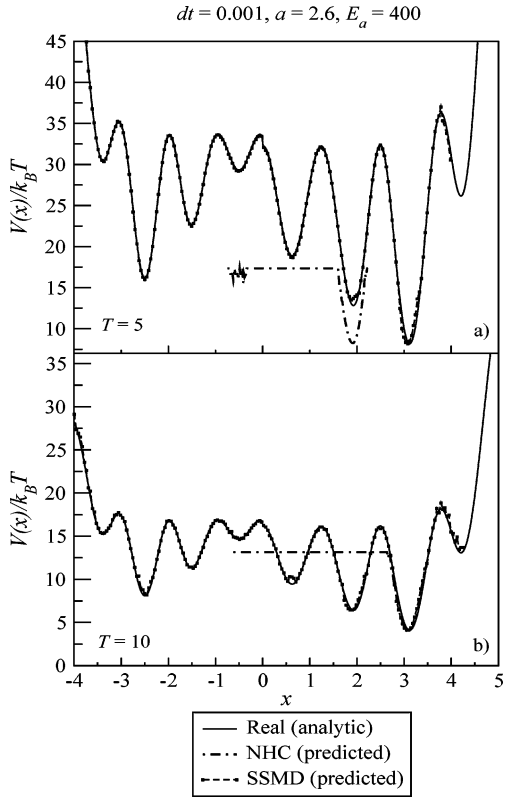


Figure 4. Predicted potential energies using NHC and the lower SSMD effective potential, $[V_{-}(x)]$, in comparison to a steeper walled analytic potential at (a) $T = 5$ and (b) $T = 10$, as described in the text.

NHC and SSMD in comparison to the analytic potential at simulation temperature $T = 5.0$, for $l = 4$ and $a = 1.3$. The MD sampling was performed with the choice of $n = 5$ in the fictitious potential. Here, different initial particle positions are used to test the SSMD method. As seen in Figure 3a,b, NHC on the actual potential $V(x)$ samples poorly as it gets trapped in one of the local minima. However, SSMD significantly increases ergodicity and predicts a result from eq 7 that resembles very closely the analytic potential.

The effectiveness of the SSMD method was further tested by choosing an analytic potential such that the walls rise more steeply ($l = 8$, $a = 2.6$, in eq 11; see also Table 1). As seen from Figure 4, this analytic potential shows more high barriers in the energy profile as compared to the example shown in Figure 3. Also, to ensure diversity in our tests and to further explore the performance of SSMD, simulations at two different temperatures are shown in Figure 4a,b. Again, the effectiveness of SSMD in sampling the rough energy landscape is observed, while standard NHC MD fails dramatically.

4. Concluding Remarks

The SSMD method has been presented in this paper and used to sample some model rough energy landscapes. It has been shown that the SSMD method is a simple and low cost computational method to ergodically sample potential energy surfaces. One possible advantage of SSMD is that it requires information primarily about global minima of the analytic potential. Also, the fictitious potential ($V_0(x)$ in SSMD) is completely general, and the choice between the upper 2-state or lower 2-state SSMD surface for the MD sampling can be made depending on the problem at hand. A mechanism to alternate between the upper 2-state and lower 2-state surface could also be implemented. In fact, in realistic situations the user may have a large number of possible choices for the SSMD fictitious potential $V_0(x)$ based on the need to overcome physical barriers in the actual potential energy surface. The choice of the off-diagonal element in SSMD should be carefully considered as well for multidimensional potentials, but here again there are many possible choices. SSMD therefore seems to be a promising and general tool to study systems with rough energy landscapes. Applications and further developments for realistic systems will be explored in the future.

Acknowledgment. This research is funded by the University of Utah Center for the Simulation of Accidental Fires and Explosions (C-SAFE), funded by the Department of Energy, Lawrence Livermore National Laboratory, under subcontract B341493. We acknowledge the Center for High Performance Computing, University of Utah for computer time support.

References

(1) Berg, B. A.; Neuhaus, T. *Phys. Rev. Lett*. **1992**, *68*, 9−12.
(2) Hansmann, U. H. E. *Chem. Phys. Lett*. **1997**, *281*, 140−150.
(3) Faller, R.; Yan, Q.; de Pablo, J. J. *J. Chem. Phys*. **2002**, *116*, 5419−5423.
(4) Lyubartsev, A. P.; Martsinovski, A. A.; Shevkunov, S. V.; Vorontsov- Velyaminov, P. N. *J. Chem. Phys*. **1992**, *96*, 1776−1783.
(5) Irbäcki, A.; Potthast, F. *J. Chem. Phys*. **1995**, *103*, 10298−10305.
(6) Stolovitzky, G.; Berne, B. J. *Proc. Nat. Sci*. **2000**, *97*, 11164−11169.
(7) Wales, D. J.; Doye, J. P. K. *J. Phys. Chem. A*. **1997**, *101*, 5111−5116.
(8) Sugita, Y.; Okamato, Y. *Chem. Phys. Lett*. **1999**, *314*, 141−151.
(9) Voter, A. F.; Montalenti, F.; Germann, T. C. *Annu. Rev. Mater. Res*. **2002**, *32*, 321−346.
(10) Voter, A. F. *J. Chem. Phys*. **1997**, *106*, 4665−4677.
(11) Voter, A. F. *Phys. Rev. B* **1998**, *57*, 13985−13988.
(12) Miron, R. A.; Fichthorn, K. A. *J. Chem. Phys*. **2003**, *119*, 6210−6216.
(13) Grubmüller, H. *Phys. Rev. E* **1995**, *52*, 2893−2906.
(14) Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci. U.S.A*. **2002**, *99*, 12562−12566.
(15) Micheletti, C.; Laio, A.; Parrinello, M. *Phys. Rev. Lett*. **2004**, *92*, 170601/1−170601/4.
(16) Rahman, J. A.; Tully, J. C. *J. Chem. Phys*. **2002**, *116*, 8750−8760.
(17) Hamelberg, D.; Mongan, J.; McCammon, J. A. *J. Chem. Phys*. **2004**, *120*, 11919−11929.
(18) Martyna, G. J.; Klein, M. L.; Tuckerman, M. *J. Chem. Phys*. **1992**, *97*, 2635−2643.
(19) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press: New York, 1987; pp 81−82.
(20) Martyna, G. J.; Tuckerman, M. E.; Tobias, D. J.; Klein, M. L. *Mol. Phys*. **1996**, *87*, 1117−1157.

CT0499175

JCTC Journal of Chemical Theory and Computation

New Effective Core Method (Effective Core Potential and Valence Basis Set) for Al Clusters and Nanoparticles and Heteronuclear Al-Containing Molecules

Nathan E. Schultz and Donald G. Truhlar*

Department of Chemistry and Supercomputer Institute, University of Minnesota, Minneapolis, Minnesota 55455-0431

Received April 26, 2004

Abstract: In previous work we have shown that the PBE0 hybrid density functional method with the MG3 all-electron basis set is an accurate method for calculating the atomization energies of small aluminum clusters (Al_2−Al_7). However, the MG3 basis set is very expensive for molecules much larger than Al_{13}; therefore, we have developed a new effective core potential (ECP) method for aluminum to reduce the cost of obtaining accurate results for nanoparticles. Our method involves a hybrid of the Stuttgart semiempirical effective core potential and the compact effective potential (CEP) potential, and it uses a newly optimized polarized valence triple-ζ basis set. The combination of the new ECP and the new polarized valence triple-ζ basis set for Al is called the Minnesota effective core (MEC) method for Al. The method was optimized with a training set of atomization energies and geometries for AlH, AlC, AlO, AlCCH, Al_2H, Al_2C, Al_2O, and Al_3 and atomization energies of three Al_{13} structures, and we tested it on six test sets composed of 20 atomization energies for systems as large as Al_{13}. We also present an improved all-electron polarized triple split basis set for oxygen, called 6-311+G(d*,p). For the test sets, the mean unsigned error (MUE) of the new method with respect to PBE0/MG3 is 0.06 eV for atomization energies and 0.007 Å for bond lengths, which constitutes a very significant improvement over the quality of the results that can be obtained with effective core potentials and valence basis sets in the literature (of the eight methods in the literature, the best previous method had average errors of 0.63 eV and 0.036 Å). We have also tested the MEC method with a variety of hybrid density functional theory, hybrid meta density functional theory, and pure GGA and meta GGA functionals and found that the average MUE, relative to each functional with all-electron basis sets, is 0.04 eV for atomization energies and 0.009 Å for bond lengths for the new effective core method and 0.16−0.20 eV and 0.013−0.033 Å for effective core potential and valence basis sets in the literature.

1. Introduction

Aluminum clusters have been the focus of several studies.[1−18] Clusters are of interest because their energetic and structural properties often differ significantly from the corresponding bulk properties. It is important to study how the properties change as the cluster size is increased into the nanoparticle range and then as the cluster properties eventually converge to the bulk limit. Unfortunately, though, experimental data are limited to small clusters[19−23] and bulk Al. Thus, theory can play an important role in studying clusters and nanoparticles in the intermediate size range.

In previous work,[18] we have begun testing electronic structure methods for Al. We tested several hybrid density functional theory[24−29] (HDFT) methods and found that the PBE0[28] method (also called PBE1PBE) with the MG3[30−36] basis set (which is a polarized triple-ζ basis set with diffuse functions) was accurate to within 0.012 eV per atom (for Al_2−Al_7). Using the PBE0 method with the MG3 basis set

10.1021/ct0400018 CCC: $30.25
Published on Web 12/21/2004

is impractical for systems that have more than about 15 atoms, because the computational effort (measured by CPU time) required for an energy calculation scales as N^4 (where N is the number of basis functions). One could reduce the size of the basis set, but that will reduce the reliability of the calculation. Another alternative would be to use an effective core method (EC method), which uses an effective core potential[37−44] (ECP) with a valence-only basis set (which we will henceforth call a valence basis set); this kind of treatment replaces the chemically inactive core electrons with an analytic potential. This reduces the CPU time required for a calculation because one does not need to expand the core orbitals in basis functions, and thus the overall size of the basis set is reduced. This represents a significant savings for Al, which has 10 core electrons ($1s^2 2s^2 2p^6$) and 3 valence electrons. Thus, for Al, the use of an EC method reduces the 13-electron problem to a 3-electron problem.

2. Effective Core Potentials

A full discussion of ECPs is beyond the scope of this paper. Overviews are available elsewhere,[38,41] but a brief overview of the basic theory is needed here to establish notation and motivate development. As stated in the Introduction, an EC method uses an effective potential to mimic the presence of the core electrons. First the nuclear charge is reduced by the number of missing core electrons. Then, a theoretically justified form of the potential centered at a given nucleus is a sum of angular-momentum-dependent terms

$$V^{eff} = V^{eff}_{L+1}(r) + \sum_{l=0}^{L}[V^{eff}_l(r) - V^{eff}_{L+1}(r)] \sum_{m=-l}^{l} |lm\rangle\langle lm| \quad (1)$$

where l is the angular momentum quantum number, m is the magnetic quantum number, L is the maximum l found in the core, $|lm\rangle\langle lm|$ is the projection operator for quantum numbers l and m, V^{eff}_l is the effective potential for angular momentum l, r is the distance from the nucleus, and V^{eff} is the total effective potential. The projection operators are present in part to represent Coulomb and exchange interactions with the missing core electrons and in part to impose the orthogonality effect of the absent core electrons on the valence electron wave function. Optionally, the effective potential can also include core-valence correlation effects.[43] In principle, the sum in eq 1 should go from $l = 0$ to $l = \infty$, but eq 1 approximates all V^{eff}_l terms for $l \geq L + 1$with V^{eff}_{L+1} since the higher-l orbitals have no orthogonality effect and have similar Coulomb interactions with the core. For Al, the maximum l found in the core is 1, so $L + 1$ is 2. V^{eff}_l is expanded, for ease of integration, as a sum of Gaussian functions

$$V^{eff}_l = \sum_{i=1}^{i_{max}} a_i r^{n_i} e^{-\alpha_i r^2} \quad (2)$$

where a_i is a negative real number if $l = L + 1$ and a positive real number if $l \leq L$, n_i is an integer between 0 and 2, and α_i is a positive real number. The physical interpretation of this potential is that it is repulsive for the angular momenta corresponding to the core orbitals, because of the orthogonality effect, and it is attractive for the angular momenta that are not present in the core, because of the screened Coulombic interaction with the missing nuclear charge as the valence electrons penetrate into the core region.

Some of the potentials that will be discussed are called ab initio ECPs, because the ECPs are extracted from an ab initio Hartree−Fock (or relativistic Hartree−Fock) atomic wave function. Ab initio ECPs are obtained from an all-electron wave function in two steps: (1) The valence orbitals from the all-electron calculation are transformed into pseudo-orbitals. There are different methods of obtaining the pseudo-orbitals, but all of the methods force the pseudo-orbitals to be smooth and nodeless and have the correct behavior in the valence region. (2) The numerical effective potential for each value of l is determined such that the eigenvalues of the pseudo-orbitals in the field of the effective core potential are equal to the eigenvalues of the all-electron valence orbital. For simulating neutral species, this should be done in a calculation on the neutral atom (as opposed to an ion with a single valence electron).

The ab initio ECPs that we will discuss in this paper are the LP, CEP,[40] LanL2,[42] Stuttgart relativistic ab initio potential,[44] and SHC.[39] In general, ECPs are developed in conjunction with specific basis sets. The LP potential is obtained by the method of Kahn et al.[38] and is available[45,46] with a double-ζ basis set. The CEP acronym stands for compact effective potential and is sometimes referred to as the SBKJC potential, which stands for Stevens, Basch, Krauss, Jasien, and Cundari. The original CEP potential was developed for Li − Ar.[40] Stevens, Basch, Krauss, and Jasien later extended the CEP potential to include elements K − Rn.[47] Cundari and Stevens extended the CEP potential further to include Ce − Lu.[48] The CEP potential is available for Al with optimized single-, double-, and triple-ζ basis sets, and these methods are denoted CEP-4G, CEP-31G, and CEP-121G, respectively. The LanL2 potential is named for Los Alamos National Laboratory, and it also referred to as the Hay and Wadt potential.[42] The LanL2 potential is available for Al with single and double-ζ basis sets; these are denoted as LanL2MB and LanL2DZ, respectively. The SHC ECP stands for shape- and Hamilitonian-consistent ECP; it was developed by Rappe et al. and is available with a double-ζ basis set for A1.[39] The Stuttgart relativistic potential (abbreviated SRP in this paper) is available with a double-ζ basis set and is named as such because it was fit to a relativistic wave function. The LanL2 potential uses a relativistic reference wave function for elements heavier than Kr, and the CEP potential also uses a relativistic wave function for elements heavier than K.

Because ab initio ECPs are fitted to SCF wave functions, relativistic effects can only be accounted for if the original SCF wave function is relativistic;[41,44] however, semiempirical ECPs can include relativistic effects if they are fitted to experimental data. The Stuttgart semiempirical potential of Igel-Mann et al.; denoted SECP, is a semiempirical ECP that was fit to the experimental energies of single-valence-electron ions,[43] which is different from the SRP potential, which was fit to a relativistic wave function.[44] Thus, the relativistic effects are implicitly included in the potential, whereas the

Table 1. Parameters in V_{GFP}

i	n_i	A_i	B_i
1	0	-7.31691×10^{-2}	2.25122
2	0	-1.06549×10^{-1}	0.73006
3	0	-3.91007×10^{-2}	0.01575
4	0	-4.08602×10^{-6}	0.00434
5	1	-2.73705×10^{-3}	0.20601
6	2	-1.95503×10^{-5}	0.06858

SRP potential includes the relativistic effects explicity. The functional form of the semiempirical Stuttgart potential[43] (and the SRP) is

$$V^{\text{eff}}(r) = \sum_{l=0}^{L+1} V_l^{\text{eff}}(r) \sum_{m=-l}^{l} |lm\rangle\langle lm| + V_{\text{pol}} \qquad (3)$$

where all notation is defined in eq 1 except for V_{pol}, which is a polarization potential. The form of the polarization potential is

$$V_{\text{pol}} = \frac{-\alpha}{2} \frac{(1 - \exp(-\lambda r^2))^2}{r^4} \qquad (4)$$

where α is the dipole polarizability of the core, and λ is an adjustable parameter. The Gaussian98[45] and Gaussian03[46] implementations of the Stuttgart semiempirical potential (which we will refer to as SECP for Stuttgart semiempirical effective core potential) do not include the polarization potential. We have fit the polarization potential to a sum of powers times Gaussians

$$V_{\text{GFP}} = \sum_{i=1}^{6} A_i r^{n_i} \exp(-B_i r^2) \qquad (5)$$

where A_i and B_i are adjustable parameters, n_i is an adjustable integer in the range 0−2, and GFP denotes Gaussian fitted polarization potential. The parameters are given in Table 1. Also, it should be noted that the Stuttgart potentials are not of the physical form specified in eq 1 because they have $V_l^{eff} = 0$ for $l \geq L + 1$. Later in the paper, we will use a modified version of the SECP potential called MSECP (modified Stuttgart effective core potential). In the MSECP, we used the $l = 2$ potential for all values of $l \geq L + 1$, and we rewrote eq 3 so that it is of the same form as eq 1; thus the MSECP is of the same form as the ab initio ECPs.

In this paper, we will present a new EC method based on a training set of dissociation energies and geometries calculated with the PBE0 method and the MG3 all-electron basis set. We parametrized a new combination of ECP and valence basis set for use with the PBE0 method to reproduce the all-electron PBE0/MG3 results. By doing this, we will obtain a physical EC method that accurately reproduces the hybrid density functional theory calculations with the all-electron basis set. If we had instead parametrized against a data set of experimental results or a data set of explicitly correlated results, the EC method would also be canceling the errors that are inherent in the PBE0 density functional method; with such a procedure it would not necessarily be a physical EC method. Thus, the EC method developed here can also be used with other electronic structure methods; it is not limited to PBE0. This will be explicitly demonstrated in section 6.

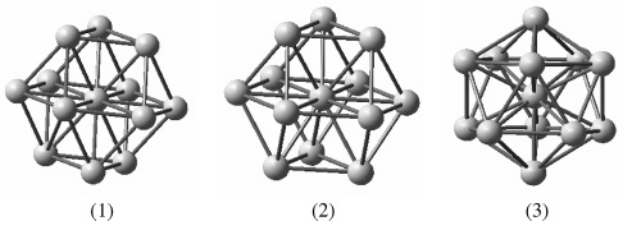


Figure 1. The face centered cubic (1), hexagonally close packed (2), and icosahedron (3).

3. Test Sets

3.1. Homonuclear Test Set. We will use six test sets in this article. The first test set is called the homonuclear test set. The homonuclear test set includes optimized Al_2 and Al_3 structures and three Al_{13} structures with fixed bond lengths. We use the ground electronic states of Al_2 and Al_3, which are $^3\Pi_u$[1,2,5,49] and $^2A_1'$.[3,4,6,17] The minimum-energy structure of Al_3 is an equilateral triangle; thus, our test set includes two bond lengths (Al_2 and the one degree of freedom in Al_3). The Al_{13} structures have face-centered cubic (FCC), hexagonally close packed (HCP), and icosahedral symmetry (see Figure 1). The bond length that we have used for the FCC, HCP, and icosahedral structures is 2.8635 Å, which is the bond length in the bulk metal. This was obtained by dividing the experimental (298 K) lattice constant, 4.04964 Å,[50] by $\sqrt{2}$.

3.2. Heteronuclear Test Set. The second test set is the called the heteronuclear test set. The molecules in the heteronuclear test set are AlH, AlC, AlO, AlCCH, Al_2H, Al_2C, and Al_2O, and they have all been optimized by the all-electron PBE0/MG3 method. The first and second row diatomics have been thoroughly studied experimentally and computationally.[51] The AlX (X = H − F) diatomics, in particular, have been thoroughly studied by Gustev et al.[52] They concluded that the ground states of AlH, AlC, and AlO are $^1\Sigma^+$, $^4\Sigma^-$, and $^2\Sigma^+$, respectively. Their conclusions agree with previous calculations[51] and our PBE0/MG3 optimizations for AlH, AlC, and AlO.

There has been little theoretical work done on the AlCCH system. Chertihin et al. have studied this system experimentally and theoretically.[53] They found that the structure of AlCCH is linear and that the isomer CCAlH is higher in energy. We have reexamined this system by optimizing different isomers with PBE0/MG3. We found that the linear AlCCH system is the minimum-energy structure, with a ground electronic state of $^1\Sigma^+$. It is interesting to note that the CCAlH systems (both singlet and triplet states) are saddle points. We also examined bent systems and π-bonded system where the Al atom is bonded to the bridge site of the C−C bond in the acetylene fragment. We also located stationary points for nonlinear systems that had C_{2v} and C_s symmetries. The minimum-energy structures are shown in Figure 2, and the results are summarized in Table 2. There is one structure with C_s symmetry that is a minimum, and there is one C_{2v} structure with one imaginary frequency (labeled C_{2v}-1*i* in

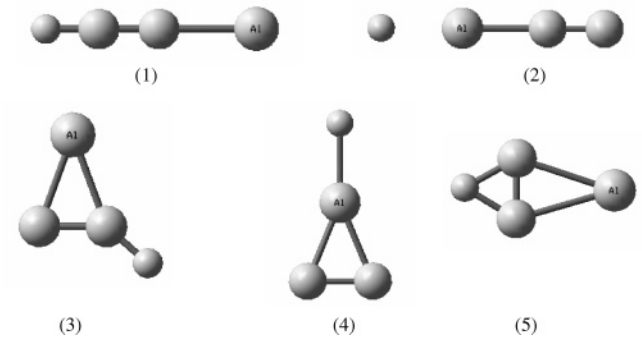


Figure 2. The AlCCH structures. The energetics are reported in Table 2 and are labeled as AlCCH(linear) (1), CCAlH(linear) (2), AlCCH(C_{2v}-2*i*) (3), AlCCH(C_{2v}-1*i*) (4), and AlCCH(C_s) (5).

Table 2. D_e Values (in eV) from All-Electron PBE0/MG3 Calculations for the AlCCH Structures Shown in Figure 1

	D_e
AlCCH(linear)[a]	
$^1\Sigma^+$	16.64
$^3\Pi$	15.26
$^3\Sigma^+$	10.92
CCAlH(linear)	
$^1\Sigma^+$	14.24
$^3\Pi$	11.92
$^3\Sigma^+$	6.04
AlCCH(C_s)	
$^1A'$	15.94
$^3A'$	14.57
AlCCH(C_{2v}-1*i*)	
$^1A_1$	15.06
$^3B_2$	12.61
AlCCH(C_{2v}-2*i*)	
$^1A_1$	14.14
$^3B_2$	12.41

[a] The $^1\Sigma^+$ state is used in the heteronuclear and small molecule test sets.

Table 2 and Figure 2) and one C_{2v} structure with two imaginary frequencies (labeled C_{2v}-2*i* in Table 2 and Figure 2). The lowest energy nonlinear structure is the C_s structure, and the C_{2v}-1*i* structure is lower in energy than C_{2v}-2*i*.

There have also been few theoretical or experimental studies of either Al_2H or Al_2C. Chertihin et al.[53] indicate that the lowest-energy structure of Al_2C is a bent structure with C_{2v} symmetry. We also found that the ground states of Al_2H or Al_2C had C_{2v} symmetry. Table 3 present the energies of optimized Al_2H and Al_2C structures. We optimized structures that have $D_{\infty h}$, $C_{\infty v}$, and C_{2v} (see Figure 3) symmetries. For Al_2H we find that the lowest energy states for the $D_{\infty h}$, $C_{\infty v}$, and C_{2v} structures are doublets. The ground state for Al_2H (C_{2v} symmetry) is the $^2B_1$ state, but the $^2A_1$ state is nearly degenerate with the $^2B_1$ state. The C_{2v} structures for Al_2H are all true minima, but the lowest energy $C_{\infty v}$ state is a saddle point and the lowest energy $D_{\infty h}$ structure is a hilltop, i.e., a stationary point with two imaginary frequencies. For Al_2C, the lowest energy structures with C_{2v} and $C_{\infty v}$ symmetries are both triplets, but the lowest energy $D_{\infty h}$ structure is a singlet. It is interesting to note that, as for

Table 3. D_e Values (in eV) for the Al_2X (X = H, C, O) Molecules that Are Shown in Figure 3[a]

	Al_2H[b]		Al_2C[c]		Al_2O[d]	
	state	D_e	state	D_e	state	D_e
$D_{\infty h}$	$^2\Pi$	3.90	$^1\Sigma_g^+$	6.12	$^1\Sigma_g^+$	10.63
			$^3\Sigma_u^+$	5.87	$^3\Pi_u$	7.42
			$^5\Sigma_g^+$	4.21	$^1\Pi_u$	7.03
$C_{\infty v}$	$^2\Pi$	4.18	$^3\Sigma$	4.57	$^1\Sigma$	7.15
	$^4\Sigma$	3.74	$^5\Pi$	4.29	$^3\Sigma$	6.08
	$^2\Sigma$	3.33	$^1\Pi$	4.10	$^1\Pi$	5.83
			$^1\Sigma$	3.91		
C_{2v}	$^2B_1$	4.51	$^3B_2$	6.95	$^3B_2$	8.02
	$^2A_1$	4.46	$^1A_1$	6.40	$^1B_2$	7.67
	$^4B_1$	2.87	$^5A_2$	5.38	$^3A^2$	7.24
	$^4A_2$	2.76	$^3A_1$	5.14		

[a] Calculated by the all-electron PBE0/MG3 method. [b] The C_{2v} ($^2B_1$) state is used in the heteronuclear and small molecule test set. [c] The C_{2v} ($^3B_2$) state is used in the heteronuclear and small molecule test set. [d] The C_{2v} ($^3B_2$) state is used in the heteronuclear and small molecule test set.

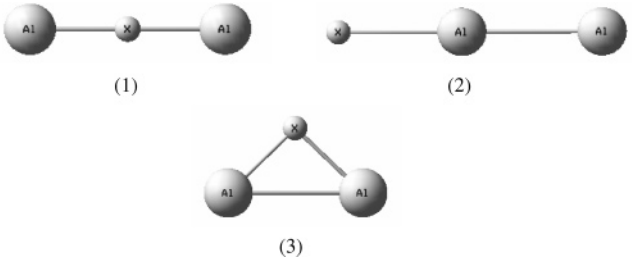


Figure 3. The $D_{\infty h}$ (1), $C_{\infty v}$ (2), and C_{2v} (3) structures of Al_2X (X = H, C, and O).

Al_2H, the lowest energy $D_{\infty h}$, $C_{\infty v}$, and C_{2v} structures of Al_2C are a hilltop, saddle point, and a true minimum, respectively. In fact, all of the C_{2v} structures are minima. However, unlike Al_2H, there are no nearly degenerate electronic states. We will use the $^2B_1$ state of Al_2H and the $^3B_2$ state of Al_2C in our training set.

The Al_2O system has been studied more thoroughly.[54−59] In all of the Al_2O studies, the lowest energy structure has been a linear (AlOAl) molecule with $D_{\infty h}$ symmetry; the $D_{\infty h}$ structure, because of symmetry, has one independent bond length (the Al−O bond). In fact, this structure is significantly more stable than the other structures. Table 3 presents the optimized Al_2O structures. The lowest-energy structure with C_{2v} symmetry is a triplet ($^3B_2$). The Al_2O systems also show no nearly degenerate electronic states, unlike the Al_2H system that has nearly degenerate electronic states ($^2B_1$ and $^2A_1$). In our training set, however, we have used the Al_2O molecule with C_{2v} symmetry instead of the $D_{\infty h}$ structure. This was done because we will ultimately be using this data set to parametrize a new ECP and valence basis set. The C_{2v} structure has an additional Al−Al bond that is not present in the $D_{\infty h}$ structure. All of the data that was used in the heteronuclear set are in the last columns of Tables 4 and 5. The other columns in Tables 4 and 5 will be explained in sections 6, and those in Table 5 will be explained in sections 4, 5, and 6.

3.3. Optimization Test Set. The test set that was used for optimizing our ECP method contains the PBE0/MG3

Table 4. D_e Values (in eV) for the Homonuclear and Heteronuclear Data Set Calculated with All-Electron PBE0/MG3 and with Valence-Electron PBE0/MSMG3/6-311+G(d,p) and PBE0/MEC/6-311+G(d*,p) Methods[f]

	MSMG3/6-311+G(d,p)[a]	MEC/6-311+G(d*,p)[b]	MG3[c]
Al_2	1.49	1.54	1.55
Al_3	3.79	3.83	3.86
AlH	2.97	2.98	2.99
AlC	3.48	3.52	3.55
AlO	4.94	5.04	5.08
AlCCH	16.51	16.57	16.64
Al_2H	4.41	4.44	4.51
Al_2C	6.87	6.89	6.95
Al_2O (C_{2v} triplet)	7.80	7.94	8.02
Al_{13}(FCC)	31.13	30.82	30.74
Al_{13}(HCP)	31.74	31.49	31.43
Al_{13}(icosahedron)	32.97	32.42	32.55
MUE[d]	0.18	0.06	
time[e]	747	473	4.3×10^6

[a] All-electron with 6-311+G(d,p) for H, C, and O and valence-electron with MSMG3 for Al. [b] All-electron with 6-311+G(d*,p) for H, C, and O and valence-electron with MEC for Al. [c] The all-electron basis set MG3 was used for all atoms. (MG3 is identical to 6-311++G(3d2f,2dp,2p) for element considered in this paper.). [d] Mean unsigned deviation from last column. [e] CPU time in seconds. [f] In the valence-electron calculations, the basis H, C, and O is 6-311G+G(d,p) for MSMG3 and 6-311+G(d*,p) for MEC.

Table 5. Equilibrium Internuclear Distance, r_e (in Å), for the Homonuclear and Heteronuclear Test Set Calculated with PBE0 and the MSMG3 and MEC Valence-Only Methods as Compared to Results Obtained with the All-Electron MG3 Basis Set

	MSMG3/6-311+G(d,p)[a]	MEC/6-311+G(d*,p)[b]	MG3[c]
Al_2	2.754	2.750	2.730
Al_3	2.508	2.513	2.507
AlH	1.662	1.673	1.664
AlC	1.964	1.970	1.964
AlO	1.620	1.621	1.617
Al−CCH	1.968	1.973	1.966
Al−AlH	2.479	2.479	2.479
AlAl−H	1.821	1.827	1.820
Al−AlC	2.798	2.786	2.792
AlAl−C	1.849	1.854	1.853
Al−AlO (C_{2v} triplet)	2.551	2.522	2.527
AlAl−O	1.736	1.737	1.730
MUE[d]	0.006	0.007	

[a] All-electron with 6-311+G(d,p) for H, C, and O and valence-electron with MSMG3 for Al. [b] All-electron with 6-311+G(d*,p) for H, C, and O and valence-electron with MEC for Al. [c] The all-electron basis set MG3 was used for all atoms. (MG3 is identical to 6-311++G(3d2f,2dp,2p).). [d] Mean unsigned deviation from all-electron results.

optimized geometries and D_e values of AlH, AlC, AlO, AlCCH, Al_2H, Al_2C, Al_2O, and Al_3, and the D_e values of the three Al_{13}s structures with bond lengths fixed at 2.8635 Å. The structural isomers and electronic states of the molecules are the same ones that are used in the heteronuclear and homonuclear test sets.

3.4. Al_{13} Test Set. The Al_{13} test set was constructed to further test the accuracy of the new ECP method. We optimized the geometry and calculated the value of D_e for FCC, HCP, and icosahedral Al_{13} with PBE0/MG3. The Al_{13} test data obtained this way is in the fourth row of Table 6; the other data in this table will be explained in section 6.

The optimum bond lengths of the FCC, HCP, and icosahedral structures are all shorter than the bulk bond length (see Table 6). The minimum energy structure is a Jahn−Teller distorted icosahedron. From one point of view, it would have been preferable to include the Al_{13} data in the optimization procedure, but the optimization procedure (described below) involves repeated evaluation of the test set, and calculating gradients of Al_{13} would have significantly increased the time required to optimize the parameters. From

Table 6. D_e (in eV) and r_e (in Å) Values for the Optimized PBE0 Calculations on the Al_{13} Structures[a]

	FCC		HCP		icosahedron		MUE	
	D_e	r_e	D_e	r_e	D_e	r_e	D_e	r_e
CEP-121G(d)	31.14	2.727	31.73	2.727	32.36	2.814	0.44	0.060
LanL2DZ(d)	30.44	2.738	31.05	2.739	31.55	2.817	1.17	0.068
LP-31G(d)	30.55	2.696	31.18	2.698	31.82	2.774	1.00	0.049
MEC	31.63	2.719	32.23	2.723	32.59	2.666	0.05	0.006
MG3	31.60	2.713	32.25	2.715	32.70	2.661		
MSMG3	31.84	2.727	32.43	2.727	33.06	2.814	0.26	0.060
SECP(d)	30.66	2.728	31.25	2.730	31.71	2.811	0.98	0.060
SHC(d)	30.28	2.712	30.90	2.756	31.14	2.838	1.41	0.073

[a] The MG3 row contains all-electron calculations, and the other rows contain valence-only calculations and their mean unsigned deviation from the MG3 row.

another point of view though, using these data as a test without including them in the training set provides a useful check on our methodology. The Al_{13} test set is especially important because Al_{13} is the smallest Al cluster in which any atom has the bulk coordination number of 12. In particular, the central atom, which participates in 12 of the 36 bonds of FCC Al_{13}, has this coordination number. It is important to test whether the effective core potential and basis set remain valid not only for binding energies but also for bond lengths for systems with this high coordination number.

3.5. Small Molecule Test Set. We have found PBE0 to be the most accurate HDFT method for pure Al compounds;[18] however, PBE0 is not necessarily the best method for other systems, such as hydrocarbons.[60,61] Since PBE0 might not be the best method for all interesting systems, we will use the small-molecule test set to test the accuracy of the MEC method with other DFT and HDFT methods and compare how well it does with competing EC methods. The small molecule test set includes the atomization energies and bond lengths of Al_2, Al_3, AlH, AlC, AlO, AlCCH, Al_2H, Al_2C, and Al_2O. The DFT methods that will be used are the Becke88-Lee-Yang-Wang model (BLYP),[62,63] modified-Perdew-Wang-Perdew-Wang-1991 model (mPWPW91),[27] PBE;[64] the HDFT methods that will be used are the Becke88 3-parameter-Lee-Yang-Wang model (B3LYP)[65,66] and modified-Perdew-Wang 1-parameter-Perdew-Wang-1991 model (mPW1PW91);[27] and the meta DFT methods that will be used are the 1-parameter Becke88-Becke95 (B1B95)[67] and a recent 1-parameter model for kinetics (BB1K).[68]

3.6. AlX Test Set. We also wanted to test how well the MEC method does with elements that are not present in the training set. The AlX test set contains the dissociation energies for AlX compounds where X = Li, Be, B, N, C, O, and F. We used the electronic states determined by Gustev et al.[52] The dissociation energies were calculated with the PBE0 HDFT method and MG3 basis set.

4. Tests of Literature ECPs and C, H, and O Basis Sets

We first tested how well the currently available EC methods do for the homonuclear test set; the results are given in Table 7. Each EC method consists of an ECP and a valence basis set. Some of the valence basis sets that were developed for the EC methods have *d* functions. The SHC and CEP methods use the same exponent for the *d* functions (0.25).[1] The LanL2DZ basis set uses a different exponent for the *d* functions (0.19).[69] To our knowledge, there are no *d* functions for the SRP or SECP, so we have also examined all of the EC methods with the standard exponents for *d* functions (0.325) for Al.[36] The standard *d* functions were used as both Cartesian *d* functions (6D) and spherical *d* functions (5D). This was done because some of the C, H, and O basis sets use the 6D formalism and some used the 5D formalism. The important result is that the choice between using 5D or 6D functions does not have a significant effect on the results. The diffuse functions that we have used are the standard diffuse functions (the *s* and *p* exponents are both 0.0318)[47] and were used with the *d* function that had the smallest error (except for SECP+V_{GFP}, which uses no *d* functions).

Table 7. Mean Unsigned Error in D_e Values (in eV), r_e Values (in Å), and CPU Times (in s) for the Homonuclear Test Set with the Literature ECPs

	no *d*[a]	literature	5D[b]	6D[c]	+ diffuse[d]
			CEP-31G		
MUE(D_e)	2.78	0.63	0.79	0.80	0.63
MUE(r_e)	0.106	0.045	0.039	0.040	0.041
time	77	271	1635	261	680
			CEP-121G		
MUE(D_e)	2.30	0.57	0.65	0.64	0.56
MUE(r_e)	0.104	0.038	0.034	0.034	0.038
time	146	464	373	489	1465
			LanL2DZ		
MUE(D_e)	2.95	0.68	0.65	0.66	0.63
MUE(r_e)	0.102	0.028	0.034	0.033	0.043
time	72	160	393	417	1610
			LP-31G		
MUE(D_e)	2.82	1.04	0.78	0.77	0.71
MUE(r_e)	0.071	0.006	0.006	0.005	0.004
time	70	209	209	196	540
			SECP		
MUE(D_e)	2.33		0.32	0.30	0.22
MUE(r_e)	0.082		0.012	0.011	0.012
time	84		328	286	801
			SECP+V_{GFP}		
MUE(D_e)	1.36		3.20	3.11	1.59
MUE(r_e)	0.007		0.077	0.076	0.013
time	87		328	335	260
			SHC		
MUE(D_e)	2.82	0.52	0.74	0.77	0.43
MUE(r_e)	0.169	0.057	0.053	0.046	0.050
time	71	291	258	255	471
			SRP		
MUE(D_e)	2.67		0.58	0.58	0.44
MUE(r_e)	0.101		0.028	0.028	0.023
time	72		719	622	844

[a] No diffuse functions. [b] The *d* functions are 5D. [c] The *d* functions are 6D. [d] Best of the previous four columns, augmented by diffuse functions.

The one EC method that shows unexpected behavior in Table 7 is the SECP+V_{GFP} EC method. This EC method has the smallest MUE for both atomization energies and bond lengths when only an *s* and *p* valence-electron basis set was used, but the errors more than double when *d* functions are used. For the other ECPs, the inclusion of *d* functions leads to significant improvements in both atomization energies and bond lengths.

The computational times reported in Table 7 are the CPU times required for a single-point calculation of icosahedral Al_{13} on a single-processor of an IBM Power4 supercomputer with a 1.3 GHz Power4 processor. These times should not be interpreted too closely because they depend on the number of SCF iterations required, which can be increased or decreased by using different initial guesses or different SCF routines. Nevertheless they do provide a qualitative indication of the cost that is needed to judge whether increasing the basis set is worth the increased cost and whether including *d* functions (except for SECP+V_{GFP}) is worth the increased

Table 8. Mean Unsigned Error for D_e (in eV) and r_e (in Å) for the Heteronuclear Test Set with Literature ECP Methods with d Functions and Various Double-ζ Basis Sets for C, H, and O

	6-31G[a]	6-31G(d)[a]	6-31G(d,p)[a]	6-31+G(d)[a]	6-31+G(d,p)[a]	6-31++G(d,p)[a]
			CEP-121G(d)			
MUE(D_e)	0.34	0.37	0.32	0.26	0.25	0.26
MUE(r_e)	0.039	0.039	0.039	0.032	0.031	0.031
			LanL2DZ(d)			
MUE(D_e)	0.39	0.27	0.19	0.30	0.29	0.31
MUE(r_e)	0.050	0.037	0.038	0.040	0.039	0.039
			LP-31G(d)			
MUE(D_e)	0.21	0.26	0.24	0.28	0.19	0.26
MUE(r_e)	0.025	0.034	0.033	0.032	0.032	0.031
			SECP(d)			
MUE(D_e)	0.28	0.17	0.16	0.18	0.17	0.19
MUE(r_e)	0.017	0.007	0.007	0.010	0.009	0.009
			SHC(d)			
MUE(D_e)	0.36	0.26	0.24	0.27	0.25	0.27
MUE(r_e)	0.053	0.042	0.041	0.045	0.045	0.045
			SRP(d)			
MUE(D_e)	0.30	0.19	0.23	0.20	0.18	0.20
MUE(r_e)	0.024	0.014	0.013	0.016	0.016	0.016

[a] Literature exponents.

computational effort. In general, the addition of diffuse functions did not lead to significant improvements in either atomization energies or bond lengths. The inclusion of diffuse functions, however, did lead to significant increases in the computational time.

The main conclusions drawn from Table 7 is that the SECP method does best for energies, and the LP-31G method does best for bond distances. Note that these are the fifth and second least expensive of the eight methods studied.

We turn next to the heteronuclear test set. For this test set we tested the literature EC methods for Al with all-electron basis sets for H, C, and O. The basis sets that we tested were as follows: (a) double-ζ, polarized valence double-ζ, and polarized augmented double-ζ: 6-31G,[70] 6-31G(d),[70,71] 6-31+G(d),[70−72] 6-31G(d,p),[70,71] 6-31+G(d,p),[70−72] 6-31++G(d,p);[70−72] and (b) triple-ζ, polarized valence triple-ζ, and polarized augmented triple-ζ: 6-311G,[73] 6-311G(d),[72,73] 6-311+G(d),[72,73] 6-311G(d,p),[73] 6-311+G(d,p),[72,73] and 6-311++G(d,p).[72,73] The double-ζ-type basis sets and the triple-ζ-type basis sets use spherical harmonic sets of five *d* basis functions. We did not reexamine all of the different permutations of Al valence basis sets, but instead, for each ECP, we used the one with the smallest errors in Table 7. For CEP, we used the triple-ζ basis set with the literature *d* functions. We used the standard *d* functions for the LanL2DZ, LP-31G, SECP, SHC, SRP, and SECP potentials. We did not use the SECP+V_{GFP} potential because of its poor performance. With these explanations, then, the methods used for Al are referred to as CEP-121G(d), LanL2DZ(d), LP-31G(d), SECP(d), SHC(d), and SHC(d). The results are in Tables 8 and 9.

None of the H, C, and O basis sets is quite as large as MG3. For H, C, and O the MG3 basis set is the same as 6-311G++G(2df,2p),[35] and for Al it is 6-311+G(3d2f).[35] Atoms heavier than Si, the *s* and *p* functions of the MG3 basis set differ from 6-311G.[35]

Tables 8 and 9 show that increasing the size of the basis set on H, C, and O does not systematically improve the results, but in general, the best H, C, and O basis sets have both polarization functions and diffuse functions. The polarization functions lowered the errors for both atomization energies and bond lengths, whereas the diffuse functions mainly lowered the errors for atomization energies. The LP-31G(d) method has the lowest errors when no polarization or diffuse functions are used on H, C, and O. The best method for the heteronuclear test set using polarized or polarized augmented basis sets for H, C, and O is 6-31G-(d,p) and SECP(d) for Al. This is encouraging because the SECP method was also the most accurate for D_e in the homonuclear test set.

5. New ECP Parameters and Basis Set Exponents

5.1. MSMG3 Method. On the basis of the tests in section 4, we chose the SECP method for further improvement, and we began improving it by defining it for all values of *l* (MSECP) as explained in section 2. Then we developed a valence electron basis set by starting with MG3 basis functions. We were able to obtain very good results for the homonuclear and heteronuclear test sets with a (3*s*4*p*1*d*) valence electron basis set for Al and various basis sets for H, C, and O. We call this method MSMG3 (modified Stuttgart semiempirical ECP with Al basis functions from MG3). The MUE in D_e and r_e (averaged now over both the homonuclear and heteronuclear test sets) decreased from 0.25 to 0.18 eV and from 0.011 to 0.006 Å, respectively.

The main drawback of the MSMG3 method is that it is computationally demanding because there are four *p* functions on Al. Therefore we tested whether the results can be improved if the Al basis functions are further optimized. We began by testing the sensitivity to the basis set exponents and found that the energies and geometries could be most

Table 9. Mean Unsigned Error for D_e (in eV) and r_e (in Å) for the Heteronuclear Test Set Calculated with the PBE0 Method and Literature ECP Methods with d Functions and Various Triple-ζ Basis Sets for C, H, and O

	6-311G[a]	6-311G(d)[a]	6-311G(d,p)[a]	6-311+G(d)[a]	6-311+G(d,p)[a]	6-311++G(d,p)[a]
			CEP-121G(d)			
MUE(D_e)	0.38	0.30	0.28	0.29	0.32	0.27
MUE(r_e)	0.041	0.034	0.032	0.035	0.033	0.033
			LanL2DZ(d)			
MUE(D_e)	0.42	0.34	0.32	0.33	0.31	0.31
MUE(r_e)	0.051	0.041	0.040	0.042	0.040	0.041
			LP-31G(d)			
MUE(D_e)	0.21	0.24	0.22	0.26	0.32	0.24
MUE(r_e)	0.024	0.031	0.030	0.030	0.030	0.030
			SECP(d)			
MUE(D_e)	0.31	0.22	0.20	0.21	0.18	0.19
MUE(r_e)	0.018	0.011	0.010	0.011	0.010	0.010
			SHC(d)			
MUE(D_e)	0.39	0.30	0.28	0.29	0.27	0.27
MUE(r_e)	0.054	0.046	0.045	0.047	0.046	0.046
			SRP(d)			
MUE(D_e)	0.31	0.23	0.21	0.21	0.20	0.20
MUE(r_e)	0.024	0.017	0.017	0.017	0.017	0.017

[a] Literature exponents.

readily improved by varying the tightest and most diffuse *s* and *p* basis functions and/or the *d* function. The errors for the compounds containing oxygen were much larger than the errors for the C and H compounds, and for this critical case we found that varying the *d* function on O leads to systematic improvements, whereas varying the diffuse functions on O did not lead improvements for compounds with O. This will be discussed in greater detail in section 6.

To take advantage of the possible improvements noted in the previous paragraph, we optimized a new basis set for Al and a new ECP. We did not want to develop a new ECP from scratch, so instead we took a linear combination of two of the already defined ECPs

$$V_{\mathrm{H-X-MSECP}} = cV_{\mathrm{X}} + (1 - c)V_{\mathrm{MSECP}} \qquad (6)$$

where X stands for either CEP, LanL2DZ, LP, and SHC and H-X-MSECP stands for hybrid X-MSECP. After some initial trials we found that a combination of the CEP and MSECP potentials led to the most improvement in D_e and r_e for both the homonuclear and heteronuclear test sets. We also tried using different linear combinations that used different values of *c* for the various *l*-components but found that the improvement obtainable in that way was marginal.

5.2. Optimization Procedure. We optimized the Al valence-electron basis set, the *d* function for O, and *c* in eq 6 with a microgenetic algorithm (μGA), which has been described elsewhere.[74−76] Briefly stated, the μGA optimizes a set of parameters by maximizing a fitness function. For this paper, the fitness function is

$$f = -\left(\frac{1}{a}\mathrm{MUE}(D_e) + \frac{1}{b}G\right) \qquad (7)$$

where MUE(D_e) is the mean unsigned error in the D_e, *G* is the root-mean-square of the gradients calculated with the EC method at the geometries of the PBE0/MG3 optimized molecules, $a = 1$ eV, and $b = 2$ eV/Å, and *f* is the unitless fitness function. The evaluation of *f* involves computing the dissociation energies for all of the molecules in the optimization test set and the gradients of the PBE0/MG3 optimized geometries in the optimization test set. This strategy allows us to parametrize against geometries without having to do an expensive geometry optimization at each step. The final parameters define what we call the MEC method, and they are given in Table 10. The parametrization was done with the 6-311++G(d,p) basis set for C, H, and O, except that the *d* function was also optimized on O. The optimized *d* function on O is 0.63924 and may be compared to the standard one (in 6-311++G(d,p)) of 1.292. Any basis set that uses the new value of the oxygen *d* exponent instead of the standard value will be denoted *d**; the *d* exponent for C is not changed. The heteroatom basis set will often be explicitly stated after the EC method. For example, PBE0/MEC/6-311+G(d*,p) means that the PBE0 HDFT method was used with the MEC method for Al atoms, and the 6-311+G(d*,p) basis set was used for non-Al atoms.

Table 10. Parameters for the MEC Method

	s exponent	*p* exponent	*d* exponent	c
Al	0.75752	1.94530	0.45580	0.461
	0.20680	0.20064		
	0.08507	0.07073		

6. Results and Discussion

All of the calculations in this paper were done with either GAUSSIAN98[45] or GAUSSIAN03,[46] except for the Al_{13} geometry optimizations which were done with NWCHEM version 4.5.[77] As mentioned in section 5, the method that we have developed (as specified in Table 10) is called the MEC (Minnesota effective core) method. The results obtained with the MEC method are given in Tables 4−6 and 11.

The MUE in D_e and r_e for both the homonuclear and heteronuclear test set with the PBE0/MEC method are given

Table 11. Mean Unsigned Error for D_e (in eV) and r_e (in Å) for the Heteronuclear Test Set Calculated with PBE0 and the MEC Valence-Electron Method for Al and Various All-Electron Basis Sets for C, H, and O

	MUE(D_e)	MUE(r_e)
6-31G[a]	0.20	0.009
6-31G(d*)[a]	0.09	0.006
6-31G(d*,p)[a]	0.08	0.006
6-31+G(d*)[a]	0.08	0.006
6-31+G(d*,p)[a]	0.06	0.006
6-31++G(d*,p)[a]	0.07	0.006
6-311G[a]	0.22	0.009
6-311G(d*)[a]	0.10	0.007
6-311G(d*,p)[a]	0.08	0.007
6-311+G(d*)[a]	0.08	0.006
6-311+G(d*,p)[a]	0.05	0.005
6-311++G(d*,p)[a]	0.05	0.005

[a] The *d*-function used for O is 0.63924 (see Table 14), and the remaining functions are unchanged from their standard values.

in Tables 4 and 5, respectfully. It can be seen that the MUE for D_e for both the heteronuclear test sets and homonuclear test sets is 0.06, which is 0.12 eV lower than PBE0/MSMG3, and it is 0.19 eV lower than the most accurate literature potential. The MUE in r_e (for the heteronuclear and homonuclear test sets) is 0.001 Å larger with PBE0/MEC than with PBE0/MSMG3, but the MUE in r_e for PBE0/MEC is 0.005 Å lower than the most accurate literature potential. In another test of the PBE0/MEC method, given in Table 6, the MUE in D_e for the Al_{13} test set is 0.04 eV, and the error in bond lengths for the Al_{13} test set is 0.006 Å.

As described above, Table 5 involves tests of the MEC model with the recommended 6-311+G(d*,p) basis for H, C, and O. We also tested the MEC model with other basis sets for H, C, and O, and the results are given in Table 11. We found that the 6-311+G(d*,p) and 6-311++G(d*,p) basis sets give the best results for D_e and r_e, and the mean unsigned errors for both of these basis sets are 0.05 eV and 0.005 Å, respectively. Removing the diffuse functions on C, H, and O altogether does not lead to a significant increase in errors for the heteronuclear test set. For example, the MUE in D_e and r_e with the 6-311G(d,p) basis set is 0.08 eV and 0.007 Å, respectively.

An issue that requires special discussion is that the *d*-function for O was reoptimized for use in conjunction with the MEC basis set. The standard value of the oxygen *d* exponent in the 6-311G(d) or 6-311+G(d) basis set is 1.292, and the standard value in the 6-31G(d) or 6-31+G(d) basis set is 0.800, whereas the recommended value of the *d* exponent for use with the MEC is 0.63924; this is denoted d*. This does not mean that the MEC is inappropriate for general use, but instead it indicates a deficiency in the 6-311+G(d) basis set for O. The MUEs in dissociation energies and bond lengths for AlO and Al_2O calculated with PBE0/MEC/MG3 (the MG3 basis set for O is 6-311+G-(2df)) are 0.06 eV and 0.003 Å, respectively, but these MUEs are 0.18 eV and 0.006 Å, respectively, if the dissociation energies and bond lengths are calculated with PBE0/MEC/6-311+G(d). In contrast, the errors for the dissociation energies and bond lengths are 0.06 eV and 0.003 Å, respectively, using PBE0/MEC/6-311+G(d*). Thus, the reoptimizatin of the *d*-function does not cancel errors in the Al potential, but rather it corrects deficiencies in the O basis set when it is reduced from a 2*df* polarization level (for which the *d* exponents are 2.584 and 0.646) to a single *d* function. Interestingly, the standard *d* exponents for the 6-311+G-

Table 12. Calculated Dissociation Energies for the AlX Training Set and MUEs for Various Basis Sets

non-Al basis	AlLi	AlBe	AlB	AlN	AlC	AlO	AlF	MUE
6-31G(d)[a]	0.88	0.80	2.17	2.45	3.48	4.93	6.72	0.07
6-31+G(d)[a]	0.89	0.81	2.14	2.54	3.50	5.02	6.79	0.03
6-311G(d)[a]	0.90	0.81	2.14	2.46	3.49	4.87	6.59	0.10
6-311+G(d)[a]	0.91	0.83	2.16	2.52	3.52	4.94	6.64	0.07
6-31G(2d)[b]	0.89	0.83	2.19	2.50	3.53	4.96	6.70	0.07
6-31+G(2d)[b]	0.89	0.84	2.18	2.59	3.56	5.08	6.76	0.03
6-311G(2d)[b]	0.90	0.83	2.17	2.52	3.53	4.97	6.69	0.06
6-311+G(2d)[b]	0.91	0.85	2.19	2.59	3.56	5.07	6.74	0.04
6-31G(1d)[c]	0.89	0.82	2.19	2.49	3.54	4.90	6.71	0.07
6-31+G(1d)[c]	0.89	0.83	2.18	2.58	3.57	5.03	6.77	0.03
6-311G(1d)[c]	0.90	0.82	2.17	2.51	3.54	4.91	6.70	0.07
6-311+G(1d)[c]	0.91	0.84	2.19	2.58	3.57	5.02	6.75	0.04
6-311+G(d*)[d]	0.91	0.83	2.16	2.52	3.52	5.04	6.64	0.06

[a] The exponents are all unchanged from those defined in the literature (see Table 14). [b] The exponents used in the *d*-functions for this basis set are the same ones used in the MG3 basis set (see Table 14). [c] The exponent used in the *d*-functions for this basis set are the same as the most diffuse *d*-exponent used in the MG3 basis set (see Table 14). [d] The exponent used in the *d*-functions used for O is 0.63924, and the remaining exponents are unchanged from their standard values.

Table 13. Calculated Dissociation Energies for the AlX Training Set and MUEs for Various Basis Sets and EC Methods

non-Al basis	AlLi	AlBe	AlB	AlN	AlC	AlO	AlF	MUE
CEP-121G(*d*)								
6-311G(d)[a]	0.87	0.72	1.92	2.05	3.20	4.27	6.08	0.39
6-311+G(d)[a]	0.88	0.72	1.92	2.10	3.20	4.34	6.15	0.36
6-311G(2d)[b]	0.88	0.73	1.95	2.23	3.30	4.56	6.32	0.27
6-311+G(2d)[b]	0.88	0.73	1.96	2.28	3.31	4.61	6.36	0.25
6-311G(1d)[c]	0.88	0.72	1.95	2.23	3.31	4.57	6.33	0.27
6-311+G(1d)[c]	0.88	0.72	1.96	2.27	3.32	4.61	6.37	0.24
LanL2DZ(*d*)								
6-311G(d)[a]	0.88	0.73	1.94	2.08	3.23	4.32	6.10	0.37
6-311+G(d)[a]	0.88	0.73	1.94	2.14	3.24	4.40	6.18	0.34
6-311G(2d)[b]	0.88	0.75	1.99	2.24	3.33	4.57	6.33	0.25
6-311+G(2d)[b]	0.89	0.74	1.99	2.30	3.34	4.64	6.39	0.23
6-311G(1d)[c]	0.88	0.73	1.99	2.23	3.34	4.58	6.34	0.25
6-311+G(1d)[c]	0.88	0.74	1.99	2.29	3.35	4.65	6.41	0.22
SECP(*d*)								
6-311G(d)[a]	0.88	0.70	1.92	2.05	3.20	4.27	6.08	0.39
6-311+G(d)[a]	0.88	0.72	1.92	2.10	3.21	4.34	6.15	0.36
6-311G(2d)[b]	0.88	0.73	1.97	2.23	3.31	4.54	6.32	0.27
6-311+G(2d)[b]	0.88	0.73	1.96	2.28	3.32	4.60	6.37	0.24
6-311G(1d)[c]	0.88	0.72	1.97	2.23	3.32	4.55	6.33	0.26
6-311+G(1d)[c]	0.88	0.73	1.96	2.27	3.33	4.61	6.38	0.24

[a] The exponents are all unchanged from those defined in the literature (see Table 14). [b] The exponents used in the *d*-functions for this basis set are the same ones used in the MG3 basis set (see Table 14). [c] The exponent used in the *d*-functions for this basis set are the same as the most diffuse *d*-exponent used in the MG3 basis set (see Table 14).

Table 14. Exponents for the *d*-Polarization Functions Used in the Non-Al Basis Sets

		d_1	d_2
Li	6-31G(d)[a]	0.20000	
	6-311G(d)[a]	0.20000	
	MG3[a]	0.10000	0.400
Be	6-31G(d)[a]	0.40000	
	6-311G(d)[a]	0.25500	
	MG3[a]	0.12800	0.510
B	6-31G(d)[a]	0.60000	
	6-311G(d)[a]	0.40100	
	MG3[a]	0.20100	0.802
N	6-31G(d)[a]	0.80000	
	6-311G(d)[a]	0.91300	
	MG3[a]	0.45700	1.826
C	6-31G(d)[a]	0.80000	
	6-311G(d)[a]	0.62600	
	MG3[a]	0.31300	1.252
O	6-31G(d)[a]	0.80000	
	6-311G(d)[a]	1.29200	
	6-311G(d*)[b]	0.63924	
	6-311+G(d*)[b]	0.63924	
	MG3[a]	0.64600	2.584
F	6-31G(d)[a]	0.80000	
	6-311G(d)[a]	1.75000	
	MG3[a]	0.87500	3.500

[a] Literature exponents. [b] Optimized in this work.

(2*df*) basis set for O, namely 2.584 and 0.646, are similar to the standard exponents for the cc-pVTZ[78] basis set for O, namely 2.314 and 0.645. Since the d* function has an exponent of 0.63924, we conclude that when one combines a polarized triple-ζ valence all-electron basis set on O with a core potential on Al, it is more important to polarize the diffuse part of the valence triple-ζ space than the tighter part.

We have used the AlX test set to test the MEC method with elements that were not included in the parametrization (specifically Li, Be, B, N, and F). Table 12 gives the dissociation energies and MUEs calculated with the PBE0 HDFT method and the MEC method for Al and 13 different basis sets for B, Be, C, F, Li, N, and O atoms. The errors are calculated by comparing to all-electron PBE0/MG3 dissociation energies. The 13 different basis sets used are 6-31G(d), 6-31+G(d), 6-311G(d), 6-311+G(d), 6-31G(2d), 6-31+G(2d), 6-311G(2d), 6-311+G(2d), 6-31G(1d), 6-31+G(1d), 6-311G(1d), 6-311+G(1d), and 6-311+G(d*,p), where (d*) is the same as (d) except for O where the exponent is changed to 0.63924. The *s* and *p* functions used in the 6-31G(2d) and 6-31G(1d), 6-31+G(2d) and 6-31+G(1d), 6-311G(2d) and 6-311G(1d), and 6-311+G(2d) and 6-311+G(1d), are the same as those used in the 6-31G(d), 6-31+G(d), 6-311G(d), and 6-311+G(d) basis sets, respectively. The *d*-functions used in the 6-31G(2d), 6-31+G(2d), 6-311G(2d), and 6-311+G(2d) are the same as the ones used in the MG3 basis set. The *d*-functions used in the 6-31G(1d), 6-31+G(1d), 6-311G(1d), and 6-311+G(1d) basis sets are the same as the most diffuse *d*-function used in the MG3 basis set basis set. The *d*-functions are presented in Table 14. Comparison of rows 11 and 12 to rows 3 and 4 in Table 12 shows that errors are significantly lower if one polarizes the diffuse part of the valence triple-ζ space than the tighter part. The MUE if the 6-311+G(1d) basis set is used is 0.04 eV. Thus, we conclude that the MEC method can be used with all of the first-row elements if the 6-311+G(1d) basis set is used.

We did some further tests to see if the need for diffuse *d* functions is specific to the MEC method or is a more general feature of Al EC methods. We have calculated atomization energies for the molecules in the AlX training set using the PBE0 HDFT method and 6 different heteronuclear basis sets (6-311G(d), 6-311+G(d), 6-311G(2d), 6-311+G(2d), 6-311G(1d), and 6-311+G(1d)) and 3 different EC methods for Al (CEP-121G(d), LanL2DZ(d), SECP(d)) and calculated the MUEs relative to PBE0/MG3 results. The results are in Table 13. It can be seen that the good performance of (1d) basis sets relative to (d) sets is not specific to the MEC method but is a much more general trend. For atoms bonding to Al, the errors are much lower if the diffuse part of valence space is polarized rather than using the usual (d) exponent.

Another test was to see how well the 6-311++G(d*,p) basis does for compounds that contain no Al atoms. We optimized the geometries and calculated the dissociation energies of BeO, CH_3OCH_3, CH_3OH, H_2CO, H_2O, and LiO within all-electron calculations with the PBE0/6-311+G(d,p), PBE0/6-311+G(d*,p), and PBE0/MG3. The MUEs, relative to PBE0/MG3, for the Be−O, C−O, Li−O, and O−H bond lengths with PBE0/6-311+G(d*,p) and PBE0/6-311+G(d,p) are 0.006 Å and 0.003 Å, respectively, and the MUEs for the atomization energies for PBE0/6-311+G(d*,p) and PBE0/6-311+G(d,p) are both 0.11 eV. The bond lengths for Be, C, H, Li, and O compounds are marginally less accurate if PBE0/6-311+G(d*,p) is used in place of the PBE0/6-311+G(d,p) basis set, whereas the errors for atomization energies are equal. However, Table 15 shows that the bond lengths and atomization energies are significantly more accurate if PBE0/6-311+G(d*,p) is used instead of PBE0/6-311+G(d,p) for Al containing compounds.

We also used the small molecule test set to test the accuracy of the MEC method and literature EC methods with functionals other than PBE0. The tests include not only hybrid functionals such as PBE0 (in particular, B3LYP[65,66] and mPW1PW91[27]) but also pure GGA functionals (BLYP,[62,63] mPWPW91,[27] and PBE[64]) and hybrid meta functionals (B1B95[67] and BB1K[68]). The MUEs for atomization energies and bond lengths are given in Table 15. The errors are the unsigned differences between the values computed with the all-electron MG3 basis set and those computed with the EC method. For the MEC method, the average MUEs for atomization energies and bond lengths for all of the functionals in Table 15 are 0.04 eV and 0.009 Å, respectively. The most accurate literature EC method for all of the functionals is SECP(d), which has an average MUE of 0.16 eV for atomization energies and 0.013 Å for bond lengths. The results show that the MEC method is useful not only for the PBE0 density functional but also that it can be used advantageously with more general electronic structure methods.

Table 15. Mean Unsigned Errors for a Variety of Density Functionals, Relative to the Results Obtained with the Same Functional with the MG3 Basis Set, in D_e (in eV) and r_e (in Å) for the Small-Molecule Test Set

	B1B95[a]	B3LYP	BB1K	BLYP	mPW1PW91[b]	mPWPW91[c]	PBE	PBE0[d]	av
				D_e					
CEP-121G(d)/6-311+G(d,p)[e]	0.22	0.18	0.24	0.15	0.23	0.20	0.20	0.21	0.20
LanL2DZ(d)/6-311+G(d,p)[e]	0.18	0.13	0.20	0.09	0.17	0.14	0.15	0.28	0.17
MEC/6-311+G(d,p)[e]	0.06	0.04	0.09	0.05	0.06	0.02	0.04	0.07	0.06
MEC/6-311+G(d*,p)[f]	0.03	0.05	0.06	0.07	0.04	0.01	0.02	0.04	0.04
SECP(d)/6-311+G(d,p)[e]	0.16	0.11	0.18	0.09	0.15	0.13	0.14	0.19	0.16
				r_e					
bond lengths									
CEP-121G(d)/6-311+G(d,p)[e]	0.042	0.026	0.043	0.022	0.034	0.031	0.033	0.034	0.033
LanL2DZ(d)/6-311+G(d,p)[e]	0.029	0.014	0.031	0.012	0.022	0.020	0.021	0.038	0.023
MEC/6-311+G(d,p)[e]	0.013	0.008	0.013	0.010	0.007	0.007	0.007	0.007	0.009
MEC/6-311+G(d*,p)[f]	0.011	0.010	0.012	0.011	0.006	0.007	0.007	0.007	0.009
SECP(d)/6-311+G(d,p)[e]	0.017	0.009	0.018	0.009	0.010	0.009	0.010	0.019	0.013

[a] Note that we used the correct B1B95 functional with 28% Hartree−Fock exchange, not the incorrect functional present in ref 46. [b] Also-called mPW1PW91 and mPW0. [c] Also-called mPWPW. [d] Also-called PBE1PBE. [e] The all-electron 6-311+G(d,p) basis set was used for H, C, and O. [f] The all-electron 6-311+G(d*,p) basis set was used for H, C, and O. The 6-311+G(d*,p) basis set is identical to the 6-311+G(d,p) basis set except for the exponent used in the O polarization functions, which was optimized in this work (see text).

Table 16. Mean Unsigned Errors in PBE0 Values of D_e (in eV) and r_e (in Å)

	quantity					
	D_e			r_e		
test set	homo	hetero	Al_{13}	homo	hetero	Al_{13}
average literature	0.62[a]	0.26[b]	1.00[c]	0.024[a]	0.030[b]	0.053[c]
best literature	0.22[d]	0.16[d]	0.44[e]	0.012[d]	0.007[d]	0.060[e]
MEC/6-31G(d,p)[f]	0.07	0.10	0.05	0.013	0.001	0.006
MEC/6-31+G(d,p)[f]	0.07	0.07	0.05	0.013	0.008	0.006
MEC/6-311+G(d,p)[f]	0.07	0.09	0.05	0.013	0.006	0.006
MEC/6-31G($2d_O$,d_C,p)[g]	0.07	0.08	0.05	0.013	0.006	0.006
MEC/6-31+G($2d_O$,d_C,p)[h]	0.07	0.06	0.05	0.013	0.006	0.006
MEC/6-311+G($2d_O$,d_C,p)[i]	0.07	0.05	0.05	0.013	0.005	0.006
MEC/6-31G(d*,p)[g]	0.07	0.08	0.05	0.013	0.006	0.006
MEC/6-31+G(d*,p)[g]	0.07	0.06	0.05	0.013	0.006	0.006
MEC/6-311+G(d*,p)[g]	0.07	0.05	0.05	0.013	0.005	0.006

[a] Average of eight methods in Table 7. [b] Average of 72 methods in Tables 8 and 9. [c] Average of 5 methods in Table 6. [d] SECP(d) for Al with 6-31G(d,p) for H, C, O. [e] CEP-121G(d) for Al. [f] The basis after the / is used for H, C, and O; the basis used for Al is the polarized triple-ζ basis set of the MEC method. The H, C, and O basis set is unchanged. [g] The basis after the / is used for H, C, and O; the basis used for Al is the polarized triple-ζ basis set of the MEC method. Two sets of *d*-polarization functions are used for O, one set of *d*-polarization functions are used for C, and one set of *p*-polarization functions are used for H. The exponents for the O *d*-polarization functions are taken from the MG3 basis set, and the polarization functions for the C and H polarization functions are taken from the 6-31G(d,p) basis sets. [h] The basis after the / is used for H, C, and O; the basis used for Al is the polarized triple-ζ basis set of the MEC method. Two sets of *d*-polarization functions are used for O, one set of *d*-polarization functions are used for C, and one set of *p*-polarization functions are used for H. The exponents for the O *d*-polarization functions are taken from the MG3 basis set and the polarization functions for the C and H polarization functions are taken from the 6-31+G(d,p) basis sets. [i] The basis after the / is used for H, C, and O; the basis used for Al is the polarized triple-ζ basis set of the MEC method. Two sets of *d*-polarization functions are used for O, one set of *d*-polarization functions are used for C, and one set of *p*-polarization functions are used for H. The exponents for the O *d*-polarization functions are taken from the MG3 basis set and the polarization functions for the C and H polarization functions are taken from the 6-311+G(d,p) basis sets. [g] The basis after the / is used for H, C, and O; the basis used for Al is the polarized triple-ζ basis set of the MEC method. The 6-311+G(d*,p) basis set is identical to the 6-311+G(d,p) basis set except the for exponent used in the O polarization functions, which was optimized in this work (see text).

7. Concluding Remarks

A summary of the improvements obtained in this paper is given in Table 16. Table 16 shows that the best literature EC method is considerably better than the average literature method (factor 1.5−3 for bond energies and factor of 2−4 for bond lengths). However, our new effective core method for Al gives another factor of about 3 improvement in bond energies and a slight (on average) improvement in bond lengths.

8. Summary

We have developed a new effective core method for Al for use with electronic structure methods. The quality of a calculation with this EC method is comparable to a calculation with the high-quality, all-electron MG3 basis set. (Note: for elements in this paper, MG3[30−36] is identical to 6-311++G(3d2f,2df,2p)). With the new MEC method, the mean unsigned errors, where the error is defined as the difference between values computed with the MG3 basis set

and the EC method, for dissociation energies and bond lengths are 0.06 eV and 0.007 Å, respectively. This is a significant improvement over previous EC methods, where the best literature method has an MUE of 0.27 eV for dissociation energies and 0.026 Å for bond lengths.

The basis set and EC can be obtained online at comp.chem.umn.edu/basissets/basis.cgi.

Acknowledgment. This work was supported in part by the Defense University Research Initiative in Nanotechnology (DURINT) through a grant managed by the Army Research Office. Computational resources were provided by the University of Minnesota Supercomputing Institute and by a Grand Challenge grant from the Molecular Science Computing Facility of the William R. Wiley Environmental Molecular Science Laboratory of Pacific Northwest Laboratory.

References

(1) Basch, H.; Stevens, W. J.; Krauss, M. *Chem. Phys. Lett.* **1984**, *109*, 212.

(2) McLean, A. D.; Liu, B.; Chandler, G. S. *J. Chem. Phys.* **1984**, *80*, 5130.

(3) Upton, T. H. *Phys. Rev. Lett.* **1986**, *56*, 2168.

(4) Basch, H. *Chem. Phys. Lett.* **1987**, *136*, 289.

(5) Bauschlicher, C. W., Jr.; Pettersson, L. G. M. *J. Chem. Phys.* **1987**, *87*, 2198.

(6) Pettersson, L. G. M.; Bauschlicher, C. W., Jr.; Halicioglu, T. *J. Chem. Phys.* **1987**, *87*, 2205.

(7) Langhoff, S. R.; Bauschlicher, C. W., Jr. *J. Chem. Phys.* **1990**, *92*, 1879.

(8) Cheng, H.-P.; Berry, R. S.; Whetten, R. L. *Phys. Rev. B* **1991**, *43*, 10647.

(9) Jones, R. O. *J. Chem. Phys.* **1993**, *99*, 1194.

(10) Martinez, A.; Vela, A.; Salahub, D. R. *J. Chem. Phys.* **1994**, *101*, 10677.

(11) Martinez, A.; Vela, A.; Salahub, D. R. *Int. J. Quantum Chem.* **1997**, *63*, 301.

(12) Akola, J.; Hakkinen, H.; Manninen, M. *Phys. Rev. B* **1998**, *58*, 3601.

(13) Baeck, K. K.; Bartlett, R. J. *J. Chem. Phys.* **1998**, *109*, 1334.

(14) Duque, F.; Mananes, A. *Eur. Phys. J. D* **1999**, *9*, 223.

(15) Rao, B. K.; Jena, P. *J. Chem. Phys.* **1999**, *111*, 1890.

(16) Geske, G.; Boldyrev, A. I. *J. Chem. Phys.* **2000**, *113*, 5130.

(17) Zhan, C. G.; Zheng, F.; Dixon, D. A. *J. Am. Chem. Soc.* **2002**, *124*, 14795.

(18) Schultz, N. E.; Staszewska, G.; Staszewski, P.; Truhlar, D. G. *J. Phys. Chem. B* in press.

(19) Fu, Z.; Lemire, Y. M. H.; Taylor, S.; Shui, J. C.; Morse, M. D. *J. Chem. Phys.* **1988**, *88*, 3524.

(20) Fu, Z.; Lemire, G. W.; Bishea, G. A.; Morse, M. D. *J. Chem. Phys.* **1990**, *93*, 8420.

(21) Fu, Z.; Russon, M.; Morse, M. D.; Armentrout, P. B. *Int. J. Mass Spec.* **2001**, *204*, 143.

(22) Cai, M. F.; Dzugan, T. P.; Bondybey, V. E. *Chem. Phys. Lett.* **1989**, *155*, 430.

(23) Harrington, J. E.; Weisshaar, J. C. *J. Chem. Phys.* **1990**, *93*, 854.

(24) Becke, A. *J. Chem. Phys.* **1993**, *98*, 5648.

(25) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Chem. Phys.* **1994**, *98*, 11623.

(26) Stephens, P. J.; Devlin, F. J.; Ashvar, C. S.; Bak, K. L.; Taylor, P. R.; Frisch, M. J. *ACS Symp. Ser.* **1996**, *629*, 105.

(27) Adamo, C.; Barone, V. *J. Chem. Phys.* **1998**, *108*, 664.

(28) Adamo, C.; Cossi, M.; Barone, V. *Theochem* **1999**, *493*, 145.

(29) Lynch, B. J.; Fast, P. L.; Harris, M.; Truhlar, D. G. *J. Phys. Chem. A* **2000**, *104*, 4811.

(30) McLean, A. D.; Chandler, G. S. *J. Chem. Phys.* **1980**, *72*, 5639.

(31) Gordon, M. S.; Binkley, J. S.; Pople, J. A.; Pietro, W. J.; Hehre, W. J. *J. Chem. Phys.* **1982**, *77*, 3654.

(32) Pietro, W. J.; Francl, M. M.; Hehre, W. J.; DeFrees, D. J.; Pople, J. A.; Binkley, J. S. *J. Am. Chem. Soc.* **1982**, *104*, 5039.

(33) Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, *80*, 3265.

(34) Spitznagel, G. W.; Clark, T.; Schleyer, P. v. R. *J. Comput. Chem.* **1987**, *8*, 1109.

(35) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. *J. Chem. Phys.* **1998**, *109*, 7764.

(36) Fast, P. L.; Sanchez, M. L.; Truhlar, D. G. *Chem. Phys. Lett.* **1999**, *306*, 407.

(37) Kahn, L. R.; Goddard, W. A., III. *J. Chem. Phys.* **1972**, *56*, 2685.

(38) Kahn, L. R.; Baybutt, P.; Truhlar, D. G. *J. Chem. Phys.* **1976**, *65*, 3826.

(39) Rappe, A. K.; Smedley, T. A.; Goddard, W. A., III. *J. Phys. Chem.* **1981**, *85*, 1662.

(40) Stevens, W. J.; Basch, H.; Krauss, M. *J. Chem. Phys.* **1984**, *81*, 6026.

(41) Krauss, M.; Stevens, W. J. *Annu. Rev. Phys. Chem.* **1984**, *35*, 357.

(42) Wadt, W. R.; Hay, P. J. *J. Chem. Phys.* **1985**, *82*, 284.

(43) Igel-Mann, G.; Stoll, H.; Preuss, H. *Mol. Phys.* **1988**, *65*, 1321.

(44) Bergner, A.; Dolg, M.; Küchle, W.; Hermann, S.; Preuss, H. *Mol. Phys.* **1993**, *80*, 1431.

(45) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; J. A. Montgomery, J.; 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.; Salvador, P.; Dannenberg, J. J.; 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.; 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. GAUSSIAN98; version A.11; Gaussian, Inc.: Pittsburgh, PA, 2001.

(46) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; 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. GAUSSIAN03; version B.01; Gaussian, Inc.: Pittsburgh, PA, 2003.

(47) Stevens, W. J.; Krauss, M.; Basch, H.; Jasien, P. G. *Can. J. Chem.* **1992**, *70*, 612.

(48) Cundari, T. R.; Stevens, W. J. *J. Chem. Phys.* **1993**, *98*, 5555.

(49) Langhoff, S. R.; Bauschlicher, C. W. *J. Chem. Phys.* **1989**, *92*, 1879.

(50) *American Institute of Physics Handbook*, 3rd ed.; Zemanksy, M. W., Gray, D. E., Eds.; McGraw-Hill Book Company: New York, 1972.

(51) Boldyrev, A. I.; Gonzales, N.; Simons, J. *J. Phys. Chem.* **1994**, *98*, 9931.

(52) Gustev, G. L.; Jena, P.; Bartlett, R. J. *J. Chem. Phys.* **1999**, 2928.

(53) Chertihin, G. V.; Andews, L.; Taylor, P. R. *J. Am. Chem. Soc.* **1994**, *116*, 3513.

(54) Hildenbrand, D. L. *Chem. Phys. Lett.* **1973**, *20*, 127.

(55) Wagner, E. L. *Theor. Chim. Acta* **1974**, *32*, 295.

(56) Dekock, R. L.; Barbachyn, M. R. *J. Inorg. Nucl. Chem.* **1981**, *43*, 2645.

(57) Dyke, J. M.; Feher, M.; Hastings, M. P.; Morris, A.; Paul, A. J. *Mol. Phys.* **1986**, *58*, 161.

(58) Masip, J.; Clotet, A.; Ricart, J. M.; Illas, F.; Rubio, J. *Chem. Phys. Lett.* **1988**, *144*, 373.

(59) Leszczynski, J.; Kwiatkowski, J. *J. Phys. Chem.* **1992**, *96*, 4148.

(60) Lynch, B. J.; Truhlar, D. G. *J. Phys. Chem. A* **2003**, *107*, 3898.

(61) Zhao, Y.; Pu, J.; Lynch, B. J.; Truhlar, D. G. *Phys. Chem. Chem. Phys.* **2004**, *6*, 673.

(62) Becke, A. *Phys. Rev. A* **1988**, *88*, 3098.

(63) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. A* **1988**, *37*, 785.

(64) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.

(65) Becke, A. *J. Chem. Phys.* **1993**, *98*, 5648.

(66) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.

(67) Becke, A. D. *J. Chem. Phys.* **1996**, *104*, 1040.

(68) Zhao, Y.; Lynch, B. J.; Truhlar, D. G. *J. Phys. Chem. A* **2004**, *108*, 2715.

(69) Höllwarth, A.; Böhme, M.; Dapprich, S.; Ehlers, A. W.; Gobbi, A.; Jonas, V.; Köhler, K. F.; Stegmann, R.; Veldkamp, A.; Frenking, G. *Chem. Phys. Lett.* **1993**, *208*, 237.

(70) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257.

(71) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213.

(72) Clark, T.; Chandrasekhar, J.; Schleyer, P. v. R. *J. Comput. Chem.* **1983**, *4*, 294.

(73) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650.

(74) Carroll, D. L. Genetic Algorithms and Optimizing Chemical Oxygen−Iodine Lasers. In *Developments in Theoretical and Applied Mechanics*; Wilson, H. B., Batra, R. C., Bert, C. W., Davis, A. M. J., Schapery, R. A., Stewart, D. S., Swinson, F. F., Eds.; The University of Alabama: Tuscaloosa, AL, 1996; Vol. XVIII, p 411.

(75) Carroll, D. L. *AIAA J.* **1996**, *34*, 338.

(76) Goldberg, D. E. *Genetic Algorithms in Search, Optimizing and Machine Learning*; Addison-Wesley: Reading, MA, 1989.

(77) Straatsma, T. P.; Apra, E.; Windus, T. L.; Dupuis, M.; Bylaska, E. J.; Jong, W. d.; Hirata, S.; Smith, D. M. A.; Hackler, M. T.; Pollack, L.; Harrison, R. J.; Nieplocha, J.; Tipparaju, V.; Krishnan, M.; Brown, E.; Cisneros, G.; Fann, G. I.; Fruchtl, H.; Garza, J.; Hirao, K.; Kendall, R.; Nichols, J. A.; Tsemekhman, K.; Valiev, M.; Wolinski, K.; Anchell, J.; Bernholdt, D.; Borowski, P.; Clark, T.; Clerc, D.; Dachsel, H.; Deegan, M.; Dyall, K.; Elwood, D.; Glendening, E.; Gutowski, M.; Hess, A.; Jaffe, J.; Johnson, B.; Ju, J.; Kobayashi, R.; Kutteh, R.; Lin, Z.; Littlefield, R.; Long, X.; Meng, B.; Nakajima, T.; Niu, S.; Rosing, M.; Sandrone, G.; Stave, M.; Taylor, H.; Thomas, G.; Lenthe, J. v.; Wong, A.; Zhang, Z. *NWChem*; version 4.5; Pacific Northwest National Laboratory: Richland, WA, 2003.

(78) Woon, D. E.; Dunning, T., Jr. *J. Chem. Phys.* **1993**, *98*, 1358.

CT0400018

JCTC Journal of Chemical Theory and Computation

Use of Block Hessians for the Optimization of Molecular Geometries

Jingzhi Pu and Donald G. Truhlar*

Department of Chemistry and Supercomputing Institute, 207 Pleasant Street S.E., Minneapolis, Minnesota 55455-0431

Received June 15, 2004

Abstract: We test a strategy for using block Hessians for transition state geometry optimizations. The block Hessian matrix is constructed by mixing a small critical block of the accurate Hessian for key atoms involved in bond breaking and forming with large noncritical blocks of the low-level Hessian. The method is tested for transition state optimizations at the MC-QCISD/3 level for five reactive systems: $H + CH_3OH$, $O + CH_4$, $OH + CH_4$, $NH_2 + CH_4$, and $H + C_2H_5OH$. When the entire low-level Hessian was used, significant oscillations were observed during the optimizations for the first four systems, whereas the transition state for the last system was optimized to a wrong structure. The block Hessian strategy efficiently removed these pathological effects of using low-level Hessians, and therefore it provides a highly reliable method for optimizing transition state structures with reduced computational cost. The method is very general, and it is especially well suited for optimizing transition state structures and equilibrium structures of large systems at very high levels of theory.

Use of molecular mechanics or electronic structure calculations to optimize molecular geometries or electronic structure calculations to optimize saddle points is one of the most common and important computational steps in modern theoretical chemistry.[1] The foundation algorithm for geometry optimization is the iterative Newton−Raphson algorithm;[2] however, most optimizations are carried out with simplified versions of this algorithm, which are often called quasi-Newton methods. A serious impediment to using full Newton−Raphson calculations is that they require a Hessian at every iteration, and Hessians are expensive, especially for the more reliable levels of electronic structure theory. Furthermore, as system size increases the number of Hessian elements increases as the square of the number (N) of atoms.

There are two main strategies in use to decrease the cost of Hessian evaluations in geometry optimization. We assume that one wants the geometry optimized at electronic structure level L. One strategy would be to use approximate Hessians at level L. For example, at a previous iteration one might have computed an accurate level-L Hessian at geometry $\mathbf{R}$. Now the iterations have advanced the geometry to $\mathbf{R}'$. One can use gradient calculations at recent iterations to approximately update the Hessian to $\mathbf{R}'$. A variation on this first strategy is to use the accurate Hessian at $\mathbf{R}$ (without updates) as an approximate Hessian at $\mathbf{R}'$. The second strategy is to use accurate Hessians at a lower level L'. For example one might use Hartree−Fock (HF) Hessians[3] to optimize a geometry at the higher level of quadratic configuration interaction with single and double excitation[4] (QCISD).

$$\begin{pmatrix} \text{CC} & \text{CN} \\ \text{NC} & \text{NN} \end{pmatrix}$$

Figure 1. The coordinates are divided into a critical (C) group and a noncritical (N) group and the Hessian matrix is blocked accordingly.

The purpose of this article is to consider a third general strategy for minimizing the Hessian cost and to demonstrate that it can be used successfully. The strategy consists of blocking the Hessian matrix, as in Figure 1, and treating the critical block or blocks at a high level and less critical blocks at a lower level. For example, consider the reaction

$$OH + C_nH_{2n+2} \rightarrow H_2O + C_nH_{2n+1} \qquad \text{(R1)}$$

If we label the acceptor oxygen atom, the transferred

10.1021/ct0400020 CCC: $30.25
Published on Web 11/02/2004

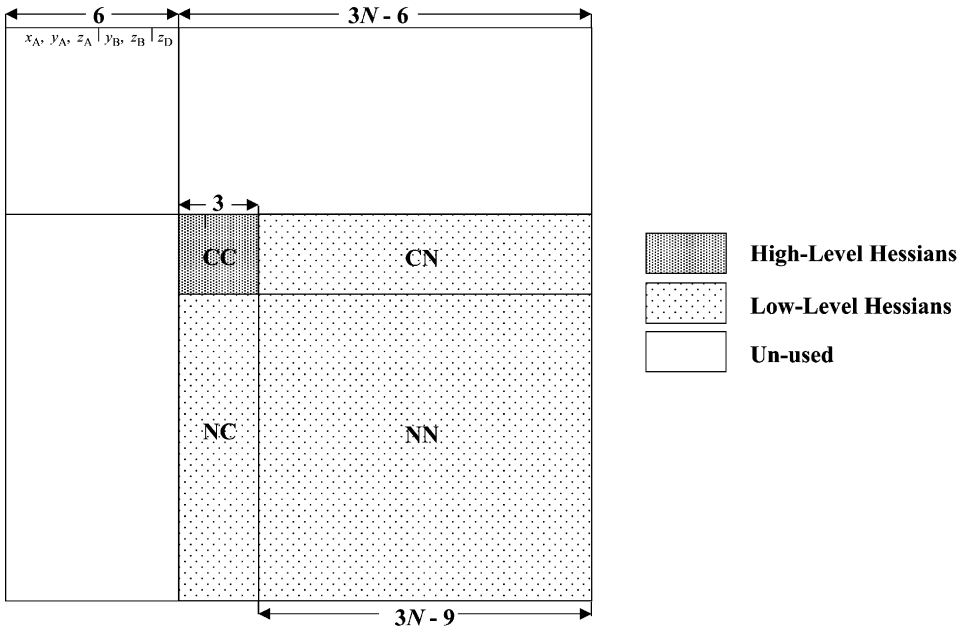


Figure 2. Another partition of the Hessian matrix where the size of the critical block is reduced from 9 × 9 to 3 × 3.

hydrogen atom, and the donor carbon atom as critical (C) and all other atoms as noncritical (N), the number of elements of the CC block is 81, the number of elements of CN and NC blocks is $162n + 54$, and the number of elements of the NN block is $81n^2 + 54n + 9$. Thus if we need high-level Hessians only for the blocks involving critical atoms (CC, CN, and NC), we will have changed the quadratic scaling to linear scaling. Already at $n = 8$, the number of Hessian elements to be evaluated at the high level is reduced by a factor of 5. Furthermore, if we only need the high-level Hessian for the CC block, the number of high-level Hessian elements required is a constant irrespective of the system size n. Thus if this strategy works (and the present paper show that it does), one should be able to optimize much larger structures at high (expensive) levels of electronic structure theory. Thus this approach has the potential to revolutionize computational strategies for electronic structure calculations on large systems.

The strategy tested here is expected to be especially useful for optimizations at very high levels where the Hessians are usually evaluated by numerical differentiation of very expensive gradients or energies. In such cases it is easy to carry out numerical differentiation of any desired subset of Hessians. For large systems one could imagine using sophisticated labeling and blocking schemes to identify the critical subset of Hessian elements. In this article, we will use the simple scheme of Figure 1 where the CC block consists of 3 Cartesian degrees of freedom for a donor atom, 3 for a transferred element, and 3 for an acceptor atom, and high-level Hessians are only calculated for this 9 × 9 critical block.

Having chosen a 9 × 9 critical block, we can, however, achieve a further efficiency by a strategic choice of the origin and the orientation of the molecule. When we use unblocked Hessians and optimize a geometry, we can place the first atom (A) at the origin ($x_A = y_A = z_A = 0$), the second atom (B) on the x axis ($y_B = z_B = 0$), and the third atom (D) to lie in the xy plane ($z_D = 0$). During the optimization these six coordinates are held fixed, and only $3N - 6$ coordinates vary. In the blocked method, we can always number the acceptor, donor, and transferred atom as the first three atoms. This partitions the 9 Cartesian coordinates of these atoms into 6 fixed critical coordinates and 3 nonfixed critical coordinates (these will be called the active coordinates). Since we do not need gradients or Hessian elements for the fixed atoms, this reduces the size of the CC block that needs to be calculated to 3 × 3. This new blocking is shown in Figure 2. To calculate the full $3N \times 3N$ Hessian at **R** by forward differentiation requires (in addition to a gradient at **R**) a total of $3N$ additional gradients, each one displaced in one of the coordinates. To calculate high-level Hessians for the 3 × 3 critical block requires only 3 gradients in addition to the one at **R**. In the implementation below, we calculate a new Hessian every third step. Thus, on average, the number of gradients required per step for calculating the Hessian is decreased from $3N + 1$ to 2.

Although the kind of strategy tested here has been mentioned as a possible approach in a pedagogical article,[5] and although the *Gaussian* program[6] allows numerical estimation of certain elements of the Hessian in internal coordinates, we know of no published examples of the kind of treatment espoused here. Furthermore the advantages of using this approach even with more straightforward and more general Cartesian optimization do not seem to be recognized by many current workers in the field. Therefore, in the

present article, we present several examples showing the power of the method.

In the first set of tests, the transition states for the following reactions are considered:

$$H + CH_3OH \rightarrow H_2 + CH_2OH \quad (R2)$$

$$O + CH_4 \rightarrow OH + CH_3 \quad (R3)$$

$$OH + CH_4 \rightarrow H_2O + CH_3 \quad (R4)$$

$$NH_2 + CH_4 \rightarrow NH_3 + CH_3 \quad (R5)$$

We carried out transition state geometry optimizations for the above four systems at the level of multicoefficient quadratic configuration interaction with single and double excitations, version 3 (MC-QCISD/3).[7,8] The MC-QCISD method belongs to the family of multicoefficient correlation methods (MCCMs),[7−11] which have been shown to be able to provide accurate atomization energies, reaction barrier heights, and transition state geometries at highly correlated levels. From a practical point of view, MC-QCISD represents the most accurate available method for which the computational effort[12] of energy calculations scales as N^6, where N is the number of atoms in the system. (Methods with steeper scaling are usually not affordable for studying medium- to large-sized systems.) However, the method used here is applicable to other levels[12] of electronic structure theory like MP2, MP4, QCISD, CCSD(T), etc. The MC-QCISD total energy and gradients are composed of a linear combination of several single-level components, i.e., HF/6-31G(d), HF/MG3S, MP2/6-31G(d), MP2/MG3S, and QCISD/6-31G(d).[7] The MG3S[13] basis set is identical to 6-311+G(2df,2p)[14] for systems containing only first row elements. In the current implementation[15] where single-level calculations are carried out using *Gaussian 98*,[6] the MC-QCISD analytical gradient is available; however, the MC-QCISD Hessian is obtained numerically due to the unavailability of an analytical Hessian for the QCISD component.

We will test the new idea by comparing the use of a high-level Hessian for a critical block to using the entire low-level Hessian. For this purpose we use a very straightforward optimization algorithm, in particular, the Newton−Raphson method[2] implemented in program MULTILEVEL.[15] To reduce the cost of calculating Hessians every Newton−Raphson step, we apply two different approaches.

In the first approach,[11] the Hessian matrix is calculated every three Newton−Raphson steps and is kept frozen until the next recalculated point. Whenever a Hessian recalculation is requested, either an entire low-level Hessian (LH) at level L or the blocked Hessian (BH, with the critical block at a high-level and the rest at a low level) is provided.

In the second approach, we test whether the block Hessian idea can be combined with Hessian update techniques to further improve the efficiency of the method. In particular, the LH or BH is only calculated at the first step and is updated for later steps during the geometry optimization. Because our strongest interest is transition structure searching, we use the Davidon-Fletcher-Powell (DFP) Hessian update scheme[16] which does not preserve a positive definite Hessian matrix. (Alternatively, one could use the improved Powell formula of Bofill,[17] which is probably the best update method for transition states.)

In the present article, the low-level L is taken as HF/6-31G(d) or AM1. The former choice has been shown to be good enough for carrying out multilevel optimizations for small stable species.[11]

The block Hessian matrix is obtained by forward differentiation of the analytical gradients for the Cartesian coordinates of critical atoms. (The use of forward differences saves a factor of 2 in computer time as compared to central differentiations.) For atom transfer reactions, such as the reactions tested in the present paper, by using the 9×9 critical block scheme described above, this algorithm involves 9 gradient calculations in total associated with 9 finite displacements of each Cartesian coordinate of the active atom

$$\frac{\partial^2 E}{\partial x_i \partial x_j} = \frac{g_{x_i}|_{x_j+\delta x} - g_{x_i}|_{x_j}}{\delta x} \quad (i, j = 1 - 9) \qquad (1)$$

where E is the total electronic energy of the system; x_i is ith Cartesian coordinate of a row or column of the critical block; g_{x_i} is the Cartesian gradient component corresponding to the coordinate x_i; and δx is a finite displacement. With the particular choice of geometry specified in the 3×3 critical block scheme, the blocked method only requires 3 gradient calculations to evaluate the high-level Hessian elements for the 3 active coordinates. In the tests presented below, we will apply the new method to a 9×9 critical block scheme and then repeat some of the calculations with the 3×3 critical block scheme.

It is worthwhile to point out that the CN and NC blocks of the Hessian matrix are also available by above numerical differentiations of the analytical gradients. However, we restrict the usage of high-level Hessian elements only to the CC block throughout the present study. This strategy is especially appealing for cases where analytical gradients are also very expensive or even analytically unavailable; in the latter case one could use a second order numerical differentiation of the energies to obtain a fairly small critical block of the accurate Hessian without calculating other gradient components to as high a precision.

The initial geometries used in the optimizations were obtained from three different levels of theory, i.e., HF/6-31G(d), B3LYP/6-31G(d), and QCISD/MG3. The optimization is considered to be complete if the maximum component of the gradient is less than 10^{-4} hartree/bohr. Table 1 lists key distances for bond breaking and forming in these initial geometries compared to the fully optimized geometries at the MC-QCISD/3 level.

We tabulate the energies and maximum gradients during the optimization process for transition states of R2 and R3 in Tables 2 and 3. It is observed a large number of cases in investigation hardly converged if the entire low-level Hessian is used. The case starting with the HF/6-31G(d) geometry is quickly converged using low-level Hessians; however, this may be fortuitous. For converged cases, Tables 2 and 3 estimate the cost in terms of the number of gradient calculations for the optimization. Among the six cases in

Table 1. Breaking and Forming Bond Distances in Initial Geometries Compared to Those in MC-QCISD/3 Optimized Geometries for Reactions R2−R6

	I[a]	II[b]	III[c]	IV[d]
R2				
C−H	1.351	1.302	1.316	1.307
H−H	0.959	1.011	0.969	0.976
R3				
C−H	1.369	1.389	1.302	1.275
H−O	1.167	1.147	1.202	1.226
R4				
C−H	1.313	1.274	1.227	1.194
H−O	1.200	1.238	1.285	1.320
R5				
C−H	1.358	1.355	1.314	1.298
H−N	1.252	1.248	1.270	1.287
R6				
C−H	1.340	1.268	nc[e]	1.282
H−H	0.969	1.054	nc[e]	1.000

[a] HF/6-31G(d) geometry. [b] B3LYP/6-31G(d) geometry. [c] QCISD/MG3 geometry. [d] Optimized geometry at the MC-QCISD/3 level. [e] nc denotes not calculated.

Tables 2 and 3, five of them display oscillations if the entire low-level Hessian (LH) is used, but using the block Hessian (BH) removes these oscillations in all cases. Although the high-level Hessian (HH) works equally as well as the BH, the cost of computing all the elements numerically makes it unfeasible for systems of large size, which prevents applying this approach to many interesting problems.

To evaluate the new BH method more systematically, we apply a statistical approach. Combining three different initial geometries, two levels of the low-level Hessian, and two Hessian update schemes for reactions R2−R5, we have 48 cases in total. The numbers of converged cases have been collected based on various conditions in Table 4, for both the LH and the BH methods. The notation X/Y represents that there are X cases successfully converged among Y cases that satisfy the conditions described in the first column of Table 4. If the entire low-level Hessian is used, only 11 of the 48 cases converge within 30 iterations. It is encouraging that the number of converged cases is significantly increased to 29 when the block Hessian strategy is applied, which is almost three times more robust compared to the method that simply uses the low-level Hessian. Table 4 indicates that the BH strategy outperforms the LH one over broad range of reactions, initial geometries, levels of theory to calculate the Hessian, and Hessian update schemes. In Table 5, we collected the statistics for R2 and R3 with the 3×3 critical block version of the BH strategy (as shown in Figure 2). Using the 3×3 critical block further improves the efficiency of the BH geometry optimization, and the stability of the algorithm is not altered.

The above tests show that the quality of the critical block of the Hessian matrix plays an important role in geometry optimizations of transition state structures. The errors of using low-level Hessian to approximate the high-level Hessian sometimes can be so significant that the optimizations oscillate wildly even if one starts from a geometry close to the high-level saddle point, for example, the QCISD/MG3 geometry (see comparisons of key bond distances in Table 1). The oscillations of energies during the optimization are not the only pathological behavior when the low-level Hessian is used; the transition state can also be optimized to

Table 2. Maximum Component of the Gradient (in au) for Transition State Optimization of $H + CH_3OH \rightarrow H_2 + CH_2OH$ at the MC-QCISD/3 Level Using Low-Level Hessians (LH), Block Hessians (BH, as Figure 1), and High-Level Hessians (HH)

	I[a]			II[b]			III[c]		
iteration	LH[d]	BH[e]	HH[f]	LH	BH	HH	LH	BH	HH
0	1.6E−2	1.6E−2	1.6E−2	9.2E−3	9.1E−3	9.1E−3	5.0E−3	5.0E−3	5.0E−3
1	2.6E−3	2.9E−3	1.0E−2	3.5E−3	1.9E−3	4.0E−4	7.0E−4	6.2E−4	2.2E−4
2	9.9E−4	1.5E−3	3.5E−2	4.3E−3	8.1E−4	1.3E−4	3.0E−4	2.7E−4	2.0E−4
3	2.1E−3	8.0E−4	6.0E−2	1.8E−3	4.2E−4	7.1E−5	1.6E−3	1.2E−4	6.2E−4
4	2.4E−4	3.8E−4	9.8E−2	1.2E−2	2.7E−4		4.8E−3	5.7E−5	2.4E−4
5	1.4E−4	2.8E−4	6.9E−2	1.0E−2	1.9E−4		1.2E−3		5.4E−5
6	5.8E−5	2.5E−4	8.8E−2	9.8E−3	1.4E−4		5.1E−3		
7		2.2E−4	4.8E−2	1.2E−2	1.1E−4		6.8E−3		
8		1.9E−4	7.2E−2	5.0E−3	9.2E−5		5.3E−3		
9		1.7E−4	1.0E−1	1.7E−2			2.6E−3		
10		1.5E−4	5.9E−2	1.5E−2			2.0E−3		
11		1.4E−4	8.4E−2	7.1E−3			1.1E−3		
12		1.2E−4	1.0E−1	2.2E−2			3.0E−3		
13		1.1E−4	6.4E−2	2.4E−2			1.1E−2		
14		9.8E−5	2.6E−2	6.9E−3			9.5E−4		
15			4.0E−2	3.3E−2			8.0E−3		
16			3.2E−2	3.6E−3			9.1E−3		
17			4.0E−3	6.2E−2			7.3E−3		
18			2.1E−3	8.1E−2			6.5E−3		
cost[g]	7	60	nc[h]	nc[h]	36	46	nc[h]	23	90

[a] HF/6-31G(d) initial geometry. [b] B3LYP/6-31G(d) initial geometry. [c] QCISD/MG3 initial geometry. [d] For the LH column, the HF/6-31G(d) level is used for the entire Hessian, and a new Hessian is computed at every third iteration; no Hessian update schemes are used. [e] Use the 9×9 critical block scheme. [f] Central differentiations of gradients are used. [g] In terms of gradient calculations. [h] nc denotes not converged within 30 iterations.

Table 3. Maximal Gradient for Transition State Optimization of $O + CH_4 \rightarrow OH + CH_3$ at the MC-QCISD/3 Level Using Low-Level Hessians (LH), Block Hessians (BH, as Figure 1), and High-Level Hessians (HH)

	I[a]			II[b]			III[c]		
iteration	LH[d]	BH[e]	HH[f]	LH	BH	HH	LH	BH	HH
0	1.5E−2	1.5E−2	1.5E−2	1.8E−2	1.8E−2	1.8E−2	5.7E−3	5.7E−3	5.7E−3
1	9.8E−3	6.4E−3	4.4E−2	1.3E−2	6.7E−3	9.3E−3	3.6E−3	2.6E−3	5.2E−4
2	7.5E−3	2.5E−3	2.4E−3	7.7E−3	3.0E−3	6.1E−3	2.2E−3	1.1E−3	9.2E−5
3	4.4E−2	1.3E−3	1.5E−3	4.4E−3	1.8E−3	6.2E−3	1.3E−3	4.6E−4	
4	7.4E−3	5.9E−4	4.0E−4	2.6E−3	8.6E−4	7.5E−4	7.2E−4	2.1E−4	
5	2.8E−2	2.7E−4	8.5E−4	1.5E−3	4.1E−4	4.0E−5	6.4E−4	9.3E−5	
6	5.8E−2	1.3E−4	3.3E−3	8.4E−4	1.9E−4		1.3E−3		
7	3.5E−2	5.9E−5	1.4E−4	6.3E−4	9.6E−5		1.4E−3		
8	6.3E−2		2.4E−5	1.1E−3			2.6E−3		
9	1.3E−1			2.0E−3			5.0E−3		
10	1.0E−1			2.0E−3			4.8E−3		
11	5.2E−2			3.8E−3			1.1E−2		
12	9.5E−2			7.6E−3			2.0E−2		
13	8.2E−2			7.5E−3			1.6E−2		
14	7.4E−2			1.4E−2			4.3E−2		
15	4.8E−2			2.8E−2			7.0E−2		
16	6.2E−2			2.6E−2			6.7E−2		
17	7.0E−2			5.7E−2			1.7E−1		
18	9.3E−2			7.6E−2			7.3E−2		
cost[g]	nc[h]	35	117	nc[h]	35	78	nc[h]	24	39

[a] HF/6-31G(d) initial geometry [b] B3LYP/6-31G(d) initial geometry [c] QCISD/MG3 initial geometry [d] For the LH column, the HF/6-31G(d) level is used for the entire Hessian, and a new Hessian is computed at every third iteration; no Hessian update schemes are used. [e] Use the 9 × 9 critical block scheme. [f] Central differentiations of gradients are used. [g] In terms of gradient calculations. [h] nc denotes not converged within 30 iterations.

Table 4. Statistics of the Convergence Cases Using the Entire Low-Level Hessian (LH) and the Block Hessian (BH) Strategy (as Figure 1)

collect by	LH	BH[a]
reaction		
R2	4/12	7/12
R3	1/12	9/12
R4	3/12	7/12
R5	3/12	6/12
initial geometry		
HF/6-31G(d)	3/16	8/16
B3LYP/6-31G(d)	2/16	9/16
QCISD/MG3	6/16	12/16
low-level Hessian		
HF/6-31G(d)	8/24	16/24
AM1	3/24	13/24
Hessian update scheme		
no update[b]	5/24	18/24
DFP	6/24	11/24
in total		
	11/48	29/48

[a] This table is based on the 9 × 9 scheme for the high-level block. [b] LH or BH Hessians are calculated every three steps and kept frozen until the next recalculation point.

Table 5. Statistics of the Convergence Cases Using the Entire Low-Level Hessian (LH) and the Block Hessian (BH) Strategy (as Figure 2)

collect by	LH	BH[a]
reaction		
R2	2/12	7/12
R3	6/12	10/12
initial geometry		
HF/6-31G(d)	2/8	4/8
B3LYP/6-31G(d)	3/8	6/8
QCISD/MG3	3/8	7/8
low-level Hessian		
HF/6-31G(d)	2/12	9/12
AM1	6/12	8/12
Hessian update scheme		
no update[b]	5/12	9/12
DFP	3/12	8/12
in total		
	8/24	17/24

[a] This table is based on the 3 × 3 scheme for the high-level block. [b] LH or BH Hessians are calculated every three steps and kept frozen until the next recalculation point.

a wrong structure. One of the examples of this type is the transition state optimization for the following reaction:

$$H + C_2H_5OH \rightarrow H_2 + C_2H_4OH \qquad \text{(R6)}$$

In Figures 3 and 4, we plot the energies as function of the optimization step for cases in which HF/6-31G(d) and B3LYP/6-31G(d) geometries were used as an initial geometry. Both Figures 3 and 4 show that using the low-level Hessian can result in a low-energy structure other than the true saddle point (it is a reactant-like structure), but using the block Hessian corrects this error straightforwardly and makes the optimizations much more stable.

The emphasis in the present article has been on saddle point optimization at high levels of electronic structure theory in medium-size molecules. When analytic Hessians are not available, computing the numerical Hessian presents a computational bottleneck for geometry optimizations if the

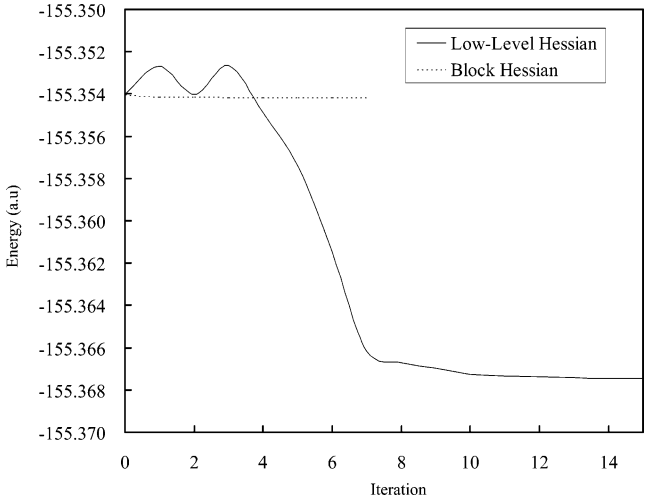


Figure 3. Energy of the transition state optimization for H + $C_2H_5OH \rightarrow H_2 + C_2H_4OH$ using the low-level Hessian at HF/6-31G(d) level or block Hessians. For the block Hessian optimization, the 9 × 9 scheme for the high-level block is used. The HF/6-31G(d) geometry is used as the initial geometry.

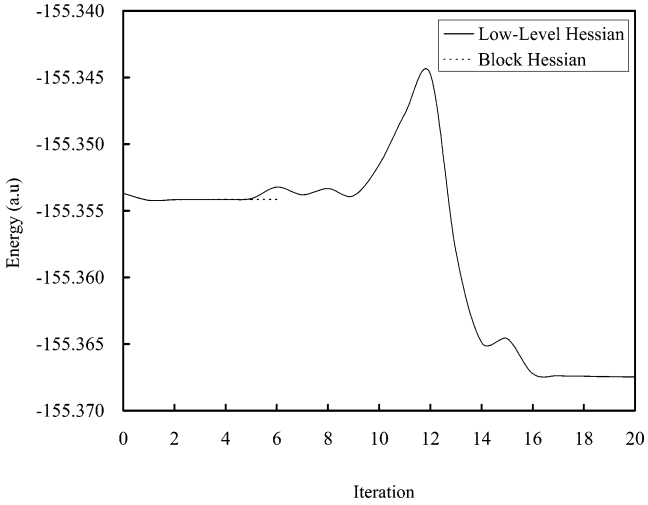


Figure 4. Energy of the transition state optimization for H + $C_2H_5OH \rightarrow H_2 + C_2H_4OH$ using the low-level Hessian at HF/6-31G(d) level or block Hessians. For the block Hessian optimization, the 9 × 9 scheme for the high-level block is used. The B3LYP/6-31G(d) geometry is used as the initial geometry.

optimization algorithm requires full Hessians. In such a case the cost is dominated by the calculations of the Hessians, at least when the present blocked Hessian method is not used. As we mentioned above, one may use the lower-level Hessian as an alternative to approximate the true Hessian during the optimization procedure. Because a low-level Hessian, especially if computed analytically and by a method with lower scaling order than the high-level Hessian, is relatively computationally inexpensive, one may hope that it provides a satisfactory solution that alleviates the bottleneck of computing numerical high-level Hessian during geometry optimizations. However, for some difficult cases, especially for transition state geometry optimizations, using a low-level Hessian is not powerful enough because the transition vector (the eigenvector associated with a negative eigenvalue) to bring the system to a first-order saddle point is very sensitive to the accuracy of the Hessian.[17] For this kind of situation, the accurate description of the key Hessian elements involved in the bond breaking and forming is crucial to direct the optimization to the correct transition state structure, and the present method may be especially useful.

Another case of interest is for very large molecules with lower levels of electronic structure theory (for example, neglect-of-diatomic-differential-overlap molecular orbital theory) where the cost is dominated not by the calculation of the Hessian but by its diagonalization. A scheme that splits the system into a reaction core and its environment has also been found to be useful in that case.[18]

The present paper has used an algorithm in Cartesian coordinates, although there are some well-known advantages to optimizing in internal coordinates. The advantages of Cartisian coordinates are also well-known,[19] namely the simplicity and generality of the algorithms used and the automatic avoidance of redundancies. In contrast, the advantage of internal coordinates is that when they are chosen physically, for example, as valence internal coordinates (bond stretches, valence angle bends, and torsions), they eliminate much of the coupling between Cartesians. Our experience, however, and that of others[19,20] is that optimization in Cartesian coordinates can be very efficient when Hessian information is used. Furthermore, Cartesians are much preferred to internal coordinates for solid-state optimizations[21] or surface science,[22] where internals can be very cumbersome, or for clusters.[20] A key result of this paper is that Cartesians can remain efficient even when only a small portion of the Hessian is computed at the high electronic structure level at which saddle point optimization is required. As mentioned above, partial Hessians can also be used with internal coordinates.

There is considerable interest in combined quantum mechanical and molecular mechanical methods for electronic structure calculations, and this is a special case of the need for multiscale[23] and multilevel[24,25] algorithms. To treat large systems efficiently, one requires such algorithms not only for the electronic structure step but also for other steps such as geometry optimizations, computational thermochemistry, and dynamics. The present article presents such a scheme suitable for geometry optimization.

References

(1) Schlegel, H. B. *Adv. Chem. Phys.* **1987**, *67*, 249.

(2) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. T. *Numerical Recipes*, 2nd ed.; Cambridge University Press: Cambridge, 1992; pp 372 ff.

(3) Pople, J. A.; Krishnan, R.; Schlegel, H. B.; Binkey, J. S. *Int. J. Quantum Chem. Symp.* **1979**, *13*, 225.

(4) Pople, J. A.; Head-Gorden, M.; Raghavachari, K. *J. Chem. Phys.* **1987**, *87*, 5968.

(5) Schlegel, H. B. In *New Theoretical Concepts for Understanding Organic Reactions*; Bertrán, J. Csizmadia, I. G., Eds.; NATO ASI Series C 267; Kluwer: Dordrecht, 1989; p 33. See p 48.

(6) *Gaussian98*, by M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, R. E. Stratmann, J. C.

Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B. B Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle, and J. A. Pople, Gaussian, Inc., Pittsburgh, PA, 1998.

(7) Lynch, B. J.; Truhlar, D. G. *J. Phys. Chem. A* **2003**, *107*, 3898.

(8) Fast, P. L.; Truhlar, D. G. *J. Phys. Chem. A.* **2000**, *104*, 6111.

(9) Tratz, C. M.; Fast, P. L.; Truhlar, D. G. *PhysChemComm* **1999**, *2*, Article 14.

(10) Fast, P. L.; Sanchez, M. L.; Truhlar, D. G. *Chem. Phys. Lett.* **1999**, *306*, 407.

(11) Rodgers, J. M.; Fast, P. L.; Truhlar, D. G. *J. Chem. Phys.* **2000**, *112*, 3141.

(12) Raghavachari, K.; Anderson, J. B. *J. Phys. Chem.* **1996**, *100*, 12960.

(13) Lynch, B. J.; Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2003**, *107*, 1384.

(14) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650.

(15) Rodgers, J. M.; Lynch, B. J.; Fast, P. L.; Zhao, Y.; Pu, J.; Chuang, Y.-Y.; Truhlar, D. G. MULTILEVEL-version 3.1; University of Minnesota, Minneapolis, 2003.

(16) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes in Fortran 77*; Cambridge University Press: Cambridge, 1992; pp 418 ff.

(17) Bofill, J. M. *Intern. J. Quantum Chem.* **2003**, *94*, 324.

(18) Turner, A. J.; Moliner, V.; Williams, I. H. *Phys. Chem. Chem. Phys.* **1999**, *1*, 1323.

(19) Baker, J.; Hehre, W. J. *J. Comput. Chem.* **1991**, *12*, 606. Baker, J. *J. Comput. Chem.* **1993**, *14*, 1085.

(20) Niesse, J. A.; Mayne, H. R. *J. Chem. Phys.* **105**, *4700*, 1996.

(21) Ke, X.; Tanaka, I. *Phys. Rev. B* **2004**, *69*, 165114/1.

(22) Ciobica, I. M.; Frechard, F.; van Santen, R. A.; Kleyn, A. W.; Hafner, J. *J. Phys. Chem. B* **2000**, *104*, 3364.

(23) Butalov, V. V.; Rubia, T. D. d. l.; Phillips, P.; Kaxiras, E.; Ghoniem, N. *Multiscale Modeling of Materials*; Materials Research Society: Pittsburgh, 1999.

(24) Corchado, J. C.; Truhlar, D. G. *ACS Symp. Ser.* **1998**, *712*, 106.

(25) Petersson, G. A. *ACS Symp. Ser.* **1998**, *677*, 237.

CT0400020

JCTC Journal of Chemical Theory and Computation

Using Hessian Updating To Increase the Efficiency of a Hessian Based Predictor-Corrector Reaction Path Following Method

H. P. Hratchian and H. B. Schlegel*

Department of Chemistry and Institute for Scientific Computing, Wayne State University, Detroit, Michigan 48202

Received June 15, 2004

Abstract: The reaction path is a key concept in the theoretical description of a chemical reaction. The intrinsic reaction coordinate is defined as the steepest descent path in mass-weighted Cartesian coordinates that connects the transition state to reactants and products on the potential energy surface. Recently, a new Hessian based predictor-corrector reaction path following algorithm was presented that is comparable to a fourth-order algorithm developed earlier. Although the method is very accurate, it is costly because second derivatives of the energy are required at each step. In this work, the efficiency of the method is greatly enhanced by employing Hessian updating. Three different updating schemes have been tested: Murtagh and Sargent, Powell-symmetric Broyden, and Bofill. Bofill's update performs the best and yields excellent speed-up.

1. Introduction

The theoretical treatment of chemical reactions invariably requires some sort of reaction path following calculation. The most common use of such a calculation is to ensure that an optimized transition state (TS) lies on a path connecting the desired reactant and product minima on the potential energy surface (PES). Additionally, accurate reaction path following methods are needed to determine the steepest descent path, or minimum energy path (MEP), so that variational transition state theory (VTST) and reaction path Hamiltonian (RPH) methods can be used to calculate reaction rate constants.[1−5] Although the steepest descent path can be considered in any coordinate system, when mass-weighted Cartesian coordinates are used the MEP is also known as the intrinsic reaction coordinate (IRC).[6]

The IRC can be determined by starting at the TS and following the steepest descent pathway down to the reactant and product minima according to

$$\frac{d\mathbf{x}(s)}{ds} = -\frac{\mathbf{g}(\mathbf{x})}{|\mathbf{g}(\mathbf{x})|} \qquad (1)$$

where s is the arc length along the path, $\mathbf{x}$ is the coordinate vector, and $\mathbf{g}$ is the gradient of the PES at $\mathbf{x}$. Because eq 1 is a stiff differential equation, care must be taken during the integration. As a result, a large number of algorithms have been developed.[7−22]

Numerical methods for integrating ordinary differential equations may be classified as either explicit or implicit. Explicit methods use information at the current point to move to the next point, while implicit methods required derivative information at the end point as well. For integrating eq 1, the differential equation defining the reaction path, common explicit algorithms include Euler's method, the Ishida-Morokuma-Komornicki (stabilized Euler) method,[11,12] Runge−Kutta, the local quadratic approximation (LQA),[15,16] and the Sun-Ruedenberg modification of LQA.[18] Some of these methods use only gradient information and are limited to rather small step sizes; others also use second derivatives (the Hessian). Methods that use the Hessian are more costly, but gain additional stability that permits somewhat larger step sizes.

Implicit methods are more difficult to implement because the gradient and possibly higher-order derivatives are necessary at the end of the step. As a result, implicit methods generally use optimization schemes to iteratively solve for

* Corresponding author e-mail: hbs@chem.wayne.edu.

10.1021/ct0499783 CCC: $30.25
Published on Web 10/21/2004

the derivatives at the end point of each step[23] and thereby tend to require multiple energy and derivative calculations for each point on the path. However, implicit methods are often able to take considerably larger steps allowing them to compensate for the cost of the additional derivative calculations. Implicit methods for IRC analysis include the Müller-Brown method (implicit Euler),[17] the second-order method of Gonzalez and Schlegel (GS2),[19,20] and higher-order methods by the same authors.[21]

In recent work, we introduced a new integrator for reaction path following.[24] Our Hessian based predictor-corrector (HPC) method provides a very accurate pathway and was originally designed to provide a useful approach for describing the reaction path for use in kinetics calculations (e.g., using VTST or RPH methods). In that work, it was shown that HPC is capable of step sizes comparable to the robust and widely used GS2 method. An attractive feature of the HPC integrator is that it only requires one evaluation of the energy and derivatives per IRC step, while the GS2 method typically requires between two and five energy and derivative evaluations per step in order to iteratively solve for the endpoint gradient. However, the original HPC algorithm requires Hessians at every step making it generally more expensive than GS2, which only requires first derivatives. In the present paper we remove this bottleneck by applying Hessian updating, and we show that this affords a very efficient and accurate means for computing IRCs. The approach here is to compute the Hessian analytically only at the TS and to update the second derivatives for the rest of the IRC integration. As we show below, using an appropriate updating scheme allows HPC step sizes similar to the popular GS2 integrator while only slightly diminishing the accuracy of the original HPC method (using all analytic Hessians). This indicates that the general HPC cost will be two to five times less than for GS2, since both methods can be used with similar step sizes and HPC with Hessian updating requires only one energy and gradient evaluation at each step. In this way, the HPC algorithm is not only useful for accurate reaction path following needed for rate constant calculations but is also efficient for those studies that require reaction path following primarily to ensure that an optimized TS lies on the pathway that connects the relevant reactant and product structures on the PES.

2. Method

In this work, we incorporate standard Hessian updating schemes into our Hessian based predictor-corrector (HPC) reaction path following algorithm, which has been previously described in detail.[24] In this section, the HPC method is briefly described, and the Hessian updating approaches are discussed.

A. Hessian Based Predictor-Corrector (HPC) Method. The HPC algorithm uses the local quadratic approximation (LQA) method of Page and McIver[15,16] for predictor steps and a modified Bulirsch-Stoer integrator[25−27] on a distance weighted interpolant (DWI) surface[28−30] for corrector steps. In this section, both pieces of the HPC method are discussed.

The LQA integrator is based upon a second-order Taylor series of the PES. Truncated at the quadratic term, the Taylor series expansion of the PES about $\mathbf{x_0}$ is given by

$$E(\mathbf{x}) = E_0 + \mathbf{g_0}^{\mathbf{t}}\Delta\mathrm{x} + \frac{1}{2}\Delta\mathrm{x}^{\mathbf{t}}\mathbf{H_0}\Delta\mathrm{x} \tag{2}$$

where $\Delta\mathbf{x}$, $\mathbf{g_0}$, and $\mathbf{H_0}$ are the displacement vector of the current position from $\mathbf{x_0}$, the gradient, and Hessian at $\mathbf{x_0}$, respectively. Taking the first derivative of eq 2 with respect to $\mathbf{x}$ gives the gradient as

$$\mathbf{g(x)} = \mathbf{g_0} + \mathbf{H_0}\Delta\mathrm{x} \tag{3}$$

Substituting eq 3 into eq 1 gives

$$\frac{\mathrm{d}\mathbf{x}(s)}{\mathrm{d}} = -\frac{\mathbf{g_0} + \mathbf{H_0}\Delta\mathrm{x}}{|\mathbf{g_0} + \mathbf{H_0}\Delta\mathrm{x}|} \tag{4}$$

In the LQA method of Page and McIver, eq 4 is integrated by introducing an independent parameter, t, such that

$$\frac{\mathrm{d}s}{\mathrm{d}t} = |\mathbf{g_0} + \mathbf{H_0}\Delta\mathrm{x}| \tag{5}$$

and

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = -[\mathbf{g_0} + \mathbf{H_0}\Delta\mathrm{x}] \tag{6}$$

The solution to eq 6 is given by

$$\mathbf{x}(t) = \mathbf{x_0} + \mathbf{A}(t)\mathbf{g_0} \tag{7}$$

where

$$\mathbf{A}(t) = \mathbf{U}\alpha(t)\mathbf{U}^{\mathbf{t}} \tag{8}$$

In eq 8, $\mathbf{U}$ is the matrix of column eigenvectors of the Hessian and $\alpha(t)$ is a diagonal matrix given by

$$\alpha_{ii}(t) = (e^{-\lambda_i t} - 1)/\lambda_i \tag{9}$$

where λ_i are the eigenvalues of the Hessian.

To integrate eq 4, one must obtain a value of t such that the desired step size ($s - s_0$) is taken. To accomplish this, iterations over successive Euler integrations of eq 5 are used. The initial value for the Euler step size, δt, is estimated by

$$\delta t = \frac{1}{N_{Euler}}\frac{(s - s_0)}{|\mathbf{g_0}|} \tag{10}$$

where N_{Euler} is the number of Euler steps to be taken. In the present implementation, $N_{Euler} = 5000$. The numerical integration of eq 5 can be carried out readily in the Hessian eigenvector space

$$\frac{\mathrm{d}s}{\mathrm{d}t} = \left(\sum_i \mathbf{g'_{0_i}}^2 e^{-2\lambda_i t}\right)^{1/2} \tag{11}$$

where

$$\mathbf{g'_0} = \mathbf{U}^{\mathbf{t}}\mathbf{g_0} \tag{12}$$

At the start of the integration, when **x** corresponds to the TS, the gradient is zero, and hence the transition vector must be used in place of $\mathbf{g_0}$. At the TS, the LQA step is equivalent to the gradient extremal step, as described by Hoffman, Nord, and Ruedenburg.[31] At the end of the LQA integration, when **x** approaches the minimum wells of the reactant and product, t heads to infinity and the LQA step is equivalent to a Newton−Raphson step, which leads to the minimum energy structure in the local quadratic region. For this reason, conservation of the desired step size, $(s - s_0)$, becomes difficult in this region. Therefore, once the calculation has moved beyond the quadratic region of the TS each LQA step is compared to a Newton−Raphson step. When the Newton−Raphson step is shorter than the LQA step, the Newton−Raphson step is taken in lieu of the LQA step. This provides convergence stability for the corrector step (see below) when the minimum well is approached.

The Bulirsch-Stoer integrator, which is very well described elsewhere,[25−27,32] is used for the HPC corrector step. Each Bulirsch-Stoer step is comprised of three basic components. First, a simple gradient based integrator is used to take multiple steps along the Bulirsch-Stoer step interval. In general, this integrator is modified midpoint; however, in our tests we found that the stiff character of eq 1 is greatly magnified by modified midpoint. Therefore, our modified version employs simple Euler integration. A detailed discussion of the causes for Euler integration being more stable than modified midpoint is available in previous work.[24] The second component of a Bulirsch-Stoer step is to describe the solution of the Euler integration as a polynomial function of step size and to extrapolate to a step size of zero (corresponding to the case where an infinite number of steps are taken). The third component consists of evaluating the error of the extrapolation to zero step size. If the error is too large, the process is repeated using more steps in the Euler integration, which in turn provides one more data set for the polynomial extrapolation. If the error is acceptable ($<1 \times 10^{-6}$ amu$^{1/2}$ bohr in the present case), then the extrapolated solution is accepted and the integration is completed.

This integrator requires a large number of function and gradient evaluations and can be quite costly if energies and derivatives are required from electronic structure methods. However, in the present algorithm the Bulirsch-Stoer integration is carried out on a surface that is fitted to energy and derivative information already available at the beginning and end of the predictor step (LQA). Once the Bulirsch-Stoer integration is completed and the LQA end point is corrected, the gradient on the fitted surface is used to take the next LQA predictor step. Since the corrected end point and predicted end point are expected to lie within the same quadratic region of the PES, the Hessian from the previous predictor end point is used for the next LQA step. The validity of this assumption has been demonstrated in previous work.[24,33,34]

In the current algorithm, the Bulirsch-Stoer integration is carried out on a DWI surface such as those described by Collins and co-workers.[30,35−37] DWI surfaces have been used in a number of diverse applications[38] and are well suited for modeling chemical PESs.[30] This fitted surface gives the interpolated energy, E_{DWI}, at a position **x** according to

$$E_{DWI} = \sum_{i=1}^{2} w_i T_i \tag{13}$$

where the summation is taken over the start (i=1) and end (i=2) points of the predictor step, about which Taylor expansions, $\{T_i\}$, are evaluated and added together in a weighted fashion, defined by the weighting functions $\{w_i\}$. In the present implementation, the two Taylor expansions are truncated at the second-order term and the weighting functions are defined as

$$w_1 = \frac{|\Delta x_2|^2}{|\Delta x_1|^2 + |\Delta x_2|^2},\ w_2 = \frac{|\Delta x_1|^2}{|\Delta x_1|^2 + |\Delta x_2|^2} \tag{14}$$

B. Hessian Updating Methods. To study larger and more interesting systems with electronic structures methods, the efficiency of the HPC algorithm must be improved. To accomplish the same goal, Hessian updating methods have been utilized in the past with great success with quasi-Newton geometry optimization methods[8,39−41] and ab initio classical trajectory calculations.[34] Furthermore, other reaction path following methods have been able to make good use of Hessian updating.[9,19−21,42] These prior successes indicate that, for HPC, Hessian updating may provide a means to decrease computational cost, especially for applications where Hessians are not needed for additional calculations at each point on the path and users intend to obtain the IRC only to ensure that an optimized TS lies on a pathway connecting reactants to products. Nevertheless, careful consideration must be given to the choice of updating scheme since the updated Hessians will not only be used to propagate the IRC via LQA, but HPC also requires Hessians to fit the DWI surface for the modified Bulirsch-Stoer corrector integration. A poorly updated Hessian may lead to an inaccurate DWI surface, severely disturb the corrector integration, and provide an inadequate gradient for the next LQA predictor step.

For minimizations, the BFGS Hessian updating scheme is preferred.[43,44] A principle reason for this preference is that the BFGS formula maintains a Hessian that is positive definite. This characteristic of the BFGS formula is achieved by employing positive weights that are functions of the current Hessian. In the case of reaction path following, the Hessian will be negative definite in some regions (i.e., near the TS) and the BFGS formula becomes ill-conditioned.[45] As a result, Hessian updating methods that have been used for TS optimization problems are more appropriate and have been considered. In this work, three different Hessian updating schemes have been incorporated with HPC to determine which, if any, provides an acceptable integration of the IRC while affording the desired increase in efficiency.

The first updating scheme used is the method of Murtagh and Sargent (MS), which is also known as the symmetric rank one formula (SR1). The MS update is given by

$$\Delta\mathbf{H}_{MS} = \mathbf{H}^{new} - \mathbf{H}^{old} = \frac{(\Delta g - \mathbf{H}^{old}\Delta x)(\Delta g - \mathbf{H}^{old}\Delta x)^t}{(\Delta g - \mathbf{H}^{old}\Delta x)^t \Delta x} \quad (15)$$

For optimization on a quadratic surface, MS updating converges to the correct Hessian without exact line searches.[43,44] However, care must be taken to avoid MS updating if the denominator of eq 15 becomes small.

Alternatively, the Powell-symmetric-Broyden (PSB) updating method can be employed. The PSB update is

$$\Delta\mathbf{H}_{PSB} = \frac{(\Delta g - \mathbf{H}^{old}\Delta x)\Delta x^t + \Delta x(\Delta g - \mathbf{H}^{old}\Delta x)^t}{\Delta x^t \Delta x} - \frac{\Delta x^t(\Delta g - \mathbf{H}^{old}\Delta x)\Delta x \Delta x^t}{(\Delta x^t \Delta x)^2} \quad (16)$$

The PSB updating scheme is free from the possibility of division by zero, unlike the MS update. Another Hessian updating approach that also avoids the division by zero problem of MS updating is Bofill's formula, which combines the MS and PSB schemes.[46] Bofill's update was devised as an alternative updating scheme for TS optimizations and is given by

$$\Delta\mathbf{H}_{Bofill} = \phi\Delta H_{MS} + (1 - \phi)\Delta\mathbf{H}_{PSB} \quad (17)$$

where

$$\phi = \frac{(\Delta x^t(\Delta g - \mathbf{H}^{old}\Delta x))^2}{\Delta x^2(\Delta g - \mathbf{H}^{old}\Delta x)^2} \quad (18)$$

An alternative form for Bofill's update has been proposed by Farkas and Schlegel.[47,48] In their approach, ϕ is given by the square root of eq 18. For HPC reaction path following, the modified Bofill update provides similar results to the standard Bofill method. As a result, this update is not explicitly considered in this work.

3. Numerical Tests

The HPC algorithm and the three Hessian updating schemes discussed above have been implemented in the development version of GAUSSIAN 03.[49] Aside from the Hessian at the TS, which is computed analytically, all Hessians are obtained by updating during the HPC reaction path calculation. Four systems have been employed for testing: HCN → HNC, $CH_3CH_2F \rightarrow CH_2CH_2 + HF$, $ClCH_3 + Cl^- \rightarrow Cl^- + CH_3Cl$, and the Diels−Alder reaction of ethene and butadiene. Calculations on the HCN rearrangement reaction have been carried out at the HF/STO-3G level of theory,[50,51] those on the four center elimination reaction have been carried out at the HF/3-21G level of theory,[52−54] and calculations on $ClCH_3 + Cl^- \rightarrow Cl^- + CH_3Cl$ and the Diels−Alder reaction have been carried out at the HF/6-31G(d) level of theory.[55−59] Because the masses of the atoms that dominate the reaction path are quite different for these four reactions, a fixed step size in mass-weighted coordinates would correspond to significantly different step sizes in nonmass-weighted coordinates. However, the trust radii for the quadratic regions of the four reactions should be similar when expressed in nonmass-weighted coordinates. Therefore, we chose a step size in nonmass-weighted coordinates at the TS, transformed it to mass-weighted coordinates, and used it for the rest of the path. In particular, we selected step sizes of 0.10 and 0.40 bohr for each of the four test reactions.

As an initial test for the accuracy of IRCs computed using HPC integration with the three updating schemes described above, we compared plots of energy vs reaction coordinate and various internal coordinates vs reaction coordinate with the HPC paths computed using all analytic Hessians. In these tests, the paths were computed from the respective TSs down to reactant and product minima, which were detected according to one of two stopping criteria: (1) the magnitude of the Cartesian gradient is less than 1.5×10^{-4} Hartree bohr^{-1} or (2) the angle between two successive steps is less than 30°.

Visual inspection indicates that in the region near the TS all three updating methods yield very good pathways when compared to the path using all analytic Hessians. Near the endpoints (i.e., the reactant and product wells), the PSB and Bofill schemes perform well, while the MS updates result in large deviations from the reference pathway for some reactions. Here, we show plots for two reactions that are representative of the set studied. Figures 1 and 2 relate to the HCN rearrangement reaction. Figures 1(a) and 2(a) give the (x,y) coordinates of the H atom, where the C−N center of mass has been placed at the origin, the C−N bond has been placed on the x-axis, the C atom has been placed on the negative side of the x-axis, and the N atom has been placed on the positive x-axis. Figures 1(b) and 2(b) show the C−N bond length as a function of the reaction coordinate. Figures 3(a) and 4(a) show energy profiles for the four center elimination reaction, $CH_3CH_2F \rightarrow CH_2CH_2 + HF$. Figures 1(c), 2(c), 3(b), and 4(b) show the errors in the pathways, given by the perpendicular distance from the paths computed with all analytic Hessians, as a function of reaction coordinate.

It is clear from Figures 1 and 3 that with a step size of 0.10 bohr PSB and Bofill updating schemes both yield very good paths. MS updating does well with the four center elimination reaction with the smaller step size but yields long C−N bonds for the HCN rearrangement. For the paths integrated with a step size of 0.40 bohr, the behavior of the three updating schemes is qualitatively similar to the results from integration with a step size of 0.10 bohr. As one might expect, slight deviations from the reference path for PSB and Bofill updating with the smaller step size become somewhat more apparent when the larger step size is used. Nevertheless, agreement of the paths found using these two updating schemes with the analytic Hessian pathway is still acceptable. The errors in the MS updating paths remain small for the four center elimination reaction but are much larger for the HCN rearrangement. As shown in Figures 1(c) and 2(c), the errors in the MS pathways for the HCN rearrange-

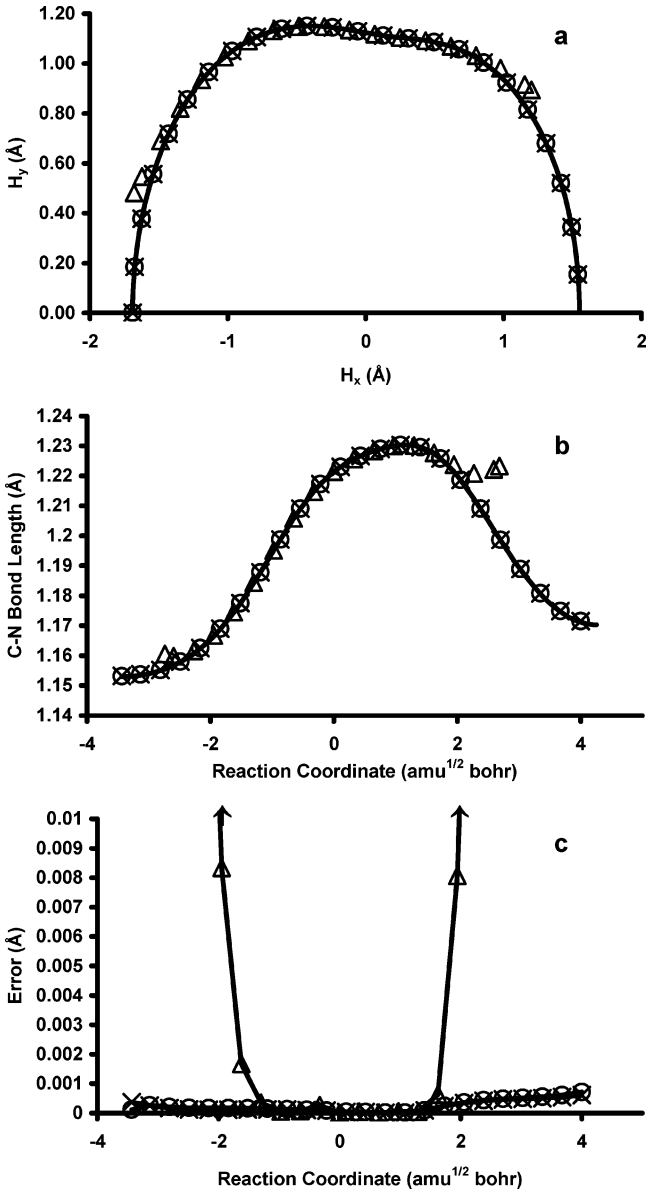


Figure 1. Reaction path for HCN → HNC with all analytic Hessians vs all updated Hessians using a step size of 0.10 bohr. (a) Coordinates of H atom relative to the C−N center of mass, (b) C−N bond length vs reaction coordinate, and (c) error, given by perpendicular distance from the all analytic Hessian path, vs reaction coordinate [HPC paths using analytic Hessians (—), MS updated Hessians (△), PSB updated Hessians (○), Bofill updated Hessians (×)]. Note: plot (c) shows solid lines connecting points for MS, PSB, and Bofill updated paths for clarity. Every third point is shown for simplicity (for all three plots).

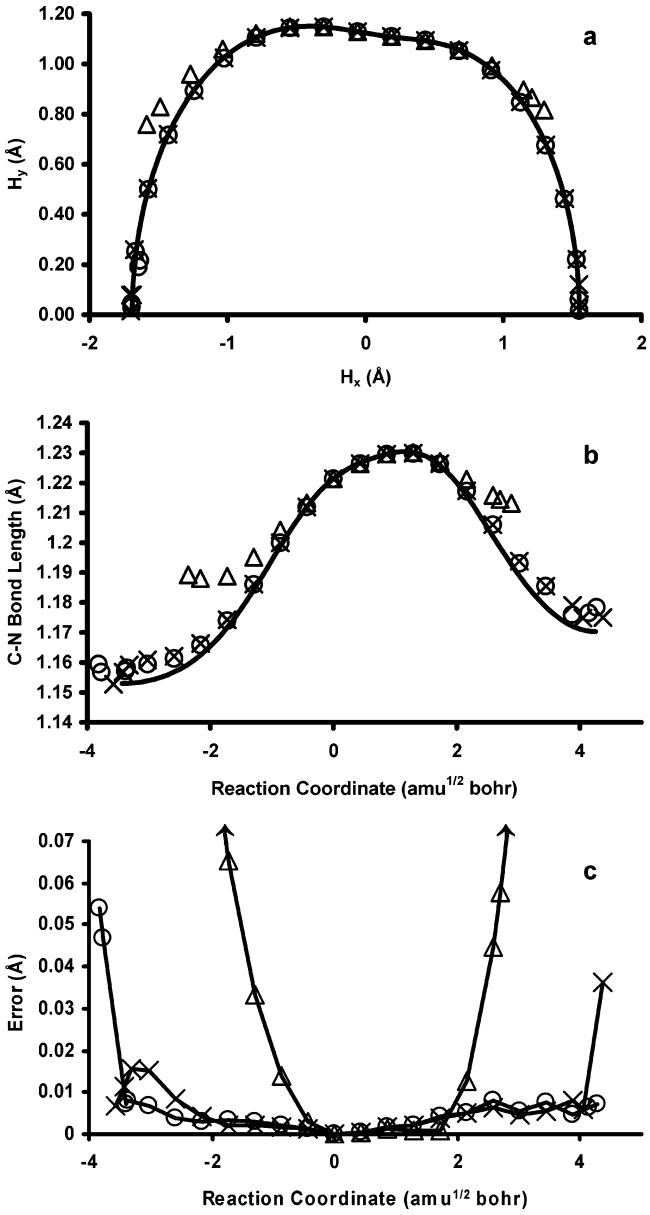


Figure 2. Reaction path for HCN → HNC with all analytic Hessians vs all updated Hessians using a step size of 0.40 bohr. (a) Coordinates of H atom relative to the C−N center of mass, (b) C−N bond length vs reaction coordinate, and (c) error, given by perpendicular distance from the all analytic Hessian path, vs reaction coordinate [HPC paths using analytic Hessians (—), MS updated Hessians (△), PSB updated Hessians (○), Bofill updated Hessians (×)]. Note: plot (c) shows solid lines connecting points for MS, PSB, and Bofill updated paths for clarity.

ment rapidly extend beyond the vertical scale of the plot as the reactant and product valleys are approached.

For both reactions, the MS paths prematurely detect minima wells and terminate. In the case of the HCN rearrangement, the calculation detects the reactant well before the H−C−N angle has reached 180°. Calculations on the four center elimination reaction end early heading toward the product when the larger step size is used and end early in the reactant direction with both step sizes. We have found that this is a common problem with MS updating, which is clearly apparent in Figures 1−4. In all cases, the calculations terminated according to the first stopping criteria listed above—the magnitude of the Cartesian gradient is less than 1.5×10^{-4} Hartree bohr^{-1}. As mentioned earlier, the gradient used in each HPC predictor step comes from the DWI gradient at the end point from the previous corrector step. Consequently, a poor Hessian can result in a local minimum on the DWI surface and artificially cause the calculation to complete.

The problems experienced using MS updating can be corrected by computing analytic Hessians every few steps.

Table 1. RMS Errors in Position (Å) for HPC Reaction Path Following Using MS, PSB, and Bofill Hessian Updating Methods with $\Delta s = 0.10$ and 0.40 bohr[a]

		Hessian updating method		
reaction	step size	MS	PSB	Bofill
HNC → HCN	0.10	3.36×10^{-2}	3.35×10^{-4}	3.05×10^{-4}
	0.40	6.96×10^{-2}	1.60×10^{-2}	1.05×10^{-2}
$CH_3CH_2F \rightarrow CH_2CH_2 + HF$	0.10	8.55×10^{-3}	2.58×10^{-3}	4.73×10^{-3}
	0.40	5.78×10^{-2}	6.56×10^{-2}	8.46×10^{-2}
$ClCH_3 + Cl^- \rightarrow Cl^- + CH_3Cl$	0.10	2.33×10^{-2}	1.09×10^{-2}	3.04×10^{-3}
	0.40	*b*	1.21×10^{-1}	3.46×10^{-2}
Diels–Alder	0.10	1.94×10^{-2}	1.38×10^{-2}	1.58×10^{-2}
	0.40	6.35×10^{-2}	8.33×10^{-2}	9.50×10^{-2}

[a] $\Delta s = 0.10$ and 0.40 bohr correspond to 0.1082 and 0.4326 amu$^{1/2}$ bohr for HNC → HCN; 0.1090 and 0.4362 amu$^{1/2}$ bohr for $CH_3CH_2F \rightarrow CH_2CH_2 + HF$; 0.3545 and 1.4182 amu$^{1/2}$ bohr for $ClCH_3 + Cl^- \rightarrow Cl^- + CH_3Cl$; and 0.2524 and 1.0098 amu$^{1/2}$ bohr for the Diels–Alder reaction. [b] The integration of eq 1 failed for this calculation.

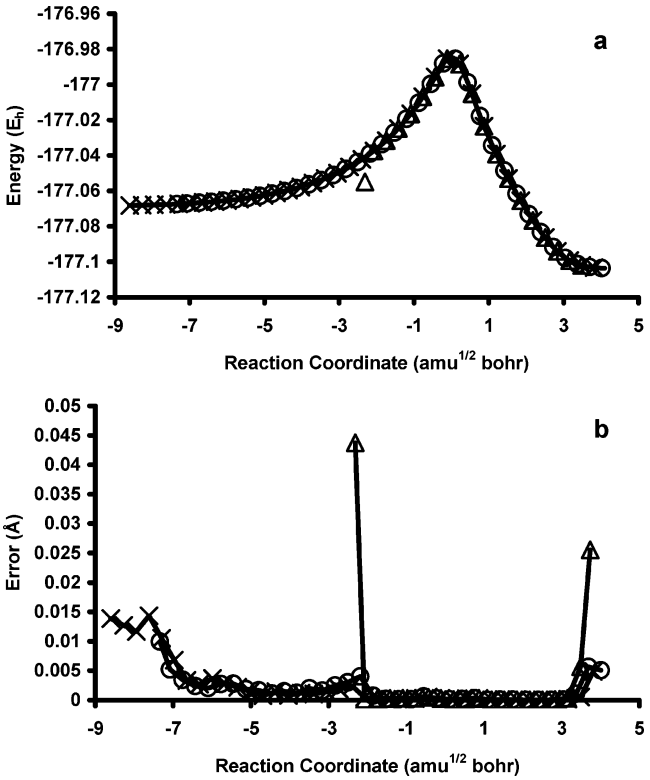


Figure 3. Reaction path for $CH_3CH_2F \rightarrow CH_2CH_2 + HF$ with all analytic Hessians vs all updated Hessians using a step size of 0.10 bohr. (a) Energy profile and (b) error, given by perpendicular distance from the all analytic Hessian path, vs reaction coordinate [HPC paths using analytic Hessians (—), MS updated Hessians (△), PSB updated Hessians (○), Bofill updated Hessians (×)]. Every third point is shown for simplicity.

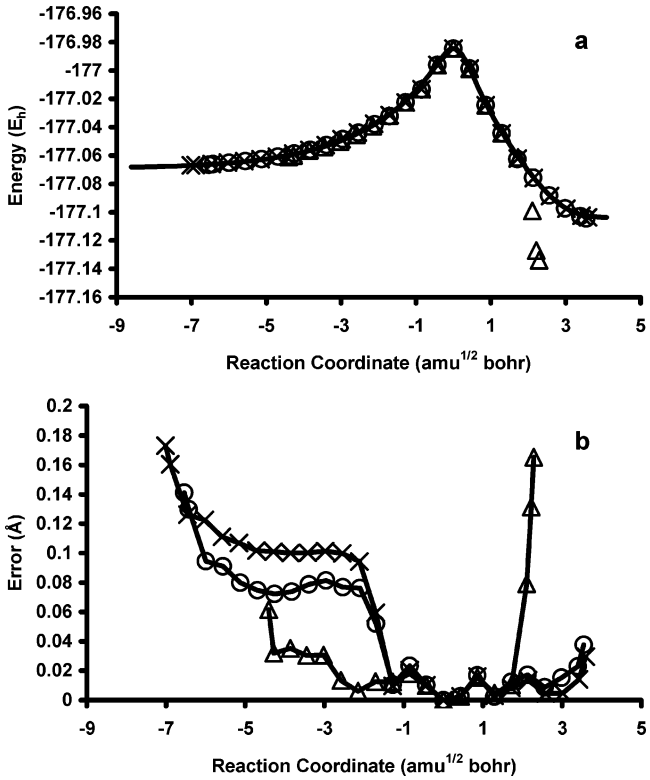


Figure 4. Reaction path for $CH_3CH_2F \rightarrow CH_2CH_2 + HF$ with all analytic Hessians vs all updated Hessians using a step size of 0.40 bohr. (a) Energy profile and (b) error, given by perpendicular distance from the all analytic Hessian path, vs reaction coordinate [HPC paths using analytic Hessians (—), MS updated Hessians (△), PSB updated Hessians (○), Bofill updated Hessians (×)].

Specifically, we have found that updating at least once every five steps provides very good agreement with the all analytic Hessian pathways. One could also recompute the Hessian whenever the rms change in matrix elements is larger than a threshold.[42] Both options have been included in this implementation of the method. Nevertheless, PSB and Bofill updating methods work well enough for the reactions considered in this paper that no analytic Hessians are required during the course of the IRC integration, except at the TS.

Although Figures 1–4 provide a qualitative measure of the accuracy, a quantitative measure of the accuracy of each updating scheme is desirable. To this end, we have computed the perpendicular distance between points on the updated Hessian and analytic Hessian pathways, which are reported in Table 1.

The inferiority of MS updating is clear from the data provided in Table 1. Indeed, for the HCN rearrangement reaction (using the smaller step size) the errors in the paths are roughly 2 orders of magnitude greater for MS than for Bofill and PSB [see Figure 1(c)]. For the four center elimination [see Figure 3(b)] and Diels–Alder reactions, all three updating schemes essentially perform the same, while the symmetric S_N2 reaction has RMS errors in position that

Table 2. Comparison of Relative CPU Time for HPC Calculations Using All Analytic Hessians vs Those Using All Updated Hessians[a]

reaction	no. of basis functions	relative time for all analytic Hessians	relative time for all updated Hessians
HNC → HCN	11	1.00	0.66
$CH_3CH_2F \rightarrow CH_2CH_2 + HF$	37	1.46	0.86
$ClCH_3 + Cl^- \rightarrow Cl^- + CH_3Cl$	59	5.51	1.77
Diels−Alder	110	61.12	7.41
CpCoMeNO → CpCoNMeO	388	5104.95	533.54

[a] All cases start with an analytic Hessian at the transition state.

considerably larger when MS and PSB updating are used as compared to Bofill updating for this step size. For the larger step size, the errors from paths using all three updating methods are much larger than the cases where the smaller step size is employed in the integration. These differences in errors are approximately 1 and 2 orders of magnitude. Additionally, the superiority of PSB and Bofill updating schemes over MS is not, in general, as clear when a larger integration step is used. In fact, for the four center elimination reaction the MS path is better than the PSB or Bofill paths, although the differences in errors for each path are small.

It is worth noting that the path errors due to Hessian updating (Table 1) are only slightly larger than the path errors previously measured for paths using all analytic Hessians at similar step sizes.[24] This indicates that using all updated Hessians only slightly decreases the accuracy of the IRC and that HPC with Hessian updating is capable of integrating the IRC using a step size that is similar to GS2. With Hessian updating, HPC requires only one energy and one gradient evaluation per step, making HPC very competitive with GS2 and other popular IRC integrators.

Since the data suggest that updating using Bofill's formula is, for most applications, superior to the other methods considered, the Bofill updating scheme has been chosen as the default scheme for HPC with Hessian updating. As previously stated, the impetus for this work is to increase the efficiency of HPC to allow its use in studies of large and interesting systems. Therefore, the relative speed-up due to updating (using the Bofill updating scheme) has been investigated. Table 2 shows relative CPU times for calculating 40 steps in the forward direction and 40 steps in the reverse direction for the HCN rearrangement, four center elimination, symmetric S_N2, and Diels−Alder reactions. The choice of the number of steps is completely arbitrary and is kept uniform among the reactions considered to provide for better cost comparison and determination of basis set size dependence. Table 2 also shows the relative cost for the insertion of NO into a Co−CH_3 σ bond, which has been previously studied by Hall and co-workers.[60−62] In this work, we have employed the B3LYP/6-311G(d,p) level of theory.[63−66] Due to the size of this system we have only calculated paths with 25 steps in the forward direction and 25 steps in the reverse direction.

As expected, updating the Hessian provides large decreases in computational cost. The degree of speed-up afforded by Hessian updating is dependent on the size of the system studied. This is expected since the bottleneck in the all analytic Hessian calculations is Hessian evaluation and the bottleneck in the updated Hessian calculations is the computation of the energy and gradient. Therefore, the speed-up due to Hessian updating will depend on the differential between the energy and Hessian calculations. For the HCN rearrangement reaction (11 basis functions) the decrease in cost for the calculation with all updated Hessians is roughly 34%, while using updated Hessians for the NO insertion reaction—the largest system studied here (388 basis functions)—provides nearly an order of magnitude speed-up.

4. Conclusions

In this paper, the efficiency of our Hessian predictor-corrector reaction path following algorithm has been improved by utilizing Hessian updating, making it an attractive alternative to other commonly used IRC integrators. In every case considered, the Hessian has been calculated analytically only at the TS. All subsequent steps have used updated Hessians. The relative performance of three popular Hessian updating schemes was investigated, and it was determined that the best method is the Bofill update. Using this updating approach, the relative CPU speed-up for HPC calculations on five different reactions was studied. For the smallest system considered, HCN rearrangement, a relative speed-up of roughly a factor of 1.3 was observed. For the largest system studied, NO insertion into a Co−CH_3 σ bond, nearly an order of magnitude speed-up was accomplished.

Acknowledgment. H.P.H. thanks the Institute for Scientific Computing at Wayne State University for support provided by a NSF-IGERT Fellowship. This work was supported by a grant from the NSF (CHE 0131157).

References

(1) Garrett, B. C.; Truhlar, D. G. 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; pp 3094−3104.

(2) Truhlar, D. G.; Garrett, B. C. *Annu. Rev. Phys. Chem.* **1984**, *35*, 159−189.

(3) Truhlar, D. G.; Garrett, B. C.; Klippenstein, S. J. *J. Phys. Chem.* **1996**, *100* (31), 12771−12800.

(4) Miller, W. H.; Handy, N. C.; Adams, J. E. *J. Chem. Phys.* **1980**, *72* (1), 99−112.

(5) Kraka, E. 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; pp 2437−2463.

(6) Fukui, K. *Acc. Chem. Res.* **1981**, *14* (12), 363−368.

(7) McKee, M. L.; Page, M. *Rev. Comp. Chem.* **1993**, 4, 35−65.

(8) Schlegel, H. B. In *Modern Electronic Structure Theory*; Yarkony, D. R., Ed.; World Scientific: Singapore, 1995.

(9) Schlegel, H. B. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., III, Schreiner, P. R., Eds.; Wiley: Chichester, 1998; pp 2432−2437.

(10) Collins, M. A. *Adv. Chem. Phys.* **1996**, *93*, 389−453.

(11) Schmidt, M. W.; Gordon, M. S.; Dupuis, M. *J. Am. Chem. Soc.* **1985**, *107*, 2585−2589.

(12) Ishida, K.; Morokuma, K.; Komornicki, A. *J. Chem. Phys.* **1977**, *66*, 2153-2156.

(13) Baldridge, K. K.; Gordon, M. S.; Steckler, R.; Truhlar, D. G. *J. Phys. Chem.* **1989**, *93*, 5107−5119.

(14) Garrett, B. C.; Redmon, M. J.; Steckler, R.; Truhlar, D. G.; Baldridge, K. K.; Bartol, D.; Schidt, M. W.; Gordon, M. S. *J. Phys. Chem.* **1988**, *92*, 1476−1488.

(15) Page, M.; McIver, J. M. *J. Chem. Phys.* **1988**, *88* (2), 922−935.

(16) Page, M.; Doubleday: C.; McIver, J. W. *J. Chem. Phys.* **1990**, *93* (8), 5634−5642.

(17) Müller, K.; Brown, L. D. *Theor. Chim. Acta* **1979**, *53*, 75−93.

(18) Sun, J. Q.; Ruedenberg, K. *J. Chem. Phys.* **1993**, *99*, 5269−5275.

(19) Gonzalez, C.; Schlegel, H. B. *J. Chem. Phys.* **1989**, *90* (4), 2154−2161.

(20) Gonzalez, C.; Schlegel, H. B. *J. Phys. Chem.* **1990**, *94* (14), 5523−5527.

(21) Gonzalez, C.; Schlegel, H. B. *J. Chem. Phys.* **1991**, *95* (8), 5853−5860.

(22) Schlegel, H. B. *J. Comp. Chem.* **2003**, *24*, 1514−1527.

(23) Gear, C. W. *Numerical Initial Value Problems in Ordinary Differential Equations*; Prentice Hall: Englewood Cliffs, 1971.

(24) Hratchian, H. P.; Schlegel, H. B. *J. Chem. Phys.* **2004**, *120* (21), 9918−9924.

(25) Bulirsch, R.; Stoer, J. *Num. Math.* **1964**, *6*, 413−427.

(26) Bulirsch, R.; Stoer, J. *Num. Math.* **1966**, *8*, 1−13.

(27) Bulirsch, R.; Stoer, J. *Num. Math.* **1966B**, *8*, 93−104.

(28) Farwig, R. *Math Comput*. **1986**, *46*, 577.

(29) Farwig, R. In *Algorithms for approximation*; Mason, J., Cox, M., Eds.: Clarendon, Oxford, 1987; p 194.

(30) Collins, M. A. *Theor. Chem. Acc.* **2002**, *108* (6), 313−324.

(31) Hoffman, D. K.; Nord, R. S.; Ruedenberg, K. *Theor. Chim. Acta* **1986**, *69* (4), 265−279.

(32) Press, W. H. *Numerical recipes in FORTRAN 77: the art of scientific computing*; Cambridge University Press: Cambridge [England]; New York, 1996.

(33) Millam, J. M.; Bakken, V.; Chen, W.; Hase, W. L.; Schlegel, H. B. *J. Chem. Phys.* **1999**, *111* (9), 3800−3805.

(34) Bakken, V.; Millam, J. M.; Schlegel, H. B. *J. Chem. Phys.* **1999**, *111* (19), 8773−8777.

(35) Bettens, R. P. A.; Collins, M. A. *J. Chem. Phys.* **1999**, *111* (3), 816−826.

(36) Thompson, K. C.; Jordan, M. J. T.; Collins, M. A. *J. Chem. Phys.* **1998**, *108* (2), 564−578.

(37) Ischtwan, J.; Collins, M. A. *J. Chem. Phys.* **1994**, *100* (11), 8080−8088.

(38) Lancaster, P.; Salkauskas, K. *Curve and Surface Fitting: An Introduction*; Academic Press: London; Orlando, 1986.

(39) Schlegel, H. B. *J. Comp. Chem.* **1982**, *3* (2), 214.

(40) Schlegel, H. B. 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; pp 1136−1142.

(41) Schlick, T. 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; pp 1142−1157.

(42) Eckert, F.; Werner, H. J. *Theor. Chem. Acc.* **1998**, *100* (1−4), 21−30.

(43) Dennis, J. E.; Schnabel, R. B. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*; Prentice Hall: Englewood Cliffs, NJ, 1983.

(44) Fletcher, R. *Practical Methods of Optimization*; Wiley: Chichester, 1987.

(45) Anglada, J. M.; Bofill, J. M. *J. Comput. Chem.* **1998**, *19* (3), 349−362.

(46) Bofill, J. M. *J. Comput. Chem.* **1994**, *15* (1), 1−11.

(47) Farkas, Ö.; Schlegel, H. B. *J. Chem. Phys.* **1999**, *111* (24), 10806−10814.

(48) Farkas, Ö.; Schlegel, H. B. *Phys. Chem. Chem. Phys.* **2002**, *4* (1), 11−15.

(49) 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.: Pittsburgh, PA, 2003.

(50) Collins, J. B.; Schleyer, P. v. R.; Binkley, J. S.; Pople, J. A. *J. Chem. Phys.* **1976**, *64*, 5142.

(51) Hehre, W. J.; Stewart, R. F.; Pople, J. A. *J. Chem. Phys.* **1969**, *51*, 2657.

(52) Binkley, J. S.; Pople, J. A.; Hehre, W. J. *J. Am. Chem. Soc.* **1980**, *102*, 939.

(53) Gordon, M. S.; Binkley, J. S.; Pople, J. A.; Pietro, W. J.; Hehre, W. J. *J. Am. Chem. Soc.* **1982**, *104*, 2797.

(54) Pietro, W. J.; Francl, M. M.; Hehre, W. J.; Defrees, D. J.; Pople, J. A.; Binkley, J. S. *J. Am. Chem. Soc.* **1982**, *104*, 5039.

(55) Ditchfield, R.; Hehre, W. R.; Pople, J. A. *J. Chem. Phys.* **1971**, *54*, 724.

(56) Gordon, M. S. *Chem. Phys. Lett.* **1980**, *76*, 163.

(57) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213.

(58) Hariharan, P. C.; Pople, J. A. *Mol. Phys.* **1974**, *27*, 209.

(59) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257.

(60) Niu, S. Q.; Hall, M. B. *Chem. Rev.* **2000**, *100* (2), 353−405.

(61) Niu, S. Q.; Hall, M. B. *J. Am. Chem. Soc.* **1997**, *119* (13), 3077−3086.

(62) Niu, S. Q.; Hall, M. B. *J. Phys. Chem. A* **1997**, *101* (7), 1360−1365.

(63) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37* (2), 785−789.

(64) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(65) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650.

(66) McLean, A. D.; Chandler, G. S. *J. Chem. Phys.* **1980**, *72*, 5639.

CT0499783

JCTC Journal of Chemical Theory and Computation

Benchmarking the Conductor-like Polarizable Continuum Model (CPCM) for Aqueous Solvation Free Energies of Neutral and Ionic Organic Molecules

Yu Takano and K. N. Houk*

Department of Chemistry and Biochemistry, University of California, Los Angeles, 607 Charles E. Young Drive East, Los Angeles, California 90095-1569

Received June 17, 2004

Abstract: The conductor-like polarizable continuum model (CPCM) using several cavity models is applied to compute aqueous solvation free energies for a number of organic molecules (30 neutral molecules, 21 anions, and 19 cations). The calculated solvation free energies are compared to the available experimental data from the viewpoint of cavity models, computational methods, calculation time, and aqueous pK_a values. The HF/6-31+G(d)//HF/6-31+G(d) and the HF/6-31+G(d)//B3LYP/6-31+G(d) with the UAKS cavities, in which radii are optimized with PBE0/6-31G(d), provide aqueous solvation effects in best agreement with available experimental data. The mean absolute deviations from experiment are 2.6 kcal/mol. The MP2/6-31++G-(d,p)//HF/6-31+G(d) with the CPCM-UAKS(HF/6-31+G(d)) calculation is also performed for the base-catalyzed hydrolysis of methyl acetate in water.

Introduction

Many chemical and biological reactions occur in water, where the polar and ionic processes are much more favorable than in the gas phase. Many efforts have been devoted to the development of methods to compute reaction barriers and energetics occurring in condensed phases with experimental accuracy.[1] Effective explicit water models become available for the description of chemical systems in liquid solution.[1] However, with high-level quantum mechanics, only a limited number of solvent molecules can be included explicitly due to the high cost of the calculations.

The goal of this work is to determine which theoretical procedure provides the most quantitative estimate of aqueous solvation effects, so that the rates of chemical and biological reactions in water can be computed accurately. One of the most successful solvation models is the conductor-like polarizable continuum model (CPCM).[2] Here we benchmark different variations of CPCM for the computation of solvation energies of neutral and ionic organic species and compare them to several other works. The CPCM method has also been applied to the computation of the alkaline hydrolysis of methyl acetate in aqueous solution.

* Corresponding author e-mail: houk@chem.ucla.edu.

Background

Dielectric continuum theories[1] are now widely used to describe hydration in conjunction with quantum mechanical calculations due to the relatively low cost of the calculation. CPCM[2] and PCM[3] are two of many successful solvation models. In their approaches, the solute interacts with the solvent represented by a dielectric continuum model. The solute molecule is embedded into a cavity surrounded by a dielectric continuum of permittivity ϵ. The accuracy of continuum solvation models depends on several factors; the most important one is the use of proper boundary conditions on the surface of the cavity containing the solute. CPCM and PCM define the cavities as envelopes of spheres centered on atoms or atomic groups: a number of cavity models have been suggested. Inside the cavity the dielectric constant is the same as in vacuo, outside it takes the value of the desired solvent. Once the cavity has been defined, the surface is smoothly mapped by small regions, called tesserae. Each tessera is characterized by the position of its center, its area, and the electrostatic vector normal to the surface passing through its center. Recently, the CPCM method has been improved and extended in GAUSSIAN03[4a] so that the cavity can be selected in a number of different ways.

10.1021/ct049977a CCC: $30.25
Published on Web 11/13/2004

In CPCM, the solvation free energy can be expressed[1a]

$$\Delta G_{solv} = \Delta G_{el} + \Delta G_{cav} + \Delta G_{dis} + \Delta G_{rep} + RT \ln\left(\frac{q_{rot,g} q_{vib,g}}{q_{rot,s} q_{vib,s}}\right) - RT \ln\left(\frac{n_{solute,g} \Lambda_{solute,g}}{n_{solute,s} \Lambda_{solute,s}}\right) + P\Delta V \quad (1)$$

ΔG_{el} is the electrostatic component of ΔG_{solv}. The G_{el} term is calculated using the CPCM self-consistent reaction field (SCRF) method.[2] The cavitation term, ΔG_{cav}, is calculated with the expression derived by Pierotti from the hard sphere theory[5] and adapted to the case of nonspherical cavities.[3b] The dispersion and repulsion terms, ΔG_{dis} and ΔG_{rep}, are computed following Floris and Tomasi's procedure,[6] with the parameters proposed by Callet and Claverie.[7] The $q_{rot,g}$, $q_{vib,g}$, $q_{rot,s}$, and $q_{vib,s}$ are denoted the microscopic partition functions for rotational and vibrational states of the solute in the gas phase and in solution, respectively; $n_{solute,g}$ and $n_{solute,s}$ are the numeral densities of solute; and $\Lambda_{solute,g}$ and $\Lambda_{solute,s}$ are the momentum partition functions. The last term, $P\Delta V$, may be neglected since its value is normally less than 10^{-3} kcal/mol.[8] The quantity, $-RT \ln(n_{solute}\Lambda_{solute})$, is a free energy correction to account for solute occupying the entire volume available in the reference state. For simple models such as isotropic solutions with no chemical association or dissociation processes, this contribution is equal to zero. The term involving the vibrational and rotational degrees of freedom, $\ln(q_{vib,g}/q_{vib,s})$ and $\ln(q_{rot,g}/q_{rot,s})$, is negligible.[1a]

The last three terms in eq 1 are neglected in the PCM and CPCM formulations.[1a] It, however, has been noted that things are actually more complex when one considers dimers or trimers held together by relatively weak interactions and for chemical association or dissociation processes.[1a]

Computational Method

The CPCM-SCRF calculations[2] at the HF/6-31+G(d) and B3LYP/6-31+G(d) levels were carried out on the stationary points to address solvation effects. A dielectric constant of 78.39 was utilized in order to simulate aqueous environment. The CPCM calculations were performed with tesserae of 0.2 Å^2 average size. All structures were optimized at the HF/6-31+G(d) and B3LYP/6-31+G(d) levels[9−11] in the gas phase and at the B3LYP/6-31+G(d) level in the aqueous environment. All stationary points were characterized by frequency calculations at the same level. All calculations were carried out with GAUSSIAN03[4a] and GAUSSIAN98.[4b] We used 1 mol L^{-1} as the standard state for both the gas phase and the solution for all thermodynamic properties.

In CPCM and PCM, the choice of cavities is important because the computed energies and properties depend on the cavity size. In this study, the UA0, UAHF, UAKS, UFF, PAULING, and BONDI cavities were used to evaluate the aqueous solvation effects using CPCM and PCM. The UA0 cavity is built up using the united atom topological model (UATM)[8a] applied on atomic radii of the universal force field (UFF).[9] By default, the UA0 model is chosen to build the cavity in GAUSSIAN03.[14] The UAHF and UAKS cavities use UATM with radii optimized for the HF/6-31G(d) and PBE0/6-31G(d)[10] levels of theory, respectively. The UAHF model is the default cavity of GAUSSIAN98. A set of the radii from UFF is used for making the UFF cavities. For the PAULING and BONDI cavities, each solute atom and group is assigned van der Waals values obtained from Pauling[12b] or Bondi[12c] atomic radii.

Table 1. Mean Absolute Deviations (MADs) of the Aqueous Solvation Free Energies of 70 Neutral and Charged Molecules at the HF/6-31+G(d)//B3LYP/6-31+G(d) Level Using CPCM with Several Cavities[a]

	total	neutral	anion	cation
UAKS	2.61	1.35	3.21	3.93
UAHF(G03)	2.84	1.10	3.92	4.93
UAHF(G98)	2.95	1.43	3.86	4.32
BONDI	3.64	2.93	3.38	5.04
PAULING	3.67	3.49	2.73	4.98
UA0	7.62	2.75	9.64	13.09
UFF	8.10	2.82	9.30	15.12

[a] MADs are shown in kcal/mol.

Test of the Reliability of CPCM

Aqueous solvation free energies for a number of organic molecules (30 neutrals, 21 anions, and 19 cations) computed using the CPCM method.[2] We investigated the calculated solvation free energies compared to the available experimental data[16] from the viewpoint of cavities, computational methods, calculation time, and aqueous pK_a values. The computed aqueous solvation effects were also compared with solvation energies computed using COSMO,[2] cluster-continuum model,[17] SM5.42R,[18] PCM,[3] and IPCM methods.[19]

Cavity Models. Table 1 summarizes the mean absolute deviations (MADs) of the aqueous solvation free energies calculated with seven cavities at the HF/6-31+G(d)//B3LYP/6-31+G(d) level[9−11] from the experimental data[16] for 70 neutral and charged molecules.

CPCM with the new cavities, UAKS and UAHF(G03), has improved accuracy in the aqueous solvation energies for a set of 70 organic species, though these methods still do not achieve the accuracy of experimental data (0.2 kcal/mol for neutral molecules[20,21] and 2 kcal/mol for ions[16]). The MADs calculated by the CPCM-UAKS and CPCM-UAHF(G03) are 2.61 kcal/mol (1.35, 3.21, and 3.93 kcal/mol for 30 neutrals, 21 anions, and 19 cations, respectively) and 2.84 kcal/mol (1.10, 3.92, and 4.93 kcal/mol), respectively. On the other hand, the CPCM-UA0 and CPCM-UFF methods fail for charged molecules with MADs of 9.64 (13.09) and 9.30 (15.12) kcal/mol for anion (cation) solutes, respectively. The PAULING cavities show the best solvation free energies (2.73 kcal/mol) for anion molecules but give the worst agreement with experiment for neutrals. For all cavities but the PAULING cavities, the calculated solvation free energies for the neutral molecules are much closer to the experimental results than those for the charged species. The large solvation energy errors for the ions is due to inadequate treatment of specific short-range interactions, probably associated with strong hydrogen bonds between the ions and first-shell water molecules. Dielectric continuum theory[1] cannot account for short-range solute−solvent interactions such as hydrogen bond. In addition, since anions and cations have aqueous solvation free energies in the range of 60−110 kcal/mol in

Table 2. Mean Absolute Deviations (MADs) of the Aqueous Solvation Free Energies of 70 Neutral and Charged Molecules at the HF/6-31+G(d)//B3LYP/6-31+G(d), B3LYP/6-31+G(d)//B3LYP/6-31+G(d), and HF/6-31+G(d)//HF/6-31+G(d) Levels Using CPCM with the UAKS Cavities[a]

	HF//HF	HF//B3LYP	B3LYP//B3LYP
total	2.56	2.61	3.32
neutral	1.14	1.35	0.88
anion	3.10	3.21	5.64
cation	4.34	3.93	4.32

[a] MADs are shown in kcal/mol.

Table 3. Mean Absolute Deviations (MADs) of the Aqueous Solvation Free Energies of 70 Neutral and Charged Molecules at the HF/6-31+G(d)//B3LYP/6-31+G(d) Level Using CPCM with the UAKS Cavities[a,b]

	water	vacuo
total	2.60	2.61
neutral	1.79	1.35
anion	3.00	3.21
cation	3.60	3.93

[a] MADs are shown in kcal/mol. [b] The geometries were optimized in the water environment and in vacuo.

contrast to neutral molecules (0−10 kcal/mol), it is hard to achieve the 1 kcal/mol level of accuracy in prediction of the solvation free energy of ions.

Computational Methods. Table 2 shows the computational method dependence of MADs of the solvation free energy for the neutral and charged solutes. The geometries were optimized in the gas phase, and the UAKS cavities were used. The MADs of the solvation free energies at the HF/6-31+G(d)//B3LYP/6-31+G(d) level become 0.7 kcal/mol smaller than those at the B3LYP/6-31+G(d)//B3LYP/6-31+G(d) level. Especially, MADs of the solvation energies for the anion solutes are improved by 2.43 kcal/mol. There is a tendency that the cavity that reduces MADs of charged species increases the MADs of the neutral species. Optimization at the HF level provides MADs similar to those at the B3LYP level, indicating that optimized geometries using HF and B3LYP are close to each other.

Calculated MADs of the neutral and charged solutes with the geometries optimized in water are shown in Table 3. The geometries optimized in water make the MADs of charged species small while those for neutrals large. As a whole, optimization in water shows MADs similar to those in vacuo, implying that the optimized structures in water are similar to those in vacuo. In Table 4, the optimized geometrical parameters for some examples in vacuo and water are reported. It is apparent that the effect of reoptimization in water on the geometrical parameters are very small. However, some charged species failed to optimize geometries in water due to the dissociation of a proton, especially using the UAKS, UAHF, and PAULING cavities (Tables S10−S15).

The 6-31G, 6-31G(d), 6-31+G(d), 6-31+G(d,p), and 6-311+G(2d,p) basis sets were utilized to investigate the dependence on basis sets for aqueous solvation free energies. Even when basis sets are enlarged up to 6-311+G(2d,p), the MAD values from the experiment are very similar as shown in Table 5, indicating that diffuse and polarization functions of basis sets hardly change the aqueous solvation effects.

Calculation Time. Figure 1A shows the CPU time required to calculate the hydration energy using the previous (GAUSSIAN98)[4b] and the present (GAUSSIAN03)[4a] versions of the code. The present version provides a remarkable decrease of computational time because of the introduction of the fast multipole method to compute the solvation charge.

In Figure 1B, we compare CPCM to PCM with respect to the CPU time for the estimate of the aqueous solvation free energies in water. Although PCM is faster than CPCM for small molecules, CPCM is faster when the molecules become larger. In the CPCM approach, the electrostatic problem related to solute−solvent interaction can be solved with a

Table 4. Optimized Geometrical Parameters and Aqueous Solvation Free Energies for Some Sample Molecules in Vacuo and in Water

	CH_3O^-			CH_3OH			$CH_3NH_3^+$	
	vacuo	water		vacuo	water		vacuo	water
ΔG_{solv}[a]	−87.69	−88.53	ΔG_{solv}[a]	−6.23	−6.43	ΔG_{solv}[a]	−71.02	−72.59
R(C−O)[b]	1.34	1.40	R(C−O)[b]	1.43	1.43	R(C−N)[b]	1.52	1.50
R(O−H)[b]	1.14	1.11	R(C−H)[b]	1.10	1.10	R(C−H)[b]	1.09	1.09
∠COH[c]	115.6	113.0	R(O−H)[b]	0.97	0.98	R(N−H)[b]	1.03	1.04
			∠COH[c]	109.0	108.8	∠CNH[c]	111.6	111.6
			∠OCH[c]	112.0	111.6	∠NCH[c]	108.2	108.2

[a] ΔG_{solv}s are shown in kcal/mol. [b] Bond lengths are shown in Å. [c] Angles are shown in degrees.

Table 5. Mean Absolute Deviations (MADs) for the Aqueous Solvation Free Energies of 70 Neutral and Charged Molecules at the HF and B3LYP Levels Using CPCM-UAKS with Five Different Basis Sets[a]

		total	neutral	anion	cation			total	neutral	anion	cation
HF	6-31G	2.86	1.74	3.66	3.76	B3LYP	6-31G	2.95	0.91	4.63	4.31
	6-31G(d)	2.78	0.93	4.08	4.28		6-31G(d)	3.17	0.68	5.16	4.90
	6-31+G(d)	2.61	1.35	3.21	3.93		6-31+G(d)	3.32	0.88	5.64	4.32
	6-31+G(d,p)	2.56	1.38	3.19	3.72		6-31+G(d,p)	3.23	0.87	5.74	4.19
	6-311+G(2d,p)	2.77	1.49	3.52	3.95		6-311+G(2d,p)	3.28	0.75	5.92	4.37

[a] The geometries were optimized in vacuo at the B3LYP/6-31+G(d) level.

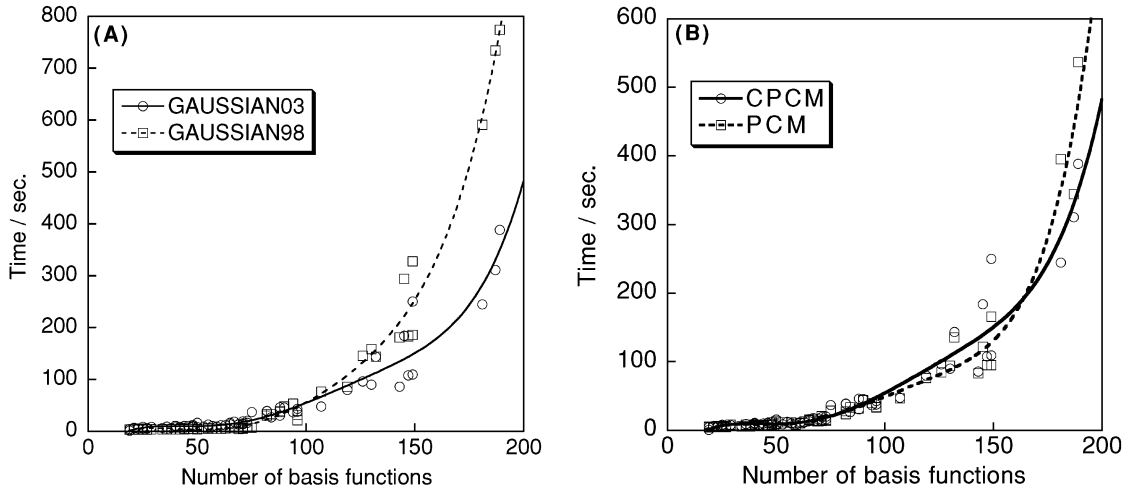


Figure 1. CPU times for the calculation of the aqueous solvation free energies at the HF/6-31+G(d) level by CPCM-UAKS using GAUSSIAN03 and GAUSSIAN98 (A) and by the CPCM-UAKS and PCM-UAKS methods using GAUSSIAN03 (B).

much simpler formalism than in PCM. This simpler formalism can be a faster estimate of the aqueous solvation free energies when larger systems are studied.

Aqueous pK_a. Aqueous pK_a values were evaluated using the B3LYP/6-31+G(d)//B3LYP/6-31+G(d) level of theory and inclusion of solvent effects at the HF/6-31+G(d) and B3LYP/6-31+G(d) levels. At present, many theoretical pK_a predictions have been reported.[17b,22] The vast majority of pK_a calculations use the direct definition shown in eq 2 with combination of the experimental data of the proton.[1c,22a−f]

$$HA \rightarrow H^+ + A^- \tag{2}$$

Since this reaction, however, involves the formation of charged species starting from neutral molecules, the procedure using eq 2 yields large errors due to imbalance of the charged and neutral species and requires the value of the experimental solvation free energy of the H^+ ion. By comparison, reactions that conserve the number of charged species are more suitable for accurate calculations of changes in solvation free energies. Pliego and Riveros used the proton-transfer reaction between the HA acids and hydroxide anion as shown in eq 3 along with the continuum-cluster method[17a] and compared their calculated values with the pK_a values calculated with the other solvation models. Since our objective is the benchmarking of the CPCM method, we made use of the reported results by Pliego and Riveros for comparison of the CPCM-UAKS and performed the pK_a calculations according to their procedure.[17b,22g,h]

$$HA + OH^- \rightarrow A^- + H_2O\ \Delta G^*_{sol} \tag{3}$$

The reaction (eq 3) conserves the number of charged species. The pK_a values were calculated according to eqs 4 and 5

$$pK_a(HA) = \frac{\Delta G^*_{sol}}{(2.303)RT} + 15.74 \tag{4}$$

$$\Delta G^*_{sol} = \Delta G^*_g + \Delta G^*_{solv}(A^-) + \Delta G^*_{solv}(H_2O) - \Delta G^*_{solv}(OH^-) - \Delta G^*_{solv}(HA) \tag{5}$$

where ΔG^*_g is the gas-phase reaction free energy (1 mol L^{-1} standard state) relative to eq 5, and $\Delta G^*_{solv}(X)$ is the solvation free energy of species X according to the Ben-Naim definition.[23]

The proton-transfer reaction as shown in eq 6 was used for the ionization of BH^+ acids

$$BH^+ + H_2O \rightarrow B + H_3O^+\ \Delta G^*_{sol} \tag{6}$$

The pK_a values for BH^+ acids can be expressed as

$$pK_a(BH^+) = \frac{\Delta G^*_{sol}}{(2.303)RT} - 1.74 \tag{7}$$

$$\Delta G^*_{sol} = \Delta G^*_g + \Delta G^*_{solv}(B) + \Delta G^*_{solv}(H_3O^+) - \Delta G^*_{solv}(H_2O) - \Delta G^*_{solv}(BH^+) \tag{8}$$

The calculated aqueous pK_a values for 27 species are listed in Table 6. CPCM provides MADs of the aqueous pK_a values of 2.83 and 2.47 pK_a units at the HF/6-31+G(d) and B3LYP/6-31+G(d) levels, respectively. The largest error occurs for $(CH_3COOC_2H_5)H^+$ and amounts to 6.32 and 6.55 pK_a units. The pK_a values estimated at the B3LYP/6-31+G(d) level are closer to experiment[16] than at the HF/6-31+G(d) level in contrast to the aqueous solvation free energy, indicating that accurate pK_a calculations require accurate gas-phase proton-transfer energies and solvation free energies.

Comparisons with Other Work. Our calculations have been also compared with other works as shown in Tables 7 and 8. We used the calculated data by Pliego and Riveros,[17] except for the CPCM[2] and COSMO[2] results. MADs of aqueous solvation free energies for 13 charged species are listed in Table 7. The higher accuracy of CPCM is observed and leads to a MAD of 3.04 kcal/mol, whereas the cluster-continuum,[17a] the COSMO,[2] the SM5.24R,[18] and the PCM methods[3] have MADs of about 10 kcal/mol. The IPCM method[19] shows the worst MAD of almost 20 kcal/mol. The failure of the cluster-continuum model can be traced to the continuum part of the calculation, specifically, the IPCM method. Improper cavities result in much larger MADs for the COSMO and PCM methods than for the CPCM method with appropriate cavity model, UAKS. PCM-UAKS provides

Table 6. Calculated and Experimental pK_a's in Water

compounds	pK_a(calc.) B3LYP[a]	pK_a(calc.) HF[b]	pK_a(exp.)	compounds	pK_a(calc.) B3LYP[a]	pK_a(calc.) HF[b]	pK_a(exp.)
CH_3CH_2O-**H**	21.80	20.93	15.90	H_2P-**H**	23.86	26.29	27
CH_3COO-**H**	5.47	6.62	4.76	C_6H_5O-**H**	12.63	13.10	9.99
CH_3O-**H**	21.12	21.54	15.50	C_6H_5S-**H**	7.31	8.18	6.62
H-Br	−12.10	−9.96	−8	CH_3CH_2OH-$\mathbf{H}^+$	−0.07	−0.38	−1.94
HCC-**H**	23.68	25.83	21.70	$(CH_3)_3N$-$\mathbf{H}^+$	11.69	11.36	9.80
HCOO-**H**	3.24	4.75	3.75	$(CH_3)_2C{=}O$-$\mathbf{H}^+$	−1.09	−1.61	−3.06
H-CN	10.28	12.51	9.22	$(CH_3COOC_2H_5)$-$\mathbf{H}^+$	−10.93	−11.16	−4.61
H-Cl	−10.79	−8.71	−6.10	$(CH_3)_2NH$-$\mathbf{H}^+$	11.89	11.58	10.73
H-F	0.68	1.40	3.18	CH_3NH_2-$\mathbf{H}^+$	10.37	10.09	10.66
H-OBr	11.37	14.93	8.60	CH_3OH-$\mathbf{H}^+$	−0.90	−1.45	−2.05
H-OCl	10.21	13.78	7.54	$(CH_3)_2O$-$\mathbf{H}^+$	−5.60	−6.13	−2.48
HO-**H**	15.74	15.84	15.74	H_3N-$\mathbf{H}^+$	9.17	8.99	9.25
HOO-**H**	16.35	19.20	11.65	$(C_6H_5COCH_3)$-$\mathbf{H}^+$	−8.32	−8.53	−3.87
HS-**H**	6.38	8.48	7.05	MAD	2.47	2.83	

[a] Electronic and aqueous solvation free energies are calculated at the B3LYP/6-31+G(d)//B3LYP/6-31+G(d) level. [b] Electronic energies are calculated at the B3LYP/6-31+G(d)//B3LYP/6-31+G(d) level, and aqueous solvation energies are estimated at the HF/6-31+G(d)//B3LYP/6-31+G(d) level.

Table 7. Comparison of the Aqueous Solvation Free Energy Calculated with CPCM,[a] Cluster-Continuum Model,[b] COSMO,[c] SM5.42R,[d] PCM,[e] and IPCM[f]

	CPCM	cluster−continuum	COSMO	SM5.42R	PCM	IPCM	Exp.
CH_3O^-	−87.69	−82.37	−81.22	−86.79	−79.98	−64.18	−95.2
Cl^-	−73.39	−67.94	−74.68	−77.05	−72.70	−61.80	−74.6
$HCOO^-$	−74.61	−63.77	−73.27	−75.22	−72.41	−60.81	−76.2
$C_2H_5O^-$	−84.93	−77.66	−77.61	−81.87	−76.70	−61.24	−91.1
OH^-	−106.57	−93.08	−96.29	−108.96	−92.14	−69.64	−105
$C_6H_5O^-$	−67.89	−68.60	*g*	−64.59	−63.78	−52.96	−71.3
SH^-	−70.37	−64.68	−72.12	−84.37	−71.04	−57.98	−71.6
CH_3S^-	−70.79	−63.87	−70.23	−78.97	−69.27	−56.58	−73.7
$CH_3CH_2OH_2^+$	−87.57	−75.71	−69.06	−74.04	−66.12	−66.48	−88.4
$CH_3NH_3^+$	−71.02	−72.69	−71.62	−76.03	−69.17	−68.83	−76.5
$CH_3OH_2^+$	−91.99	−84.64	−74.62	−79.82	−71.16	−72.24	−93.1
H_3O^+	−108.59	−101.87	−90.09	−92.51	−83.64	−88.77	−110.2
NH_4^+	−80.36	−79.38	−81.43	−87.03	−77.42	−77.56	−85.2
MAD	3.04	8.91	9.15	7.49	11.27	19.46	

[a] HF/6-31+G(d)//B3LYP/6-31+G(d) with UAKS cavities. [b] MP2/6-31+G(2df,2p)//HF/6-31+G(d,p). IPCM were used for the continuum part (ref 17a). [c] HF/6-31G(d)//B3LYP/6-31+G(d) with Klamt cavity model (SCRF=COSMORS). [d] HF/6-31G(d)//HF/6-31+G(d,p) (ref 17a). [e] HF/6-31G(d,p)//HF/6-31+G(d,p). The spheres defining the cavity were taken to be 1.2 times the van der Waals radii (ref 17a). [f] MP2/6-31+G(d,p)//HF/6-31+G(d,p). An isodensity of 0.0004 was used (ref 17a). [g] Not converged.

Table 8. Comparison of the Aqueous pK_a Values Calculated with CPCM,[a] Cluster-Continuum Model,[b] SM5.42R,[c] and PCM[d]

	CPCM	cluster− continuum	SM5.42R	PCM	exp.		CPCM	cluster− continuum	SM5.42R	PCM	exp.
CH_3CH_2O-**H**	21.80	16.08	25.93	16.52	15.90	C_6H_5O-**H**	12.63	7.19	16.98	4.50	9.99
CH_3COO-**H**	5.00	2.57	8.25	−3.27	4.76	CH_3CH_2OH-$\mathbf{H}^+$	−0.07	−4.16	0.17	1.74	−1.94
CH_3O-**H**	21.12	16.08	25.40	17.32	15.50	CH_3NH_2-$\mathbf{H}^+$	10.37	15.30	23.56	26.11	10.66
HCOO-**H**	2.83	1.03	4.30	−5.11	3.75	CH_3OH-$\mathbf{H}^+$	−0.90	−1.73	0.64	1.53	−2.05
H-Cl	−10.79	−9.25	−5.40	−13.47	−6.10	H_3N-$\mathbf{H}^+$	9.17	11.73	23.38	23.61	9.25
HS-**H**	6.38	5.69	2.33	0.35	7.05	MAD	2.19	2.06	6.20	6.91	

[a] B3LYP/6-31+G(d)//B3LYP/6-31+G(d) with UAKS cavities. [b] MP2/6-31+G(2df,2p)//HF/6-31+G(d,p). IPCM were used for the continuum part (ref 17b). [c] HF/6-31G(d)//HF/6-31+G(d,p) (ref 17b). [d] HF/6-31G(d,p)//HF/6-31+G(d,p). The spheres defining the cavity were taken to be 1.2 times the van der Waals radii (ref 17b).

MADs of 5.28 kcal/mol for 13 charged species. The IPCM method provides a very appealing approach to the problem of the cavity size. Unfortunately, besides the convergence problems that sometimes plague these methods, it seems that a single value for the density threshold is not adequate to reproduce experimental data, so that the problem of linking the atomic size to the molecular context remains.

Recently, Thompson, Cramer, and Truhlar presented a new continuum solvation model (SM5.43R) and compared it to the other solvation models including SM5.42R and CPCM.[20]

The MADs calculated by the SM5.43R, SM5.42, CPCM-UAHF(G03), and CPCM-UAHF(G98) were 0.51, 0.54, 1.07, and 1.08 kcal/mol for 257 neutrals and 4.65, 4.83, 4.57, and 5.90 kcal/mol for 47 ions, respectively. SM5.43R provides highly accurate aqueous solvation free energies for neutral species compared to experimental data. We cannot directly compare our calculated results to theirs, because they did not report what species were used for the calculation of the aqueous solvation energies, nor did they report the calculated aqueous solvation free energies. However, from Table 7, the MADs calculated by the CPCM-UAKS (3.04 kcal/mol) is a factor of 2.5 smaller than those calculated by the SM5.42R (7.49 kcal/mol) for 13 ions. Furthermore, they reported that the CPCM-UAHF(G98) gives smaller MADs than both the SM5.43R and SM5.42R for 47 ions. We found that a MAD calculated by CPCM-UAKS (3.5 kcal/mol) is smaller than that calculated by CPCM-UAHF(G98) (4.1 kcal/mol) for 40 ions (Table 1). These comparisons indicate that CPCM-UAKS is expected to give prediction of more accurate aqueous solvation energies for ions than SM5.42R and SM5.43R.

Table 8 lists the aqueous pK_a values for 11 species. The cluster-continuum model is the most reliable for the estimate of aqueous pK_a values and has a MAD of 2.06 pK_a units. Despite the lack of explicit water molecules, the CPCM method is also reasonable with an aqueous pK_a MAD of 2.19, but SM5.42R and PCM have much larger MADs (about 6 pK_a units). The worst computed aqueous pK_a values come from SM5.42R and PCM methods. Compared with the PCM method, the CPCM method provides much better calculated aqueous pK_a values because of the cavities used.

Application of the CPCM for Chemical Reactions

The base-catalyzed hydrolysis of methyl acetate in water[24,25] was investigated with the CPCM method. The hydrolysis of esters in basic solution stands out as one of the most studied reactions in chemistry because of its common occurrence in many organic and biochemical processes. In the base-catalyzed ester hydrolysis in water, with isolated methyl acetate and hydroxide ion (**Reactants**) as the starting materials, the approach of hydroxide ion to methyl acetate forms a tetrahedral intermediate (**Int**) through a transition state (**TS1**). Another barrier (**TS2**) is crossed to form the products, methanol and acetate ion (**Products**), via nonbarrier proton transfer.

Guthrie evaluated the free energy changes in this reaction using thermochemical and kinetic data.[24a] The calculated free energies for the five stationary points (**Reactants**, **TS1**, **Int**, **TS2**, and **Products**) have been compared to them.

The free energies for the base-catalyzed hydrolysis reaction of methyl acetate were calculated at the MP2/6-31++G(d,p)//HF/6-31+G(d) level. All the stationary points were obtained by full geometry optimizations in vacuo and characterized by harmonic frequency analysis. Zero point energies and thermal corrections at 298 K (scaled by 0.91)[26] were included in the reported energies. Solvation energies were computed using the CPCM-UAKS method at the HF/6-31+G(d) level using GAUSSIAN03.[4a] We used 1 mol L^{-1} as the standard state for both the gas phase and the solution for all thermodynamic properties.

Table 9. Calculated and Experimental Gibbs Free Energies for the Stationary Points of the Base-Catalyzed Hydrolysis of Methyl Acetate[a,b]

	ΔG(MP2)[c]	ΔG(exp.)
Reactants	0.0	0.0
TS1	21.3	18.5
Int	15.1	10.0
TS2	26.5	17.4
Products	−13.5	−14.4

[a] All Gibbs free energies relative to the isolated reactants. [b] All Gibbs free energies are shown in kcal/mol. Standard state of 1 mol L^{-1} for both gas phase and water phase thermodynamic properties. Correction factors ($-RT$ ln 2) were included for the enantiomeric minima and the transiton state structures. [c] MP2/6-31++G(d,p)//HF/6-31+G(d).

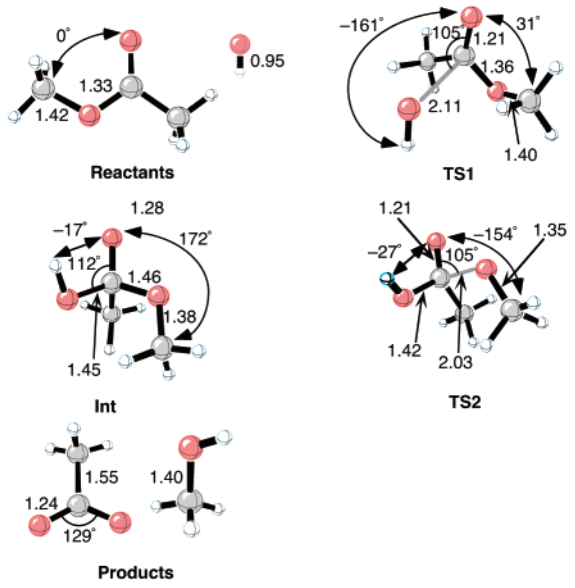


Figure 2. HF/6-31+G(d) geometries of **Reactants**, **TS1**, **Int**, **TS2**, and **Products** in the base-catalyzed hydrolysis of methyl acetate. Distances are in angstroms.

Table 9 shows the calculated and experimental free energies[24a] for the stationary points of the base-catalyzed hydrolysis of methyl acetate in water. The HF/6-31+G(d) geometries of the stationary points in the base-catalyzed hydrolysis of methyl acetate are illustrated in Figure 2. The calculated relative Gibbs free energy for **Products** reproduces the experimental value to 1 kcal/mol. In contrast, those for **TS1** and **Int** were overestimated by 2.8 and 5.1 kcal/mol. In the base-catalyzed hydrolysis of *N*,*N*-dimethylacetamide, Massova and Kollman estimated the free energy difference of 28.7 kcal/mol between the intermediate and the reactants using the PCM-UAHF(G98).[27] This values is about 10 kcal/mol larger than the experimental value of 19 kcal/mol.[28] Pliego and Riveros calculated an activation barrier (17.6 kcal/mol) close to experiment (18.6 kcal/mol),[24a] but this was because there was an equally high error in the calculations of the TS and reactants:[25e] the calculated solvation free energy of the hydroxide ion, −92.5 kcal/mol, gives an especially large deviation from experiment (−105.0 kcal/mol).[16a] CPCM-UAKS provides the reasonable solvation free energies

for the hydroxide ion (−105.5 kcal/mol). Compared to other methods, CPCM-UAKS gives an improved estimate of the activation barrier of hydrolysis but still overestimates it. An entropic correction in solution may be required for further improvement, because the CPCM total solute energy does not include an entropic change in solution as shown in eq 1. The ester hydrolysis involves association and dissociation processes, and the solute in this case is a transition state composed of two or more loosely bound molecules. The much larger deviation (about 10 kcal/mol) of **TS2** is likely due to the involvement of an explicit water molecule, which is proposed by previous theoretical works.[25] The accurate evaluation of the relative free energy for **TS2** must include an explicit water molecule.

Conclusions

Benchmarks of different variations of CPCM for the computation of solvation energies of neutral and ionic organic species have been performed and compared to other work in the literature. The CPCM-UAKS method provides the aqueous solvation free energies in agreement with experimental data and with improved computational times compared to other cavity methods. The mean absolute deviations from experiment are 2.6 kcal/mol. The largest solvation energy errors are obtained for anions and cations.

Acknowledgment. We are grateful to the National Science Foundation for financial support of this research. We also thank the National Computational Science Alliance under Grant MCA93S015N and the UCLA Advanced Technology Services for computational resources. We thank Fernando R. Clemente, Christophe Allemann, Andrew S. Dutton, Patrick R. McCarren, and Paul Ha Yeon Cheong for helpful discussions. Y.T. thanks the Japan Society for the Promotion of Science for a Research Fellowship for Young Scientists Abroad.

Supporting Information Available: Aqueous solvation free energies for 30 neutrals, 21 anions, and 19 cations (Tables S1−S27) and an example of an input file (Figure S1). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027−2094. (b) Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, *99*, 2161−2200. (c) Chipman, D. M. *J. Phys. Chem. A* **2002**, *106*, 7413−7422. (d) Chipman, D. M. **2003**, *118*, 9937−9942.

(2) (a) Klamt, A.; Schüürmann, G. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799. (b) Andzelm, J.; Kölmel, C.; Klamt, A. *J. Chem. Phys.* **1995**, *103*, 9312−9320. (c) Barone, V.; Cossi, M. *J. Phys. Chem. A* **1998**, *102*, 1995−2001. (d) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. Comput. Chem.* **2003**, *24*, 669−681.

(3) (a) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117−129. (b) Barone, V.; Cammi, R.; Tomasi, J. *Chem. Phys. Lett.* **1996**, *255*, 327−335. (c) Cossi, M.; Scalmani, G.; Rega, N.; Barone, V. *J. Chem. Phys.* **2002**, *117*, 43−54. (d) Barone, V.; Improta, R.; Rega, N. *Theor. Chem. Acc*. **2004**, *111*, 237−245.

(4) (a) 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.; Lyengar, 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.; Camml, 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.04; Gaussian, Inc.: Pittsburgh, PA, 2003. (b) 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.; 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.; 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.9; Gaussian, Inc.: Pittsburgh, PA, 1998.

(5) Pierotti, R. A. *Chem. Rev.* **1976**, *76*, 717−726.

(6) (a) Floris, F. M.; Tomasi, J. *J. Comput. Chem.* **1989**, *10*, 616−627. (b) Floris, F. M.; Tomasi, J.; Pascual-Ahuir, J. L. *J. Comput. Chem.* **1991**, *12*, 784−791.

(7) Caillet, J.; Claverie, P.; Pullman, B. *Acta Crystallgr.* **1978**, *B34*, 3266−3272.

(8) Ben-Naim, A.; Marcus, Y. *J. Chem. Phys.* **1984**, *81*, 2016−2027.

(9) (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. (c) Becke, A. D. *J. Chem. Phys*. **1993**, *98*, 5648−5652.

(10) (a) Mclean, A. D.; Chandler, G. S. *J. Chem. Phys.* **1980**, *72*, 5639−5648. (b) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650−654.

(11) (a) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. v. R. *J. Comput. Chem.* **1983**, *4*, 294−301. (b) Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, *80*, 3265−3269.

(12) (a) Barone, V.; Cossi, M.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3210−3221. (b) *Handbook of Chemistry and Physics*; Weast, R. C., Ed.; Chemical Rubber: Cleveland, OH, 1981. (c) Bondi, A. *J. Chem. Phys.* **1964**, *68*, 441−451.

(13) Rappé, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A., III; Skiff, W. M. *J. Am. Chem. Soc.* **1992**, *114*, 10024−10035.

(14) GAUSSIAN03 online manual of the PCM and CPCM: http://www.gaussian.com/g_ur/k_scrf.htm

(15) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158−6170.

(16) (a) Pliego, J. R., Jr.; Riveros, J. M. *Phys. Chem. Chem. Phys.* **2002**, *4*, 1622−1627. (b) Williams, R. taken from: http://icg.harvard.edu/~chem206/Fall_2001/lectures/18_Acid_Based-_Properties_of_Organic_Molecules/pKa_compilation.pdf.

(17) Cluster-continuum is a hybrid approach that combines gas-phase clustering by explicit solvent molecules and solvation of the cluster by the dielectric continuum: (a) Pliego, J. R., Jr.; Riveros, J. M. *J. Phys. Chem. A* **2001**, *105*, 7241−7247. (b) Pliego, J. R., Jr.; Riveros, J. M. *J. Phys. Chem. A* **2002**, *106*, 7434−7439.

(18) (a) Li, J.; Hawkins, G. D.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. *Chem. Phys. Lett.* **1998**, *288*, 293−298. (b) Giesen, D. J.; Hawkins, G. D.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. *Theor. Chem. Acc.* **1997**, *98*, 85−109. (c) Chambers, C. C.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem.* **1996**, *100*, 16385−16398. (d) Cramer, C. J.; Truhlar, D. G. *Science* **1992**, *256*, 213−217.

(19) Foresman, J. B.; Keith, T. A.; Wiberg, K. B.; Snoonian, J.; Frisch, M. J. *J. Phys. Chem.* **1996**, *100*, 16098−16104.

(20) Thompson, J. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. A* **2004**, *108*, 6532−6542. In this paper, the authors mentioned that they used the default setting of GAUSSIAN98 and GAUSSIAN03 and that the cavity was built by the UAHF method. They, however, performed the CPCM calculation with older version of GAUSSIAN03 (revision A.01). It should be noted that the later and latest versions of GAUSSIAN03 (revision B and revision C) build the cavity by the UA0 method. If the reader would like to use the cavity built by the UAHF method in the CPCM or PCM calculation, it is necessary to use the option, RADII=UAHF as shown in Supporting Information (Figure S1).

(21) (a) Rohschneider, L. *Anal. Chem.* **1973**, *45*, 1241−1247. (b) Park, J. H.; Hussam, A.; Couasnon, P.; Friz, D.; Carr, P. W. *Anal. Chem.* **1987**, *59*, 1970−1976.

(22) (a) Liptak, M. D.; Gross, K. C.; Seybold, P. G.; Feldgus, S.; Shields, G. C. *J. Am. Chem. Soc.* **2002**, *124*, 6421−6427. (b) Liptak, M. D.; Shields, G. C. *J. Am. Chem. Soc.* **2001**, *123*, 7314−7319. (c) Kallies, B.; Mitzner, R. *J. Phys. Chem. B* **1997**, *101*, 2959−2967. (d) Lopez, X.; Schaefer, M.; Dejaegere, A.; Karplus, M. *J. Am. Chem. Soc.* **2002**, *124*, 5010−5018. (e) Klici'c, J. J.; Friesner, R. A.; Liu, S.-Y.; Guida, W. C. *J. Phys. Chem. A* **2002**, *106*, 1327−1335. (f) Fu, Y.; Liu, L.; Li, R.-Q.; Liu, R.; Guo, Q.-X. *J. Am. Chem. Soc.* **2004**, *126*, 814−822. (g) Klamt, A.; Eckart, F.; Diedenhofen, M.; Beck, M. E. *J. Phys. Chem. A* **2003**, *107*, 9380−9386. (h) Almerindo, G. I.; Tondo, D. W.; Pliego, J. R., Jr. *J. Phys. Chem. A* **2004**, *108*, 166−171.

(23) Ben-Naim, A. *J. Phys. Chem.* **1978**, *82*, 792−803.

(24) (a) Guthrie, J. P. *J. Am. Chem. Soc.* **1973**, *95*, 6999−7003. (b) Bender, M. L. *Chem. Rev.* **1960**, *60*, 53−113. (c) Jencks, W. P. *Chem. Rev.* **1972**, *72*, 705−718. (d) *Ester Formation and Hydrolysis*; Bamford, C. H., Tipper, C. F. H., Eds.; Elsevier Science: Amsterdam, 1972; Vol. 10.

(25) (a) Hæffner, F.; Hu, C.-H.; Brinck, T.; Norin, T. *J. Mol. Struct. (THEOCHEM)* **1999**, *459*, 85−93. (b) Zhan, C.-G.; Landry, D. W.; Ornstein, R. L. *J. Am. Chem. Soc.* **2000**, *122*, 2621−2627. (c) Zhan, C.-G.; Landry, D. W.; Ornstein, R. L. *J. Phys. Chem. A* **2000**, *104*, 7672−7678. (d) Pliego, J. R., Jr.; Riveros, J. M. *Chem. Eur. J.* **2002**, *8*, 1945−1953. (e) Pliego, J. R., Jr.; Riveros, J. M. *J. Phys. Chem. A* **2004**, *108*, 2520−2526.

(26) (a) Bauschlicher, C. W.; Partridge, H. *J. Chem. Phys.* **1995**, *103*, 1788−1791. (b) Wong, M. W. *Chem. Phys. Lett.* **1996**, *256*, 391−399. (c) Scott, A. P.; Radom, L. *J. Phys. Chem.* **1996**, *100*, 16502−16513.

(27) Massova, I.; Kollman, P. A. *J. Phys. Chem. B* **1999**, *103*, 8628−8638.

(28) Guthrie, J. P. *J. Am. Chem. Soc.* **1974**, *96*, 3608−3615.

CT049977A

JCTC Journal of Chemical Theory and Computation

Adsorption of Water Molecules on Flat and Stepped Nickel Surfaces from First Principles

Daniel Sebastiani* and Luigi Delle Site*

Max-Planck-Institute for Polymer Research, Ackermannweg 10, 55128 Mainz, Germany

Received August 25, 2004

Abstract: We present an ab initio density functional study of the adsorption of a series of water oligomers (molecule, dimer and trimer) on nickel surfaces with and without step defects. We investigate the preferred adsorption geometries and adsorption energies and analyze the binding mechanisms by means of electronic density difference maps. Special attention is devoted to the incremental adsorption process, i.e., the way additional molecules attaches to an already adsorbed water. In agreement with recent findings, we show that the first water molecule is bound to the surface with an energy of about 0.2−0.4 eV, i.e., with up to twice the strength of a hydrogen bond. In contrast to this, subsequent water molecules increase the total adsorption energy by typically 0.5 eV. However, electron density difference considerations indicate that this additional attraction is *not* due to the interaction of the new molecule with the surface but mediated by the first water molecule. The interaction of the additional molecule with the surface appears even to be repulsive. We discuss the implications of these findings for the wetting properties of transition metal surfaces.

1. Introduction

The interaction of water with metal surfaces is of tremendous importance for industrial applications and of very high relevance in surface chemistry. The fundamental mechanisms of the initial phases of this adsorption process has been the subject of a variety of recent experimental and theoretical investigations[1−6] and is still a matter of intense controversy.[7−9]

There is evidence that the water molecules can approach the surfaces in both hydrogen-up and hydrogen-down orientations. Depending on the metal species, a water monolayer can be formed through simultaneous binding of all water molecules to the surface, e.g. in the case of Pt-(111), or alternatively as a partially dissociated layer, which is the case for Ru(0001).[2,4]

The deposition of isolated water molecules has already been studied on several flat metallic surfaces, such as Ru(0001), Rh(111), Pd(111), Pt(111), Cu(111), and Ag(111).[10−12] An interesting mechanism for dimer diffusion through a combined proton tunneling and molecular rotation scheme has been proposed recently.[13]

Most of these studies deal with the interaction of either a single water molecule on a metallic slab or the adsorption of a highly symmetric water monolayer. A high degree of order facilitates the computational modeling within periodic boundary conditions, because the unit cell can be kept small.

In this work, we study more extended systems, which allow for significantly more geometrical relaxation. Our aim is to focus on a realistic description of the initial steps of wetting, going beyond a single adsorbed molecule but without imposing a complete coverage of the surface. For this purpose, we look at the adsorption of a sequence of water oligomers on nickel. Starting from an isolated molecule, we investigate a water dimer as well as a trimer, paying particular attention to the influence of hydrogen bonding on the adsorption energy and the structure of the adsorbate. In particular, we compare the incremental adsorption energy due to the addition of a second and third molecule.

Inspired by previous studies,[14−17] we look not only at a perfect surface but also at the simplest possible defect, a one-dimensional step. This is realized by using a surface in the

* Corresponding authors fax: +49-6131-379-100; e-mail: sebastia@mpip-mainz.mpg.de (D.S.) and fax: +49-6131-379-100; e-mail: dellsite@mpip-mainz.mpg.de (L.D.S.).

10.1021/ct049955o CCC: $30.25
Published on Web 11/13/2004

Table 1. Energetical and Structural Data for the Various Water Oligomers and Surfaces[a]

	ΔE	d_{Ni-O}	$\Delta_{xy}(O)$	$\Delta_z(Ni)$	α	$d_{O\cdots O}$
$(H_2O)_2$	0.22					
$Ni_{flat}\cdot(H_2O)_{1,atop}$	0.242	2.26	0.14	0.17	5	
$Ni_{flat}\cdot(H_2O)_{1,bridge}$	0.10	3.02	0.02	0.16	−10	
$Ni_{step}\cdot(H_2O)_{1,atop,trans}$	0.403	2.12	0.04	0.10	4	
$Ni_{step}\cdot(H_2O)_{1,bridge,trans}$	0.249	2.46	0.03	0.08	8	
$Ni_{step}\cdot(H_2O)_{1,atop,cis}$	0.397	2.10	0.07	0.05	2	
$Ni_{step}\cdot(H_2O)_{1,bridge,cis}$	0.137	2.51	0.05	0.03	0	
$Ni_{flat}\cdot(H_2O)_{2,atop}$	0.675	2.12	0.16	0.19	22	2.71
$Ni_{step}\cdot(H_2O)_{2,atop}$	0.900	2.09	0.13	0.08	28	2.69
$Ni_{flat}\cdot(H_2O)_{3,atop}$	1.26	2.08	0.2	0.25	34	2.72/2.81
$Ni_{flat}\cdot(H_2O)_{3,bridge}$	1.09	2.24	0.2	0.11	72	2.69/2.7

[a] Adsorption energies (ΔE) are in eV and are always computed relative to isolated water molecules according to eq 1, such as to include also the hydrogen bonding energy. The Ni−O bond lengths (d_{Ni-O}), the vertical displacements of the binding nickel atom ($\Delta_z(Ni)$), and the lateral displacements of the binding oxygen atom from its optimal position ($\Delta_{xy}(O)$) are given in Å, and the angle between the molecular plane of the adsorbing water molecule and the surface (α) is in degrees. A negative angle means that the hydrogens are pointing toward the surface. For oligomers, the distance between the oxygen atoms (d_{O-O}) is also shown in Å.

(221) direction. The adsorption-enhancing effect of such steps has also been shown very recently by experimental studies of germanium deposition on silicon surfaces.[18]

2. Computational Details

We have modeled the flat and stepped surfaces within density functional theory (DFT) as two-dimensional slabs under periodic boundary conditions. We have chosen hexagonal supercells containing 3 (111)-layers for the (221) step and 4 layers for the flat surface, with nickel supercells of 2×2 atoms for the water monomer on the flat surface, 3×3 for the water dimer and trimer on the flat, and 1×2 for the monomer and dimer on the step surface. In all cases, the slabs were separated in the third dimension by about 10 Å of vacuum. Throughout this work, a plane-wave cutoff of 60Ry has been used,[19] together with Troullier-Martins pseudopotentials[20] in the Kleinman-Bylander scheme for hydrogen and oxygen as well as a special pseudopotential by Lee[21] for nickel. All calculations have been done with the PBE exchange-correlation functional.[22] The electronic structure has been computed using the implementation of the free energy functional of Alavi.[23,24] In this approach, a finite-temperature propagator is diagonalized iteratively for the electronic degrees of freedom, ensuring a Fermi distribution function for the occupation levels of the Kohn−Sham orbitals. We have used a $4 \times 4 \times 1$ and a $3 \times 3 \times 1$ k-point mesh within the Monkhorst-Pack scheme[25] to sample the Brillouin zone for the step and flat surfaces, respectively. Using this computational setup, we could reproduce the recently published results of Michaelides et al.[26] for the adsorption of a highly structured water monolayer with a numerical error of less than 50 meV.

Geometry optimizations were done until the atomic forces dropped below a threshold of $2\cdot10^{-3}$atomic units. This level is sufficiently strict for the considered systems; further optimization changes the total energy by negligible amounts only.[27] We have also computed an estimator for the artificial energy lowering due to the periodicity of our simulation cell. An isolated water molecule in our standard box has a total energy which is about 0.01 eV lower than that in a box of twice the lattice constant.

All adsorption energies have been computed with respect to single isolated water molecules in the same simulation box (thus eliminating the dipole interaction error):

$$\Delta E = E[\mathrm{Ni}\cdot(\mathrm{H_2O})_n] - E[\mathrm{Ni}] - nE[\mathrm{H_2O}] \quad (1)$$

This definition also allows to obtain directly the incremental adsorption energies for an additional water molecule, which would be more involved when considering the adsorption energy of the water cluster as an entity (i.e. when taking $E[(H_2O)_n]$ instead of $nE[H_2O]$).

We have further computed electron density difference maps for selected energetically favorable configurations, showing the rearrangement of the density between the surface-adsorbed $Ni\cdot(H_2O)_n$ complex and the fragments. For visualizing these maps, we plot color-coded slices with the projected electron density differences $\delta\rho$ defined as

$$\delta\rho(x, z) = \int \mathrm{d}y(\rho^{\mathrm{Ni}\cdot(\mathrm{H_2O})n}(\mathbf{r}) - \rho^{\mathrm{Ni}}(\mathbf{r}) - \rho^{(\mathrm{H_2O})n}(\mathbf{r})) \quad (2)$$

Here, in contrast to the definition of the energy difference, the density plots compare the adsorbed system with the water oligomer cluster and the isolated surface. In this way, the density displacements due to the formation of hydrogen bonds between the water molecules are not plotted, while the modifications of these densities due to the adsorption become visible. The density difference maps were computed in orthorhombic cells obtained by doubling the original hexagonal unit cells and by cutting out a suitable orthorhombic subpart of at least the size of the original setup. This was necessary for a proper visualization of the densities with the program MOLEKEL.[28]

3. Results

3.1. Monomer. The geometry for three typical adhesion sites ("top", "bridge", and "hollow") of a single water molecule on a flat nickel surface was optimized as well as "top" and "bridge" sites on a 221-surface which represents a step defect. Pictograms illustrating the top and bridge geometries on the flat and stepped surfaces are shown in Figures 1 and 2. The corresponding adsorption energies for these configurations which exhibit a (local) minimum of the potential energy

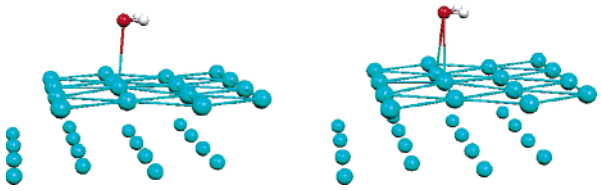


Figure 1. Pictogram of the two adsorption sites for the flat nickel surface which constitute local minima of the potential energy surface: top (left) and bridge (right). The hollow structure turns out to converge toward the top configuration. Only two layers of nickel atoms are shown: the first surface layer is drawn with bonds, the second layer as spheres only.

surface are shown in Table 1. The energies of the top site differ from those of the bridge position by typically a factor of 2 in favor of the top configuration. Since the hollow system relaxed to the top site, it was not considered further. Despite the tetrahedral location of the lone pairs of the oxygen atom, the geometry optimization yields a flat arrangement of the molecule, so that the protons are found at the same distance from the surface as the oxygen (for both flat and step cases).

We have explicitly checked a possible vertical adsorption geometry for the flat surface on the hollow site and a tilted one (with one OH-bond parallel to the surface and the other OH normal to it) on the atop site. For both initial geometries, the optimization yields a flat orientation, and in the case of the hollow site, the oxygen atom also moves over to the atop site. Intermediate adsorption energy values from the optimization process are in the area of 0.1 eV during the turning process. This indicates that when the water molecule approaches the surface vertically, the adsorption strength is roughly half as large as in the parallel orientation. These results are in full agreement with the orientations found by Ranea, Michaelides, and others[10,11,13] for water configurations on various other metallic surfaces.

Similar to the recently studied case of an adsorbed benzene molecule,[14] the adsorption of a water molecule is significantly stronger on the surface with the step defect than on the flat one. The adsorption energy for the latter is roughly one hydrogen bond (cf. the $(H_2O)_2$ value in Table 1), while the step provides about twice that attraction. There exists a cis and a trans orientation (both shown in Figure 2) for the step defect, but their adsorption energy is almost the same. Thus, the top position on the step surface reaches the highest energy value, which is also approximately equal to the typical hydrogen bond energy of a 4-fold-coordinated liquid water molecule.

The electron density difference according to eq 2 for both the flat and step surfaces is shown in Figure 3. The plot represents the density of the aggregate minus the sum of the densities of the isolated surface and the water molecule at the top site (trans configuration for the step).

The formation of a weak bond between the surface nickel atom and the oxygen is clearly visible through the displacement in electronic density (dark green and blue regions). To some extent, we also find additional density on top of the water molecule, while relatively little is removed from the central area around the oxygen. Most of the electronic density

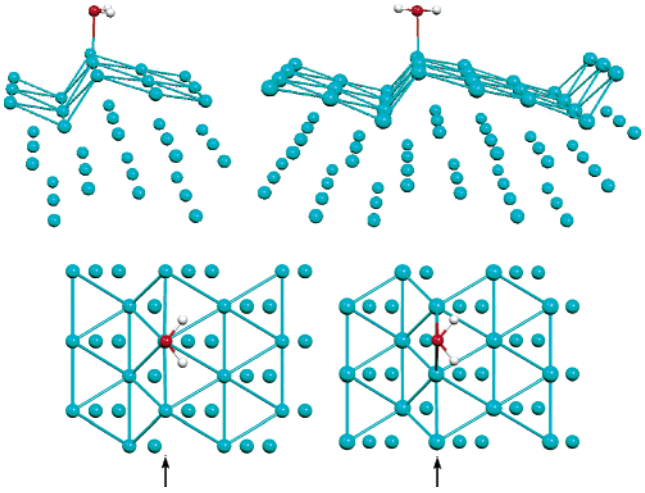


Figure 2. Pictogram of the two adsorption sites for the stepped nickel surface: top/cis (upper left), top/trans (upper right). A comparison with the bridge site is given below, from a top view: top/cis (lower left), and bridge/cis (lower right). As in Figure 1, the atoms of the top layer are drawn with bonds, those of the deeper layers as spheres only.

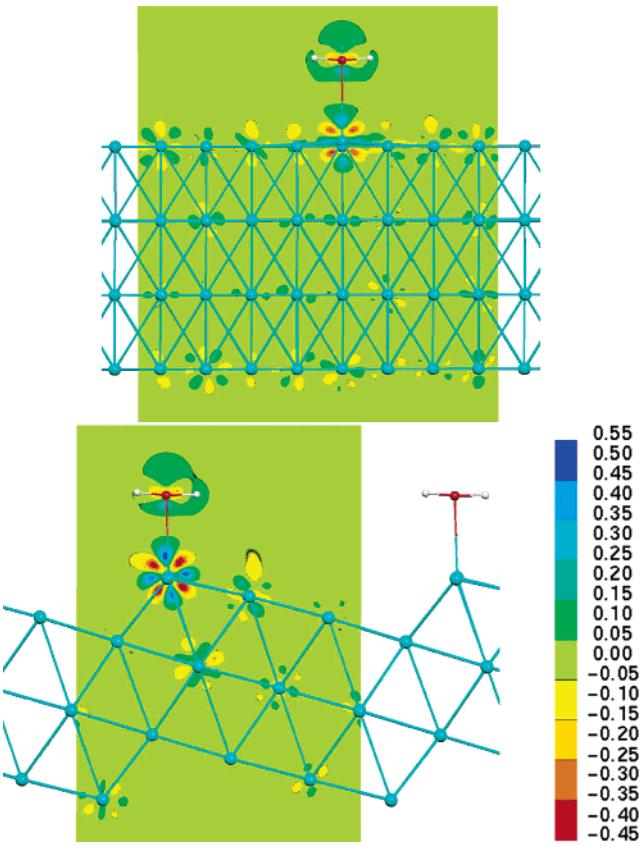


Figure 3. Density difference plots for the adsorption of a water molecule on the flat (top plot, atop configuration) and the step surfaces (bottom plot, atop trans configuration). The scale is given in units of e/Å^2.

is taken from the bonding nickel atom, which is strongly polarized, and its first neighbors.

Similar to the case of an adsorbed benzene molecule,[14] the polarization of the nickel atom which is bonded to the water molecule is significantly stronger on the step surface

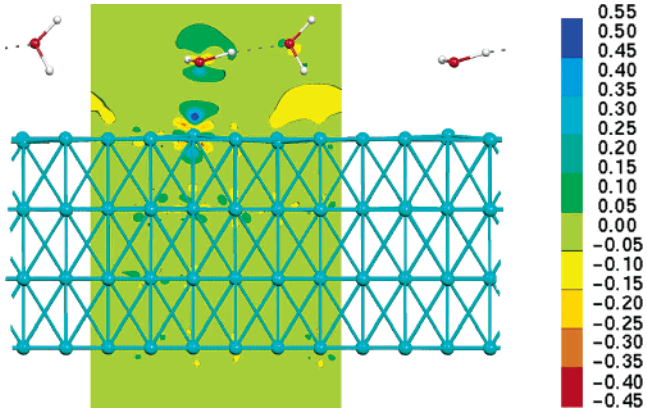


Figure 4. Electron density difference plot for the adsorption of a water dimer on the flat surface.

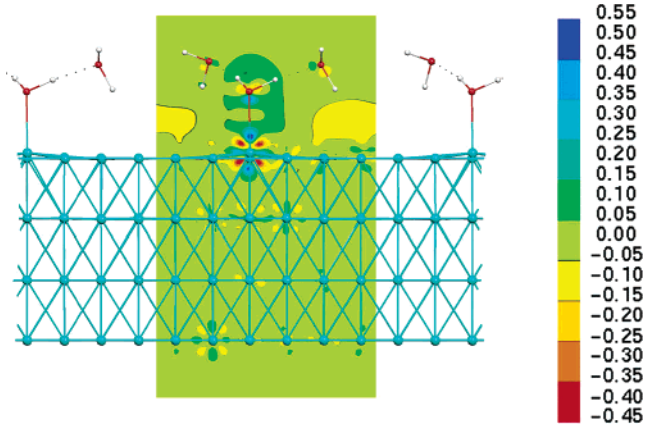


Figure 5. Electron density difference plot for the adsorption of a water trimer on the flat surface.

than on the flat one. This effect also translates into the higher adsorption energy for the latter.

3.2. Dimer. The second step in water adsorption on the surface is the attachment of a second water molecule to the first one. For this purpose, the geometries of a water dimer on the nickel surfaces (flat and with the step defect) have been optimized. Since the monomer adsorption is energetically significantly more favorable on the top site than in the bridge position, only the atop configuration has been considered.

The computed Ni−O bond distances and adsorption energies are shown in Table 1. They are taken relative to isolated water molecules in order to have a common reference for all systems. On both the flat and the step surfaces, the attachment of an additional water molecule to the first one through a hydrogen bond yields an additional 0.43 eV and 0.5 eV, respectively. These energies are about twice as large as it would be expected for a standard hydrogen bond (cf. the water dimer in Table 1), leading again to an adsorption strength that is comparable to that of a water molecule in liquid water. Especially on the step defect, two water molecules attach with an energy that is equivalent to four hydrogen bonds, while still possessing two hydrogen bond acceptor sites (one on each oxygen) and two dangling donor protons.

Thus, the dimer adsorption on the metal surface can energetically compete with the solvation of the second molecule in liquid water, even though the optimized cluster on the surface is not directly comparable to the situation in liquid water due to the high dynamics of the hydrogen bond network at finite temperature.

For the flat surface, a part of this increased energy probably stems from the decreased Ni−O bond distance compared to the water monomer, whereas the step surface does not show this effect.

The analysis of the electronic density difference for the dimer adsorption on the flat surface is shown in Figure 4. According to eq 2, the isolated water dimer is taken as reference system, to suppress the charge difference due to the water−water hydrogen bond and to show rather how much this hydrogen bond is changed due to the adsorption on the surface. The plot reveals that the bonding mechanism of the first water molecule is essentially the same as for the monomer, except that the charge displacement around the first water is now less symmetric than before. There is a region of strongly increased electron density leading to a Ni−O bond and an additional charge increase on top of the first water molecule. Furthermore, the amount of electronic charge density which is found on top of the bonding Ni atom is significantly stronger than for the water monomer (deep blue color in Figure 4 compared to light blue in Figure 3).

It is interesting to note that the second water molecule does not bind directly to the surface, it is even repelled from it. The second oxygen is not accumulating any electronic density toward the metal surface, and there is a distinguishable region of decreased electron density (yellow color coding) below the hydrogen which points toward the surface. In contrast to this, the hydrogen bond between the two water molecules becomes slightly stronger than in the isolated dimer, as seen by the polarization of the H-bond accepting oxygen.

3.3. Water Trimer on the Flat Surface. The last point in our investigation focuses on the adsorption of a third water molecule onto H_2O dimer on the flat nickel surface. As before, the energy and the bonding distance of the first water are shown in Table 1 for both the bridge and the top sites. The energy difference between bridge and top geometries is conserved upon adsorption of further molecules, implying that the secondary water molecules do not feel any significant influence from the adsorption site of the first one.

It is surprising that the third water increases the total adsorption energy by almost 0.6 eV, thus practically doubling the value of the dimer. This can be explained only in parts by the decreased Ni−O bond length directly. The density difference map of the trimer, which is shown in Figure 5 for the top site, reveals a highly increased electronic charge density in the Ni−O bond region, at the expense of the areas below the dangling protons of the secondary water molecules. Together with a very high polarization of the binding nickel atom as well as slightly stronger hydrogen bonds between the waters, this indicates a much stronger binding of the first oxygen atom.

It is interesting to note that this increased binding of the first water is practically not related to any of its geometric properties but is rather due to the mere presence of secondary water molecules, which constitute a kind of a first solvation

shell. These additional water molecules perturb and repel the electron density at the metal surface in the neighborhood of the initial molecule in such a way that the polarization of the bonding nickel is significantly increased. This phenomenon is already visible for the adsorbed dimer but even stronger in the presence of a third water molecule.

4. Conclusion

We have presented an investigation of the energetic details and the electronic mechanism of the adsorption of water oligomers on nickel surfaces with and without a step defect. The data shown indicates an increased binding strength on the step and in "top" geometries. The findings are in good agreement with previous theoretical and experimental results for similar systems,[6] where an energetic enhancement of water adsorption was found along step defects on platinum surfaces.

Further, we could show by means of electronic density difference maps that additional water molecules tend to strengthen the nickel−oxygen bond. This effect leads to a significant stabilization of the binding of the first water molecule and to strongly increased binding energies of the dimer and trimer complexes. Not surprisingly, the adhesion of water to nickel surfaces is much weaker than of aromatic molecules such as benzene or phenol,[14,29,30] but the adsorption energies can definitively compete with those found in liquid water. The second and third water molecule increased the total binding energy by the equivalent of more than two hydrogen bonds each, which correspond to the average binding energy per molecule of standard 4-fold coordinated water.

By means of our study, the initial steps of aqueous wetting of transition metal surfaces can be understood on the basis of electronic effects that govern molecular adsorption.

Acknowledgment. We have profited from fruitful exchange and very helpful discussions with A. Alavi. We also would like to thank P. Blümler for critically reading the manuscript.

References

(1) Henderson, M. A. *Surf. Sc. Rep.* **2002**, *46*, 1−308.

(2) Feibelman, P. J. *Science* **2002**, *295*, 99−102.

(3) Menzel, D. *Science* **2002**, *295*, 58−59.

(4) Ludwig, R. *Angew. Chem., Int. Ed. Engl.* **2003**, *42*, 3458−3460.

(5) Mitsui, T.; Rose, M. K.; Fomin, E.; Ogletree, D. G.; Salmeron, M. *Science* **2002**, *297*, 1850−1852.

(6) Grecea, M. L.; Backus, E. H. G.; Riedmüller, B.; Eichler, A.; Kleyn, A. W.; Bonn, M. *J. Phys. Chem. B* **2004**, *108*, 12575−12582.

(7) Meng, S.; Xu, L. F.; Wang, E. G.; Gao, S. *Phys. Rev. Lett.* **2002**, *89*, 176104.

(8) Feibelman, P. *Phys. Rev. Lett.* **2003**, *91*, 059601.

(9) Meng, S.; Xu, L. F.; Wang, E. G.; Gao, S. *Phys. Rev. Lett.* **2003**, *91*, 059602.

(10) Michaelides, A.; Ranea, V. A.; de Andres, P. L.; King, D. A. *Phys. Rev. Lett.* **2003**, *90*, 216102.

(11) Ranea, V. A.; Michaelides, A.; Ramírez, R.; Vergès, J. A.; de Andres, P. L.; King, D. A. *Phys. Rev. B* **2004**, *69*, 205411.

(12) Cerdá, J.; Michaelides, A.; Feibelman, M.-L. B. P. J.; Mitsui, T.; Rose, M.; Fomin, E.; Salmeron, M. *Phys. Rev. Lett.* **2004**, *93*, 116101.

(13) Ranea, V. A.; Michaelides, A.; Ramírez, R.; de Andres, P. L.; Vergés, J. A.; King, D. A. *Phys. Rev. Lett.* **2004**, *92*, 136104.

(14) Delle Site, L.; Sebastiani, D. *Phys. Rev. B* **2004**, *70*, 115401.

(15) Dahl, S.; Logadottir, A.; Egeberg, R. C.; Larsen, J. H.; Chorkendorff, I.; Törnqvist, E.; Nørskov, J. K. *Phys. Rev. Lett.* **1999**, *83*, 1814.

(16) Mavrikakis, M.; Bäumer, M.; Freund, H. J.; Nørskov, J. K. *Catal. Lett.* **2002**, *81*, 153.

(17) Shah, V.; Li, T.; Baumert, K. L.; Cheng, H. S.; Sholl, D. S. *Surf. Sc.* **2003**, *537*, 217.

(18) Hannon, J. B.; Copel, M.; Stumpf, R.; Reuter, M.; Tromp, R. *Phys. Rev. Lett.* **2004**, *92*, 216104.

(19) "The plane wave cutoff convergence has been checked on a water dimer geometry optimization at 60Ry and at 100Ry. The resulting hydrogen bond energies are 0.2400eV and 0.2398eV, which represents a difference that lies within the numerical noise of the method. The OH···O distance changes by about 1%, while covalent bond lengths vary even less."

(20) Troullier, N.; Martins, J. L. *Phys. Rev. B* **1991**, *43*, 1993.

(21) Lee, M. Ph.D. Thesis; Cambridge University: 1995.

(22) Perdew, J.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.

(23) Alavi, A.; Kohanoff, J.; Parrinello, M.; Frenkel, D. *Phys. Rev. Lett.* **1994**, *73*, 2599.

(24) Silvestrelli, P. L.; Alavi, A.; Parrinello, M.; Frenkel, D. *Europhys. Lett.* **1996**, *33*, 551−556.

(25) Monkhorst, H. J.; Pack, J. D. *Phys. Rev. B* **1973**, *13*, 5188.

(26) Michaelides, A.; Alavi, A.; King, D. A. *Phys. Rev. B* **2004**, *69*, 113404.

(27) "The value of $2 \cdot 10^{-3}$ Hartree/bohr corresponding to 0.1 eV/Å for the force convergence describes the maximum component of the forces on all atoms. The commonly cited root-mean-square (rms) force was always below 10^{-3} Hartree/bohr. This remaining force usually arises from metal atoms which are moving on a highly complex (non-parabolic) potential energy hypersurface which is very difficult to optimize. Further optimization does not change the adsorption picture significantly. However, a test calculation has been performed, in which the threshold for the convergence criterion of the geometry optimization was lowered to $3 \cdot 10^{-4}$ atomic units. The resulting change in the total energy was less than 0.05eV."

(28) Flükiger, P.; Lüthi, H. P.; Portmann, S.; Weber, J. Molecular visualization program MOLEKEL, version 4.0, Swiss Center for Scientific Computing, Manno, Switzerland, 2000.

(29) Delle Site, L.; Abrams, C. F.; Alavi, A.; Kremer, K. *Phys. Rev. Lett.* **2002**, *89*, 156103.

(30) Delle Site, L.; Alavi, A.; Abrams, C. F. *Phys. Rev. B* **2003**, *67*, 193406.

CT049955O

JCTC Journal of Chemical Theory and Computation

An Aromaticity Scale Based on the Topological Analysis of the Electron Localization Function Including σ and π Contributions

Juan C. Santos,*,§ Juan Andres,† Arie Aizman,§ and Patricio Fuentealba‡

Departamento de Química, Universidad Técnica Federico Santa Maria, Casilla 110V, Valparaíso, Chile, Departament de Ciències Experimentals, Universitat Jaume I, Apartat 224, 12080, Castelló, Spain, and Departamento de Física, Universidad de Chile, Las Palmeras 3425, Santiago, Chile

Received September 30, 2004

Abstract: In this work, the average bifurcation value of the electron localization function (ELF) of both σ (ELFσ) and π (ELFπ) contributions was used to construct an aromaticity scale for chemical compounds. We have validated the scale with a series of well-known molecules and then used it to evaluate global aromaticity on aluminum based clusters, which present σ aromaticity and π antiaromaticity. The proposed scaled predicts an overall antiaromatic character for the Al_4^{4-} moiety.

Introduction

The concept of aromaticity and, by extension, antiaromaticity, has general acceptance among the chemical community. However, the criteria used to validate the aromaticity or antiaromaticity of a chemical compound are controversial,[1−8] due to the fact that a well-established definition of these concepts has yet to be presented in a quantitative way for general use.[2−21]

Currently there is an interesting discussion about the net aromaticity of a chemical species when it presents both, σ aromaticity and π antiaromaticity.[22] The all metal cluster Al_4^{4-} is a classical example where this phenomenon appears to be present. The Al_4^{4-} ring in $Li_3Al_4^-$ is not square. It has two slightly different bond lengths, and it contains 4π electrons. These characteristics plus the shape of the molecular orbitals have been used to classify this species as antiaromatic.[23] On the other hand however, this analysis has been presented as an incomplete view of the electronic properties of this compound, because, although the cluster is π antiaromatic, it presents characteristics of aromaticity in its σ system.[24] While the different studies[23,24] agree in the σ aromatic and π antiaromatic character of the ion, the differences appear in the global characterization of the species as aromatic or antiaromatic.

The lack of a unique scale of aromaticity is an important limiting factor in this debate. Scales based on structure or energetic behavior such as the harmonic oscillator model of aromaticity (HOMA)[25,26] or the aromatic stabilization energy (ASE),[2] respectively, have been proposed. Other aromaticity criteria based on the analysis of electron delocalization[12,13] or associated to measurable response properties[20] have also been used. These scales have been constructed thinking in traditional organic aromatic rings, and it is not clear that they can be applied to all metal clusters. Recently, the nucleus independent chemical shift (NICS) index,[5] based on magnetic properties, has been extensively used in the theoretical literature. Extensions and modifications[5,6] as well as criticism[16,19] of this methodology have been published.

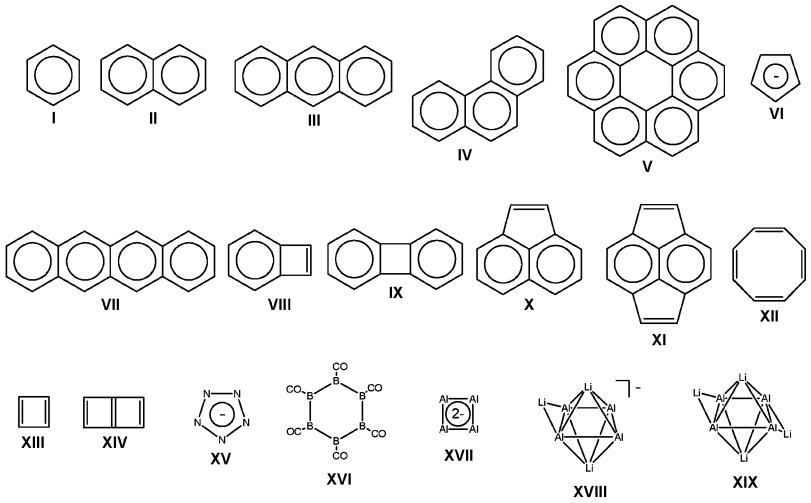
An alternative approach, based on properties of the electron density probed by the electron localization function (ELF)[27] of Becke and Edgecombe,[28] has been introduced to understand aromaticity.[14,15,29] A separation of the ELF into its σ and π components was shown to provide a useful scheme to discuss σ and π character in a molecular system.[30] The ELF is defined in terms of the excess of local kinetic energy density due to the Pauli exclusion principle, $T(\rho(r))$, and the Thomas-Fermi kinetic energy density, $T_h(\rho(r))$

$$\mathrm{ELF} = \left[1 + \left(\frac{T(r)}{T_h(r)}\right)^2\right]^{-1}$$

* Corresponding author e-mail: jucasa@circonio.ciencias.uchile.cl.
† Universitat Jaume I.
‡ Universidad de Chile.
§ Universidad Técnica Federico Santa Maria.

10.1021/ct0499276 CCC: $30.25
Published on Web 11/20/2004

Chart 1. Molecules Analyzed



Its numerical values are conveniently mapped on the interval (0,1) facilitating its analysis and interpretation. According to the interpretation of the ELF, a region of the space with a high value of ELF corresponds to a region where it is more probable to localize an electron or a pair of electrons. Hence, local maximum (attractors) of the ELF characterize these regions, and the volume enclosed by all gradient lines which end up at one attractor is called a basin. At low values of the ELF, the volume enclosed by the respective isosurface may contain more than one attractor, and eventually, for a sufficiently low value, all the attractors will be contained in the isosurface. When the isosurface goes to larger values of the ELF, the basins begin to split, and, finally, the respective isosurface shows all the basins separated. This process is conveniently followed by means of a bifurcation diagram.[31] The bifurcation points have been interpreted as a measure of the interaction among the different basins and, chemically, as a measure of electron delocalization.[31,32] Here, this scheme will be used to construct a σ and π aromaticity scale. This scale will be validated with a series of well-known molecules (**I-XIV**) and two recently proposed aromatic molecules N_5^- (**XV**)[33] and B_6CO_6 (**XVI**)[34] and then used on the new aluminum based clusters, **XVII**, whose aromatic properties have recently been described,[35−38] and **XVIII** and **XIX** where the net aromaticity of both systems has been a matter of recent controversy.[22−24]

We will show that the average between the two bifurcation values provides a useful measure of the global aromatic character of a molecular system.

Computational Details

The molecules in Chart 1 have been studied. All molecular geometries were optimized at the B3LYP/6-31G(d) level of theory using the GAUSSIAN98[39] package of programs. The π and σ orbitals were separated to form the π and σ densities, respectively. The separated ELF, ELF_π, and the ELF_σ, were constructed using the TopMod[40] software. The visualization of the isosurfaces and the bifurcation points were done using the Vis5d[41] program.

Results and Discussion

The $\sigma-\pi$ separation of the ELF rest upon the additive character of the kinetic energy and reflects the symmetry properties of the molecular system. It is expected that the main characteristics of the π cloud should be reflected in the ELF_π. In particular, a high value of bifurcation implies that the minimum in the ELF is high and the respective basins do not separate each other. Therefore, in this situation the basins are correlated, and one can expect that the higher the bifurcation value the more aromatic the system will be.

Table 1. Bifurcation Values of ELF_σ and ELF_π and Their Average for All Molecules Analyzed[a]

molecules	ELF_σ	ELF_π	average
I	0.76	0.91	0.84
II	0.76	0.78	0.77
III	0.77	0.70	0.74
IV	0.75	0.64	0.70
V	0.76	0.75	0.76
VI	0.75	0.82	0.79
VII	0.74	0.69	0.72
VIII	0.76	0.72	0.74
IX	0.76	0.74	0.75
X	0.76	0.75	0.76
XI	0.77	0.71	0.74
XII	0.73	0.35	0.54
XIII	0.79	0.11	0.45
XIV	0.78	0.15	0.47
XV	0.81	0.78	0.80
XVI	0.68	0.85	0.77
XVII	0.88	0.99	0.94
XVIII	0.86	0.08	0.47
XVIII	0.86	0.41	0.64
XIX	0.87	0.24	0.56

[a] The two ELFπ bifurcation values of compound XVIII correspond to the separation of the basins over and under the plane of the Al_4 ring, respectively.

Table 1 displays the bifurcation ELF_σ and ELF_π values and their average for all the molecules considered in this study. There are some clear trends worth commenting. For the traditional π aromatic molecules (**I** to **XI**), the sigma bifurcation occurs at ELF_σ values around 0.75 with small deviations, and the ELF_π bifurcation values correspond perfectly with the aromatic degree of the molecular systems as was previously discussed.[21] Notice that the antiaromatic molecules (**XII** to **XIV**) have a remarkable low bifurcation

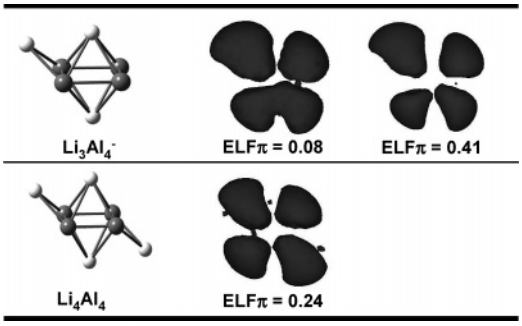


Figure 1. π System bifurcation of the Al_4^{4-} ring in $Li_3Al_4^-$ and Li_4Al_4.

ELF_π value. On the other hand, the metallic systems (**XVIII** and **XIX**) and the N_5^- ion (**XV**) present a high value of the bifurcation ELF_σ which is related to their σ aromaticity. For the last systems, the ELF_π bifurcation values show great variations thereby reflecting the fact that some of them are π antiaromatic (**XVIII** and **XIX**). The opposite character of the π and σ parts of the $Li_3Al_4^-$ and Li_4Al_4 metallic systems has raised the question about the global aromatic or antiaromatic character of these clusters. In an attempt to bring a new perspective to the point, the average value of the two bifurcations for each molecule has been calculated. They are displayed in the third column of Table 1. For the classic organic compounds, one can clearly observe that the aromatic molecules present values higher than 0.70 (**I** to **XI**), and the antiaromatic ones have values around 0.55 or lower (**XII** to **XIV**).

Using this range of values to classify the new systems, **XV**, **XVI**, and **XVII** are clearly aromatic. The situation is different for the $Li_3Al_4^-$ ion (**XVIII**). It presents two different types of ELF_π bifurcation as can be seen in Figure 1. The first bifurcation (0.08) corresponds to the separation of the basins over the plane of the ring, and the other bifurcation (0.41) is associated with the separation of the basins under the plane of the ring. This is surely due to the asymmetric capping of the lithium atom. Depending on which value one takes, the average is 0.47 or 0.64. In both cases, the molecule does not enter into the category of aromatic in the proposed scale. However, to provide additional evidence, the symmetric Li_4Al_4 cluster (**XIX**) has also been studied. The ELF_π bifurcation is now well defined, as can be seen in Figure 1, and has a low value of 0.24. The ELF_σ bifurcation occurs at a value of 0.87, and the average is 0.56, similar to the value obtained for cycleoctatetraene (**XII**), an antiaromatic molecule. It is interesting to note that this value is similar to the average between the two bifurcation values for the $Li_3Al_4^-$ ion (**XVIII**).

Conclusions

The average value of the ELF_σ and ELF_π bifurcations can be used to construct a general scale to measure the aromaticity of a molecular system. It works well for known organic and metallic aromatic and antiaromatic systems. The proposed scale predicts an overall antiaromatic character for the controversial Al_4^{4-}-based clusters, built from σ aromatic and π antiaromatic contributions.

Acknowledgment. This work has been supported by Universidad Técnica Federico Santa María grant UTFSM 130423, MIDEPLAN, through Millennium Nucleus for Applied Quantum Mechanics and Computational Chemistry P02-004-F and FONDECYT Grants 1010649 and 1030548. J.C.S. thanks Generalitat Valenciana and Fundación Bancaixa-Universitat Jaume I, Spain, for financial support.

References

(1) Zhou, Z.; Parr, R. G. *J. Am. Chem. Soc*. **1989**, *111*, 7371.

(2) Minkin, V. I.; Glukhovtev, M. N.; Simkin, B. Y. *Aromaticity and Antiaromaticty, Electronic and Structural Aspects*; John Wiley and Sons: New York, 1994.

(3) Katritzky, A. R.; Karelson, M.; Sild, S.; Krygowski, T. M.; Jug, K. *J. Org. Chem*. **1998**, *63*, 5228.

(4) Schleyer, P. v. R.; Freeman, P. K.; Jiao, H.; Goldfuss, B. *Angew. Chem., Int. Ed. Engl*. **1995**, *16*, 337.

(5) Schleyer, P. v. R.; Maerker, C.; Dransfeld, A.; Jiao, H.; Hommes, N. *J. Am. Chem. Soc*. **1996**, *118*, 6317.

(6) Cyrañski, M.; Krygowski, T.; Katritsky, A.; Schleyer, P. v. R. *J. Org. Chem*. **2002**, *67*, 1333.

(7) Schleyer, P. v. R.; Manoharan, M.; Wang, Z.; Kiran, B.; Jiao, H.; Puchta, R.; Hommes, N. *Org. Lett*. **2001**, *3*, 16, 2465.

(8) Schleyer, P.v. R. *Chem Rev*. **2001**, *101*, 1115.

(9) Fleischer, U.; Kutzelnigg, W.; Lazzeretti, P.; Muhlenkamp, V. *J. Am. Chem. Soc*. **1994**, *116*, 5298.

(10) Balawander, R.; Komorowsky, L.; De Proft, F.; Geerlings, P. *J. Phys. Chem. A* **1998**, *102*, 9912.

(11) De Proft, F.; Geerlings, P. *Chem. Rev*. **2001**, *101*, 1451.

(12) Poater, J.; Fradera, X.; Duran, M.; Sola, M. *Chem. Eur. J*. **2003**, *9*, 400.

(13) Poater, J.; Fradera, X.; Duran, M.; Sola, M. *Chem. Eur. J*. **2003**, *9*, 1113.

(14) Chesnut, D. B.; Bartolotti, L. J. *Chem. Phys*. **2000**, *253*, 1.

(15) Fuster, F.; Savin, A.; Silvi, B. *J. Phys. Chem. A* **2000**, *104*, 852.

(16) Sakai, S. *J. Phys. Chem. A* **2002**, *106*, 10370.

(17) Sakai, S. *J. Phys. Chem. A* **2002**, *106*, 11526.

(18) Rassat, A. *Phys. Chem. Chem. Phys*. **2004**, *6*, 232.

(19) Lazzeretti, P. *Phys. Chem. Chem. Phys*. **2004**, *6*, 217.

(20) Steiner, V.; Fowler, P. W. *Phys. Chem. Chem. Phys*. **2004**, *6*, 261.

(21) Krygowski, T. M.; Cyranski, M. K. *Chem. Rev*. **2001**, *101*, 1385.

(22) Ritter, S. *Chem. Eng. News* **2003**, *81*, 23.

(23) Kuznetsov, A.; Birch, K.; Boldyrev, A.; Li, X.; Zhai, H.; Wang, L. *Science* **2003**, *300*, 622.

(24) Chen, Z.; Corminboeuf, C.; Heine, T.; Bohmann, J.; Schleyer, P. v. R. *J. Am. Chem. Soc*. **2003**, *125*, 13930.

(25) Kruszewski, J.; Krygowski, T. M. *Tetrahedron Lett*. **1972**, 3839.

(26) Krygowski, T. M.; Cyranski, M. K. *Tetrahedron* **1996**, *52*, 1713.

(27) Silvi, B.; Savin, A. *Nature* **1994**, *371*, 683.

(28) Becke, A.; Edgecombe, K. *J. Chem. Phys*. **1990**, *92*, 5397.

(29) Silvi, B. *Phys. Chem. Chem. Phys*. **2004**, *6*, 256.

(30) Santos, J. C.; Tiznado, W.; Contreras, R.; Fuentealba, P. *J. Chem. Phys*. **2004**, *120*, 1670.

(31) Savin, A.; Silvi, B.; Colonna, F. *Can. J. Chem*. **1996**, *74*, 1088.

(32) Savin, A.; Becke, A.; Flad, D.; Nesper, R.; Preuss, H.; Schnering, H. V. *Angew. Chem., Int. Ed. Engl*. **1991**, *30*, 409.

(33) Vij, A.; Pavlovich, J.; Wilson, W.; Vij, V.; Christie, K. *Angew. Chem., Int. Ed. Engl*. **2002**, *41*, 16.

(34) Wu, H.; Jiao, H.; Wang, Z.; Scheleyer, P. v. R. *J. Am. Chem. Soc*. **2003**, *125*, 4428.

(35) Li, X.; Kuznetsov, A.; Zhang, H.; Boldyrev, A.; Wang, L. *Science* **2001**, *291*, 859.

(36) Kuznetsov, A.; Boldyrev, A.; Li, X.; Wang, L. *J. Am. Chem. Soc*. **2001**, *123*, 8825.

(37) Boldyrev, A.; Kuznetsov, A. *Inorg. Chem*. **2002**, *41*, 532.

(38) Zhan, C.; Zheng, F.; Dixon, D. *J. Am. Chem. Soc*. **2002**, *124*, 14795.

(39) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A.; Stratmann, R. E. Jr.; 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.; 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.; 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.9; Gaussian Inc.; Pittsburgh, PA, 1998.

(40) Noury, S.; Krokidis, X.; Fuster, F.; Silvi, B. TopMoD Package; Universite Pierre et Marie Curie, 1997.

(41) Hibbard, B.; Kellum, J.; Paul, B. Vis5d 5.1, Visualization Project; University of Winsconsin-Madison Space Science and Engineering Center (SSEC); B. Hibbard and D. Santek, Proc. IEEE Visualization ’90, 1990; p 129.

CT0499276

JCTC Journal of Chemical Theory and Computation

A Simple Mathematical Model for the Cooperative and Competitive Substituent Effects Found in the Cope Rearrangements of Phenyl-Substituted 1,5-Hexadienes

David A. Hrovat and Weston Thatcher Borden*

Department of Chemistry, Box 351700, University of Washington, Seattle, Washington 98195-1700

Received September 28, 2004

Abstract: A simple mathematical model is derived for the dependence of the substituent effects in the Cope rearrangements of substituted 1,5-hexadienes on the interallylic bond lengths in the transition structures. The model qualitatively reproduces the cooperative substituent effects that have been found in the Cope rearrangements of 2,5-diphenyl- and 1,3,4,6-tetraphenyl-1,5-hexadiene and the competitive substituent effects that have been found in the Cope rearrangement of 1,3,5-triphenyl-1,5-hexadiene. B3LYP/6-31G* calculations have been performed in order to assess more quantitatively the performance of the model, and the results of these calculations are discussed.

The effects of substituents on $\Delta H^{\ddagger}$ for the Cope rearrangement[1] can be either cooperative or competitive.[2] For example, as shown in Table 1, a single phenyl substituent at C2 lowers $\Delta H^{\ddagger}$ for the Cope rearrangement of 1,5-hexadiene[3] by 4.2 kcal/mol,[4] but a second phenyl group, attached to C5, lowers the experimental enthalpy of activation by 8.0 kcal/mol, nearly twice as much.[4,5] Clearly, there is a strong cooperative effect between phenyl substituents at C2 and C5 on lowering the barrier height.

Phenyl groups at C1, C3, C4, and C6 of 1,5-hexadiene also give rise to a strongly cooperative substituent effect. When one phenyl group is attached to C1 and another to C3, the barrier to the Cope rearrangement is decreased by 3.0 kcal/mol from that for the unsubstituted molecule.[6] If the phenyl substituent effects on the Cope rearrangement were additive, augmentation of the phenyl groups at C1 and C3 by another pair at C4 and C6 would be anticipated to lower $\Delta H^{\ddagger}$ by another 3.0 kcal/mol. However, the measured effect of the additional pair of phenyl groups is three times larger, lowering $\Delta H^{\ddagger}$ for the Cope rearrangement by an additional 9.2 kcal/mol.[7] As in the case of phenyl substituents at C2 and C5, the effect of phenyl substituents at C1, C3, C4, and C6 of 1,5-hexadiene is obviously cooperative.

* Corresponding author e-mail: borden@unt.edu. Address correspondence to Department of Chemistry, University of North Texas, Box 305070, Denton, TX 76203-5070.

In contrast to the cooperative substituent effects described above, placement of phenyl groups at C1, C3, and C5 of 1,5-hexadiene results in a competitive substituent effect.[6] As already noted, a single phenyl group at C2(C5) lowers $\Delta H^{\ddagger}$ by 4.2 kcal/mol, and phenyl groups at C1 and C3 lower $\Delta H^{\ddagger}$ by 3.0 kcal/mol. However, the simultaneous presence of phenyl groups at C1, C3, and C5 lowers $\Delta H^{\ddagger}$ by 5.7 kcal/mol, 1.5 kcal/mol *less* than the 7.2 kcal/mol expected on the basis of substituent effect additivity.

Table 1 shows that B3LYP/6-31G* calculations give activation enthalpies for the Cope rearrangement of 1,5-hexadiene[8] and of phenyl-substituted derivatives[2b] that are in very good agreement with experiment. This agreement lends credibility to the calculated dependence in Table 1 of the interallylic distance, *R*, in Cope transition structures (TSs) on the number and placement of radical stabilizing, phenyl substituents. Phenyl groups at C2 and C5, which can stabilize diradical structure **A** in Figure 1, are predicted to shorten *R*, whereas phenyls at C1, C3, C4, and C6, which stabilize diradical structure **C**, are predicted to lengthen *R*.

Direct experimental evidence that the Cope rearrangement has a variable TS, in which the bond lengths change with the presence and placement of radical-stabilizing substituents, comes from the secondary kinetic isotope effects, measured by Gajewski and Conrad.[9] Doering has coined the term "chameleonic" to describe the Cope TS and the ability of

10.1021/ct049929q CCC: $30.25
Published on Web 12/07/2004

Table 1. Effect of Phenyl Substituents on the Calculated and Experimental Activation Enthalpies (kcal/mol) for the Cope Rearrangements of Phenyl Substituted 1,5-Hexadienes and on the Differences (Å) between the Calculated Interallylic Distance, R, in Each TS and R_0 in the TS for Unsubstituted 1,5-Hexadiene

substituents	method[a]	$\Delta\Delta H^{\ddagger}$ (calc.)	$\Delta\Delta H^{\ddagger}$ (exp.)	$R - R_0$[b]
none	B3LYP	0[c]	0[d]	0[e]
2-phenyl	B3LYP	−2.8		−0.128, −0.144
2-phenyl	UB3LYP	−2.9		−0.188, −0.265
2-phenyl	UB3LYP	−4.0[f]	−4.2[g]	−0.366
1,3-diphenyl	B3LYP	−3.0	−3.0[h]	0.253
2,5-diphenyl	B3LYP	−8.1		−0.171, −0.285
2,5-diphenyl	UB3LYP	−8.4		−0.126, −0.298
2,5-diphenyl	UB3LYP	−11.9[f]	−12.2[g,i]	−0.389
1,3,5-triphenyl	B3LYP	−4.0	−5.7[h]	0.148, 0.141
1,3,4,6-tetraphenyl	B3LYP	−14.1	−12.2[j]	0.684

[a] The 6-31G* basis set was used for all of the calculations in this table. [b] Difference for, respectively, the forming bond in the TS and the breaking bond, in the TS. Just one difference between the interallylic bond lengths R and R_0 is given for those TSs that have a plane of symmetry. [c] Enthalpy relative to $\Delta H^{\ddagger} = 33.2$ kcal/mol. [d] Enthalpy relative to the value of $\Delta H^{\ddagger} = 33.5$ kcal/mol, reported in ref 3. [e] $R_0 = 1.965$ Å. [f] Enthalpy, relative to the reactant, of the diradicaloid intermediate. [g] Reference 4. [h] Reference 6. [i] Reference 5. [j] Reference 7.

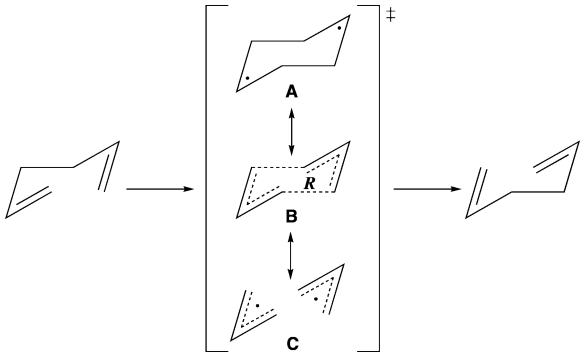


Figure 1. Depiction of possible TSs for the Cope rearrangement. In the two diradical extremes bond making either precedes (**A**) or lags behind (**C**) bond breaking, while in TS **B** bond making and bond breaking occur synchronously. The interallylic distance (R) is expected to increase in going from TS **A** to TS **C**.

the bond lengths in it to altar in response to radical stabilizing substituents.[5]

In this paper we present a simple mathematical model for the dependence of the energy of the Cope TS on the presence and placement of radical stabilizing substituents. This model is helpful for analyzing how the chameleonic character of the Cope TS gives rise to the cooperative and competitive substituent effects that are evident in the values of $\Delta\Delta H^{\ddagger}$ in Table 1.

A Simple Mathematical Model for the Effects of Substituents

The optimal value of R, the interallylic distance in the TS for the Cope rearrangement of a substituted 1,5-hexadiene, can be viewed as a compromise between maximizing the magnitude of the substituent stabilization energy of the TS (ΔE_{subst})[10] and minimizing the energetic cost (ΔE_{dist}) of distorting the interallylic bond lengths in the TS from the optimal value of R_0 (= 1.965 Å at the B3LYP/6-31G* level)[8] for the Cope rearrangement of unsubstituted 1,5-hexadiene. Thus, the optimal value of R maximizes the magnitude of the net TS stabilization energy (ΔE_{stab}),[10] which is the sum of ΔE_{dist} and ΔE_{subst}.

$$\Delta E_{\text{stab}} = \Delta E_{\text{dist}} + \Delta E_{\text{subst}} \tag{1}$$

The simplest possible mathematical model for finding this value of R assumes that the energy of the C_{2h} structure in the chair Cope rearrangement of unsubstituted 1,5-hexadiene increases quadratically with the difference between R and R_0. The energy required to distort the interallylic distance from R_0 is then given by

$$\Delta E_{\text{dist}} = k_{\text{dist}}(R - R_0)^2/2 \tag{2}$$

A mathematical expression for ΔE_{subst} can be derived by considering the effect of phenyl substituents at, for instance, C1 and C3 on the energy of the high-symmetry chair structure. Since radical stabilizing groups at C1 and C3 lower the energy of structure **C** in Figure 1, they will have their greatest effect at values of $R > R_0$, where the wave function becomes increasingly like that for two allyl radicals. Conversely, the energy lowering provided by the phenyl substituents at C1 and C3 will decrease in size with decreasing values of R. Therefore, the leading term in the expression for the effect of the interallylic distance on the stabilization of the high-symmetry chair structure by phenyl groups at C1 and C3 must be linear in R

$$\Delta E_{\text{subst}} - \Delta E^0_{\text{subst}} = -k_{\text{subst}}(R - R_0) \tag{3}$$

where $\Delta E^0_{\text{subst}}$ is the stabilization provided by the phenyl substituents at R_0, and k_{subst} is positive.

Thus, at least for values of R around R_0, the dependence on R of the energy of the high-symmetry species in the chair Cope rearrangement of 1,3-diphenyl-1,5-hexadiene can be written as

$$\Delta E_{\text{stab}} - \Delta E^0_{\text{stab}} = \Delta E_{\text{dist}} + \Delta E_{\text{subst}} - \Delta E^0_{\text{subst}}$$

$$= k_{\text{dist}}(R - R_0)^2/2 - k_{\text{subst}}(R - R_0) \tag{4}$$

where $\Delta E^0_{\text{stab}} = \Delta E^0_{\text{subst}}$ is the substituent stabilization energy at R_0.

Differentiating eq 4 with respect to R, the high-symmetry chair structure of minimum energy (i.e. the TS) for the Cope rearrangement of 1,3-diphenyl-1,5-hexadiene is predicted to occur at

$$R - R_0 = k_{\text{subst}}/k_{\text{dist}} \tag{5}$$

Substituting the expression for $R - R_0$ in eq 5 back into eq 4, the energy difference between the TS and the chair

geometry at $R = R_0$ is given by

$$\Delta E_{stab} - \Delta E^0_{stab} = \Delta E_{dist} + \Delta E_{subst} - \Delta E^0_{subst}$$
$$= k_{dist}(k_{subst}/k_{dist})^2/2 - k_{subst}(k_{subst}/k_{dist})$$
$$= k_{subst}{}^2/2k_{dist} - k_{subst}{}^2/k_{dist}$$
$$= -k_{subst}{}^2/2k_{dist} \quad (6)$$

Eq 6 predicts that at the optimal value of R

$$\Delta E_{stab} - \Delta E^0_{stab} = -\Delta E_{dist} = (\Delta E_{subst} - \Delta E^0_{subst})/2 \quad (7)$$

This relationship between ΔE_{dist} and $(\Delta E_{subst} - \Delta E^0_{subst})/2$ is analogous to that in the virial theorem,[11a] with ΔE_{dist} and $(\Delta E_{subst} - \Delta E^0_{subst})$ in eq 7 corresponding, respectively, to kinetic and potential energy.

If a second pair of phenyl substituents is added to 1,3-diphenyl-1,5-hexadiene at C4 and C6, the second term in eq 4 is multiplied by a factor of 2. The TS for the Cope rearrangement is then predicted to have an interallylic distance, R, such that

$$R - R_0 = 2k_{subst}/k_{dist} \quad (8)$$

Upon increasing R from R_0 to $R_0 + 2k_{subst}/k_{dist}$, the energy of the C_{2h} structure for the Cope rearrangement of 1,3,4,6-tetraphenyl-1,5-hexadiene is predicted to decrease by

$$\Delta E_{stab} - \Delta E^0_{stab} = k_{dist}(2k_{subst}/k_{dist})^2/2 - 2k_{subst}(2\,k_{subst}/k_{dist})$$
$$= -2k_{subst}{}^2/k_{dist} \quad (9)$$

The net stabilization energy in eq 9 is four times larger in magnitude than that in eq 6 for the chair TS for the Cope rearrangement of 1,3-diphenyl-1,5-hexadiene. Thus, if ΔE^0_{stab} in eqs 6 and 9 is small, the simple mathematical model in eq 4 correctly predicts the existence of the cooperative substituent effect in 1,3,4,6-tetraphenyl-1,5-hexadiene, whereby the four phenyl substituents lower $\Delta H^{\ddagger}$ for the Cope rearrangement by four times more than the pair of phenyl substituents in 1,3-diphenyl-1,5-hexadiene.[7]

The origin of the cooperative effect, predicted by the model, can be easily seen in eq 9. Doubling the number of phenyl substituents doubles the number of radical stabilizing groups that are positioned to stabilize structure **C** in Figure 1. In addition, doubling the number of phenyl substituents is predicted by eq 8 to make the difference between R and R_0 in the TS twice as large for 1,3,4,6-tetraphenyl-1,5-hexadiene (eq 8) as for 1,3-diphenyl-1,5-hexadiene (eq 5), thus doubling the stabilization of the TS that is provided by each phenyl group. Consequently, adding phenyl substituents to C4 and C6 of 1,3-diphenyl-1,5-hexadiene is predicted to quadruple $\Delta E_{subst} - \Delta E^0_{subst}$.

The doubling of $R - R_0$ in the TS for the Cope rearrangement of 1,3,4,6-tetraphenyl-1,5-hexadiene is also predicted to quadruple ΔE_{dist}, which depends quadratically on $R - R_0$. Since doubling the number of phenyl substituents quadruples both $\Delta E^R_{subst} - \Delta E^0_{subst}$ and ΔE_{dist}, $\Delta E_{stab} - \Delta E^0_{stab}$ in eq 9 for the Cope rearrangement of 1,3,4,6-tetraphenyl-1,5-hexadiene is four times larger than $\Delta E_{stab} - \Delta E^0_{stab}$ in eq 6 for the Cope rearrangement of 1,3-diphenyl-1,5-hexadiene.

Although the simple mathematical model, embodied in eq 4, apparently predicts the size of the cooperative substituent effect in the Cope rearrangement of 1,3,4,6-tetraphenyl-1,5-hexadiene[2b,7] the model fails to predict correctly the calculated value of R in the TS for this reaction. The model predicts that $R - R_0$ in the TS for the Cope rearrangement of 1,3,4,6-tetraphenyl-1,5-hexadiene (eq 8) should be twice as large as $R - R_0$ in the TS for the Cope rearrangement of 1,3-diphenyl-1,5-hexadiene (eq 5). However, Table 1 shows the B3LYP/6-31G* ratio of $R - R_0$ values actually is 0.684/0.253 = 2.70, 35% higher than expected.

This failure of the model could reside in the assumption that ΔE_{dist} is quadratic over the difference of nearly 0.7 Å between R_0 and R in the TS for the Cope rearrangement of 1,3,4,6-tetraphenyl-1,5-hexadiene. However, another possibility is that relief of steric repulsion between the pairs of phenyl groups at C1−C6 and C3−C4 in the C_{2h} chair TS geometry has the effect of making k_{subst} for 1,3,4,6-tetraphenyl-1,5-hexadiene more than a factor of 2 larger than k_{subst} for 1,3-diphenyl-1,5-hexadiene. Results of calculations, described in the next section, indicate that both effects contribute to the ratio of $R - R_0$ values in the TSs for the Cope rearrangements of these two dienes being 35% higher than expected.

The simple model in eq 4 also fails to predict correctly the ratio of $R - R_0$ values in the TSs for the Cope rearrangements of 2-phenyl- and 2,5-diphenyl-1,5-hexadiene. Phenyl groups at C2 and C5 provide stabilization for structure **A** in Figure 1. Consequently, they have their greatest stabilizing effect at values of $R < R_0$, where the wave function in the TS becomes increasingly like that for cyclohexane-1,4-diyl. As a result, the sign of the right-hand side of eq 3 is positive for 2-phenyl- and 2,5-diphenyl-1,5-hexadiene. Therefore, for the Cope rearrangement of 2-phenyl-1,5-hexadiene, eq 3 must be replaced by

$$\Delta E_{subst} - \Delta E^0_{subst} = k'_{subst}(R - R_0) \quad (10)$$

where, like k_{subst}, $k'_{subst} > 0$.

The second phenyl group in 2,5-diphenyl-1,5-hexadiene makes the expression for $\Delta E_{subst} - \Delta E^0_{subst}$ a factor of 2 greater than the expression in eq 10 for 2-phenyl-1,5-hexadiene. Thus, replacing k_{subst} in eqs 4 and 8 by $-k'_{subst}$, the model again predicts a ratio of 2.0 for $R - R_0$ in the Cope TS for these two dienes. However, the UB3LYP/6-31G* ratio is only −0.389/−0.366 = 1.06.

The problem here is certainly the assumption in eq 2 that ΔE_{dist} is quadratic. If the ratio of $R - R_0$ values really were 2.0, the UB3LYP value of R = 1.599 Å in the TS for the Cope rearrangement of 2-phenyl-1,5-hexadiene would give R = 1.233 Å in the TS for the Cope rearrangement of 2,5-diphenyl-1,5-hexadiene. This predicted bond length is about 0.3 Å shorter than the typical length of a C−C single bond.

If a Morse, rather than a quadratic potential is used, the energy for compressing a C−C single bond to a length substantially below 1.54 Å increases exponentially with

decreasing R.[11b] Consequently, the assumption of a quadratic potential for ΔE_{dist} in eq 2 greatly underestimates the energetic cost of decreasing R in the TS for the Cope rearrangement of 2,5-diphenyl-1,5-hexadiene below the value of $R = 1.599$ Å in the TS for the Cope rearrangement of 2-phenyl-1,5-hexadiene.

Despite the failure of the simple model in eq 4 to predict the size of R in the TS for the Cope rearrangement of 2,5-diphenyl-1,5-hexadiene, the model does predict the existence of the cooperative effect of the pair of phenyl groups on the enthalpy of activation.[4,5] Assuming for 2-phenyl-1,5-hexadiene that

$$R - R_0 = -k'_{subst}/k_{dist} \tag{11}$$

then

$$\Delta E_{stab} - \Delta E^0_{stab} = k_{dist}(-k'_{subst}/k_{dist})^2/2 + k'_{subst}(-k'_{subst}/k_{dist})$$
$$= -k'^2_{subst}/2k_{dist} \tag{12}$$

Since UB3LYP/6-31G* calculations find that R is nearly the same in the TSs for the Cope rearrangements of 2-phenyl- and 2,5-diphenyl-1,5-hexadiene, ΔE_{dist} is approximately the same for both TSs. However, due to the presence of two phenyl groups in 2,5-diphenyl-1,5-hexadiene, the term for $\Delta E_{subst} - \Delta E^0_{subst}$ is a factor of 2 larger in the TS for the latter reaction than in the TS for the former. Therefore, the expression for $\Delta E_{stab} - \Delta E^0_{stab}$ in the latter reaction is

$$\Delta E_{stab} - \Delta E^0_{stab} = k_{dist}(-k'_{subst}/k_{dist})^2/2 + 2k'_{subst}(-k'_{subst}/k_{dist})$$
$$= -3k'^2_{subst}/2k_{dist} \tag{13}$$

Comparison of eq 13 with eq 12 results in the prediction that $\Delta E_{stab} - \Delta E^0_{stab}$ should be a factor of 3 larger for the Cope rearrangement of 2,5-diphenyl-1,5-hexadiene than for the Cope rearrangement of 2-phenyl-1,5-hexadiene. Assuming again that ΔE^0_{stab} is small, this prediction is consistent with both the experimental and UB3LYP values of $\Delta\Delta H^{\ddagger}$ in Table 1.

The mathematical reason for the prediction, that $\Delta E_{stab} - \Delta E^0_{stab}$ should be a factor of 3 larger for the Cope rearrangement of 2,5-diphenyl-1,5-hexadiene than for the Cope rearrangement of 2-phenyl-1,5-hexadiene, is that in eq 12 the first term, which gives ΔE_{dist}, cancels half of the second term, which gives $\Delta E_{subst} - \Delta E^0_{subst}$. In eq 13 the term for ΔE_{dist} is the same as in eq 12; but the term for $\Delta E_{subst} - \Delta E^0_{subst}$ is twice as large, which makes the difference between these two terms three times larger in eq 13 than in eq 12.

In physical terms, the second phenyl group in 2,5-diphenyl-1,5-hexadiene provides the same amount of TS stabilization as the lone phenyl group in 2-phenyl-1,5-hexadiene. However, because the interallylic distances, R, in the two TSs are nearly the same, the stabilization of the TS by the second phenyl group comes without the energetic cost of $(\Delta E_{stab} - \Delta E^0_{stab})/2$, paid by the first phenyl group, for distorting R from R_0 in the TS for the Cope rearrangement of unsubstituted 1,5-hexadiene.

The simple mathematical model in eq 4 also accounts for the competitive phenyl substituent effects, computed for[2b] and found in[5] the Cope rearrangement of 1,3,5-triphenyl-1,5-hexadiene. Combining eqs 3 and 10 gives the expression for $\Delta E_{subst} - \Delta E^0_{subst}$ in the TS for the Cope rearrangement of this compound as

$$\Delta E_{subst} - \Delta E^0_{subst} = (k'_{subst} - k_{subst})(R - R_0) \tag{14}$$

Substituting this expression for $\Delta E_{subst} - \Delta E^0_{subst}$ into eq 4 and maximizing the magnitude of ΔE_{stab} gives

$$R - R_0 = (k_{subst} - k'_{subst})/k_{dist}$$
$$= k_{subst}/k_{dist} - k'_{subst}/k_{dist} \tag{15}$$

According to eq 15, $R - R_0$ in the TS for the Cope rearrangement of 1,3,5-triphenyl-1,5-hexadiene should be equal to the algebraic sum of $R - R_0$ in the TS for 1,3-diphenyl-1,5-hexadiene in eq 5 and $R - R_0$ in the TS for 2-phenyl-1,5-hexadiene in eq 11. Using the B3LYP results in Table 1,[12] $R - R_0 = 2.110 - 1.965 = 0.145$ Å is, indeed, reasonably close to the sum of $0.253 - 0.136 = 0.117$ Å.

Upon increasing the interallylic distance from $R_0 = 1.965$ Å to $R = R_0 + (k_{subst} - k'_{subst})k_{dist}$, the net stabilization of the TS for the Cope rearrangement of 1,3,5-triphenyl-1,5-hexadiene is given by

$$\Delta E_{stab} - \Delta E^0_{stab} = -(k_{subst} - k'_{subst})^2/2k_{dist} \tag{16}$$

$\Delta E_{stab} - \Delta E^0_{stab}$ in eq 16 is smaller in size than the sum of the net stabilization energies of the TSs for the Cope rearrangements of 1,3-diphenyl-1,5-hexadiene (eq 6) and 2-phenyl-1,5-hexadiene (eq 12) by

$$-(k_{subst} - k'_{subst})^2/2k_{dist} + k'^2_{subst}/2k_{dist} + k^2_{subst}/2k_{dist} = k_{subst}k'_{subst}/k_{dist} \tag{17}$$

Thus, eq 17 predicts the existence of the competitive substituent effect that is computed for and found in the Cope rearrangement of 1,3,5-triphenyl-1,5-hexadiene.

In physical terms the existence of this effect is due to the fact that phenyls at C1 and C3 provide the most stabilization of the high-symmetry chair structure for the Cope rearrangement at $R > R_0$, whereas a phenyl group at C5 provides the most stabilization at $R < R_0$. At the compromise value of R in the TS for the Cope rearrangement of 1,3,5-triphenyl-1,5-hexadiene, the phenyl groups at C1 and C3 provide less stabilization than they do in the TS for the Cope rearrangement of 1,3-diphenyl-1,5-hexadiene, and the phenyl group at C5 provides less stabilization than it does in the TS for the Cope rearrangement of 2-phenyl-1,5-hexadiene. This is the origin of the competitive substituent effect in the Cope rearrangent of 1,3,5-triphenyl-1,5-hexadiene.

Additional Computational Tests of the Mathematical Model

The simple mathematical model in eq 4 predicts successfully the existence of the cooperative and competitive phenyl substituent effects, seen in Table 1, but only if the differences between $\Delta E_{stab} - \Delta E^0_{stab}$ values, derived from the model, can

Table 2. Dissection of the Effects of Phenyl Substituents on Lowering the Energy of the TS for the Chair Cope Rearrangement of 1,5-Hexadiene[a]

substituents	$R - R_0$ (Å)	ΔE_{dist}	$\Delta E_{subst} - \Delta E^0_{subst}$[b]	$\Delta E_{stab} - \Delta E^0_{stab}$	ΔE^0_{stab}[b]	ΔE_{stab}
2-phenyl[c]	−0.366	3.3	−6.2	−2.9	−1.7	−4.6
2-phenyl[d]	0.145	0.7	1.4	2.1	−1.7	0.4
2,5-diphenyl[c]	−0.389	3.7	−12.9	−9.2	−3.7	−12.9
1,3-diphenyl	0.253	2.0	−4.0	−2.0	−0.7	−2.7
1,3-diphenyl[d]	0.145	0.7	−2.4	−1.7	−0.7	−2.4
1,3,5-triphenyl	0.145	0.7	−1.1	−0.4	−2.8	−3.2
1,3,4,6-tetraphenyl	0.684	12.0	−23.7	−11.7	−1.4	−13.1
1,3,4,6-tetraphenyl[e]	0.253	2.0	−9.3	−7.3	−1.4	−8.7

[a] Energies (kcal/mol) were obtained from B3LYP/6-31G* calculations in the manner described in the text. [b] $\Delta E^0_{subst} = \Delta E^0_{stab}$. [c] Values are from UB3LYP calculations. [d] Calculations performed at, $R = 2.110$ Å, the interallylic bond distance in the TS for Cope rearrangement of 1,3,5-triphenyl-1,5-hexadiene. [e] Calculations performed at $R = 2.218$, the interallylic bond distance in the TS for Cope rearrangement of 1,3-diphenyl-1,5-hexadiene.

be equated with the $\Delta\Delta H^{\ddagger}$ values in Table 1. The apparent success of the model thus suggests that substituent-induced lengthening or shortening of the interallylic bond lengths must be the dominant mode by which phenyl substituents stabilize the Cope TS and thus lower the enthalpy of activation for the Cope rearrangement.

$\Delta E_{stab} - \Delta E^0_{stab} \approx \Delta E_{stab}$ in eqs 6, 9, 12, 13, and 16 requires that, at least for phenyl substituents, $\Delta E_{stab} \gg \Delta E^0_{stab}$. This means that the net substituent stabilization energies must be much larger at the TS geometries than at R_0, the interallylic bond length in the TS for the Cope rearrangement of unsubstituted 1,5-hexadiene.

To test whether ΔE_{stab} is, in fact, much larger than ΔE^0_{stab}, we carried out electronic structure calculations. We also computed ΔE_{dist}, the difference between the energy of the Cope TS for unsubstituted 1,5-hexadiene at R and at R_0. From the difference between ($\Delta E_{stab} - \Delta E^0_{stab}$) and ΔE_{dist} we were able, using eq 4, to obtain ($\Delta E_{subst} - \Delta E^0_{subst}$). We were thus able to test the predicted virial-theorem-like relationship in eq 7 between $\Delta E_{stab} - \Delta E^0_{stab}$, ΔE_{dist}, and $\Delta E_{subst} - \Delta E^0_{subst}$.

As in our previous studies of Cope substituent effects,[2] the electronic structure calculations utilized Becke's 3-parameter functional,[13] in conjuction with the correlation functional of Lee, Yang, and Parr[14] and the 6-31G* basis set.[15] The B3LYP/6-31G* calculations were performed with the Gaussian 98 suite of programs.[16]

The B3LYP values of ΔE_{dist}, $\Delta E_{subst} - \Delta E^0_{subst}$, $\Delta E_{stab} - \Delta E^0_{stab}$, ΔE^0_{stab} ($=\Delta E^0_{subst}$), and ΔE_{stab} are given in Table 2 for the Cope rearrangements of 2-phenyl, 1,3- and 2,5-diphenyl-, 1,3,5-triphenyl-, and 1,3,4,6-tetraphenyl-1,5-hexadiene. Because the values of ΔE_{stab} in Table 2 are uncorrected for differences in zero-point energies or integrated heat capacities, these values are similar to but not exactly the same as the corresponding B3LYP values of $\Delta\Delta H^{\ddagger}$ in Table 1.

A number of observations can be made from inspection of Table 2. First, at R_0, the interallylic bond length in the TS for the parent Cope rearrangement, a phenyl group at C2 provides $\Delta E^0_{stab} = 1.7$ kcal/mol of stabilization, which is slightly more than twice the stabilization furnished by a *pair* of phenyl groups at C1 and C3. The larger amount of stabilization provided by a single phenyl group at C2 supports the conjecture, based on the relative enthalpies of the two diradical extremes in Figure 1,[2a] that the TS for the Cope rearrangement of unsubstituted 1,5-hexadiene resembles cyclohexane-1,4-diyl (**A**) more than two allyl radicals (**C**).[1,3]

The values of ΔE^0_{stab} for 2,5-diphenyl, 1,3,5-triphenyl, and 1,3,4,6-tetraphenyl substituents are each close to being the sum of the appropriate ΔE^0_{stab} values for 2-phenyl and 1,3-diphenyl substituents. It is noteworthy that, even for the phenyl substituents in 2-phenyl- and 2,5-diphenyl-1,5-hexadiene, ΔE^0_{stab} is only about one-third to one-quarter, respectively, of ΔE_{stab}, the total net lowering by the substituents of the energy of the Cope TS, relative to the energy of the reactants.

The fact that substituents provide a comparatively small amount of stabilization at $R_0 = 1.965$ Å was first found for cyano, rather than for phenyl substituents.[2a] This finding is consistent with the results of Staroverov and Davidson. They used local spin analysis to show that not only does the TS for the parent Cope rearrangement have little diradical character[17a] but also that, at the geometry of this TS, radical stabilizing substituents do little to enhance the diradical character of the wave function.[17b]

Another observation, which can be made from the results in Table 2, is that at the optimal values of R for the intermediate in the Cope rearrangement of 2-phenyl-1,5-hexadiene and for the TSs in the Cope rearrangements of 1,3-diphenyl- and 1,3,4,6-tetraphenyl-1,5-hexadiene, $\Delta E_{stab} - \Delta E^0_{stab} \approx -\Delta E_{dist} \approx (\Delta E_{subst} - \Delta E^0_{subst})/2$. These are the relationships predicted by eq 6 under the assumption that ΔE_{dist} is quadratic in $R - R_0$. The least-squares plot in Figure 2 shows how well the equality between $\Delta E_{stab} - \Delta E^0_{stab}$ and $-\Delta E_{dist}$ holds for the Cope rearrangements of these three dienes.

As discussed in the previous section, at small values of R, ΔE_{dist} is not quadratic in $R - R_0$, and the intermediates in the Cope rearrangments of 2-phenyl- and 2,5-diphenyl-1,5-hexadiene have UB3LYP values of R that differ by only 0.023 Å. Because the value of R in the two intermediates is nearly the same, the value of $\Delta E_{subst} - \Delta E^0_{subst} = -12.9$ kcal/mol for the UB3LYP diradical intermediate in the Cope rearrangement of 2,5-diphenyl-1,5-hexadiene is only about a factor of 2 larger than $\Delta E_{subst} - \Delta E^0_{subst} = -6.2$ kcal/mol for the UB3LYP diradical intermediate in the Cope rearrangement of 2-phenyl-1,5-hexadiene.

The fact that R is nearly the same in the two diradical intermediates has another consequence—upon adding a second phenyl group to C5 of the diene, the approximate doubling of $\Delta E_{subst} - \Delta E^0_{subst}$ occurs with an increase in ΔE_{dist} of only 0.4 kcal/mol. Therefore, upon addition of a

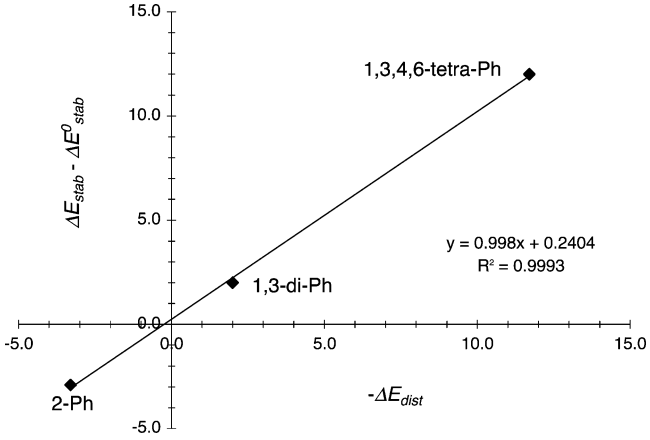


Figure 2. Linear least-squares plot of $\Delta E_{stab} - \Delta E^0_{stab}$ vs $-\Delta E_{dist}$ for 2-phenyl-, 1,3-diphenyl-, and 1,3,4,6-tetraphenyl-1,5-hexadiene.

second phenyl group at C5, the 6.3 kcal/mol increase in the size of the net stabilization energy, $\Delta E_{stab} - \Delta E^0_{stab}$, is only 0.4 kcal/mol less than the 6.7 kcal/mol increase in the size of $\Delta E_{subst} - \Delta E^0_{subst}$.

Because addition of a phenyl group to C5 of 2-phenyl-1,5-hexadiene increases the magnitude of $\Delta E_{stab} - \Delta E^0_{stab}$ in the diradical intermediate by 6.3 kcal/mol, $\Delta E_{stab} - \Delta E^0_{stab} = -9.2$ kcal/mol for 2,5-diphenyl-1,5-hexadiene is more than a factor of 3 larger in magnitude than $\Delta E_{stab} - \Delta E^0_{stab} = -2.9$ kcal/mol for 2-phenyl-1,5-hexadiene. However, as expected, the ratio of ΔE^0_{stab} values for the two dienes is only about two. Therefore, the ratios of their ΔE_{stab} values in Table 2 and their $\Delta\Delta H^{\ddagger}$ values in Table 1 are each a little less than a factor of 3.

When a pair of phenyl substituents is added to C1 and C3 of 1,5-hexadiene, the value of R in the TS increases by 0.253 Å. However, when a second pair of phenyls is added to C4 and C6 of 1,3-diphenyl-1,5-hexadiene, the value of R in the TS increases by 0.431 Å, a 70% larger increase than for addition of the first pair of phenyl groups. Consequently, in the TS for the Cope rearrangement of 1,3,4,6-tetraphenyl-1,5-hexadiene the substituent stabilization energy, $\Delta E_{subst} - \Delta E^0_{subst} = -23.7$ kcal/mol, is not a factor of 4, but a factor of about 6 larger in size than $\Delta E_{subst} - \Delta E^0_{subst} = -4.0$ kcal/mol in the TS for the Cope rearrangement of 1,3-diphenyl-1,5-hexadiene.

When a second pair of phenyl groups is added to C4 and C6 of 1,3-diphenyl-1,5-hexadiene, but the interallylic distance is frozen at the value of $R = 2.218$ Å in the TS for the Cope rearrangement of the disubstituted diene, Table 2 shows that $\Delta E_{subst} - \Delta E^0_{subst} = -9.3$ kcal/mol. This increase in the magnitude of $\Delta E_{subst} - \Delta E^0_{subst}$ is 30% more than the factor of 2 that is expected for doubling the number of phenyl substituents at a fixed value of R. It seems likely that the 1.3 kcal/mol of extra stabilization comes from a reduction in the steric repulsion between the pairs of phenyl groups at C1−C6 and C3−C4 upon increasing R by 0.253 Å from $R_0 = 1.965$ Å in the TS for the parent Cope rearrangement to $R = 2.218$ Å in the TS for the Cope rearrangement of 1,3-diphenyl-1,5-hexadiene.

The hypothesis, that steric repulsion between phenyl groups is relieved by lengthening R in the TS for the Cope rearrangement of 1,3,4,6-tetraphenyl-1,5-hexadiene, is also consistent with the fact that $R - R_0 = 0.684$ Å in the TS for this reaction. As already noted, this change in interallylic bond lengths is a factor of 2.7 greater than $R - R_0 = 0.253$ Å in the TS for the Cope rearrangement of 1,3-diphenyl-1,5-hexadiene, rather than the expected factor of 2.0.

However, the fact that ΔE_{dist} is not quadratic out to $R = 2.649$ Å also contributes to the ratio of the changes in interallyic bond lengths being larger than the ratio of 2.0 expected from eq 8. Table 2 shows that $\Delta E_{dist} = 2.0$ kcal/mol for lengthening R by 0.253 Å from $R_0 = 1.965$ in the TS for the Cope rearrangement of unsubstituted 1,5-hexadiene to $R = 2.218$ Å in the TS for the Cope rearrangement of 1,3-diphenyl-1,5-hexadiene. If ΔE_{dist} were quadratic in $R - R_0$, $\Delta E_{dist} = 14.6$ kcal/mol would be expected for lengthening R by 0.684 Å from $R_0 = 1.965$ to $R = 2.649$ Å in the TS for the Cope rearrangement of 1,3,4,6-tetraphenyl-1,5-hexadiene. However, as Table 2 shows, the actual value of $\Delta E_{dist} = 12.0$ kcal/mol in the latter TS.

Upon increasing the interallylic distance from $R = 2.218$ Å to $R = 2.649$ Å, ΔE_{dist} increases by 10.0 kcal/mol, and this energy increase cancels all but 4.4 kcal/mol of the increase of 14.4 kcal/mol in the magnitude of ΔE_{subst}, the substituent stabilization energy. Therefore, of the 9.7 kcal/mol increase in the magnitude of $\Delta E_{stab} - \Delta E^0_{stab}$ that results from adding a pair of phenyl groups to C4 and C6 of the TS for the Cope rearrangement of 1,3-diphenyl-1,5-hexadiene, 5.3 kcal/mol (55%) comes from the two additional phenyl groups increasing the magnitude of ΔE_{subst} at $R = 2.218$ Å, and the remainder comes from the 4.4 kcal/mol increase in the size of the difference between ΔE_{subst} and ΔE_{dist} as R increases from 2.218 Å to 2.649 Å.

The data in Table 2 are useful for a detailed understanding of not only the cooperative substituent effects, which have been both observed in and computed for the Cope rearrangements of 2,5-diphenyl- and 1,3,4,6-tetraphenyl-1,5-hexadiene (Table 1), but also of the competitive substituent effect in the Cope rearrangement of 1,3,5-triphenyl-1,5-hexadiene.[5] Table 2 shows that at $R - R_0 = -0.366$ Å the C2 phenyl group in 2-phenyl-1,5-hexadiene provides a net stabilization of $\Delta E_{stab} - \Delta E^0_{stab} = -2.9$ kcal/mol, and at $R - R_0 = 0.253$ Å the C1 and C3 phenyl groups in 1,3-diphenyl-1,5-pentadiene provide a net stabilization of $\Delta E_{stab} - \Delta E^0_{stab} = -2.0$ kcal/mol. However, as discussed in the previous section, the Cope rearrangement of 1,3,5-triphenyl-1,5-pentadiene occurs at a compromise TS geometry with $R - R_0 = 0.145$ Å. At this TS geometry, neither the phenyl groups at C1 and C3 nor the phenyl group at C5 provide as much net stabilization as these phenyl groups furnish in the TSs for the Cope rearrangements of, respectively, 1,3-diphenyl-1,5-pentadiene at $R - R_0 = 0.253$ Å and 2-phenyl-1,5-hexadiene at $R - R_0 = -0.366$ Å.

Table 2 reveals that at $R - R_0 = 0.145$ Å the C2 phenyl group in 2-phenyl-1,5-hexadiene provides 2.1 kcal/mol less net stabilization than at $R_0 = 1.965$ Å and 5.0 kcal/mol less net stabilization than at $R - R_0 = -0.366$ Å. At $R = 2.110$ Å the C1 and C3 phenyl groups in 1,3-diphenyl-1,5-pentadiene provide 1.7 kcal/mol more stabilization that at

$R_0 = 1.965$ Å but 0.3 kcal/mol less stabilization than at $R - R_0 = 0.253$ Å.

Thus, assuming additivity, one might expect the three phenyl groups in 1,3,5-triphenyl-1,5-pentadiene would provide 5.3 kcal/mol less net TS stabilization than the total amount they furnish in the Cope rearrangements of 1,3-diphenyl-1,5-pentadiene and 2-phenyl-1,5-hexadiene. In fact, Table 2 shows that $\Delta E_{stab} = 3.2$ kcal/mol for the Cope rearrangement of 1,3,5-triphenyl-1,5-hexadiene, a value that is 4.1 kcal/mol smaller in magnitude than the sum of the ΔE_{stab} values for the Cope rearrangements of 2-phenyl-1,5-pentadiene and 1,3-diphenyl-1,5-pentadiene.

This difference between the ΔE_{stab} values in Table 2 is 1.1 kcal/mol larger than the difference of $4.0 + 3.0 - 4.0 = 3.0$ kcal/mol between the calculated $\Delta\Delta H^{\ddagger}$ values in Table 1 and 2.6 kcal/mol larger than the difference of $4.2 + 3.0 - 5.7 = 1.5$ kcal/mol between the experimental $\Delta\Delta H^{\ddagger}$ values.

The biggest contributor to making the difference of 3.0 kcal/mol between the calculated $\Delta\Delta H^{\ddagger}$ values 1.5 kcal/mol larger than the difference between the experimental $\Delta\Delta H^{\ddagger}$ values is the fact that the calculated value of $\Delta\Delta H^{\ddagger}$ for the Cope rearrangement of 1,3,5-triphenyl-1,5-hexadiene is 1.7 kcal/mol smaller in size than the experimental value. Nevertheless, the calculated and experimental $\Delta\Delta H^{\ddagger}$ values in Table 1 both show that the substituent effects in the Cope rearrangement of this compound are competitive.

Summary and Conclusions

The simple mathematical model in eq 4 assumes that radical stabilizing substituents attached to 1,5-hexadiene prefer interallylic bond lengths, R, in the TS for the Cope rearrangement that allow the substituents to provide the greatest amount of stabilization energy (ΔE_{subst}) for the TS and that the leading term in ΔE_{subst} is linear in R. The model also assumes that lengthening or shortening these bonds from the preferred value, R_0, in the TS for the Cope rearrangement of unsubstituted 1,5-hexadiene results in an energy increase (ΔE_{dist}) that is quadratic in the difference between R and R_0. Minimization of the net stabilization energy, $\Delta E_{stab} = \Delta E_{dist} + \Delta E_{subst}$, gives the expression in eq 5 for the optimal value of R in the TS for the Cope rearrangement of a substituted 1,5-hexadiene.

At this value of R, eq 6 predicts a relationship between $\Delta E_{subst} - \Delta E^0_{subst}$, the difference between the substituent stabilization energies at R and R_0, and ΔE_{dist} that is reminiscent of the relationship between potential and kinetic energy in the virial theorem. The results of B3LYP/6-31G* calculations show that the predicted relationship between $\Delta E_{subst} - \Delta E^0_{subst}$ and ΔE_{dist} does, in fact, hold in the TSs for the Cope rearrangement of several phenyl-substituted 1,5-hexadienes.

B3LYP calculations also show that the effect of phenyl substituents on stabilizing the Cope TS is much smaller at R_0 than at the optimal value of R. Therefore, the major mechanism by which phenyl substituents lower $\Delta H^{\ddagger}$ for the Cope rearrangement is to increase the magnitude of ΔE_{subst} by distorting the interallylic bond lengths in the TS from R_0 to R. The dependence of ΔE_{subst} on R is what gives rise to the cooperative substituent effects, both calculated and observed, in the Cope rearrangements of 2,5-diphenyl- and 1,3,4,6-tetraphenyl-1,5-hexadiene and to the competitive substituent effect, both calculated and observed, in the Cope rearrangement of 1,3,5-triphenyl-1,5-hexadiene.

Acknowledgment. We thank the National Science Foundation for supporting the research described in this paper.

References

(1) Reviews of the Cope rearrangement: (a) Gajewski, J. J. *Hydrocarbon Thermal Isomerization;* Academic Press: New York, 1981; pp 166−176. (b) Borden, W. T.; Loncharich, R. J.; Houk, K. N. *Annu. Rev. Phys. Chem.* **1988**, *39*, 213. (c) Houk, K. N.; Li, Y.; Evanseck, J. D. *Angew. Chem., Int. Ed. Engl.* **1992**, *31*, 682. (d) Houk, K. N.; Gonzalez, J.; Li, Y. *Acc. Chem. Res*. **1995**, *28*, 81. (e) Borden, W. T.; Davidson, E. R. *Acc. Chem. Res.* **1996**, *29*, 67. (f) Wiest, O.; Montiel, D. C.; Houk, K. N. *J. Phys. Chem. A* **1997**, *101*, 8378. (g) Borden, W. T. *Mol. Phys*. **2001**, *100*, 337.

(2) (a) Hrovat, D. A.; Beno, B. R.; Lange, H.; Yoo, H.-Y.; Houk, K. N.; Borden, W. T. *J. Am. Chem. Soc.* **1999**, *121*, 10529 (b) Hrovat, D. A.; Chen, J.; Houk, K. N.; Borden, W. T. *J. Am. Chem. Soc.* **2000**, *122*, 7456.

(3) Doering, W. von E.; Toscano, V. G.; Beasley, G. H. *Tetrahedron* **1971**, *27*, 5299.

(4) Dewar, M. J. S.; Wade, L. E. *J. Am. Chem. Soc.* **1977**, *99*, 4417.

(5) Roth, W. R.; Lennartz, H.-W.; Doering, W. von E.; Birladeanu, L.; Guyton, C. A.; Kitagawa, T. *J. Am. Chem. Soc.* **1990**, *112*, 1722.

(6) Doering, W. von E.; Wang, Y. *J. Am. Chem. Soc.* **1999**, *121*, 10112.

(7) Doering, W. von E.; Birladeanu, L.; Sarma, K.; Blaschke, G.; Scheidemantel, U.; Boese, R.; Benet-Bucholz, J.; Klärner, F.-G.; Gehrke, J.-S.; Zinny, B. U.; Sustmann, R.; Korth, H.-G. *J. Am. Chem. Soc.* **2000**, *122*, 193.

(8) (a) Weist, O.; Black, K. A.; Houk, K. N. *J. Am. Chem Soc*. **1994**, *116*, 10336. (b) Jiao, H.; Schleyer, P. Von R. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 334.

(9) Gajewski, J. J.; Conrad, N. D. *J. Am. Chem. Soc.* **1979**, *101*, 6693.

(10) The substituent stabilization energy (ΔE_{subst}) and the net stabilization energy (ΔE_{stab}) are both negative numbers.

(11) See, for example: (a) Pitzer, K. S. *Quantum Chemistry*; Prentice Hall: Englewood Cliffs, NJ, 1953; pp 426−8. (b) Pitzer, K. S. *Quantum Chemistry*; Prentice Hall: Englewood Cliffs, NJ, 1953; p 56.

(12) Since the values of R computed for the TSs for the Cope rearrangements 1,3-diphenyl- and 1,3,5-triphenyl-1,5-hexadiene are B3LYP values, it is more appropriate to use the B3LYP value of $R = 1.829$ Å in the TS for the Cope rearrangement of 2-phenyl-1,5-hexadiene, rather than the UB3LYP value of $R = 1.599$ Å in the UB3LYP intermediate.

(13) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(14) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(15) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213.

(16) 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.; Rega, N.; Salvador, P.; Dannenberg, J. J.; 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.; 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.11.4; Gaussian, Inc., Pittsburgh, PA, 2002.

(17) (a) Staroverov, V. K.; Davidson, E. R. *J. Am. Chem. Soc.* **2000**, *122*, 186. (b) Staroverov, V. K.; Davidson, E. R. *J. Am. Chem. Soc.* **2000**, *122*, 7377.

CT049929Q

JCTC Journal of Chemical Theory and Computation

Computational and Spectroscopic Studies of Re(I) Bipyridyl Complexes Containing 2,6-Dimethylphenylisocyanide (CNx) Ligand

Stanislav R. Stoyanov,[†] John M. Villegas,[†] Arvin J. Cruz,[†] Loranelle L. Lockyear,[‡] Joseph H. Reibenspies,[§] and D. Paul Rillema*,[†]

Department of Chemistry, Wichita State University, 1845 N. Fairmount Street, Wichita, Kansas 67260-0051, Department of Chemistry, Bethany College, 421 N. First Street, Lindsborg, Kansas 67456-1897, and Department of Chemistry, Texas A & M University, P.O. Box 30012, College Station, Texas 77842-3012

Received August 18, 2004

Abstract: Density Functional Theory (DFT) calculations produce optimized geometries of the complexes $[Re(CO)_3(bpy)Cl]$ (**1**), $[Re(CO)_3(bpy)(py)](CF_3SO_3)$ (**2**), $[Re(CO)_3(bpy)(CNx)](CF_3SO_3)$ (**3**), and $[Re(CO)(bpy)(CNx)_3](CF_3SO_3)$ (**4**), where bpy = 2,2′-bipyridine, py = pyridine, and CNx = 2,6-dimethylphenylisocyanide in their ground and lowest-lying triplet states. The ground-state optimized geometry for the cation of $[Re(CO)_3(bpy)(CNx)](CF_3SO_3)$ (**3**) results in a Re−C (CNx) bond length of 2.10 Å, a Re−C (CO) bond length trans to CNx of 2.01 Å, and a Re−C (CO) bond length cis to CNx of 1.96 Å which compares favorably to the single-crystal analysis of a Re−C (CNx) bond length of 2.074(4) Å, a Re−C (CO) bond length trans to CNx of 1.971(4) Å, and Re−C (CO) bond length cis to CNx of 1.932(4) Å. The majority of the singlet excited-state energies calculated using Time-dependent Density Functional Theory (TDDFT) and Conductor-like Polarizable Continuum Model (CPCM) are metal-ligand-to-ligand charge transfer (MLLCT) states and are in good agreement with the UV−vis spectral energies for the complexes in ethanol. The complexes exhibit emission both at room temperature and at 77 K except **4** which is only emissive at 77 K. The 77 K emission lifetimes range from 3.9 μs for **1** to 8.8 μs for **3**. The emissive lowest-lying triplet state is a ^{3}MLLCT state for complexes **1**−**3** but a triplet ligand-to-metal charge transfer (^{3}LMCT) state for complex **4**. The electronic, electrochemical, thermodynamic, HOMO−LUMO, and emitting-state energy gaps as well as the emission lifetimes increase in the order **1** < **2** < **3.** A ^3d−d excited- state, which is located above the ^{3}LMCT state, accounts for the loss of room-temperature emission for complex **4**.

Introduction

Density Functional Theory (DFT) is very useful for interpretation of electrochemical and spectroscopic results. A linear relationship between the HOMO−LUMO energy gap and the electrochemical redox potentials for a series of isoelectronic Ru(II) diimine complexes was reported from our laboratory.[1] DFT calculations on the ground and metal-to-ligand charge transfer (MLCT) states of a series of Re(I) tricarbonyl complexes were used by others for the investigation of the excited-state geometries and electronic structures.[2−4] Time-dependent density functional theory (TDDFT) calculated MLCT states and UV−vis spectra correlations were also reported for Re(I) complexes containing the ligand azophenine.[4] The TDDFT method when used alone treats molecules in the gas phase; therefore, the method does not always give the right electronic excited-state energies in solution.[5,6a]

We have reported that combining the TDDFT with the conductor-like polarizable continuum model (CPCM) yields calculated singlet excited-state energies that correlate quite

* Corresponding author e-mail: paul.rillema@wichita.edu.
[†] Wichita State University.
[‡] Bethany College.
[§] Texas A & M University.

10.1021/ct049956g CCC: $30.25
Published on Web 11/05/2004

well with the UV−vis absorption energies of $[Ru(bpy)_2(CNx)Cl]^+$ in a series of seven solvents of varied polarity.[6b] According to other reports, the tandem use of TDDFT and CPCM has produced dramatic changes in the energies and the assignments of the singlet excited-states for Ru(II) and Os(II) polypyridyl complexes.[7] It is the primary method used in our study.

The synthesis, photochemistry, and computational studies of 2,6-dimethylphenylisocyanide (CNx) complexes of Re(I) and Ru(II) have been recently reported by our group.[8] Generally, Re(I) diimine tricarbonyl complexes are *facial*, MLCT emitters and have broad and structureless emission bands which are sensitive to changes in the nature of the environment.[9−11] Thus variations in the structure of the non-carbonyl bidentate as well as the ancillary ligands could produce considerable effects on luminescence energies, lifetimes, and quantum yields. Here we report computational, electronic absorption, and excited-state emission studies of a series of Re(I) bipyridine complexes, $[Re(CO)_3(bpy)Cl]$ (**1**), $[Re(CO)_3(bpy)(py)](CF_3SO_3)$ (**2**), $[Re(CO)_3(bpy)(CNx)](CF_3SO_3)$ (**3**), and $[Re(CO)(bpy)(CNx)_3](CF_3SO_3)$ (**4**). The photophysical properties of the two new complexes containing the CNx ligand, $[Re(CO)_3(bpy)(CNx)]^+$, and $[Re(CO)(bpy)(CNx)_3]^+$ are compared to similar properties of two more conventional complexes, $[Re(CO)_3(bpy)Cl]$ and $[Re(CO)_3(bpy)(py)]^+$. The study is focused on the electronic effect of the CNx ligand on the molecular geometry and UV−vis and IR spectra of the $Re(CO)_3(bpy)$ moiety using computational and experimental methods.

Experimental Section

Materials. The ligand 2,6-dimethylphenylisocyanide was purchased from Fluka. $Re(CO)_5Cl$, 2,2′-bipyridine, and pyridine were purchased from Aldrich. Optima grade methanol was purchased from Fischer Scientific, while acetonitrile was purchased from Sigma-Aldrich. AAPER Alcohol and Chemical Company was the source of absolute ethanol. Ethanol and methanol were used in a 4:1 (v/v) mixture to prepare solutions for the emission and emission lifetime studies. Elemental analyses were obtained from Desert Analytics Laboratory, Tucson, AZ.

Instrumentation and Physical Measurements. UV−vis spectra were obtained using a Hewlett-Packard model 8452A diode array spectrophotometer. The IR spectra were acquired using a Nicolet Avatar 360 FT-IR spectrophotometer. Proton NMR spectra were obtained using a Varian Mercury 300 FT-NMR spectrometer. An EG&G PAR model 263A potentiostat/galvanostat was used to obtain the cyclic voltammograms. The measurements were carried out in a typical H-cell using a platinum disk working electrode, a platinum wire counter electrode, and a Ag/AgCl reference electrode in acetonitrile. The supporting electrolyte used was 0.1 M tetrabutylammonium hexafluorophosphate (TBAH). Ferrocene was added for reference.

The sample preparation for emission studies involved dissolving a small amount of sample (∼2 mg) in the appropriate solvent, and the absorbance of the solution was measured. The concentration of the solution was altered in order to achieve an absorbance of about 0.10 at lowest energy transition. Such a concentration provides enough material for data acquisition but excludes self-quenching processes. A 3−4 mL aliquot of the solution was then placed in a 10 mm diameter Suprasil (Heraeus) nonfluorescent quartz-tube equipped with a tip-off manifold. The sample was then freeze−pump−thaw degassed for at least three cycles (to approximately 75 milliTorr) removing any gases from the sample. The manifold was then closed, and the sample was allowed to equilibrate at room temperature. The solvent evaporation was assumed to be negligible; therefore, the concentrations were assumed to remain constant throughout this procedure. The corrected emission spectra were collected using a Spex Tau3 Fluorometer.

The emission quantum yields were then calculated using eq 1, where ϕ_x is the emission quantum yield of the sample and ϕ_{std} is the emission quantum yield for the standard ($[Ru(bpy)_3]^{2+}$), A_{std} and A_x represent the absorbance after degassing the standard and the sample, respectively, while I_{std} and I_x are the integrals of the emission envelope of the standard and the sample, respectively.[12]

$$\phi_x = (A_{std}/A_x)(I_x/I_{std})\phi_{std} \quad (1)$$

The excited state lifetimes were determined by exciting the sample at the lowest energy transition (usually MLCT) using an OPOTEK optical parametric oscillator pumped by a frequency tripled Continuum Surlite Nd:YAG laser run at ∼20 mJ/10 ns pulse. The oscilloscope control and data curve fitting analysis was accomplished using the Origin 6.1 program by OriginLab Corporation. The excited state lifetime experiments were conducted as previously published.[13]

X-ray Crystallography Data Collection. A Bausch and Lomb 10× microscope was used to identify a suitable colorless plate from a representative sample of crystals of $[Re(bpy)(CO)_3(CNx)](PF_6)$ grown by slow evaporation of ethanol. The crystal was coated in a cryogenic protectant (paratone) and was then fixed to a loop which in turn was fashioned to a copper mounting pin. The mounted crystal was then placed in a cold nitrogen stream (Oxford) maintained at 110 K.

A BRUKER SMART 1000 X-ray three-circle diffractometer was employed for crystal screening, unit cell determination, and data collection. The goniometer was controlled using the SMART software suite, version 5.056 (Microsoft NT operating system). The sample was optically centered with the aid of a video camera such that no translations were observed as the crystal was rotated through all positions. The detector was set at 5.0 cm from the crystal sample (CCD-PXL-KAF2, SMART 1000, 512 × 512 pixel). The X-ray radiation employed was generated from a Mo sealed X-ray tube ($K_\alpha = 0.70173$ Å with a potential of 50 kV and a current of 40 mA) and filtered with a graphite monochromator in the parallel mode (175 mm collimator with 0.5 mm pinholes).

Dark currents were obtained for an exposure time of 30 s. A rotation exposure was taken to determine crystal quality and the X-ray beam intersection with the detector. The beam intersection coordinates were compared to the configured coordinates, and changes were made accordingly. The rotation exposure indicated acceptable crystal quality, and the unit cell determination was undertaken. Sixty data frames were taken at widths of 0.3° with an exposure time of 10 s. Over 200 re-

Table 1. Crystal Data for $[Re(bpy)(CO)_3(CNx)](PF_6)$, **3**

formula	$C_{22}H_{17}F_6N_3O_3PRe$
formula weight	702.56
crystal size, mm	0.30 × 0.20 × 0.05
crystal system	monoclinic
space group	P2(1)/n
a, Å	8.9203(18)
b, Å	12.925(3)
c, Å	21.104(4)
α, deg	90
β, deg	90.521(13)
γ, deg	90
V, Å^3	2433.2(8)
Z	4
density (calculated), g cm^{-1}	1.918
absorption coefficient, mm^{-1}	5.135
F(000)	1342
θ range for data collection, deg	1.85 to 28.30
index ranges	−11≤ h≤ 11, −17≤ k≤16, −20≤ l≤ 28
reflections collected	22815
independent reflections	5719 [R(int) = 0.0287]
completeness to θ	94.7%
max. and min. transmission	0.7833 and 0.3080
goodness-of-fit on F2	1.184
final R indices [I > 2σ(I)]	R1 = 0.0279, wR2 = 0.0748
R indices (all data)	R1 = 0.0319, wR2 = 0.0868
largest difference peak and hole, e Å^{-3}	2.435 and −0.681

flections were centered, and their positions were determined. These reflections were used in the auto-indexing procedure to determine the unit cell. A suitable cell was found and refined by nonlinear least squares and Bravais lattice procedures and reported here in Table 1. The unit cell was verified by examination of the *hkl* overlays on several frames of data, including zone photographs. No super-cell or erroneous reflections were observed.

After careful examination of the unit cell, a standard data collection procedure was initiated. This procedure consisted of a collection of one hemisphere of data using omega scans, involving the collection of 1201 0.3° frames at fixed angles for ϕ, 2θ, and χ ($2\theta = -28°$, $\chi = 54.73°$), while varying omega. Each frame was exposed for 30 s and contrasted against a 30 s dark current exposure. The total data collection was performed for duration of approximately 13 h at 110 K. No significant intensity fluctuations of equivalent reflections were observed.

After data collection, the crystal was measured carefully for size, morphology, and color. These measurements are reported in Table 1.

(1) Preparation of *fac*-[Re(bpy)(CO)$_3$Cl]. The synthesis of the complex was carried out following a previously published procedure.[14a,b] Yield 1.34 g (94%). Anal. Calcd for $ReC_{13}H_8N_2O_3Cl$: C, 33.81; H, 1.75; N, 6.07. Found: C, 34.00; H, 1.65; N, 6.04. IR (KBr pellet): 2022, 1890, 1653, 1600, 1559, 1471, 1444, 1314, 1245, 1070, 764, 732, 647, 536 cm^{-1}. ^{1}H NMR (DMSO): δ ppm 7.76 (dd, 2H, J = 0.9, 6.8 Hz), 8.35 (dd, 2H, J = 1.5, 7.7 Hz), 8.77 (d, 2H, J = 8.4 Hz), 9.02 (dd, 2H, J = 0.9, 5.5 Hz).

(2) Preparation of *fac*-[Re(bpy)(CO)$_3$(py)](CF$_3$SO$_3$). The synthesis of the complex was carried out following a previously published procedure.[14b,c] Yield 0.25 g (88%). Anal. Calcd for $ReC_{19}H_{13}N_3O_6FS$: C, 34.86; H, 2.00; N, 6.42. Found: C, 34.92; H, 1.82; N, 6.43. IR (KBr pellet): 2026, 1906, 1603, 1491, 1473, 1448, 1279, 1260, 1147, 1030, 776, 735, 700, 637, 573, 534 cm^{-1}. ^{1}H NMR (DMSO): δ ppm 7.42 (dd, 2H, J = 1.2, 7.2 Hz), 7.90 (dd, 2H, J = 1.2, 6.5 Hz), 7.95 (t, 1H, J = 7.8), 8.39 (m, 4H), 8.69 (d, 2H, J = 8.4 Hz), 9.31 (dd, 2H, J = 1.2, 5.7 Hz).

(3) Preparation of *fac*-[Re(bpy)(CO)$_3$(CNx)](CF$_3$SO$_3$). A 0.19 g sample of CNx (1.4 mmol) was added to a 100 mL round-bottomed flask containing 0.83 g of $[Re(CO)_3(bpy)(CF_3SO_3)]$ (1.4 mmol) in 50 mL of absolute ethanol. The mixture was refluxed for 6 h. The solvent was evaporated in vacuo until about 5 mL of the solution was left in the flask. The remaining solution was added dropwise into 100 mL of ether with constant stirring. The dark yellow precipitate was removed by filtration and dried in a vacuum. Yield: 1.00 g (98%). Anal. Calcd for $ReC_{23}H_{17}N_3SF_3O_6$: C, 39.10; H, 2.42; N, 5.95. Found: C, 40.40; H, 2.44; N, 5.68. IR (KBr pellet): 2173, 2036, 1955, 1603, 1473, 1443, 1279, 1224, 1152, 1033, 1025, 790, 638, 488 cm^{-1}. ^{1}H NMR (DMSO): δ ppm 1.98 (s, 6H), 7.17 (d, 2H, J = 7.8 Hz), 7.28 (t, 1H, J = 6.9 Hz), 7.85 (dd, 2H, J = 1.2, 6.8 Hz), 8.43 (dd, 2H, J = 1.5, 8.1 Hz), 8.86 (d, 2H, J = 7.8 Hz), 9.16 (d, 2H, J = 4.2 Hz).

(4) Preparation of *mer*-[Re(bpy)(CO)(CNx)$_3$](PF$_6$). A 0.10 g sample of mer-$[Re(CNx)_3(CO)_2Cl]$[15] (0.15 mmol) was added to 0.05 g of $AgPF_6$ (0.20 mmol) and 0.03 g of 2,2′-bipyridine (0.18 mmol) in a 125 mL round-bottomed flask. About 75 mL of ethanol was added, and the solution was refluxed in the dark overnight. The solution was filtered to remove the solid AgCl, and the filtrate was evaporated to dryness. An orange solid was recovered, washed with ether, and dried under vacuum. Yield 0.11 g (81%). Anal. Calcd for $ReC_{38}H_{35}N_5OPF_6$: C, 50.21; H, 3.88; N, 7.71. Found: C, 50.41; H, 4.00; N, 7.90. IR (KBr pellet): 2072, 1885, 1658, 1590, 1468, 1439, 842, 774, 557, 491 cm^{-1}. ^{1}H NMR (DMSO): δ ppm 2.43 (s, 18H), 7.29 (m, 9H), 7.48 (dd, 2H, J = 0.9, 5.7 Hz), 7.97 (dd, 2H, J = 1.8, 8.1 Hz), 8.40 (dd, 2H, J = 0.9, 8.1 Hz), 8.70 (dd, 2H, J = 1.8, 4.5 Hz).

Results

Synthesis. The preparation of the complexes followed the scheme shown in Figure 1. Complexes **1** and **2** were prepared by modifying published procedures.[14] Complex **1** was synthesized by reacting $[Re(CO)_5Cl]$ with bpy in ethanol, and it served as the starting material for complexes **2** and **3**. The chloro ligand was first removed from the coordination sphere by reaction of $[Re(bpy)(CO)_3Cl]$ with $AgCF_3SO_3$ to precipitate AgCl generating the intermediate $[Re(bpy)(CO)_3(CF_3SO_3)]$. Complex **2** was produced when py replaced $CF_3SO_3^-$ in the coordination sphere. Similarly, upon addition of CNx to a solution containing $[Re(bpy)(CO)_3(CF_3SO_3)]$, complex **3** was obtained. The PF_6^- salt of complex **3** was also prepared by first reducing the volume of the solvent using a rotary evaporator and then adding NH_4PF_6 to precipitate the product.

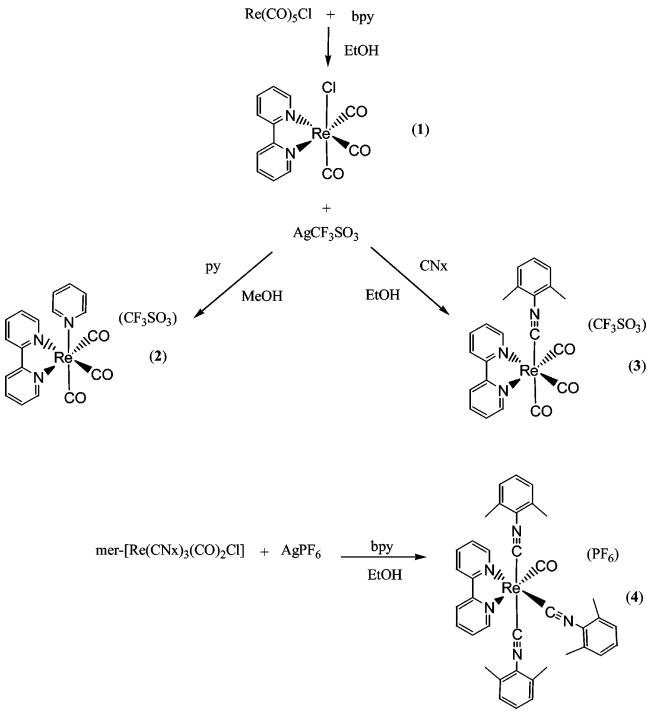


Figure 1. Schematic diagram of the synthesis of the complexes.

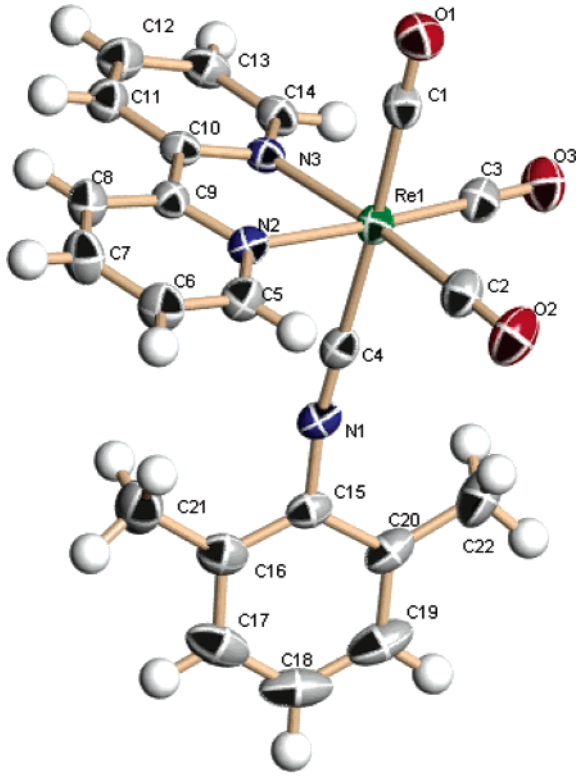


Figure 2. ORTEP diagram of $[Re(bpy)(CO)_3(CNx)]^+$, **3**.

Complex **4** was prepared by a different route. The precursor complex, *mer*-$[Re(CNx)_3(CO)_2Cl]$, was prepared using a modification of a previously published procedure.[15] This complex was refluxed with $AgPF_6$ and bpy in ethanol in the dark overnight to produce the desired product. The complexes were characterized by IR, ^1H NMR, and elemental analysis. The structure of the PF_6^- salt of complex **3** was determined by X-ray crystallography.

Table 2. Selected Bond Length (Å) and Angles (deg) for $[Re(bpy)(CO)_3(CNx)](PF_6)$, **3**

Re(1)−C(1)	1.971(4)	C(1)−Re(1)−C(4)	177.33(15)
Re(1)−C(2)	1.932(4)	C(2)−Re(1)−N(3)	173.56(15)
Re(1)−C(3)	1.927(4)	C(3)−Re(1)−N(2)	174.05(15)
Re(1)−C(4)	2.074(4)	O(1)−C(1)−Re(1)	178.5(4)
Re(1)−N(2)	2.171(3)	O(2)−C(2)−Re(1)	178.0(4)
Re(1)−N(3)	2.169(3)	O(3)−C(3)−Re(1)	177.5(4)
N(1)−C(4)	1.161(5)	N(1)−C(4)−Re(1)	175.8(3)
O(1)−C(1)	1.136(5)	C(4)−N(1)−C(15)	171.2(4)
O(2)−C(2)	1.151(5)		
O(3)−C(3)	1.147(5)		

X-ray Crystal Structure Determination of *fac*-$[Re(CO)_3(bpy)(CNx)](PF_6)$ (3). The structure of **3** was determined by X-ray crystallography and is shown in the ORTEP diagram of Figure 2. The crystal data are listed in Table 1, while selected bond distances and angles for the complex are given in Table 2.

The complex adopted a distorted octahedral geometry with the equatorial plane defined by the bipyridine ligand and two carbonyl groups trans to it and the axial ligands defined as a carbonyl group with a CNx ligand opposite it. The angles of the trans ligands at the metal center were 177.33(15)° for C(1)−Re(1)−C(4), 173.56(15)° for C(2)−Re(1)−N(3), and 174.05(15)° for C(3)−Re(1)−N(2). The Re−N bond lengths were similar within experimental error, 2.17 ± 0.01 Å, as well as the Re−C (CO) bond distances, 1.93 ± 0.02 Å, in the equatorial positions. However, the Re−C (CO) bond length of CO trans to the CNx ligand was considerably longer by ∼0.04 Å compared to the other two Re−C (CO) bond distances. The longest Re−C bond distance was 2.074(4) Å to the CNx ligand.

Electronic Absorption Studies. The electronic absorption properties of the complexes were studied at room temperature in 4:1 ethanol/methanol as solvent. Molar absorption coefficients (ϵ) were determined from Beer's Law plots using at least five dilution points. The probable assignments of these bands were made on the basis of the documented optical transitions of similar Re(I) complexes[9,16,17] and the computational assignment of the singlet excited-states (vide infra). The results are listed in Table 3.

The lowest energy transition of the complexes was assigned as metal−ligand-to-ligand charge transfer (MLLCT), while those at higher energies were assigned as the intraligand $\pi \rightarrow \pi^*$ transitions. The MLLCT absorption peaks of the series occurred in the following order: **4** (26 200 cm^{-1}) < **1** (26 900 cm^{-1}) < **3** (27 000 cm^{-1}) < **2** (28 600 cm^{-1}). In addition to the optical transition for **3** located at 27 000 cm^{-1}, another one was found at 30 300 cm^{-1}. It is important to note that since the MLLCT band occurs as a broad shoulder, the exact position of the band maximum as well as the extinction coefficient were subject to error.

The intraligand $\pi \rightarrow \pi^*$ transitions were located in the UV region of the spectrum. As expected, the molar absorptivities of the electronic transitions associated with the CNx ligand[8b] (33 000 and 42 000 cm^{-1}) increased as the number of coordinated CNx ligands increased from one in **3** to three in **4**.

Electrochemical Studies. The redox potentials of the series of complexes were determined using cyclic voltam-

Table 3. Experimental[a] Electronic Transitions and Calculated[b,c] Excited State Energies of Re(I) Complexes

complex	E_{exp}, $\times 10^3$ cm^{-1}, ϵ (M^{-1} cm^{-1})	E_{calc}, $\times 10^3$ cm^{-1}	assignment
[Re(bpy)(CO)$_3$Cl]	26.9 (3900)	25.0	MLLCT
	34.2 (15 200)	34.1	LC ($\pi \rightarrow \pi^*$)
	41.7 (16 400) (s)[d]		LC ($\pi \rightarrow \pi^*$)
[Re(bpy)(CO)$_3$(py)](CF$_3$SO$_3$)	28.6 (4400)	27.4	MLLCT
	31.2 (13 300)		MLLCT
	32.7 (12 300)	32.8	MLLCT[e]
	37.9 (18 800)	37.4	MLLCT[e]
	40.0 (19 200)	39.7	MLLCT
[Re(bpy)(CO)$_3$(CNx)](CF$_3$SO$_3$)	27.0 (2300) (s)[d]	26.8	MLLCT
	31.6 (22 700)	31.5	MLLCT
	32.9 (21 800)	32.6	MLLCT
	38.5 (49 300)	37.9	MLLCT[e]
[Re(bpy)(CO)(CNx)$_3$](PF$_6$)	26.2 (7000) (s)[d]	26.7	MLLCT
	29.1 (22 000) (s)[d]	29.0	MLLCT
	32.7 (53 000) (s)[d]	32.8	MLLCT
	34.0 (59 000)	34.3	MLLCT
	42.0 (45 000)		LC ($\pi \rightarrow \pi^*$)

[a] In 4:1 (v/v) ethanol:methanol. [b] In ethanol. [c] Only singlet excited-states with $f > 0.05$ are considered. [d] s = shoulder. [e] Mixed state.

Table 4. Electrochemical Properties of the Complexes in CH$_3$CN at Room Temperature

complex	$E_{1/2\ (ox)}$, V[a]	$E_{1/2\ (red)}$, V[a]
[Re(bpy)(CO)$_3$Cl][b]	1.32[d]	−1.35
[Re(bpy)(CO)$_3$(py)](CF$_3$SO$_3$)	1.78[d]	−1.18[b]
[Re(bpy)(CO)$_3$(CNx)](CF$_3$SO$_3$)	1.99[d]	−1.22
[Re(bpy)(CO)(CNx)$_3$](PF$_6$)	1.30[e]	
[Ru(bpy)$_3$]$^{2+}$[(c)]	1.27[f]	−1.31
		−1.50
		−1.77

[a] Potential in volts vs SSCE (scan rate = 250 mV/s). [b] Data from ref 14. [c] Data from ref 13. [d] Irreversible oxidation wave. [e] Quasi-reversible oxidation wave. [f] RuIII/II redox couple.

metry. The electrochemical properties of the complexes are listed in Table 4. Irreversible oxidation waves attributed to the Re(I) → Re(II) process increased in potential in the order **1** < **2** < **3** for [Re(CO)$_3$(bpy)Cl] (**1**), [Re(CO)$_3$(bpy)(py)]$^+$ (**2**), and [Re(CO)$_3$(bpy)(CNx)]$^+$ (**3**). The single reduction observed for **1**−**3** was attributed to the reduction of the bpy ligand.

The electrochemical behavior of [Re(CO)(bpy)(CNx)$_3$]$^+$ (**4**) differed from the others. It underwent a quasi-reversible oxidation ($\Delta E_p = 0.09$ V) at 1.30 V attributed to the Re(II/I) couple. Further, a reduction normally associated with coordinated bpy was absent. The oxidation of **4** was shifted to a lower potential compared to the other three complexes.

Emission Properties and Excited-State Lifetimes. The emission properties and excited-state lifetimes of the complexes were studied both at room temperature and at 77 K. The results are given in Table 5. Complexes **1**−**3** were emissive both at room temperature and at 77 K. Complex **4** however was nonemissive at room-temperature but was highly emissive at 77 K.

The emission lifetimes (τ_{em}) increased in the order **1** < **2** < **3** both at room temperature and at 77 K as Cl in **1** was replaced with py in **2** and CNx in **3**. The emission quantum yield (ϕ_{em}) also increased from 0.0052 for **1** to 0.10 for **2** and 0.27 for **3**. It is worthy to note that ϕ_{em} determined at room temperature for complex **3** was more than three times higher than ϕ_{em} of the standard complex [Ru(bpy)$_3$]$^{2+}$ under the same experimental conditions. The ϕ_{em} of complex **2** was also slightly higher than the standard under the same conditions. At 77 K, complex **4** produced an intense, structureless emission peak centered at 20 600 cm^{-1}.

Computational Section

The singlet ground-state geometries of the complexes **1**−**4** were optimized in the gas phase using the B3LYP[18] functional of the Gaussian '03[19] program package. The Stuttgart-Dresden (SDD) ECP[20] was used for the Re core potentials. The {(8s7p6d)/[6s5p3d]}-GTO was applied for the valence shell of Re together with the all-electron 6-311G*

Table 5. Calculated ^{3}MMLCT State Energies[a] and Emission Properties of the Complexes at 77 K and Room Temperature[b,f]

complexes	E_{calc}	E_{exp} 77 K	E_{exp} RT	τ_{em}, μs 77 K	τ_{em}, μs RT	φ_{em} [d] RT
[Re(bpy)(CO)$_3$Cl][c]	20.6	19.2	16.8	3.9	0.037	0.0052
[Re(bpy)(CO)$_3$(py)](CF$_3$SO$_3$)[c]	21.4	20.2	18.0	5.4	0.33	0.10
[Re(bpy) (CO)$_3$(CNx)](CF$_3$SO$_3$)	21.9	22.2	19.3	8.8	1.22	0.27
		20.9				
		19.6 (s)[e]				
		18.0 (s)[e]				
[Re(bpy)(CO)(CNx)$_3$](PF$_6$)	21.4	20.6		6.3		

[a] In ethanol. [b] In 4:1 ethanol/methanol. [c] Emission properties were also reported in CH$_2$Cl$_2$ in refs 9 and 14. [d] Relative to [Ru(bpy)$_3$]$^{2+}$ (ref 13). [e] (s) = shoulder. [f] Energies in $\times 10^3$ cm^{-1}.

Table 6. Selected Geometry Parameters of **3** Based on Calculated Singlet Ground and Lowest-Lying Triplet State Geometries and X-ray Crystallography

source	Re−N_2, Å	Re−C_4, Å	Re−C_1, Å	Re−C_2, Å	N_1−C_4, Å	C_1−O_1, Å	C_2−O_2, Å	Re−C_4−N_1, deg	Re−C_1−O_1, deg	C_4−N_1−C_{15}, deg
X-ray	2.171(3)	2.074(4)	1.971(4)	1.932(4)	1.161(5)	1.136(5)	1.151(5)	175.8(3)	178.5(4)	171.2(4)
singlet	2.23	2.10	2.01	1.96	1.17	1.14	1.15	177.8	179.6	178.4
triplet	2.16	2.09	2.06	1.99	1.17	1.13	1.14	179.1	179.9	178.7

Table 7. Calculated Triplet Excited-States of Complexes **1**−**4** in Ethanol Based on the Lowest-Lying Triplet State Geometry[a]

state	*f*	$\psi_o \rightarrow \psi_v$	type	E_{VER}
		Complex **1**		
1	0.00	H-2 → H (1.0)	Re_d, CO → Re_d, Cl	22.9
2	0.00	H-3 → H (1.0)	bpy → Re_d, Cl	25.8
3	0.01	H-1 → H (1.0)	Re_d, Cl → Re_d, Cl	27.0
4	0.02	L → L+2 (0.8)	bpy → bpy	32.6
		L → L+4 (0.6)	bpy → del	
		Complex **2**		
1	0.00	H-2 → H (1.0)	Re_d, CO → Re_d, CO	24.8
2	0.00	H-3 → H (1.0)	bpy → Re_d, CO	25.5
3	0.02	H-1 → H (1.0)	Re_d, CO → Re_d, CO	26.7
4	0.00	L → L+1 (1.0)	bpy → py	32.1
		Complex **3**		
1	0.00	H-2 → H (1.0)	Re_d, CO → Re_d, CNx	25.8
2	0.00	H-4 → H (1.0)	bpy → Re_d, CNx	26.2
3	0.04	H-1 → H (0.9)	Re_d → Re_d, CNx	26.5
4	0.00	H-3 → H (1.0)	CNx → Re_d, CNx	31.9
		Complex **4**		
1	0.00	H-3 → H (1.0)	CNx → Re_d, CNx	21.4
2	0.00	H-2 → H (0.8)	Re_d, CNx → Re_d, CNx	22.0
		H-1 → H (0.6)	Re_d, CNx → Re_d, CNx	
3	0.00	L → L+1 (1.0)	bpy → CNx	24.2
4	0.01	H-1 → H (0.8)	Re_d, CNx → Re_d, CNx	26.0
		H-2 → H (0.5)	Re_d, CNx → Re_d, CNx	

[a] E_{VER} is the energy of the vertical transition in $\times 10^3$ cm^{-1}, *f* is the oscillator strength, and ψ_o and ψ_v are the occupied and the virtual orbitals that define the transition. The transition type is determined based on the change in the spatial distribution from occupied to virtual orbital. The absolute value of the transition coefficient for each transition is given in parentheses. H = HOMO, L = LUMO, and del = delocalized. (See text for calculation details.)

basis set[21] for Cl, O, N, C, and H atoms. Selected parameters of the optimized geometry of complex **3** are presented in Table 6. The optimized geometries of the complexes are listed in Supporting Information Table S1.

Nonequilibrium TDDFT[22]/CPCM[23] calculations were employed to produce a number of singlet excited-states[24] of complexes **1**−**4** in ethanol based on the singlet ground-state geometry optimized in the gas phase.[25] The output contained information for the excited-state energies and oscillator strengths (*f*) and a list of the excitations that give rise to each excited state, the orbitals involved as well as the wave function coefficients of the excitations. The singlet excited-states of the four complexes are presented in Figure 3 as vertical bars with height equal to the extinction coefficient calculated from the oscillator strength.[5,6]

The lowest-lying triplet state geometries of the four complexes were calculated using unrestricted B3LYP in the gas phase. The spin contamination from states of higher multiplicity was low. The value of $\langle S^2 \rangle$ was 2.010 for **1**, 2.012 for **2**, 2.018 for **3**, and 2.017 for **4**. The energies of the lowest-lying triplet states were higher than these of the corresponding ground states by 20 600 cm^{-1} for **1**, 21 400 cm^{-1} for **2**, 21 900 cm^{-1} for **3**, and 16 800 cm^{-1} for **4** (Table 7 and Figure 4). The lowest-lying triplet states were ^{3}MLLCT states and featured single HOMO and LUMO occupancy.

A number of triplet excited-states were computed based on the ^{3}MLLCT state geometry for the four complexes. The four low-lying triplet excited-states are listed in Table 7 and shown in Figure 3, even if the *f* value is low. These excited states were used for the interpretation of the temperature-dependent emission properties of the complexes.

Vibrational analysis was performed using B3LYP for the ground-state optimized geometries of the four complexes. The frequencies of the most intense vibrations are presented in Table 8.

Table 8. Calculated and Experimental Vibrational IR Frequencies (cm^{-1}) for the Four Complexes[a]

1		**2**		**3**		**4**		
calc.	exp.	calc.	exp.	calc.	exp.	calc.	exp.	assignment
				2173	2173	2084	2072[c]	a[b]
2037	2022	2061	2026	2065	2036	1950	1885	b[b]
1964	sh	1989	sh	2013	sh			b[b]
1939	1890	1977	1906	1992	1955			b[b]
		1608	1603	1597	1603	1603	1590	c[b]
1466	1471	1484	1448	1467	1473	1468	1468	d[b]
757	764	765	776	763	790	761	774	e[b]
634	647	629	637	623	638	589	557	f[b]
				496	488	488	491	g[b]

[a] The calculated values were factored by 0.975. [b] a = CN stretch, b = CO stretch, c = CNx or py ring breathing modes, d = bpy ring breathing modes, e = out-of-plane H (bpy) vibrations, f = symmetric CO bending, and g = CNx vibrations. [c] Calculated vibrational frequencies at 2154 and 2099 cm^{-1} correspond to shoulders near the experimental peak at 2072 cm^{-1}.

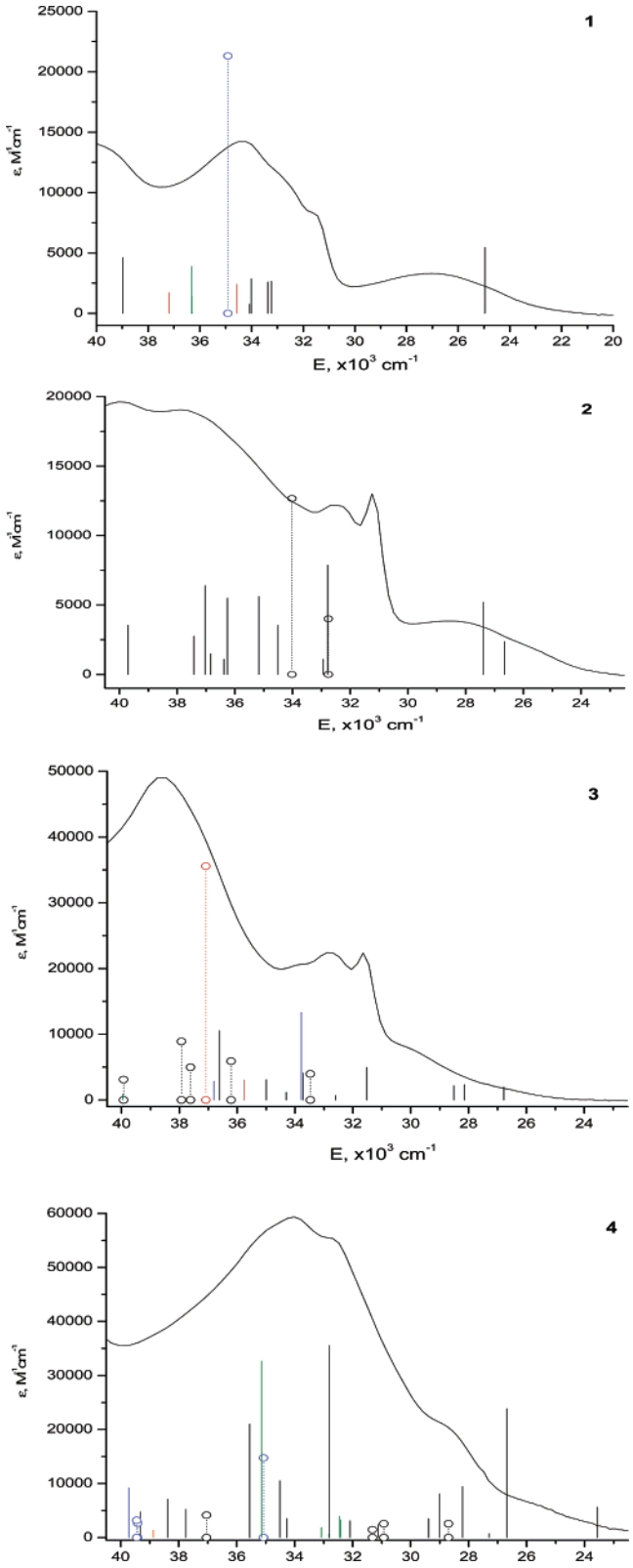


Figure 3. Experimental absorption spectra of **1**−**4** and calculated singlet excited-states. The excited states are shown as vertical bars with height equal to the extinction coefficient.[5a] Black = MLLCT, green = LLCT, blue = $\pi \rightarrow \pi^*$, red = MCDCT, orange = Re_d, CNx → Re_s and (○·····○) = mixed excited-state.

Discussion

X-ray Structure. As noted in the results section, $[Re(CO)_3(bpy)(CNx)]^+$ (**3**) adopted a distorted octahedral geometry.

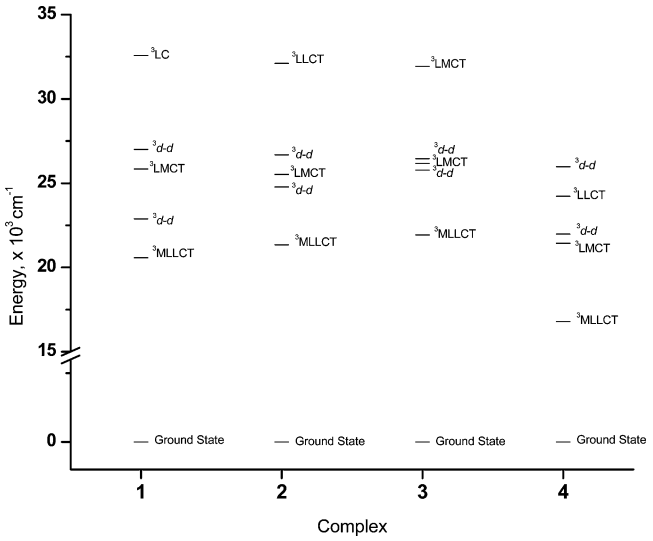


Figure 4. Triplet excited-state energy diagram for complexes **1**−**4**.

The increase in Re−C (CO) bond distance of 0.04 Å compared to the other two Re−C (CO) bond distances is a result of the trans effect resulting from the strong σ donating character of the CNx ligand to the metal center. The bond elongation of CO trans to the CNx ligands of 0.04 Å was analogous to that reported for the $[Ru(bpy)_2(CNx)_2](PF_6)_2$ by our group where the Ru−N (bpy) trans to CNx was 0.03 Å longer than for Ru−N (bpy) cis to CNx.[8a]

In a related X-ray crystallographic study, the Re−C (L) distance in *cis*-$[ReCl(CO)_3L_2]$ (L = 2,6-diisopropyl-4-acetyleno-phenylisocyanide)[26] was 2.092(5) Å, or 0.02 Å longer than the Re−C (CNx) distance found for $[Re(CO)_3(bpy)(CNx)](PF_6)$. The Re−C (CO trans to L) distance of 1.973(6) Å was the same as Re−C (CO trans to CNx) in this report. However for Re−C with CNx trans to CNx, the bond distance for $[Re(CNx)_5Cl]$ was only 2.019(7) Å[8b] compared to 2.074(4) Å for $[Re(CO)_3(bpy)(CNx)]^+$. The Re−C (CNx) distance was longer by 0.12 Å than the Ru−C (CNx) distance in $[Ru(bpy)_2(CNx)_2](PF_6)_2$.[8a] This may be due to the larger ionic radius of Re(I).

Molecular Orbitals. The molecular orbital energy diagram for the four complexes in ethanol is shown in Figure 5. The HOMO, HOMO-1, and HOMO-2 of complexes **1**−**4** contained 50% or higher Re_d character. The remaining contribution was from the nonimine ligands. Additionally, for complex **1**, the HOMO and HOMO-1 contained equal contributions (∼20% for each) from the CO and Cl moieties, whereas HOMO-2 contained 23.5% CO but less than 1% Cl character. The HOMO, HOMO-1, and HOMO-2 for complex **2** contained ∼20% CO character and negligible imine ligand contributions. For complex **4**, the HOMO, HOMO-1, and HOMO-2 contained more than 23% CNx ligand contributions while the CO contributions were low. The HOMO of complex **3** contained ∼32% CNx ligand contribution, whereas the HOMO-1 and HOMO-2 contained 18% or higher CO character (the CNx ligand contribution was less than 7%). The LUMOs contained 80% or higher bpy ligand π^* character (Figure 6). The LUMOs+1 were located on the bpy and py ligands for complexes **1** and **2**,

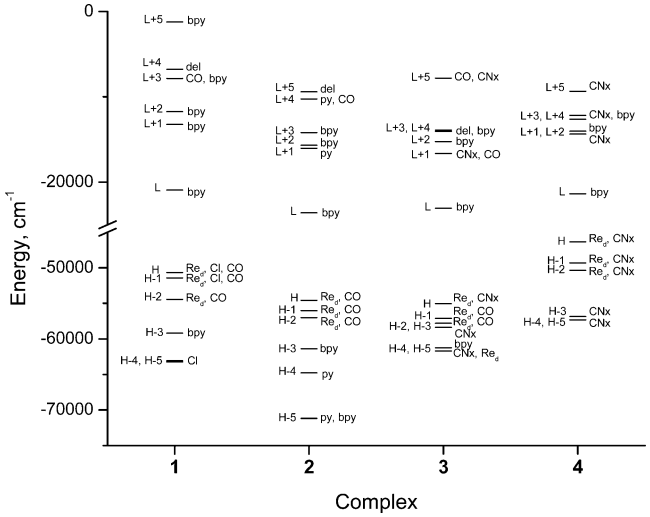


Figure 5. Molecular orbital energy diagram for six occupied and six virtual frontier orbitals of **1**−**4** in the singlet ground state in ethanol.

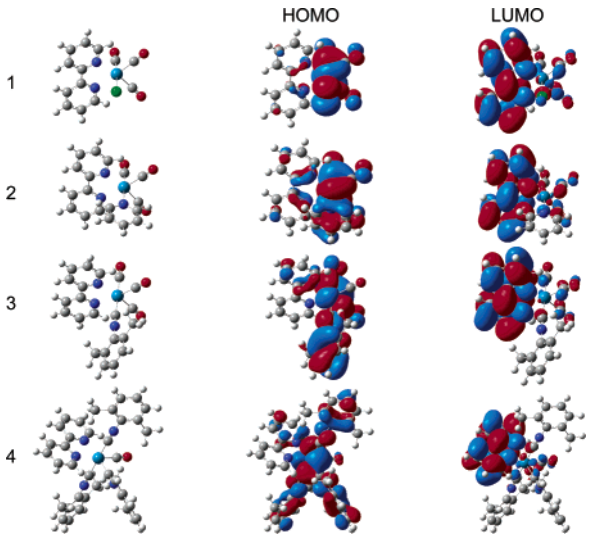


Figure 6. HOMO and LUMO schematic diagrams for complexes **1**−**4**.

respectively, and on the CNx ligands for complexes **3** and **4**. The percent orbital contributions are listed in Supporting Information Table S2.

The schematic diagrams of the HOMO and the LUMO for the four complexes are shown in Figure 6. The HOMOs contained more than 50% Re$_d$ character and the remaining contributions were almost evenly distributed among the nonimine ligands, whereas the LUMOs were on the bpy ligand. The HOMO−LUMO energy difference increased from 29 800 cm^{-1} for complex **1**, to 31 000 cm^{-1} for **2**, and 32 000 cm^{-1} for **3**. The HOMO−LUMO energy gap for complex **4** was only 25 000 cm^{-1}.

Geometry Optimization of 3. Selected parameters of the optimized geometry of complex **3** in the singlet ground and lowest-lying triplet states are listed in Table 6, next to the results of the X-ray structure determination. The singlet ground-state bond lengths and bond angles presented are in general agreement with the experimental values. The Re−N distances were overestimated by ∼0.05 Å and the Re−C (CO and CNx) distances by ∼0.03−0.04 Å. The $C_4-N_1-C_{15}$ angle was 178.4° in comparison to the 171.2° found experimentally. The elongation of the Re−C (CO trans to CNx) bond relative to Re−C (CO cis to CNx) was confirmed computationally. It was due to the electron-donating effect of the isocyanide ligand (vide infra). The deviation of the experimental value from 180° could be a result of crystal packing effects. These results obtained with one of the largest basis sets appear satisfactory despite the shortcomings associated with the calculation of the metal−ligand bond lengths using B3LYP functional.[5,6a]

The geometry of complex **3** in the ^{3}MLLCT state was slightly different from the ground-state geometry. The differences listed in Table 6 can be interpreted using the molecular orbital schematic diagram shown in Figure 6. The Re−N_2 distance decreased from 2.23 to 2.16 Å in the triplet relative to the singlet ground state. This bond shortening could be due to the mixing between the Re dπ HOMO and the bpy π* LUMO in the lowest-lying triplet state as reported for *fac*-[Re(bpy)(CO)$_3$(4-ethylpyridine)]$^+$.[2a] The Re−C_1 and Re−C_2 distances increased by 0.05 Å and 0.03 Å, respectively. The bonding character of the HOMO with respect to the Re−C (CO) bonds could account for these bond elongations in the triplet state. The bonding and antibonding character was determined by visual examination of the phases of the molecular orbital for each diagram. The phases are related to the spatial distributions of alpha (α) and beta (β) electron densities[27] shown in red and blue colors, respectively.

Mulliken Charges. Mulliken charges were calculated for the singlet ground-state geometry, and the solvent effect of ethanol was accounted for using CPCM. The changes in the Mulliken charge of the metal atom can be used for the evaluation of the electron-donating power of the ligands.[8a] The Re charge lowering can be expressed as follows: $Re^{calc} = Re^{1-\Sigma nL}$, where n is the number of coordinated atoms from ligand L. The relative electron-donating power of CNx was 0.08 as determined from the 0.54 charge on Re[8b] for [Re(CNx)$_6$]$^+$ using the following mathematical relationship, $Re^{0.54} \sim Re^{1-6\times0.08}$. Analogously, the electron-donating power of CO was determined to be 0.11 based on a residual charge of 0.34 on Re[28] for [Re(CO)$_6$]$^+$. For [Re(CO)$_3$(bpy)(CNx)] the donating power of each coordinated N atom was determined to be −0.05 from the following: $Re^{0.81} \sim Re^{1-3\times0.11-3(-0.05)}$. The electron donating power of the ligands for Re(I) can be arranged as follows CO > CNx > N. When these electron donating powers are used to calculate the predicted Re charge produced for complexes **3** and **4**, the values obtained were in good agreement with the DFT computed values. For complexes **3** and **4** the above approach produced $Re^{1-3\times0.11-1\times0.08-2\times(-0.05)} = Re^{0.69}$ and $Re^{1-1\times0.11-3\times0.08-2\times(-0.05)} = Re^{0.75}$ compared to 0.67 and 0.73, respectively, from DFT. The Re charge in complex **1** was 0.56. However, Mulliken charge analysis of the donating power of the chloro ligand for complex **1** was not performed due to concerns of properly accounting for the negative charge on the chloro ligand by the CPCM method.

The donating power of CNx and N (imine) is greater for Ru(II) (CNx = 0.32; N = 0.1)[8a] than for Re(I) (CNx = 0.08; N = −0.05), but in both cases the contribution of CNx >

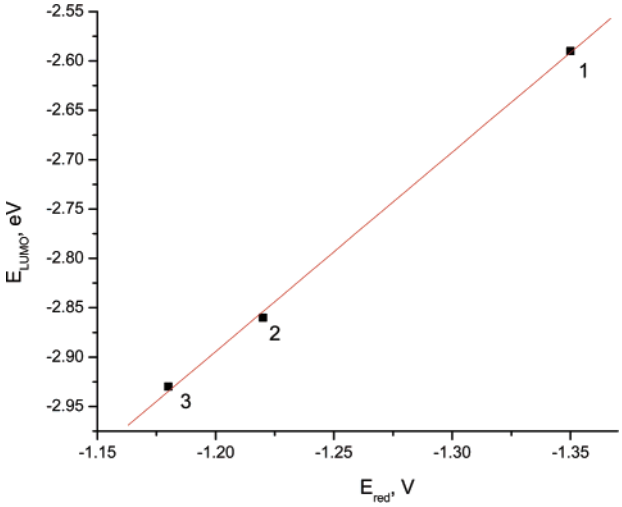


Figure 7. Linear dependence of $E_{1/2(red)}$ and the LUMO energies for complexes **1**−**3**.

N. Also the contribution of each to Re(I) is smaller than to Ru(II). Perhaps the differences are due to the initial charges of the complexes, but more examples are needed to unravel the trends.

Electrochemical Behavior. The electronic effects of the CNx ligand were evident in the electrochemical results as listed in Table 4. The redox potentials of importance are those derived from the processes involving the HOMOs which contained primarily Re dπ orbitals and the LUMOs which contained the π^* orbital located on the bpy ligand. Hence, oxidation involves the removal of an electron from Re (HOMO) and reduction involves the addition of an electron to the bpy π^* orbital (LUMO).

The reduction potentials were assigned based on the energies and spatial distributions of the LUMO (Figure 5). The potential for the single reduction of the bpy ligand increased from −1.35 V for **1** to −1.22 V for **3** and −1.18 V for **2**. The reduction potentials of complexes **1**−**3** were linearly dependent on the DFT calculated LUMO energies as shown in Figure 7. The slope of the line was −2.0 (±0.1) compared to −2.32 reported for a series of 10 isoelectronic Ru(II) diimine complexes.[1]

The oxidation potentials however, which varied considerably across the series, were all irreversible and increased from **1** to **3**. The energy gap between the HOMO and the LUMO increased in the order **1** < **2** < **3**. The potential differences ($E_{ox} - E_{red}$) follow the same order as the room-temperature emission maxima. The results indicated the parallel relationship between the thermodynamic and electronic energy gaps.[29,30] The fact that **3** has the highest energy gap compared to **1** and **2** indicated the strong influence of the CNx ligand on the system.

Singlet Excited-States and Electronic Absorption Spectra. The singlet excited-states of complexes **1**−**4** with $f >$ 0.01 are shown in Figure 3 as vertical bars with height equal to the extinction coefficient. The bars are presented in colors that correspond to the type of singlet excited-state as follows: black = MLLCT, green = LLCT, blue = $\pi \rightarrow \pi^*$, orange = Re_d,CNx → Re_s, and red = MCDCT (metal-to-complex-delocalized charge transfer). These assignments were made based on the major contributing excitation. The singlet excited-states had contributions from several excitations. For those presented with solid bars there was one major contributing excitation (with a transition coefficient for the major excitation being higher than the transition coefficient of the other excitations by more than 0.2). For some singlet excited-states, however, there was more than one contributing excitation with high transition coefficients. Singlet excited-states that contained contributions from several excitations with transition coefficients that were within 0.2 of the major excitation transition coefficient are assigned as mixed singlet excited-states and are presented with the vertical bars symbolized as ○·····○.

The excited states of complex **1** were mostly of MMLCT type. All of the MLLCT states were associated with transitions that originated from the HOMO, HOMO-1, and HOMO-2 that contained Re_d, Cl, and CO contributions. The excited state at 34 900 cm^{-1} with the highest ϵ involved equal contributions from bpy $\pi \rightarrow \pi^*$ and an MLLCT state. Two degenerate Cl → bpy excited-states were computed at 36 300 cm^{-1}. The singlet excited-states of complex **2** were assigned as Re_d,CO → bpy and Re_d,CO → py states, or MLLCT states. Excited states 4 and 7 (shown with the notation ○·····○ in Figure 3) contained significant bpy $\pi \rightarrow \pi^*$ contributions. The majority of the singlet excited-states of complex **3** were MLLCT states. In addition, there were two $\pi \rightarrow \pi^*$ states, one involving the bpy ligand and the other involving the CNx ligand. The singlet excited-state at 39 900 cm^{-1} was due to three excitations—a CNx ligand $\pi \rightarrow \pi^*$, MMLCT, and a CNx → bpy with transition coefficients of 0.4, 0.3, and 0.3, respectively. This state was labeled as $\pi \rightarrow \pi^*$ according to the assignment of the major excitation and is shown in Figure 3 as a blue vertical line symbolized as ○·····○. For complex **4**, the singlet excited-states below 32 000 cm^{-1} were mainly of MLLCT character. The singlet excited-state at 35 200 cm^{-1} was assigned as LLCT (CNx → bpy). Two CNx $\pi \rightarrow \pi^*$ states were calculated at 39 400 and 39 700 cm^{-1}. The higher CNx ligand contribution to the frontier molecular orbitals and the smaller HOMO−LUMO gap for complex **4** compared to **1**−**3** resulted in an increase in the number of singlet excited states calculated in the energy range 28 000−32 000 cm^{-1}. These excited states correlated well with the experimental peak broadening and the merging of the LC and MLLCT peaks in complex **4**. The singlet excited-states are listed in Supporting Information Table S3.

The calculated singlet excited-state energies correlated well with the experimental UV−vis peaks in the same solvent (Figure 3 and Table 3). For complex **1**, the MLLCT singlet excited-state at 25 000 cm^{-1} is 1900 cm^{-1} lower in energy relative to the position of the broad UV−vis peak at 26 900 cm^{-1}, whereas the bpy $\pi \rightarrow \pi^*$ excited-state at 34 100 cm^{-1} is 100 cm^{-1} lower than the most intense experimental peak at 34 200 cm^{-1}. For complex **2** the MLLCT excited-state at 27 400 cm^{-1} was 1200 cm^{-1} lower in energy compared to the UV−vis peak at 28 600 cm^{-1}. The mixed excited-states at 32 800 and 37 400 cm^{-1} were within 500 cm^{-1} from the UV−vis peaks at 32 700 and 37 900 cm^{-1}, respectively. These states were assigned as MLLCT but contained significant $\pi \rightarrow \pi^*$ character. For complex **4**, the singlet excited-states of 32 800 cm^{-1} and 34 300 cm^{-1} were blue-

shifted by 100 cm^{-1} and 300 cm^{-1} relative to the experimental peaks at 32 700 and 34 000 cm^{-1}, respectively. The excited states were of MLLCT type. For complexes **2**−**4**, the agreement between the calculated and experimental data was remarkable. The experimental peaks at energies higher than 40 000 cm^{-1} were assigned as $\pi \rightarrow \pi^*$ states according to the conventional assignment. Simulation of the singlet excited-states into Gaussian line shapes and subsequent integration, following a procedure previously reported[5a] (not shown), did not correlate well with the UV−vis spectrum curvature. The use of the TDDFT/CPCM method is known to produce optical energies in good agreement with the experimental absorption spectra. The oscillator strengths calculated using TDDFT/CPCM method were higher than both the experimental values determined in the respective solvent and the calculated values in the gas phase for a series of ruthenium(II) polypyridyl complexes containing CNx ligand.[6b,8a] In another study, the *f* values computed in 2-methyltetrahydrofurane solution using the TDDFT/PCM method were also significantly higher than the experimental values in the same solvent and the calculated values in the gas phase.[31]

Triplet Excited-States. Four triplet excited-states for each of the complexes **1**−**4** were computed in ethanol based on the ^{3}MLLCT geometry and are listed in Table 7. The major excitations in these states had transition coefficients higher than 0.8. The three lowest-lying triplet excited-states for complexes **1**−**3** involved excitations to the rhenium and nonimine ligand centered HOMO (Figure 6). These states were pure, with the transition coefficients of the major excitations of 1.0. Excited state 4 for complexes **1** and **2** involved excitations from the LUMO (bpy) to LUMO+2 (bpy) and LUMO+1 (py), respectively. In complex **4** each of the triplet excited-states 2 and 4 involved two excitations of the same type.

The relative energies of the triplet excited-states are presented in the schematic diagram in Figure 4. The excited states that involved occupied and virtual molecular orbitals with more than 50% Re contribution were assigned as ^3d-d states, whereas states that involved ligand-to-metal charge-transfer excitations, like states 2 for complexes **1**−**3**, were assigned as ^{3}LMCT states. The energies of the ^{3}MLLCT states increase from 20 600 cm^{-1} for **1**, 21 400 cm^{-1} for **2**, and 21 900 cm^{-1} for **3** following the same trend as the experimental emission energies of 19 200 cm^{-1} for **1**, 20 200 cm^{-1} for **2**, and 22 200 cm^{-1} for **3** at 77 K. Above the ^{3}MLLCT states were ^3d-d states and ^{3}LMCT states for complexes **1**−**3**. The ^{3}MLLCT state of complex **4** was only 16 800 cm^{-1} above the ground state, and the ^{3}LMCT state at 21 400 cm^{-1} was only 200 cm^{-1} lower than the experimental emission energy at 77 K. The ^{3}LMCT is then assigned as the emitting state for complex **4**. The thermal population of the ^3d-d state at 22 000 cm^{-1} accounts for the loss of room-temperature emission of complex **4**.

The ^3d-d and ^{3}LLCT states were obtained via single electron vertical excitations from the ^{3}MLCT states and the excited-state energies reported were not the minima.[32] The ^3d-d transitions are symmetry forbidden[33] and have very low oscillator strengths (Table 7). The role of the ^3d-d states in quenching room-temperature emission is discussed in the succeeding section.

Recent study suggests that the addition of diffuse functions to double-ζ basis sets can improve the effectiveness of DFT for the calculations of reaction and conformation energies of butadiene and 1,2-ethanediol.[34a] In a TDDFT/PCM study the 6-311++G** basis set is utilized for a more complete description of the mixing between the valence and Rydberg excite states of acrolein.[34b] Further studies on the effects of the addition of diffuse functions to double and triple-ζ basis sets for the calculation of excited-state energies are needed and have the potential of improving the results presented here.

Vibrational Analysis. The vibrational frequencies for the four complexes were calculated, factored[35] by 0.975, and listed in Table 8. The most distinct and intense calculated frequencies were correlated with the experimental results. The C≡N and C═O stretching modes were among the most pronounced peaks. Excellent agreement was obtained between the calculated and the experimental values of the C≡N stretching frequencies for complexes **3** and **4**. For complex **4**, there were three C≡N stretching modes calculated but only the one at 2072 cm^{-1} was resolved experimentally. The calculated frequencies at 2154 and 2099 cm^{-1} corresponded to the shoulders on both sides of the experimental peak at 2072 cm^{-1}. Three C═O stretching frequencies were calculated for each of the complexes **1**−**4** and correlated with the experimental values. The assignment of the three C═O frequencies for each of the complexes **1**−**3** was limited by the experimental resolution. Both the experimental and the calculated C═O stretching modes shifted to higher frequencies in the order **1** < **2** < **3**. The higher vibrational frequency of **3** was anticipated based on the electron-donating power of the CNx ligand. The calculated values were higher than the experimental by 30 cm^{-1}−75 cm^{-1}. The calculated ring breathing and symmetric C═O bending modes correlated well with the experimental values. In the low energy range of the spectrum an intense vibration was located. According to the computational results this involved the vibration of the CNx ligand with respect to the rest of the complex. The vibrational frequency of CNx for **3** was computed at 496 cm^{-1}, compared to the experimental value of 488 cm^{-1}. For complex **4**, three CNx specific vibrations were expected, but only one was resolved experimentally. The coordinated CNx ligand vibrations occurred ∼20 cm^{-1} lower in frequency than the corresponding modes in $[Ru(bpy)_2(CNx)Cl]^+$, $[Ru(bpy)_2(CNx)(py)]^{2+}$, and $[Ru(bpy)_2(CNx)_2]^{2+}$.[8a] This shift is likely due to the heavier Re atom.

Emission Properties and Excited-State Lifetimes. The room-temperature emission spectra of complexes **1**−**3** compared to $[Ru(bpy)_3]^{2+}$ are shown in Figure 8A. The 77 K emission spectra of complexes **1**−**4** are presented in Figure 8B. Complex **4** was nonemissive at room temperature. The emission maxima of the complexes at 77 K are blue-shifted compared to that at room temperature.

Figure 9 shows a model diagram that will aid in discussing the observed emission properties and excited-state lifetimes of the complexes in the series. It shows the three states involved—ground state, ^{3}MLCT, and ^{3}LC (triplet ligand-centered state) both at room temperature and at 77 K. In this case, the ^{3}MLCT state is taken to be essentially temperature dependent. Hence, the energy gap between the

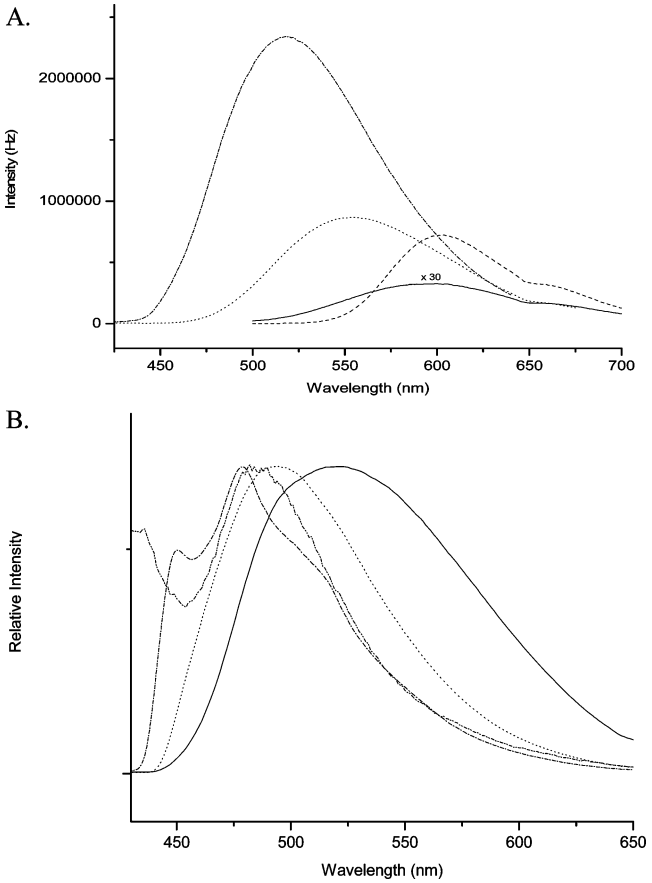


Figure 8. Emission spectra of complexes **1** (—), **2** (···), **3** (- · -), and $[Ru(bpy)_3]^{2+}$ (- - -) at room temperature (A) and **1** (—), **2** (···), **3** (- · -), and **4** (- ·· -) at 77 K (B) in 4:1 (v/v) ethanol: methanol.

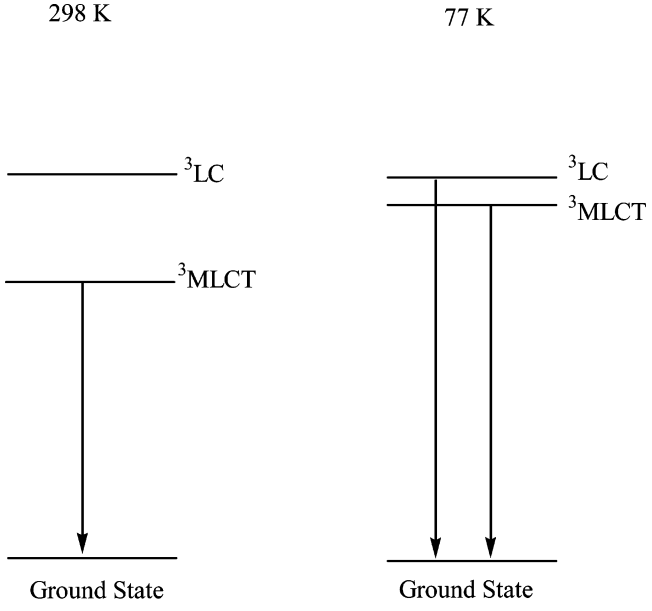


Figure 9. Energy level diagram showing the emitting state of the complexes in the series.

emitting triplet state at room temperature is smaller than at 77 K resulting in the observed blue shift of the emission maxima at 77 K relative to room temperature. At room temperature, the emitting state is the ^{3}MLCT, which is equilibrated with respect to the solvent environment and geometry of the complex. The reorientation of the solvent cage is fast compared to the rate of deactivation of the emitting state. The trend of increasing emission energies and excited-state lifetimes in the order **3** > **2** > **1** is in accord with the energy gap law. The thermodynamic and the electronic energy gaps increase parallel to each other in the above order as has been previously reported.[17]

Conclusion

Four Re(I) bipyridyl complexes were investigated using spectroscopic methods and DFT calculations. The results were correlated and revealed the following: (1) The electronic, electrochemical, thermodynamic, HOMO−LUMO, and emitting state energy gaps as well as the emission lifetimes and ^{3}MLLCT energies increased in the order **1** < **2** < **3**. (2) The electron-donating power of the CNx ligand was evaluated, and its effect on the computed optimized geometry parameters, singlet excited states, and C═O vibrational frequencies was correlated with the experimental results from X-ray crystallography and UV−vis and IR spectroscopies. (3) The three complexes (**1**−**3**) were emissive both at room temperature and at 77 K. (4) The complexes showed reversible reduction but irreversible oxidation waves. (5) The LUMO energies were linearly dependent on the reduction potentials of complexes **1**−**3**. (6) Complex **4** was nonemissive at room temperature but showed intense emission at 77 K with a lifetime of 6.3 μs which is longer than 5.4 μs for complex **2** but shorter than the 8.8 μs for **3**. A low-lying ^3d-d state was responsible for the loss of room-temperature emission in **4**. The complex showed a reversible oxidation but no reduction wave.

Acknowledgment. We thank the support of National Science Foundation under Grant No. EIA-0216178 and Grant No. EPS-0236913, matching support from the State of Kansas and the Wichita State University High Performance Computing Center, the Wichita State University Office of Research Administration, the Department of Energy, and Parker Fellowships (J.M.V. and S.R.S.).

Supporting Information Available: Crystallographic data for $[Re(bpy)(CO)_3(CNx)](PF_6)$ in CIF format, the optimized geometries (Table S1), the percent orbital contributions (Table S2), and the calculated singlet excited-state energies of the four complexes (Table S3). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Stoyanov, S. R.; Villegas, J. M.; Rillema, D. P. *Inorg. Chem.* **2002**, *41*, 2941.

(2) Dattelbaum, D. M.; Martin, R. L.; Schoonover, J. R.; Meyer, T. J. *J. Phys. Chem. A* **2004**, *108*, 3518−3526. (b) Dattelbaum, D. M.; Omberg, K. M.; Hay, J. P.; Gebhart, N. L.; Martin, R. L.; Schoonover, J. R.; Meyer, T. J. *J. Phys. Chem. A* **2004**, *108*, 3527−3536.

(3) Yang, L.; Ren, A.-M.; Feng, J.-K.; Liu, X.-J.; Ma, Y.-G.; Zhang, M.; Liu, X.-D.; Shen, J.-C.; Zhang, H.-X. *J. Phys. Chem. A* **2004**, *108*, 6797−6808. (b) Dyer, J.; Blau, W. J.; Coates, C. G.; Creely, C. M.; Gavey, J. D.; George, M. W.; Grills, D. C.; Hudson, S.; Kelly, J. M.; Matousek, P.; McGarvey, J. J.; McMaster, J.; Parker, A. W.; Towrie, M.; Weinstein, J. A. *Photochem. Photobiol. Sci.* **2003**, *2*, 542−554.

(4) Frantz, S.; Rall, J.; Hartenbach, I.; Schleid, T.; Zalis, S. Kaim, W. *Chem.-A Eur. J.* **2004**, *10*, 149−154.

(5) Monat, J. E.; Rodriguez, J. H.; McCusker, J. K. *J. Phys. Chem. A* **2002**, *106*, 7399. (b) Rodrigues, J. H.; Wheeler, D. E.; McCusker, J. K. *J. Am. Chem. Soc.* **1998**, *120*, 12051.

(6) Stoyanov, S. R.; Villegas, J. M.; Rillema, D. P. *Inorg. Chem.* **2003**, *42*, 7852. (b) Stoyanov, S. R.; Villegas, J. M.; Rillema, D. P. *Inorg. Chem. Commun.* **2004**, *7*, 838−841.

(7) Guillemoles, J.-F.; Barone, V.; Joubert, L.; Adamo, C. *J. Phys. Chem. A* **2002**, *106*, 11345.

(8) Villegas, J. M.; Stoyanov, S. R.; Wei, H.; Lockyear, L. L.; Reibenspies, J.; Rillema, D. P. *Inorg. Chem.* **2004**, *43*, 6383−6396. (b) Villegas, J. M.; Stoyanov, S. R.; Reibenspies, J.; Rillema, D. P. *Organomet.* **2004**, submitted.

(9) Wrighton, M.; Morse, D. L. *J. Am. Chem. Soc*. **1974**, *96*, 998. (b) Giordano, P. J.; Wrighton, M. S. *J. Am. Chem. Soc.* **1979**, *101*, 2888.

(10) Lees, A. *J. Chem. Rev.* **1987**, *87*, 711. (b) Giordano, P. J.; Fredericks, S. M.; Wrighton, M. S.; Morse, D. L. *J. Am. Chem. Soc*. **1978**, *100*, 2257.

(11) Leasure, R. M.; Sacksteder, L.; Nesselrodt, D.; Reitz, G. A.; Demas, J. N.; DeGraff, B. A. *Inorg. Chem*. **1991**, *30*, 3722.

(12) Jovanovic, B.; Manojlovic-Muir, Lj.; Muir, K. W. *J. Chem. Soc., Dalton Trans.* **1972**, 1178.

(13) Villegas, J. M.; Stoyanov, S. R.; Rillema, D. P. *Inorg. Chem.* **2002**, *41*, 6688.

(14) Giordano, P. J.; Fredericks, S. M.; Morse, D. L.; Wrighton, M. S. *J. Am. Chem. Soc*. **1978**, *100*, 2257. (b). Caspar, J. V.; Meyer, T. J. *J. Phys. Chem.* **1983**, *87*, 952−957. (c) Luong, J. C. Ph.D. Thesis, Massachusetts Institute of Technology, 1981.

(15) Treichel, P. M.; Williams, J. P. *J. Organomet. Chem.* **1977**, *135*, 39−51.

(16) Wrighton, M. S.; Geoffroy, G. L. *Organometallic Photochemistry*; Academic Press: New York, 1979; Chapter 2. (b) Luong, J. C.; Nadjo, L.; Wrighton, M. S. *J. Am. Chem. Soc*. **1978**, *100*, 5790.

(17) Wallace, L.; Rillema, D. P. *Inorg. Chem.* **1993**, *32*, 3836−3843.

(18) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098. (b) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648. (c) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785. (d) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200.

(19) 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.03; Gaussian, Inc.: Pittsburgh, PA, 2003.

(20) Andrae, D.; Hauessermann, U.; Dolg, M.; Stoll, H.; Preuss, H. *Theor. Chim. Acta* **1990**, *77*, 123.

(21) McLean, A. D.; Chandler, G. S. *J. Chem. Phys.* **1980**, *72*, 5639. (b) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650.

(22) Stratmann, R. E.; Scuseria, G. E.; Frisch, M. J. *J. Chem. Phys* **1998**, *109*, 8218. (b) Bauernschmitt, R.; Ahlrichs, R. *Chem. Phys. Lett.* **1996**, *256*, 454. (c) Casida, M. E.; Jamorski, C.; Casida, K. C.; Salahub, D. R. *J. Chem. Phys.* **1998**, *108*, 4439.

(23) Cossi, M.; Barone, V. *J. Chem. Phys.* **2001**, *115*, 4708. (b) Barone, V.; Cossi, M. J. *Phys. Chem. A* **1998**, *102*, 1995. (c) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. Comput. Chem.* **2003**, *24*, 669.

(24) The CPCM is designed to account for the bulk physical properties of the solvent. It does not account for specific solvent−solute interactions. The TDDFT is known to perform well for the computing of charge transfer excited-states between closely spaced moieties.

(25) Geometry optimization in solvents was not achieved. Partial optimizations (change in distance of less than 0.001 Å and change in angles of less than 0.01°) followed by TDDFT/CPCM calculation produced excited-state energies that were not in better agreement with the experimental excited-state energies than the excited-state energies based on the gas-phase optimized geometry.

(26) Yang, L.; Cheung, K.-K.; Mayr, A. *J. Organomet. Chem.* **1999**, *585*, 26−34.

(27) Foresman, J. B.; Frisch, Æ. *Exploring Chemistry with Electronic Structure methods*, 2nd ed.; Gaussian Inc.: Pittsburgh, PA, 1996; pp 206, 215.

(28) The geometry of $[Re(CO)_6]^+$ was optimized using the procedure described in the Computational section and is available in Supporting Information Table S1.

(29) Juris, A.; Campagna, S.; Bidd, I.; Lehn, J. M.; Zeissel, R. *Inorg. Chem*. **1988**, *27*, 4007.

(30) Della Ciana, L.; Dressick, W. J.; Sandrini, D.; Maestri, M.; Ciano, M. *Inorg. Chem*. **1990**, *29*, 2792.

(31) Wan, J.; Ren, Y.; Wu, J.; Xu, X. *J. Phys. Chem. A* **2004**, *43*, 9453−9460.

(32) The triplet excited-states were calculated based on the lowest-lying triplet state geometry because according to Kasha's rule this state would be the emitting state. Thus the triplet excited-states were determined based on the most stable triplet geometry. Triplet excited-states calculated based on the singlet ground-state geometry are multiplicity forbidden ($f = 0$).

(33) Drago, R. S. *Physical Methods for Chemist*, 2nd ed.; Saunders College Publishing: Orlando, Florida, 1992; p 123.

(34) Lynch, B. J.; Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2003**, *107*, 1384−1388. (b) Aquilante, F.; Barone, V.; Roos, B. *J. Chem. Phys.* **2003**, *119*, 12323−12333.

(35) The correction factor is within the range reported for correction of B3LYP computed vibrational frequencies for positive transition metal complex ions as outlined in the following: Zhou, M.; Andrews, L.; Bauschlicher, C. W., Jr. *Chem. Rev.* **2001**, *101*, 1931.

CT049956G

JCTC Journal of Chemical Theory and Computation

Modeling the Morphology and Phase Stability of TiO_2 Nanocrystals in Water

A. S. Barnard,*,† P. Zapol,†,‡ and L. A. Curtiss‡

Center for Nanoscale Materials and Materials Science and Chemistry Divisions, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, Illinois 60439

Received August 6, 2004

Abstract: The potential of titanium dioxide nanoparticles for advanced photochemical applications has prompted a number of studies to analyze the size, phase, and morphology dependent properties. Previously we have used a thermodynamic model of nanoparticles as a function of size and shape to predict the phase stability of titanium dioxide nanoparticles, with particular attention given to the crossover of stability between the anatase and rutile phases. This work has now been extended to titanium dioxide nanoparticles in water, to examine the effects of various adsorption configurations on the equilibrium shape and the phase transition. Density functional calculations have been used to accurately determine surface energies and surface tension of low index hydrated stoichiometric surfaces of anatase and rutile, which are presented along with a brief outline of the surface structure. We have shown that morphology of TiO_2 nanocrystals is affected by the presence of water, resulting in variations in the size of the (001) and (00$\bar{1}$) truncation facets in anatase, and a reduction in the aspect ratio of rutile nanocrystals. Our results also highlight that the consideration of hydrated nanocrystal surfaces is necessary to accurately predict the correct size dependence of the anatase to rutile phase transition.

I. Introduction

Titanium dioxide (TiO_2) is an important accessory oxide mineral[1] used widely in science and technology.[2−5] Nanoparticles of this material are proving to be highly suitable for advanced photochemical applications,[6] especially interfacing with organic molecules[7] including DNA.[8] While the nanoscale dimension is instrumental in facilitating many new technologies, the size, phase, and morphology have been found to be critical parameters in determining their suitability for particular applications.[9−13]

Although macroscopically the rutile phase is more thermodynamically stable than the anatase phase[14] (at ambient pressures and temperatures), anatase has been found to be a majority product of industrial sol-gel, and aerosol syntheses of TiO_2,[3] and is common in nanoscale natural and synthetic samples.[5,15,16] Gribb and Banfield,[3] and later Zhang and Banfield,[4] found that the synthesis of nanocrystalline TiO_2 consistently resulted in anatase nanoparticles, which transformed to rutile upon reaching a particular size (<14 nm). The transformation from anatase to rutile has been observed under different experimental conditions depending upon parameters such as temperature[17,18] and size. This implies that the transition energetics are closely coupled with the particle size[19] and that anatase is in fact the most stable polymorph at the nanoscale.

It has been proposed by a number of authors that the anatase to rutile phase transformation is not only dependent on grain size but also on impurities,[20−22] reaction atmosphere,[23,29] and synthesis conditions.[24−28] Yang et al.[27] showed that synthesis conditions (chemicals/peptizing agents) affect the crystallinity and phase transition temperature. In this vein, Zaban et al.[28] noted that the surface structure of TiO_2 is affected by the preparation conditions; and Ahonen et al.[29] observed that anatase synthesized in air transformed to rutile at 973 K, but anatase synthesized in nitrogen transformed to rutile at 1173 K. It has also been shown that nanocrystalline TiO_2 may be phase selected by careful control of

* Corresponding author e-mail: amanda.barnard@anl.gov.
† Center for Nanoscale Materials, Argonne National Laboratory.
‡ Materials Science and Chemistry Divisions, Argonne National Laboratory.

10.1021/ct0499635 CCC: $30.25
Published on Web 12/15/2004

the particle size as well as other experimental conditions.[30−32] In addition to the control of the size of anatase nanocrystals, the shape may also be manipulated,[31,33,34] which may in turn enhance the adsorption properties of the nanocrystals by increasing the effective area of preferred surface facets.

We suggest that nanocrystal morphology is also a factor affecting the phase stability[40] as, contrary to numerous misconceptions, anatase nanoparticles are not necessarily spherical.[35] For example, high-resolution transmission electron microscopy (HRTEM) micrographs of Penn and Banfield[36] clearly show that the tetragonal bipyramidal morphology persists down to 3−5 nm in diameter. Similarly, the morphology of larger (∼20 nm) rutile nanocrystals can be discerned from the TEM images of Aruna et al.[37]

As part of an ongoing study, we have previously used a thermodynamic model[38] based on the free energy of (arbitrary) nanocrystals as a function of size and shape to determine the minimum energy morphology of anatase and rutile at the nanoscale and to examine the phase stability of faceted TiO_2 nanocrystals,[39] as a function of surface hydrogenation.[40] The model predicted that a bifrustum Wulff construction[42] for anatase nanocrystals, and a bitetragonal bipyramidal Wulff construction of rutile nanocrystals, which became more squat as the coverage of hydrogen was increased.

Further, the results of our study predict that (at low temperatures) the anatase to rutile phase transition size also depends on the surface hydrogenation. For clean surfaces (vacuum conditions) this phase transition is predicted to occur at an average diameter of approximately 9.3 nm for anatase nanocrystals.[39] This transition point slightly decreases to 8.9 nm when the surface bridging oxygens are H-terminated but increases significantly to 23.1 nm when both the bridging oxygens and undercoordinated titanium atoms of the surface trilayer are H-terminated.[40] As an extension of our previous work, the present study uses the same model to examine the relative phase stability of nanoscale anatase and rutile in water. This is considered to be of vital importance, since TiO_2 nanoparticles are often produced and stored in solution.

The model takes as input the geometry of the nanocrystal morphology, the surface free energy (γ), and the surface tension (σ). The surface tension has a small but important effect on the calculation of the anatase to rutile phase transition.[43] We have generated a consistent set of surface energy and surface tension values for water terminated stoichiometric low index (1×1) surfaces of both anatase and rutile, using ab initio methods. Both molecular and dissociative adsorption configurations have been considered, and the relative stability of each configuration has been compared with the results of other authors where possible.

II. Methodology

The surface structure and energetics of the low index surfaces of rutile and anatase were investigated by comparing highly accurate first principles calculations of the total energy of two-dimensional slabs with the corresponding three-dimensional bulk lattice structures. The slabs were generated by the addition of a 10 Å vacuum layer in the crystallographic plane of interest and then terminating the 'cleaved' surfaces with either molecular or dissociative H_2O, in a complete monolayer ($\theta = 1$). Both the bulk and surface slabs were relaxed prior to calculation of the total energies.

The first principles calculations have been carried out using Density Functional Theory (DFT) within the Generalized-Gradient Approximation (GGA), with the exchange-correlation functional of Perdew and Wang (PW91).[44] This has been implemented via the Vienna Ab initio Simulation Package (VASP),[45,46] which spans reciprocal space with a plane-wave basis up to a kinetic energy cutoff of 270 eV. We have used the Linear Tetrahedron Method (LTM) with a $4 \times 4 \times 4$ Monkhorst-Pack k-point mesh, for both the initial relaxations of the TiO_2 slabs, and the final calculation of surface energies and surface tensions. Although this choice of k-mesh results in some superfluous k-points in the nonperiodic direction of the surface slabs, it was found that the inclusion of these k-points is more consistent with the LTM.

The electronic relaxation technique used here is an efficient matrix-diagonalization routine based on a sequential band-by-band residual minimization method of single-electron energies,[49,50] with direct inversion in the iterative subspace, whereas the ionic relaxation involves minimization of the Hellmann-Feynman forces. During the relaxations we have used ultrasoft, gradient-corrected Vanderbilt-type pseudopotentials (US−PP)[47,48] and real-space projected wave functions (to decrease the computational cost) and have relaxed to a convergence of 10^{-4} eV. The following (final) energy calculations were then performed using the Projected Augmented Wave (PAW) potentials,[51] with a basis set increased to a cutoff of 350 eV and reciprocal-space projected wave function (to improve accuracy), also to a convergence of 10^{-4} eV. PAW potentials are generally considered to be more accurate than the ultrasoft pseudopotentials,[52] since the radial cutoffs (core radii) are smaller than the radii used for the US pseudopotentials, and the fact that the PAW potentials reconstruct the exact valence wave function with all nodes in the core region (all electron).

III. TiO_2 Surface

The surface science of the titanium dioxide polymorphs has been investigated,[53] including numerous studies investigating the surface of H_2O-terminated stoichiometric anatase[54,56−58] and rutile.[57−60] However, as a consistent set of surface energies for both phases (calculated using the same theoretical technique and convergence criteria) is required to provide a suitable input for the phase stability model, we have undertaken our own calculations on the surfaces of interest. Thus, the surface free energy and surface tension have been determined for the structurally relaxed low index anatase and rutile surfaces.

The relaxed surface structures are also presented, for the purposes of comparison with other studies (where available) for the lowest energy adsorption geometries. This will be followed by the presentation of the calculated adsorption energies, surface energies, and surface tensions.

A. Anatase Surfaces. In the following description each atom in the outermost TiO_2 trilayer has been labeled (in Figure 1) according to species, with the subscript denoting the *atomic* layer with respect to the vacuum (layer (1) is the

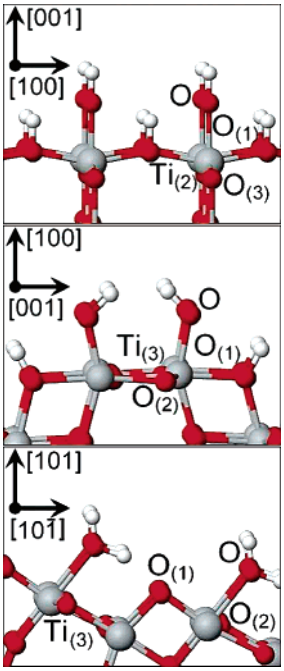


Figure 1. The relaxed anatase hydrated (001), (100), and (101) surfaces. The (001) and (100) surface exhibit dissociative adsorption, whereas the (101) surface exhibits molecular adsorption. The atoms occupying the atomic layers (denoted by the subscript) of the upper most trilayer labeled according to Table 1.

outermost layer of the surface TiO_2 trilayer). The oxygen associated with the adsorbate is denoted simply as O.

The hydrated surface trilayers of the anatase (001) surface are shown at the top of Figure 1. The clean surface contains 5-fold coordinated ($Ti_{(2)}$) atoms and 2-fold ($O_{(1)}$) and 3-fold coordinated ($O_{(3)}$) oxygens. Dissociative adsorption was found to be energetically favorable on this surface, with the H and OH terminations oriented perpendicular to the surface after relaxation. Compared with the clean surface,[39] the outward displacement of the oxygen atoms in the upper trilayer was found to decrease in the presence of water. This is most significant in the case of the $O_{(1)}$ bridging oxygen. In contrast the outward displacement of the $Ti_{(2)}$ was found to increase slightly in the presence of water. The final $Ti_{(2)}-OH$ bond length was just under the equatorial Ti−O bond length of anatase, and the $O_{(1)}-H$ bond length was found to be slightly longer than the O−H bond length in water (calculated to be 0.98 Å).

Calculated relaxations of these atoms are given in Table 1 along with the $O_{(1)}-H$ and $Ti_{(2)}-OH$ bond lengths. The displacement of selected atoms on the anatase (001) surface was examined by Bredow and Jug[58] using the semiempirical SINDO1 method and model clusters. The authors found that the bridging $O_{(1)}$ atoms relaxed outward by 0.12 Å, the $Ti_{(2)}$ atom relaxed outward by 0.16 Å, and the $Ti_{(2)}-OH$ bond length was 1.85 Å, when a $4 \times 4 \times 3$ cluster was used. Although the structure of the (001) surface was examined by Vittadini et al.,[54] they did not treat a complete monolayer of dissociatively adsorbed water. However, for a coverage of $\theta = 0.5$ (where the $O_{(1)}-H$ terminations were removed) Vittadini et al.[54] reported $Ti_{(2)}-OH$ bonds lengths between 1.74 and 1.93 Å.

Table 1: Comparison of Displacements (in Å), Normal to the Surface, of Atoms in the Uppermost Trilayer of the Hydrated and Clean[39] Anatase Surfaces, along with the $Ti-OH_2$ and O−H Bond Lengths

surface	label	clean[39]	hydrated
(001)	$O_{(1)}$	0.20	0.03
dissociative	$Ti_{(2)}$	0.04	0.09
	$O_{(3)}$	0.05	0.02
	Ti−OH		1.93
	$O_{(1)}-H$		1.01
(100)	$O_{(1)}$	0.18	−0.02
dissociative	$O_{(2)}$	0.04	0.01
	$Ti_{(3)}$	−0.16	0.10
	Ti−OH		1.87
	$O_{(1)}-H$		0.99
(101)	$O_{(1)}$	0.06	−0.01
molecular	$O_{(2)}$	0.28	0.10
	$Ti_{(3)}$	−0.12	−0.05
	$Ti-OH_2$		2.28

Like the (001) surface, the clean (100) surface contains 5-fold coordinated $Ti_{(3)}$ atoms and 2-fold $O_{(1)}$ and 3-fold coordinated $O_{(2)}$ atoms. When covered with a monolayer of dissociated water, the surface (shown in the center of Figure 1) was found to undergo an outward relaxation of the $O_{(2)}$ and $Ti_{(3)}$ atoms and a small inward relaxation of the $O_{(1)}$ atoms. The $Ti_{(3)}-OH$ bond length was found to be 1.87 Å and the $O_{(1)}-H$ bond length was 0.99 Å (see Table 1). The most interesting aspect of this surface was the bending of the O−H bonds of the OH groups toward each other, with a H−H distance of 2.6 Å.

The clean (101) surface contains 5-fold $Ti_{(3)}$ and 2-fold $O_{(1)}$ and 3-fold $O_{(2)}$ coordinated atoms within the first trilayer, with a characteristic saw-tooth profile perpendicular to the (010) direction. In this case molecular adsorption was found to be energetically preferred (see section 3.3) as indicated in the lower image of Figure 1. Following relaxation, both $O_{(1)}$ and $Ti_{(3)}$ atoms were found to contract inward, and the $O_{(2)}$ was found to relax outward (see Table 1). The $Ti_{(3)}-OH_2$ bond length of 2.28 Å matches the length of 2.28 Å reported by Vittadini et al.[54] for this bond.

The anatase (110) surface was found to be structurally unstable in the presence of water, both molecular and dissociated. The surface was found to deteriorate into a disordered structure. This structural change involved the breaking of symmetry as the 4-fold coordinated surface Ti atoms shifted from their lattice positions, and the desorption of OH and H_2O groups, sometimes involving $O_{(1)}$ atoms rather than the O atoms of the water molecules. Hence, the resulting surface could no longer be characterized as anatase, rendering it unsuitable for use in the phase stability model presented in the next section.

B. Rutile Surfaces. The (100), (110), and (011) surfaces are considered by applying the labeling convention used in the previous section for the anatase surfaces and are listed in Table 2. The corrugated (100) surface contains 5-fold coordinated $Ti_{(2)}$ atoms, with chains of 2-fold coordinated $O_{(1)}$ bridging atoms in the upper most atomic layer. This surface was found to prefer dissociative adsorption (as shown

Table 2: Comparison of Displacements (in Å), Normal to the Surface, of Atoms in the Uppermost Trilayer of the Hydrated and Clean[39] Rutile Surfaces, along with the Ti−OH_2 and O−H Bond Lengths

surface	label	clean[39]	hydrated
(100)	$O_{(1)}$	0.16	0.13
dissociative	$Ti_{(2)}$	0.06	0.02
	$O_{(3)}$	0.15	0.16
	Ti−OH		1.90
	$O_{(1)}$−H		0.99−1.01
	H-bonds (O···H)		2.01−2.03
(011)	$O_{(1)}$	0.09	0.17
dissociative	$Ti_{(2)}$	0.05	0.19
	$O_{(3)}$	0.14	0.25
	Ti−OH		2.14−2.16
	$O_{(1)}$−H		1.01−1.08
	H-bonds (O···H)		1.44−1.47
(110)	$O_{(1)}$	−0.19	0.08
molecular	$Ti_{(2)}$	0.40	0.35, −0.36
	$O_{(3)}$	0.24	−0.05
	Ti−OH_2		2.30
	H-bonds (O ···H)		1.66−1.67

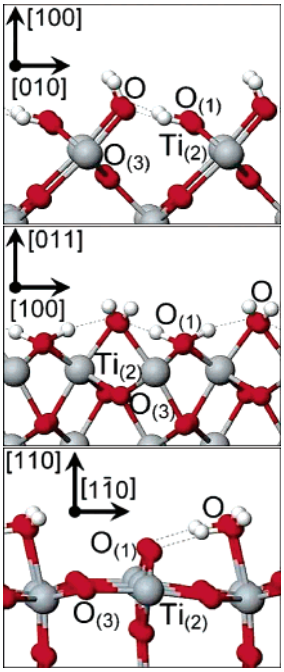


Figure 2. The relaxed rutile hydrated (100), (110), and (011) surfaces. The (100) and (011) surface exhibit dissociative adsorption, whereas the (110) surface exhibits molecular adsorption. The atoms occupying the atomic layers (denoted by the subscript) of the upper most trilayer labeled according to Table 2.

at the top of Figure 2), which is generally considered to be the case in other experimental and theoretical studies.[53] In general, the (outward) relaxation of the hydrated surface was found to be similar to the clean surface.[39] The $Ti_{(2)}$−OH bond length of 1.90 Å was found to be just under the equatorial Ti−O bond length of bulk rutile, and the $O_{(1)}$−H bond length was found to be just over the O−H bond length in water.

The hydrogen atoms terminating the bridging $O_{(1)}$ atoms were also found to form weak hydrogen bonds with the O atoms of the OH groups, with an O···H length of 2.01−2.03 Å. The issue of how well the PW91 exchange-correlation functional reproduces hydrogen bonding was briefly investigated by examining the relaxed geometry and binding energy of a water dimer and comparing with results obtained using second-order Møller−Plesset Perturbation theory (MP2) and the 6-311+G(3df,2p) basis set.[55] The PW91−PAW (MP2) optimized structure was found to have an O−O distance of 2.89 Å (2.90 Å) and an O···H hydrogen bond length of 1.90 Å (1.94 Å), with a binding energy of 0.177 eV (0.234 eV). The good agreement between the PW91−PAW and MP2 results indicates that the method used here in the calculation of surface properties is capable of describing H-bonding reasonably well.

The (011) surface (center of Figure 2) has a ridged-like structure and was also found to prefer dissociative adsorption. The upper trilayer was found to relax outward approximately 0.1 Å more than the clean surface. The $Ti_{(2)}$−OH_2 bonds were longer than the rutile apical and equatorial Ti−O bond lengths. Similarly, the $O_{(1)}$−H bond lengths were considerably longer than the O−H bond in water. Like the (100) surface, hydrogen bonds were observed between the hydrogen atoms connected to the bridging oxygens and the oxygen atoms of the OH groups but with a considerably reduced length of 1.44−1.47 Å.

Finally, the relaxed hydrated (110) surface (see lower image of Figure 2), containing inequivalent Ti atoms lying in a centered rectangular arrangement, was found to prefer molecular adsorption. There is some disagreement in the literature as to whether water adsorbs on this surface molecularly or dissociatively, but in most cases experimental investigations indicate molecular adsorption.[53] Other theoretical studies have concluded that molecular adsorption is most probable,[61,62] and it has also been suggested that molecular and dissociative water may coexist on the rutile (110) surface.[59]

Even though our results indicate molecular adsorption, the structure of the surface is somewhat distorted by the presence of water, as shown in the lower image of Figure 2. On the clean surface, the $Ti_{(2)}$ atoms denoted as connected to the bridging oxygen (center of the image) are 6-fold coordinated, while the exposed $Ti_{(2)}$ atoms at the sides of the image are 5-fold coordinated. These $Ti_{(2)}$ atoms become inequivalent (see Table 2) on the hydrated surface, with the $Ti_{(2)}$ atoms bound to the bridging $O_{(1)}$ relaxing outward, and the $Ti_{(2)}$ atoms bound to the water molecules relaxing inward.

This layer distortion has been observed before[60,62] and is in part due to the accommodation of H-bonding between the bridging $O_{(1)}$ atoms and hydrogens in the water molecules. The length of these H-bonds of 1.66−1.67 Å is in good agreement with the result 1.6−1.8 Å calculated by Langel[62] using Car Parinello Molecular Dynamics (CPMD), and the value of 1.61 Å is calculated by Ferris and Wang using the self-consistent field (SCF) method and 3-21G basis set but is considerably less that the 2.22 Å calculated by Menetrey et al.[60] using PW91−USPP.

Like the anatase (110) surface, the rutile (001) surface was found to be unstable (with respect to a disordered surface structure) in the presence of water, exhibiting bond breaking

and symmetry breaking upon relaxation. This instability has (like the anatase (110) surface) been attributed to the 4-fold coordinated Ti atoms on the surface.

C. Surface Energetics. Although both molecular and dissociative adsorption geometries were examined on the surfaces described above, the previous section only outlines the details of the energetically preferred configuration. The distinction as to which geometry constituted the energetically preferred configuration was made by comparing the adsorption energies using the expression

$$E_{ad} = \frac{1}{N_{H_2O}}(E^{surface}_{H_2O} - E^{surface}_{clean} + N_{H_2O}E_{H_2O}) \quad (1)$$

where N_{H_2O} is the total number of H_2O units on the surface, $E^{surface}_{clean}$ is the total energy of the relaxed surface without water, E_{H_2O} is the total energy of a free water molecule, and $E^{surface}_{H_2O}$ is the total energy of the water covered surface. The adsorption energies for all surfaces are given in Table 3, along with the results of other authors for comparison.

The value of the surface free energy γ was calculated from the total energy of the bulk (E^{bulk}_N) and surface ($E^{surface}_N$) slabs using the expressions

$$\gamma = \frac{G}{A} \quad (2)$$

and

$$G = \frac{1}{2}(E^{surface}_N - E^{bulk}_N - N_{H_2O}\mu_{H_2O}) \quad (3)$$

where G is the free energy of the slab, A is the area of the surface, and N is the number of TiO_2 units in the (stoichiometric) cell. To account for the surface hydration, μ_{H_2O} is the chemical potential of water, calculated using[63]

$$\mu_{H_2O} = E_{H_2O} + \frac{h\nu}{2} + k_B T\left[\ln\left(\frac{PV}{k_B T}\right)\right] \quad (4)$$

where k_B is Boltzmann's constant, T, P, and ν are the temperature, pressure, and vibrational frequencies of water in the reservoir, and V is the quantum volume

$$V = \left(\frac{h^2}{2\pi m k_B T}\right)^{3/2} \quad (5)$$

Experimental values for ν were used,[64] and the chemical potential was calculated at ambient temperature and pressure (298.15 K and 101.33 kPa).

The value of the surface tension σ was obtained using the expression

$$\sigma = \frac{\partial G}{\partial A} \approx \frac{\Delta G}{\Delta A} \quad (6)$$

By applying a two-dimensional uniform dilation in the plane of the surface (including optimization of all internal parameters) and calculating the free energy G as shown in eq 3 for each area, the change in free energy (ΔG) was found for a set of area dilations (ΔA). After plotting these results an estimate of the surface tension was obtained from the slope. The results of these calculations are contained within Tables 4 and 5 for anatase and rutile, respectively.

Table 3: Comparison of the Adsorption Energy E_{ad} (eV) as Indicated in Eq 1, Compared with the Hartree−Fock (HF) Results of Fahmi and Minot,[57] the DFT GGA Results of Vittadini et al.,[54] the DFT GGA Results of Lindan et al.,[59] and the DFT GGA Results of Menetrey et al.[60]

	this study		reference 57		reference 54	
anatase	dissoc.	molec.	dissoc.	molec.	dissoc.	molec.
(001)	−0.45	0.23	−1.248	−0.682		
(100)	−0.29	−0.24				
(101)	−0.48	−0.56			−0.44	−0.72

	this study		reference 59		reference 60	
rutile	dissoc.	molec.	dissoc.	molec.	dissoc.	molec.
(100)	−0.57	unstable[a]				
(011)	−0.98	−0.43				
(110)	−0.27	−0.82	−0.91	−0.99	−1.14	−1.03

[a] Proton transfer was observed during relaxation, resulting in dissociative geometry, indicating that this configuration was unstable with respect to $H^+ + OH^-$ dissociation.

Table 4: Comparison of Surface Free Energy γ and Surface Tension σ (in J/m^2), for the Clean[39] Hydrated Low Index Surfaces of Anatase

		clean[39]		hydrated	
surface	adsorption	γ	σ	γ	σ
(001)	dissociative	0.51	2.07	1.55	−0.37
(100)	dissociative	0.39	0.60	1.13	−0.59
(101)	molecular	0.35	0.51	1.03	0.45

Table 5: Comparison of Surface Free Energy γ and Surface Tension σ (in J/m^2), for the Clean[39] and Hydrated Low Index Surfaces of Rutile

		clean[39]		hydrated	
surface	adsorption	γ	σ	γ	σ
(100)	dissociative	0.60	0.95	1.57	0.61
(011)	dissociative	0.95	1.50	1.79	1.36
(110)	molecular	0.47	1.25	1.08	0.92

The relative stability of the low index anatase surfaces may be discerned by comparing the values of γ listed in Table 4. The thermodynamic sequence (101) < (100) < (001) is the same for the clean and hydrated surfaces; however, the surface tension σ varies considerably. On the hydrated (001) and (100) surfaces σ is negative, indicating a tendency for this surface to expand (rather than contract). If present on the surface of a nanocrystal it would produce a tensile dilation in the direction of the surface normal, rather than a contraction. In general the surface tension of the anatase surface was found to decrease in the presence of water.

Previously we have obtained the order of (110) < (100) < (011) for the clean rutile surfaces,[39] and once again, this is the same for the hydrated surfaces. The σ of the rutile surfaces (like anatase) was found to decrease when termi-

nated by water, but in this case all tensions remained compressive.

IV. Phase Stability of Faceted TiO_2 Nanocrystals in Water

Previously we have shown that for a given nanoparticle of material *x*, the free energy may be expressed (to first order) as a sum of contributions from the particle bulk and surfaces,[38−41] such that

$$G_x^o = G_x^{\text{bulk}} + G_x^{\text{surface}} \tag{7}$$

The free energy of formation of a nanocrystal G_x^o is defined in terms of the surface energy γ_{xi} for each surface *i*, weighted by the factors f_i, such that $\sum_i f_i = 1$.

Hence,

$$G_x^o = \Delta_f G_x^o + \frac{M}{\rho_x}(1 - e)[q\sum_i f_i \gamma_{xi}] \tag{8}$$

where $\Delta_f G_x^o$ is the standard free energy of formation of the bulk (macroscopic) material, *M* is the molar mass, ρ_x is the density, and *e* is the volume dilation induced by the surface tension (which cannot be ignored at the nanoscale). In general, the surface-to-volume ratio *q* and the weighting factors f_i must be calculated explicitly for each shape and the facet therein. In this model the size dependence is introduced not only by the surface-to-volume ratio *q* but also by the reduction of *e* as the crystal grows larger. The shape dependence is also introduced by *q* as well as the weighted sums of the surface energies and the surface tensions, corresponding to the surfaces present in the particular morphology of interest.

In general, the volume dilation due to the surface tension may be approximated using the Laplace-Young equation[38] for the effective pressure

$$P_{\text{eff}} = \frac{2\sigma_x}{R} \tag{9}$$

where *R* is the average radius of the particle, so that (with the compressibility $\beta = 1/B_0$)

$$e = \frac{2\beta\sigma_x}{R} \tag{10}$$

The surface tension is approximated by summing over the (weighted) surface tension of the crystallographic surfaces present on the nanocrystal

$$\sigma_x = \sum_i f_i \sigma_{xi} \tag{11}$$

where σ_{xi} is defined in eq 6 and listed in Tables 3 and 4 for $x =$ anatase ($\mathscr{A}$) and rutile ($\mathscr{R}$), respectively. The values of β for anatase and rutile were previously calculated[39] by fitting energy versus volume curves to the Vinet equation of state.[65] The resulting values for anatase of 190 GPa and rutile of 218 GPa were in good agreement with the experimental values of 179 GPa measured by Arlt et al.[66] and of 211 GPa measured by Gerward and Olsen,[67] respectively.

Previously it has been determined that the Laplace-Young description of the pressure is suitable in the case of faceted nanocrystals and that the edge and corner effects are limited over a diameter of approximately 2 nm.[38−40] Therefore, the surface energies and surface tensions for each surface facet *i*, along with the molar mass of TiO_2, the density of anatase and rutile ($\rho_{\mathscr{A}} = 3.893$ g/cm^3 and $\rho_{\mathscr{R}} = 4.249$ g/cm^3), and the standard free energies of formation[68] ($\Delta_f G^o_{\mathscr{A}}$ and $\Delta_f G^o_{\mathscr{R}}$) are all that is required to compare the phase stability of faceted TiO_2 nanocrystals over this size.

It is important to note that for the reasons mentioned above (and outlined in refs 38 and 39) this model is not applicable for small nanoparticles but rather to the 2 nm to submicron range. Therefore, for small nanoparticles less than 2 nm it is still preferable to examine each morphology explicitly by undertaking suitable calculations of isolated nanoparticles (e.g.: DFT or Tight Binding).

At large sizes, in the range of 75−100 nm, the free energy of the surfaces is less than 10^{-4} J/mol, making the energetics of the surface less significant and other bulk effects more significant. For example, in this size regime the macroscopic (bulk) strain is as important as the surface strain, and the entopic effects of the bulk will be as important as the surface entropy. The present model does include the bulk strain but does not explicitly include other macroscopic thermodynamic arguments. For this reason, it is best applied to particles in the range 2−100 nm.

All of the calculations in the present study have been performed at $T = 0$, so that G_x^o is equivalent to the enthalpy of formation. It has been shown by Zhang and Banfield[4] that the change in the surface free energies with temperature is of the order of 10^{-4} J/m^2, so it has been assumed here that variations in the equilibrium morphology of anatase and rutile nanocrystals due to temperature effects will be negligible.

A. Predicting Nanomorphology. The standard method for determining the equilibrium morphology of a material is to generate the Wulff construction[42] using the surface energies. However, as the Wulff construction does not take into account the effects of surface tension it is possible that the morphologies of nanocrystals may deviate from this shape. Using the model described above, we have investigated this possibility by optimizing the nanoparticle shape, as a function of size.

Beginning with the Wulff constructions as the initial case, the morphology of anatase and rutile nanocrystals were defined in terms of two independent length parameters **A** and **B**, as shown in Figures 3 and 4. In the case of anatase, this shape may vary depending upon the energetic relationship between {101} and {001} forms. In either case, the side of this bipyramidal form is denoted **A**. In nature, anatase crystals often exhibit a truncated bipyramid, or bifrustum, with square facets in the (001) and (00$\bar{1}$) planes (displayed as the interior solid in Figure 3). The side of this 'truncation' facet is denoted **B**. The degree of truncation may therefore be described by the size of **B** with respect to **A** (where $0 \leq \mathbf{B} \leq \mathbf{A}$).

The Wulff construction for rutile predicts a tetragonal prism bounded by {110} surfaces and terminated by a pair of tetragonal pyramids bounded by {011} surfaces (see

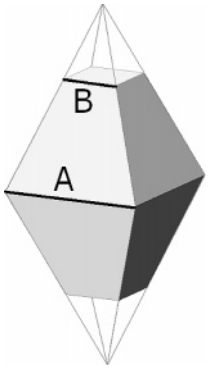


Figure 3. An example of an anatase tetragonal {101} bipyramid is shown in outline, and an example of the bifrustum formed by addition of (001) and (00$\bar{1}$) truncation facets is shown as the interior solid. The side lengths labeled **A** and **B** are used to define the degree of truncation.

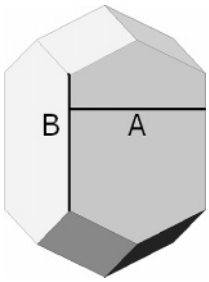


Figure 4. An example of the rutile tetragonal {110} prism terminated by a pair of tetragonal {011} pyramids. The side lengths labeled **A** and **B** are used to define the aspect ratio of the crystal.

Figure 4). The side of the tetragonal {110} prism is denoted **A**. In nature, the shape of rutile may vary from long acicular crystals to a short blocky habit. Therefore, the length of the tetragonal {110} prism is denoted **B**. The shape of the rutile crystal (the aspect ratio) may be described by the length of **B** with respect to **A** (where $0 \leq \mathbf{B} < \infty$).

By defining all the geometric parameters such as the volume and surface area of the various facets in terms of the ratio **B**/**A**, the energy was minimized with respect to this new variable. The numerical minimization was performed using a conjugate gradient scheme, in an attempt to find a value of **B**/**A** that produces a shape that is lower in energy than the Wulff construction as a function of size. This procedure was performed for 'clean' and hydrated anatase and rutile.

In the case of nanoscale anatase, shapes lower in energy than the Wulff construction were identified, as shown in Figure 5. Figure 5(a) shows a plot of the optimized **B**/**A** for sizes $\mathbf{A} = 2$ to 300 nm, and Figure 5(b) shows the length of **B** for **A** in the same range. At very small sizes (in a vacuum and in water), the model predicts that small truncation facets of $\mathbf{B}/\mathbf{A} = 0.32$ and 0.30, even though the Wulff constructions (with same energies) predict values of $\mathbf{B}/\mathbf{A} = 0.47$ and 0.45. The size of these facets with respect to the overall size of the nanocrystal fluctuates somewhat, especially in the case of the clean surfaces for nanocrystals less than $\mathbf{A} =$ 100 nm. It is also important to note that as the model predicts that the size of the truncation facet will be slightly different

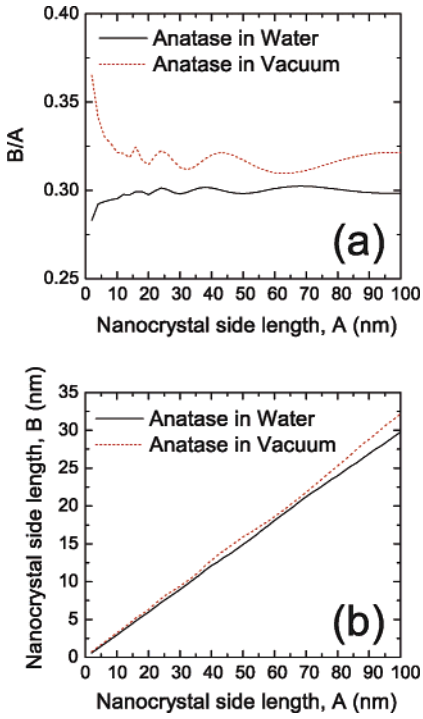


Figure 5. (a) Plot of the optimized ratio **B**/**A** and (b) length of **B** for each **A** for anatase nanocrystals in a vacuum and in water, with a side length $\mathbf{A} = 2$ to 100 nm. The facet edges **A** and **B** are defined in Figure 3.

for clean and hydrated nanocrystals. In both cases, as the anatase crystals increase in size and the effects of surface tension diminish, the shape converges to that of the Wulff construction which remains the macroscopically lowest energy shape.

In the case of rutile, no lower energy shape was found, indicating that the Wulff construction dominates even at the nanoscale. It is important to note, however, that the optimization of rutile shape is more complicated than anatase since **B** has no upper bound. The length of **B** may therefore increase rapidly with respect to **A**, producing long needlelike crystals with a higher energy per TiO_2 unit than the Wulff construction. These nano-rod shaped crystals represent local minima in the shape-energy surface and are sampled due to the fact that the shape optimization described above acts only on the *open* form of the tetragonal {110} prism. During these optimizations the energy associated with the terminating tetragonal {011} pyramids remains constant.

The final shapes predicted by the model outlined above for anatase and rutile nanocrystals in a vacuum and in water (with a side length $\mathbf{A} \leq 100$ nm) are shown in Figure 6. These results predict that water reduces the (001) and (00$\bar{1}$) truncation facets at the apexes of the anatase nanocrystals and reduced the aspect ratio of the rutile nanocrystals.

B. Anatase to Rutile Phase Transition in Water. Finally, for each of the clean and hydrated nanocrystals shown in Figure 6, the value of G^{o}_{ℓ} and $G^{o}_{\wp}$ were calculated and plotted as a function of the number of TiO_2 units, by using the appropriate surface energies from Tables 4 and 5. In each case, the point of intersection of the free energies of anatase and rutile identifies the phase transition. These plots are shown

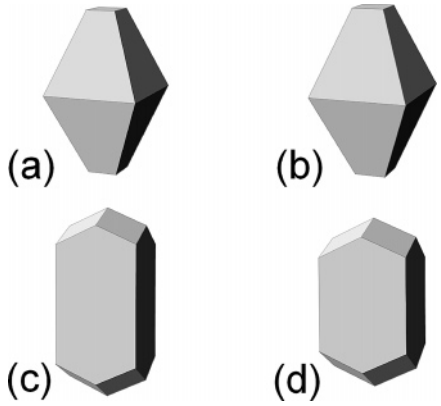


Figure 6. The morphology predicted in the present study for anatase (a) in a vacuum and (b) in water via optimization of the nanocrystal shape using the thermodynamic model described above and rutile (c) in a vacuum and (d) in water using the Wulff construction. These shapes correspond to nanocrystals with a side length (**A**) less than 100 nm.

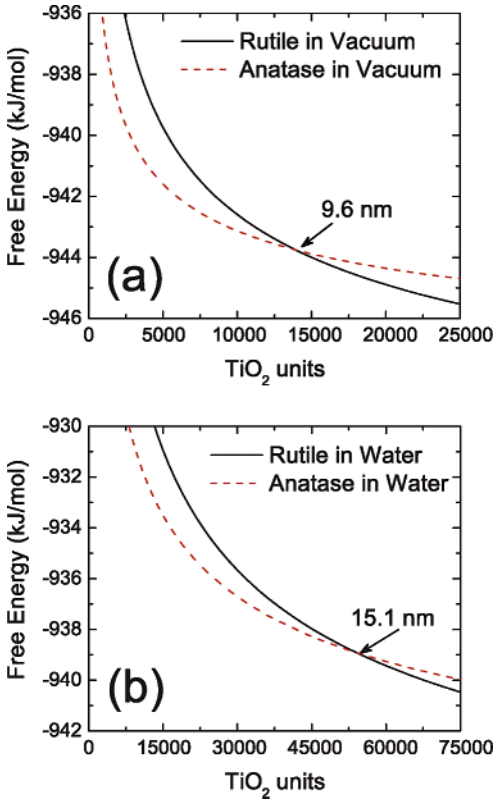


Figure 7. Free energy as a function of number of TiO_2 units for anatase and rutile (a) in a vacuum[39] and (b) in water, calculated using the shapes given in Figure 6 and the values of γ and σ from Tables 3 and 4. The intersection points indicate the phase transition.

in Figure 7 for nanoscale TiO_2 with (a) clean and (b) hydrated surfaces. The phase stability of TiO_2 in a vacuum has been reported before[39] and is included here for the purposes of comparison.

By comparing these results it is clear that the phase stability of nano-anatase is greater in water than in a vacuum. The intersection point for clean nanocrystals occurs at ~12 600 TiO_2 units, corresponding to an anatase nanocrystal with average diameter of ~9.6 nm. For nanocrystals in water the intersection point occurs at ~52 580 TiO_2 units, corresponding to an anatase nanocrystal with an average diameter of ~15.1 nm. Experimentally, the transition point for hydrothermal samples at ~650−800 K has been predicted to be at approximately 11.4−17.6 nm but has been found to decrease with temperature.[3]

Previously, a similar increase in the phase transition size of nano-anatase was predicted when the anatase and rutile surfaces were saturated with hydrogen.[40] At a full monolayer coverage of hydrogen the phase transition was predicted to be at ~196 900 TiO_2 units, corresponding to an anatase nanocrystal with an average diameter of ~23.1 nm. This is not the same as in water, indicating that it is not simply the termination of undercoordinated surface sites that affect the stability of nanoscale TiO_2, but also what these sites are terminated with. These striking results illustrate that the chemical environment plays an important part in phase stability and that the theoretical model used here is quite capable of describing these subtleties.

V. Conclusions

We have presented results of a thermodynamic model for the free energy of nanocrystals as a function of size and shape to illustrate the effects of water on the equilibrium shape and phase stability of anatase and rutile at the nanoscale. Using a complete set of values for the surface energy and surface tension of low index stoichiometric surfaces calculated using DFT GGA with the PAW potential method, we have shown that the morphology of TiO_2 nanocrystals is affected by the presence of water, resulting in variations in the size of the (001) and (00$\bar{1}$) truncation facets in anatase, and a reduction in the aspect ratio of rutile nanocrystals.

Our results for hydrated nanocrystals also predict (at low temperatures) an anatase to rutile phase transition size of ~15.1 nm that is in good agreement with experiment,[3] highlighting that the consideration of appropriate surface passivation of nanocrystal surfaces is necessary to accurately predict the correct size dependence of the anatase to rutile phase transition. Further work is currently underway to examine the effects of pH on the phase transition size and shape.

Acknowledgment. This work has been supported by the U.S. Department of Energy BES-Chemical Sciences, under Contract W-31-109-ENG-38. Computational resources for this project have been supplied by Argonne National Laboratory − Laboratory Computing Resource Center, Pacific Northwest National Laboratory Molecular Science Computing Facility and the U.S. Department of Energy National Energy Research Scientific Computing Center.

References

(1) Banfield, J. F.; Bischoff, B. L.; Anderson, M. A. *Chem. Geol.* **1993**, *110*, 211−231.

(2) Elder, S. H.; Cot, F. M.; Su, Y.; Heald, S. M.; Tyryshkin, A. M.; Bowman, M. K.; Gao, Y.; Joly, A. G.; Balmer, M. L.; Kolwaite, A. C.; Magrini, K. A.; Blake, D. M. *J. Am. Chem. Soc.* **2000**, *122*, 5138−5146.

(3) Gribb, A. A.; Banfield, J. F. *Am. Mineral.* **1997**, *82*, 717−728.

(4) Zhang, H.; Banfield, J. F. *J. Mater. Chem.* **1998**, *8*, 2073−2076.

(5) Navrotsky, A.; Kleppa, O. J. *J. Am. Ceram. Soc.* **1967**, *50*, 626.

(6) Zhang, W. F.; Zhang, M. S.; Yin, Z.; Chen, Q. *Appl. Phys. B* **2000**, *70*, 261−265.

(7) Rajh, T.; Chen, L. X.; Lucas, L.; Liu, T.; Thurnauer, M. C.; Tiede, D. M. *J. Phys. Chem. B* **2002**, *106*, 10543−10552.

(8) Paunesku, T.; Rajh, T.; Wiederrecht, G.; Maser, J.; Vogt, S.; Stojićević, N.; Protić, M.; Lai, B.; Oryhon, J.; Thurnauer, M.; Woloschak, G. *Nature Mater.* **2003**, *2*, 343−346.

(9) Chen, L. C.; Rajh, T.; Jäger, W.; Nedeljkovic, J. M.; Thurnauer, M. C. *J. Synchrotron. Rad.* **1999**, *6*, 445−447.

(10) Rajh, T.; Nedeljkovic, J. M.; Chen, L. C.; Poluektov, O.; Thurnauer, M. C. *J. Phys. Chem. B* **1999**, *103*, 3515−3519.

(11) Zhang, H.; Penn, R. L.; Hamers, R. J.; Banfield, J. F. *J. Phys. Chem. B* **1999**, *103*, 4656−4662.

(12) Jang, K. D.; Kim, S.-K.; Kim, S.-J. *J. Nanoparticle Res.* **2001**, *3*, 141−147.

(13) Bullen, H.; Garrett, S. *Nano Lett.* **2002**, *2*, 739−745.

(14) Muscat, J.; Swamy, V.; Harrison, N. M. *Phys. Rev. B* **2002**, *65*, 224112.

(15) Mitsuhashi, T.; Kleppa, O. J. *J. Am. Ceram. Soc.* **1979**, *62*, 356.

(16) Zhang, H. Z.; Banfield, J. F. *J. Phys. Chem. B* **2000**, *104*, 3481−3487.

(17) Gourma, P. I.; Mills, M. J. *J. Am. Ceram. Soc.* **2001**, *84*, 619−622.

(18) Li, Y.; White, T.; Lim, S. H. *Rev. Adv. Mater. Sci.* **2003**, *5*, 211.

(19) Navrotsky, A. *Thermochemistry of Nanomaterials in Nanoparticles the Environment, Reviews in Mineralogy Geochemistry*; Banfield, J. F., Navrotsky, A., Eds.; Mineralog. Soc. Am.: 2001; Vol. 44.

(20) Arroyo, R.; Córdoba, G.; Padilla, J.; Lara, V. H. *Mater. Lett.* **2002**, *54*, 397−402.

(21) Arbiol, J.; Cerdà, J.; Dezanneau, G.; Cirera, A.; Peiró, F.; Cornet, A.; Morante, J. R. *J. Appl. Phys.* **2002**, *92*, 853−861.

(22) Francisco, M. S. P.; Mastelaro, V. R. *Chem. Mater.* **2002**, *14*, 2514−2518.

(23) Kittaka, S.; Matsuno K.; Takahara, S. *J. Solid State. Chem.* **1997**, *132*, 447−450.

(24) Zhang, H. Z.; Banfield, J. F. *J. Mater. Res.* **2000**, *15*, 437−448.

(25) Yoshinaka, M.; Hirota, K.; Yamaguchi, O. *J. Am. Ceram. Soc.* **1997**, *80*, 2749−2753.

(26) Okada, K.; Yamamoto, N.; Kameshima, Y.; Yasumori, A. *J. Am. Ceram. Soc.* **2001**, *84*, 1591−1596.

(27) Yang, J.; Mei, S.; Ferreira, J. M. F. *J. Am. Ceram. Soc.* **2000**, *83*, 1361−1368.

(28) Zaban, A.; Aruna, S. T.; Tirosh, S.; Gregg, B. A.; Mastai, Y. *J. Phys. Chem. B* **2000**, *104*, 4130−4133.

(29) Ahonen, P. P.; Kauppinen, E. I.; Joubert, J. C.; Deschanvres, J. L.; Van Tendeloo, G. *J. Mater. Res.* **2000**, *14*, 3938−3948.

(30) Zhang, H.; Finnegan, M.; Banfield, J. F. *Nano Lett.* **2001**, *1*, 81−85.

(31) Sugiyama, M.; Okazaki H.; Koda, S. *Jpn. J. Appl. Phys.* **2002**, *41*, 4666−4674.

(32) Sugimoto, T.; Zhou, X.; Muramatsu, A. *J. Colloidal. Int. Sci.* **2003**, *259*, 43−52.

(33) Sugimoto, T.; Okada, K.; Itoh, H. *J. Colloidal. Int. Sci.* **2003**, *193*, 140−143.

(34) Sugimoto, T.; Zhou, X.; Muramatsu, A. *J. Colloidal. Int. Sci.* **2003**, *259*, 53−61.

(35) Gao, Y.; Elder, S. A. *Mater. Lett.* **2000**, *44*, 228−232.

(36) Penn, R. L.; Banfield, J. F. *Geochim. Cosmochim. Acta* **1999**, *63*, 1549−1557.

(37) Aruna, S. T.; Tirosh S.; Zuban, A. *J. Mater. Chem.* **2000**, *10*, 2388.

(38) Barnard, A. S.; Zapol, P. *J. Chem. Phys.* **2004**, *121*, 4276−4283.

(39) Barnard, A. S.; Zapol, P. *Phys. Rev. B* **2004**, *70*, 235403.

(40) Barnard, A. S.; Zapol, P. *J. Phys. Chem. B* **2004**, *49*, 18435.

(41) Reference deleted in press.

(42) Wulff, G. *Z. Kristallogr. Mineral* **1901** *34*, 449.

(43) Lu, H. M.; Zhang, W. X.; Jiang, Q. *Adv. Eng. Mater.* **2003**, *5*, 787−788.

(44) Perdew, J.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244.

(45) Kresse, G.; Hafner, J. *Phys. Rev. B* **1993**, *47*, RC558.

(46) Kresse, G.; Hafner, J. *Phys. Rev. B* **1996**, *54*, 11169.

(47) Kresse, G.; Furthmüller, J. *Comput. Mater. Sci.* **1996**, *6*, 15−50.

(48) Wood, D. M.; Zunger, A. *J. Phys. A* **1985**, *18*, 1343−1359.

(49) Vanderbilt, D. *Phys. Rev. B* **1990**, *41*, 7892.

(50) Kresse, G.; Hafner, J. *J. Phys.: Condens. Matter.* **1994**, *6*, 8245−8257.

(51) Blöchl, P. E. *Phys. Rev. B* **1994**, *50*, 17953.

(52) Kresse, G.; Joubert, D. *Phys. Rev. B* **1999**, *59*, 1758.

(53) Diebold, U. *Surf. Sci. Rep.* **2003**, *48*, 53−229.

(54) Vittadini, A.; Selloni, A.; Rotzinger, F. P.; Grätzel, M. *Phys. Rev. Lett.* **1998**, *81*, 2954.

(55) Redfern, P. C.; Zapol, P.; Curtiss, L. A.; Rajh, T.; Thurnauer, M. C. *J. Phys. Chem. B* **2003**, *107*, 11419−11427.

(56) Fahmi, A.; Minot, C. *Surf. Sci.* **1994**, *304*, 343−359.

(57) Bredow, T.; Jug, K. *Surf. Sci.* **1995**, *327*, 398−408.

(58) Lindan, P. J. D.; Harrison, N. M.; Gillian, M. J. *Phys. Rev. Lett.* **1998**, *80*, 762.

(59) Menetrey, M.; Markovits, A.; Minot, C. *Surf. Sci.* **2003**, *524*, 49−62.

(60) Redfern, P. C. **2004**, unpublished.

(61) Shapovalov, V.; Stefanovich, E. V.; Truong, T. N. *Surf. Sci. Lett.* **2002**, *498*, L103−L108.

(62) Langel, W. *Surf. Sci.* **2002**, *496*, 141−150.

(63) Rushbrook, G. S. *Introduction to Statistical Mechanics*; Clarendon Press: Oxford, 1957.

(64) *Handbook of Chemistry and Physics*, 84th ed.; Lide, D. R., Ed.; CRC Press: Boca Raton, 2003−2004.

(65) Vinet, P.; Rose, J. H.; Ferrante, J.; Smith, J. R. *J. Phys: Condens. Matter*. **1989**, *1*, 1941−1963.

(66) Arlt, T.; Bermejo, M.; Blanco, M. A.; Gerward, L.; Jiang, J. Z.; Staun Olsen, J.; Recio, J. M. *Phys. Rev. B* **2000**, *61*, 14414.

(67) Gerward, L.; Olsen, J. S. *J. Appl. Crystallogr.* **1997**, *30*, 259−264.

(68) The experimental free energies of formation of anatase (rutile) $\Delta_f G^o_{\mathscr{A}}$ ($\Delta_f G^o_{\mathscr{R}}$) taken from the JANAF tables have been applied here, such that $\Delta_f G^o_{\mathscr{A}} = -9.491471 \times 10^{-2}$ mJ/mol ($\Delta_f G^o_{\mathscr{R}} = -9.539962 \times 10^{-2}$ mJ/mol). Chase, M. W.; Davies, C. A.; Downey, J. R.; Frurip, D. J.; McDonald, R. A.; Syverud, A. N. *J. Phys. Chem. Ref. Data* **1985**, *14* (Suppl. 1), 1680.

CT0499635

JCTC Journal of Chemical Theory and Computation

Normal-Mode Analysis of Circular DNA at the Base-Pair Level. 1. Comparison of Computed Motions with the Predicted Behavior of an Ideal Elastic Rod

Atsushi Matsumoto,[†,‡] Irwin Tobias,[†] and Wilma K. Olson*[,†]

Department of Chemistry and Chemical Biology, Rutgers, The State University of New Jersey, Wright-Rieman Laboratories, 610 Taylor Road, Piscataway, New Jersey 08854-8087, and Quantum Bioinformatics Group, Center for Promotion of Computational Science and Engineering, Japan Atomic Energy Research Institute, 8-1 Umemidai, Kizu, Kyoto 619-0215, Japan

Received August 31, 2004

Abstract: We have extended a newly developed approach to study the low-frequency normal modes of mesoscopic fragments of linear DNA in order to investigate the dynamics of closed circular molecules of comparable size, i.e., a few hundred base pairs. We have added restraint energy terms and a global minimization step to treat the more complicated, spatially constrained duplex in terms of the intrinsic conformation and flexibility of the constituent base-pair "step" parameters. Initial application of the methodology to the normal modes of an ideal closed circular DNA molecule—which is naturally straight in its relaxed open linear state, inextensible, and capable of isotropic bending and independent twisting at the base-pair level—matches theoretical predictions of elastic rod dynamics. The energy-optimized closed circular states and the types of low frequency motions follow expected behavior, with (1) uniform twist density and uniform energy density in the minimum energy state; (2) a near-zero frequency torsional mode with "free" rotation about the global helical axis; (3) higher-order torsional modes accompanied by global rocking motions and pure in-plane and out-of-plane bending motions in the torsionally relaxed circle; and (4) mixed modes of bending when the chain is supercoiled (over- or undertwisted). Furthermore, the computed changes in normal-mode frequencies with imposed supercoiling or with variation of chain length are virtually identical to theoretically predicted values.

Introduction

The elastic properties of DNA are revealed in its dynamic structure. Experiments that probe the dynamics of DNA are usually interpreted in terms of the motions of a spatially homogeneous, naturally straight, elastic rod without distinguishing chemical features, and any local sequence-dependent behavior is folded into three global constants tied to the overall bending, twisting, and stretching of the molecule. By contrast, computations of the dynamics of DNA are often highly detailed, allowing for the movement of each atom in the double helical structure and the surrounding layers of solvent. The level of detail in such studies precludes the simulation of mesoscopic pieces of DNA of a few hundred base pairs, and conversely, computational treatment of mesoscopic DNA fragments necessitates some loss of information about chemical fine structure.

Recently, we reported a new computational approach for studying the dynamic properties of relatively long, linear DNA molecules without losing track of the local conformational features.[1] Specifically, we identify the low frequency normal modes that underlie the global bending, twisting, and stretching of defined polymeric sequences. Application of the method to representative DNA chain models reveals subtle relationships between sequence and collective polymeric motions. For example, the appropriate spacing of highly deformable pyrimidine-purine dimer steps in phase

* Corresponding author phone: (732)445-3993; fax: (732)445-5958; e-mail: olson@rutchem.rutgers.edu.

[†] Rutgers, The State University of New Jersey.

[‡] Japan Atomic Energy Research Institute.

10.1021/ct049950r CCC: $30.25
Published on Web 12/21/2004

with the 10-fold double helical repeat induces a mesoscopic bending anisotropy that is conducive to DNA loop formation.

The computations have now been extended so that it is possible to study more complicated DNA systems, where the structure of the double helix is restrained by long-range physical or chemical forces. In this paper, we report the dynamic properties of closed circular DNA, one of the simplest systems that can be studied with the new approach. We compare our numerical results with the theoretically predicted dynamical properties of a circular DNA modeled as an ideal, inextensible elastic rod.

Methods

We take advantage of a newly developed computational method for carrying out normal-mode analyses of long, linear DNA molecules in terms of the constituent base-pair step parameters.[1] The approach taken here is identical in all respects to the earlier work, except for an additional energy term needed to keep the ends of the chain in place and an initial energy minimization step used to incorporate this anchoring constraint. We thus omit further description of the methodology and refer the reader to the literature[1] for additional information. We focus instead on the independent variables and potential function used in the present calculations, the energy minimization process, and the theoretical principles used to determine the normal modes of DNA.

Independent Variables. Normal-mode analyses of proteins and nucleic acids are usually performed in Cartesian or dihedral angle space. The set of individual atomic coordinates used to specify molecular structure in the former case has two important computational advantages: (1) the simple form of the equations of molecular motion and (2) the detailed information about local conformational change that can be obtained from the analysis. The applicability of Cartesian-based studies is limited, however, by the large number of variables needed to describe molecular structure. The use of dihedral angles as independent variables partially overcomes this limitation, with a reduction in the number of conformational parameters realized by setting the chemical bond lengths and valence angles to equilibrium values and taking only the rotatable dihedral angles as independent variables. The contribution of the latter parameters to overall polymeric motions is dominant even if all chemical degrees of freedom are considered. Therefore, as long as one is interested in the global properties of a macromolecule, normal-mode analysis in dihedral angle space can be quite effective.[2−4]

In the case of nucleic acids, there is an even simpler way to describe three-dimensional molecular motion. Because each base or base pair can be approximated as a rigid body, molecular structure can be described in terms of the relative positions and orientations of complementary bases or successive base pairs. Six rigid-body parameters are needed to specify the relative position of each pair of rigid bodies. In this study, we treat the base pairs as rigid objects and use the six base-pair step parameters—Tilt, Roll, Twist, Shift, Slide, Rise[5]—as independent variables. The description of nucleic acid structure is not complete, however, without specification of the sugar−phosphate backbone. We therefore treat each DNA strand as a chain of nucleotide 5′-monophosphates, with each residue fixed in the B-form and related to its sequential neighbors by a given set of local base-pair step parameters. This treatment drastically reduces the number of variables needed to describe helical structure and is key to the present study of mesoscopic-length DNA circles. The step parameters are defined according to the formulation of El Hassan and Calladine,[6] and the backbone is incorporated by superposition of a 5′-nucleotide fragment from the canonical B-DNA fiber diffraction model.[7] The local chemical environment is implicitly treated in terms of the range of allowed dimeric deformations (see below). The ensemble properties deduced from the normal modes are independent of the surrounding medium, e.g., solvent viscosity.

Force Field. Minimization of the potential energy is carried out prior to normal-mode analysis and is the most time-consuming part of the calculation. The simple form of the energy, made up of the internal dimer step energies and an external ring-closure restraint term, accelerates the energy minimization step.

The energy of each dimer step, E_d, is expressed as a sum of elastic contributions over the six base-pair step parameters[8]

$$E_d = \frac{1}{2}\sum_{i=1}^{6}\sum_{j=1}^{6} f_{ij}(\theta_i - \theta_i^u)(\theta_j - \theta_j^u) \qquad (1)$$

where the θ_i ($i = 1, 2, ..., 6$) are the instantaneous values of the base-pair step parameters—Tilt, Roll, Twist, Shift, Slide, Rise—of a given dinucleotide step and the θ_i^u are the equilibrium values of the parameters in an undeformed, linear B-DNA reference state. The total internal energy is a sum of the E_d values for the n_B base-pair steps of the cyclic molecule. Here, to facilitate comparison of the computed normal modes with the theoretically predicted fluctuations of an elastic rod, we consider an ideal, naturally straight B-DNA homopolymer in which each dimer adopts an identical equilibrium rest state. In this model, the planes of neighboring base pairs are perfectly parallel ($\theta_1^u = \theta_2^u = 0°$), the equilibrium Twist is fixed at $\theta_3^u = 36°$, and the sequential displacement is restricted to Rise, i.e., $\theta_4^u = \theta_5^u = 0$ Å; $\theta_6^u = 3.4$ Å.

The elastic constants in eq 1, f_{ij}, are similarly chosen to mimic the known properties of an ideal rod. The variation of individual "step" parameters is thus assumed to be independent of one another, and the f_{ij} of cross terms ($i \neq j$) are set equal to zero. As a result, only self-product terms remain in eq 1. An energy of $k_BT/2$ is assigned to each such term at thermal equilibrium. Thus, the elastic constants f_{ii} are described in terms of the mean-square fluctuations, i.e., $f_{ii} = k_BT/\langle\Delta\theta_i^2\rangle$, where k_B is the Boltzmann constant and T the temperature in Kelvin. The expression of the f_{ii} in this form is convenient for calibrating equilibrium properties of the DNA model,[9] including the persistence length, a measure of the distance over which segments of the polymer remain directionally correlated.[10] [The scaling of the local energy in terms of the thermal fluctuations of individual base-pair step parameters should be not be confused with the normal mode deformations of the DNA as a whole, which are described below in terms of global energy changes of $k_BT/2$ or less.] The energy change in the sugar−phosphate back-

bone associated with the variation of base-pair step parameters is implicitly considered in the potential, validating the incorporation of fixed backbone units noted above. Bending deformations are assumed to be isotropic and are calibrated such that the persistence length of the linear chain is 500 Å,[9] i.e., $\langle \Delta \text{Tilt}^2 \rangle^{1/2} = \langle \Delta \text{Roll}^2 \rangle^{1/2} = 4.7°$. The variation in dimeric twist, $\langle \Delta \text{Twist}^2 \rangle^{1/2}$, is set to 4° based on estimates of the fluctuations of helical twist in supercoiled DNA involving considerations of the residual writhe in closed circular structures.[11−13] The force constants of the displacement variables—$\langle \Delta \text{Shift}^2 \rangle^{1/2} = \langle \Delta \text{Slide}^2 \rangle^{1/2} = \langle \Delta \text{Rise}^2 \rangle^{1/2} = 0.1$ Å—are assigned values large enough to preclude stretching so that the model is directly comparable to the representation of DNA used in the theory of an ideal inextensible elastic rod.[14,15] Calculations based on the model can thus be compared with predictions of the theory. The choice of local force constants produces a linear polymer with global elastic properties characterized by a bending rigidity A of 2.1×10^{-19} erg-cm and a twisting modulus C of 2.9×10^{-19} erg-cm. The latter values are obtained from the computed normal-mode frequencies of the linear molecule.[1] Effects of sequence, i.e., the known sequence-dependent fine structure and deformability of DNA,[8] are not considered in the theory and are thus ignored in the present calculations. The assumed DNA variability is, nevertheless, comparable to the range of conformational distortions seen in high-resolution crystal structures but is broader than the local angular fluctuations deduced from direct measurements of short-time chain dynamics.[16,17] As discussed previously,[1] the potential energy of DNA is not a smooth quadratic function. Whereas the equilibrium bending properties of DNA reflect transitions between multiple discrete minima, the apparent stiffness of the DNA probed in the time-dependent experiments seemingly arises from molecules trapped in a single low-energy state.

The restraint energy, E_r, is given by a sum over N distance restraints

$$E_r = \sum_{k=1}^{N} C_k (d_k - d_k^\circ)^2 \qquad (2)$$

where C_k is an arbitrarily chosen spring constant, d_k is the instantaneous distance between two points on which a distance restraint is placed, and d_k° is the desired separation distance. Restraints that force the overlap of the origins and the x- and z-axes of terminal base pairs are used to generate DNA circles. Because the restraint energy is close to zero after energy minimization, the dimeric contributions from eq 1 dominate the total energy.

Energy Minimization. The normal modes of a molecule are collective, low-energy motions of the system as a whole in the vicinity of its minimum energy configuration. The deformations in global structure are described by harmonic fluctuations of individual conformational parameters from their rest values at the energy minimum. In the case of linear DNA, the minimum energy state is self-evident from the equilibrium rest states, i.e., θ_i^u values, of sequential dimers. The minimum energy configuration of circular DNA, however, cannot be determined a priori because of the additional long-range distance restraint(s). Therefore, numerical calculations must be performed to find the state of lowest energy. Both conjugate gradient and Newton−Raphson methods are used in the minimization process, the former approach introduced at the first stage of optimization and the latter technique once the decrease of energy per iteration becomes small. Determination of the requisite first and second derivatives of the energy with respect to the base-pair step parameters, however, is not straightforward as the restraint energy (eq 2) is a function of long-range intramolecular distances. We make use of analytical expressions for the derivatives developed by Gō and co-workers[18,19] to treat the normal mode dynamics of a system of two molecules, each of which moves in dihedral angle space. The set of rigid body parameters used here to relate neighboring base-pair planes is identical in form to the variables previously used to describe the relative global positions and orientations of different molecules. The second derivative of the energy function and the molecular configuration at the energy minimum are used as input for normal-mode analysis.

Global Reference System. Each base pair of the energy-minimized DNA circle has a fixed coordinate frame defined in accordance with established guidelines.[20] The sequence of vectors connecting the origins of base-pair coordinate frames forms an approximately circular pathway in the minimum energy state. The step parameters used as independent variables in the normal-mode calculations, while useful for analyzing structural changes of DNA at the base-pair level, are less convenient for describing the global structural changes of a closed, circular DNA. We thus introduce a second coordinate frame on each base pair with one axis $\mathbf{n}^\circ$ directed from the origin toward the center of the circle, a second axis $\mathbf{b}^\circ$ oriented perpendicular to the plane of the circle, and the third axis $\mathbf{t}^\circ$ chosen to form a right-handed coordinate system ($\mathbf{t}^\circ = \mathbf{n}^\circ \times \mathbf{b}^\circ$). These three axes constitute a right-handed Frenet-Serret triad ($\mathbf{n}^\circ$, $\mathbf{b}^\circ$, $\mathbf{t}^\circ$), with elements corresponding respectively to the principal normal, binormal, and unit tangent vector at points on the DNA circle. For further analysis, we define a fourth axis $(\mathbf{O}_{\text{BP}} - \mathbf{P}) \times \mathbf{t}^\circ$, where $\mathbf{O}_{\text{BP}}$ is the origin of the base-pair coordinate frame at a given hydrogen-bonded nucleotide unit in an arbitrary DNA structure, and $\mathbf{P}$ is a point on the base-pair plane. This axis is useful for analyzing torsional motions, where a given point on each base-pair plane rotates around $\mathbf{t}^\circ$. By analyzing the fluctuations of the molecule with respect to these axes, we obtain an image of overall global motion for each normal mode that complements the picture of local structural deformations provided by the base-pair step parameters.

Analytical Treatment. Given the linear differential equations which appear in the analytical theory of the normal modes of circular rings formed from naturally straight, inextensible, symmetric, elastic rods,[14,15] the distortions from the circular configuration (the solutions of these equations) can be written as a linear combination of terms, each of which is proportional to the time-dependent function $e^{\pm i\omega t}$, where ω is the frequency. Furthermore, the fact that the axial curve is closed requires that the solutions be periodic, so that their dependence on arc length s (along the axial curve)

is similarly determined. That is, each of the terms in the linear combination must also be proportional to $e^{\pm in\kappa^o s}$, where κ^o is the curvature of the circle and $n = 0, 1, 2, \ldots$. (The wavelength λ_n of a normal mode is simply related to the integer n by the equation $\lambda_n = 2\pi/n\kappa^o$.) Introduction of these substitutions leads to a system of homogeneous linear equations with unknown constant amplitudes. The determinant made up of the coefficients of these amplitudes is a function of Ω, κ^o, ΔTw^o, n, and ω^2, where ΔTw^o is the excess twist (of the circular equilibrium structure as a whole) and $\Omega = C/A$ is the ratio of the global twisting modulus to the global bending modulus. The determinant takes the form of a polynomial cubic in ω^2, $M_3\omega^6 + M_2\omega^4 + M_1\omega^2 + M_0$, with the coefficients M_3, M_2, M_1, and M_0 given by

$$M_3 = 2(\kappa^o)^{-10}(1 + (\kappa^o)^2 + n^6(\kappa^o)^4 - 2n^4(\kappa^o)^2((\kappa^o)^2 - 1) + n^2((\kappa^o)^4 - (\kappa^o)^2 + 1))$$

$$M_2 = -n^2(\kappa^o)^{-8}\begin{pmatrix} n^6(\kappa^o)^4(\Omega + 4) + 2n^4(\kappa^o)^2(\Omega + 2 - 5(\kappa^o)^2) \\ + n^2(4(\kappa^o)^2(2(\kappa^o)^2 - 1) + \Omega(1 + (\kappa^o)^2 - 3(\kappa^o)^4)) \\ + 2(\kappa^o)^4(\Omega - 1) + \Omega(3(\kappa^o)^2 + 1) \end{pmatrix}$$

$$M_1 = n^4(n^2 - 1)(\kappa^o)^{-4}\begin{pmatrix} -(\Omega - 2)(\kappa^o)^2 + 2n^4(\Omega + 1)(\kappa^o)^2 \\ -n^2(4(\kappa^o)^2 + \Omega((\kappa^o)^2 - 2)) \\ -2\Omega^2(\kappa^o)^2(\Delta Tw^o)^2(n^2 - 1) \end{pmatrix}$$

$$M_0 = -n^6(n^2 - 1)^2\Omega(n^2 - 1 - (\Delta Tw^o)^2\Omega^2) \qquad (3)$$

The normal-mode frequency ω for given values of the four parameters Ω, κ^o, ΔTw^o, and n is obtained by setting the polynomial equal to zero and solving the cubic equation in (ω^2), i.e., the roots of the polynomial determine the normal-mode frequencies of the elastic ring for that choice of parameters.

For each value of n, there are three distinct frequencies. The case n equal to 0 or 1 is special. For these values of n, both M_1 and M_0, the only two of the four coefficients that depend on ΔTw^o, are zero. It follows that the six mode frequencies associated with these two values of n are independent of ΔTw^o. Of these six frequencies, five are zero, and the nonzero frequency is given by $(M_2/M_3)^{1/2}$. Finally, the circular symmetry inherent in the problem gives rise to a 2-fold degeneracy: in general there are two modes associated with each distinct normal-mode frequency.

For the case of the torsionally relaxed circle, where $\Delta Tw^o = 0$, the coefficients M_i in the dispersion relation are such that it is possible to give explicit expressions for the square of each of the three frequencies associated with a given n. The square of the frequency $\omega_{ip}^2(n)$ of the two degenerate in-plane modes for that n is

$$\omega_{ip}^2(n) = \frac{n^2(n^2 - 1)^2}{(R^o)^2((n^2 - 1)^2 + (R^o)^2(n^2 + 1))} \qquad (4)$$

and the square of the two remaining frequencies, those of the out-of-plane modes, $\omega_{op}^2(n)$, are

$$\omega_{op}^2(n) = \frac{n^2(\Omega(n^2 + (R^o)^2 + 2) + 2(n^2 - 1))}{4(R^o)^2(n^2 + (R^o)^2)} \pm \frac{n^2\sqrt{(\Omega(n^2+(R^o)^2+2) + 2(n^2 - 1))^2 - 8\Omega(n^2 - 1)(n^2 + (R^o)^2)}}{4(R^o)^2(n^2 + (R^o)^2)} \qquad (5)$$

where R^o $(= 1/\kappa^o)$ is the radius of the circle. In these expressions R^o is measured in units of half the radius r of the elastic rod, about 5 Å in the case of DNA, and the frequency is measured in units of $(4/r^3)(A/\pi\rho)^{1/2}$. Here ρ is the density of the material from which the rod is made.

Results and Discussion

Minimum Energy Circular Configurations. We start the energy minimization from an open, perfectly straight DNA molecules with ideal B-DNA step parameters ($\theta_1 = \theta_2 = 0°$, $\theta_3 = 36°$, $\theta_4 = \theta_5 = 0$ Å, $\theta_6 = 3.4$ Å) at all dimeric units. The optimization balances two competing terms, i.e., the elastic energy in eq 1, which acts to preserve the B-DNA equilibrium geometry ($\theta_1^u = \theta_2^u = 0°$, $\theta_3^u = 36°$, $\theta_4^u = \theta_5^u = 0$ Å, $\theta_6^u = 3.4$ Å) of individual base-pair steps, and the end-to-end restraint energy in eq 2, which forces chain cyclization. The energy-minimization process is divided into two steps. Minimization is initially performed with restraints that force the overlap of the origins and normals (z-axes) of terminal base pairs and then is carried out with an additional restraint on the x-axes of the same residues. This two-step procedure guarantees successful ring closure and also ensures that the energy-minimized circular DNA configuration is torsionally relaxed, i.e., with the desired number of helical turns.

Using the energy-minimized structure as an initial conformation, we then create a series of slightly perturbed 200 base pair (bp) closed circular DNA molecules by changing the intrinsic Twist θ_3^u in eq 1, to study the effects of over- and undertwisting on the structure and dynamics of circular DNA. The uniform variation of θ_3^u mimics the known effects of temperature and added salt on DNA helical twisting.[21−23] The potential energies found upon minimization are plotted versus the intrinsic Twist θ_3^u in Figure 1. A parabolic function of the form $y = a(x - x_0)^2 + y_0$, which fits the potential energy plot, is also depicted in the figure. The minimum energy value y_0, the Twist angle x_0 at the minimum, and the coefficient a of the parabolic function are summarized in Table 1.

Mechanical Constants. The minimum energy state and curvature in Figure 1 can be used as follows to deduce mechanical constants that describe the global bending and twisting rigidity of an ideal, inextensible DNA elastic rod. Given that the potential energy of the rod arises from its overall bending and twisting, the minimum energy in the parabolic function in Figure 1 corresponds to the torsionally relaxed state of DNA, where the energy contribution comes purely from bending. Furthermore, the bending energy of an intrinsically straight rod closed into a circle is $2\pi^2A/L$, where A is the bending rigidity and L is the length of the rod.[24] The bending rigidity is therefore easily calculated from the minimum energy values as $A = y_0(L/2\pi^2)$ and is also

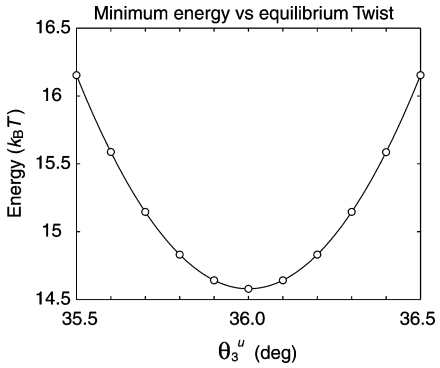


Figure 1. Minimized energy of a 200 bp closed circular DNA molecule, which is naturally straight in its equilibrium rest state and governed by an ideal elastic potential, plotted as a function of the intrinsic Twist, θ_3^u. The parabolic function, $y = a(x - x_0)^2 + y_0$, fitted to computed points of minimum energy (open circles), is depicted by a solid line. See Table 1 for numerical values of the fitted constants.

Table 1. Details of the Parabolic Function, $y = a(x - x_0)^2 + y_0$, Which Fits the Minimum Energy Plot in Figure 1, and Derived Mechanical Constants of a 200 bp DNA Circle, Which Is Naturally Straight in Its Equilibrium Rest State and Subject at the Level of Neighboring Base-Pair Steps to an Ideal Elastic Potential

x_0 (deg)	y_0 (k_BT)	a (k_BT/deg^2)	bending rigidity[a] (10^{-19} erg-cm)	twisting rigidity[a] (10^{-19} erg-cm)
36	14.579	6.298	2.066 (2.098)	2.892 (2.879)

[a] Mechanical constants calculated on the basis of the normal-mode frequencies of the corresponding linear DNA chain are reported in parentheses.

reported in the table. Similarly, the curvature $2a$ of the parabolic function in Figure 1 is related to the twisting rigidity C by $a = Cn_B^2/2L$, where n_B is the number of base-pair steps. Cyclization of a linear DNA of n_B bp with intrinsic Twist $x = \theta_3^u$ introduces additional torsional stress, $-n_B(x - x_0)$, where x_0 is the equilibrium Twist of the torsionally relaxed molecule. The increase in twist raises the elastic energy by $Cn_B^2(x - x_0)^2/2L$,[24] corresponding to an increment of $a(x - x_0)^2$ in the computed energy (Figure 1). Therefore, the global twisting rigidity is easily calculated and also reported in Table 1. The values of the bending and twisting rigidities of a linear 200 bp DNA molecule with the same conformational features are included for comparison in the table. The latter constants, which are derived from the computed normal-mode bending and twisting frequencies of an open chain subject to the same dimeric potential,[1] are remarkably close to the values calculated from the minimum energy of the circular molecule.

It should be noted that the above expressions for A and C are only valid for the current simplified DNA force field, where the Twist in the minimum energy structure is nearly constant at every base-pair step. In the case of real DNA,[8] the uptake of Twist in the minimum energy state is not uniform, with local variations associated with intrinsic bending, correlations between base-pair parameters, etc.[25,26] The preceding calculation of mechanical constants does not hold in such instances.

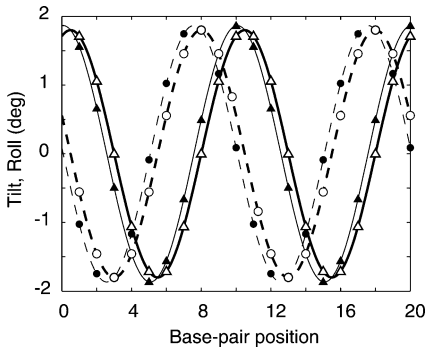


Figure 2. Sequential variation of Tilt° and Roll° angles (Δ and o symbols, respectively) along the contour of a 200 bp torsionally relaxed DNA circle which is naturally straight in its equilibrium rest state and subject to an ideal elastic potential. Cosine functions fitted to the data (see Table 2) are represented respectively by thick solid and dashed curves. Values of Tilt and Roll (filled-in Δ and o symbols) at the moment when the potential energy is raised by $k_BT/(2 \times 10^4)$ along the lowest frequency ($n = 0$) torsional mode are also plotted. Cosine functions fitted to the data are represented by thin solid and dashed curves, respectively. Because of the repetitive pattern of conformational variation, data are shown for only the first 20 steps of the chain.

Table 2. Base-Pair Step Parameters, in Degrees, at the mth Dimer Step of an Ideal, Inextensible[a] Circular DNA Molecule of n_B Base Pairs and Intrinsic Twist θ_3^u in the Minimum Energy State

n_B	θ_3^u	Tilt°	Roll°	Twist°
160	36	2.25 cos(36(m − 0.50))	2.25 cos(36(m − 0.50) + 90)	36
180	36	2.00 cos(36(m − 0.50))	2.00 cos(36(m − 0.50) + 90)	36
200	35.5	1.80 cos(36(m − 0.74))	1.80 cos(36(m − 0.74) + 90)	36
200	36	1.80 cos(36(m − 0.50))	1.80 cos(36(m − 0.50) + 90)	36
200	36.5	1.80 cos(36(m − 0.12))	1.80 cos(36(m − 0.12) + 90)	36
220	36	1.64 cos(36(m − 0.50))	1.64 cos(36(m − 0.50) + 90)	36
240	36	1.50 cos(36(m − 0.50))	1.50 cos(36(m − 0.50) + 90)	36

[a] (Shift°, Slide°, Rise°) = (0 Å, 0 Å, 3.4 Å).

Smooth Bending. Figure 2 illustrates the uptake of bending along the contour of the 200 bp torsionally relaxed DNA circle in its minimum energy state. Because the same conformational pattern is repeated every 10 bp, step parameters in the minimum energy structure, which are henceforth termed Tilt°, Roll°, Twist°, ..., are reported for only a fragment of the chain, here the first 20 bp steps. [These "static" values should not be confused with the "dynamic" fluctuations of step parameters which are reported in later sections.] As expected from the theory,[24] the values of Twist° are constant (36°) in the minimum energy state and therefore not shown. The sinusoidal changes in Tilt° and Roll° follow the characteristic pattern of "smooth" DNA bending,[27−29] with one parameter attaining a maximum or minimum value when the other is zero. Details of the best-fit cosine functions, which are plotted in the figure, are summarized in Table 2. As is clear from the expressions, the sequential changes of Tilt° and Roll° differ in phase by 90°. The curves that fit Tilt° and Roll° values in DNA with different values of intrinsic Twist, θ_3^u, are also summarized in the table. The

latter curves, reported for chains of 200 bp, are also related to each other by a 90° phase shift.

In a later section, we study the dependence of normal-mode frequencies on the length of DNA. For this reason, we have constructed torsionally relaxed DNA circles of different chain lengths (160, 180, 220, 240 bp) using the same ideal force field. The values of Tilt° and Roll° in the minimum energy state are again well approximated by single cosine curves, which are summarized in Table 2. Step parameters, in degrees, at the *m*th dimer steps of the minimum energy structures fit the following general form if n_B, the number of base pairs, is a multiple of 10:

$$\text{Tilt}^\circ = \left(\frac{360}{n_B}\right)\cos(36(m - 0.5))$$

$$\text{Roll}^\circ = \left(\frac{360}{n_B}\right)\cos(36(m - 0.5) + 90)$$

$$\text{Twist}^\circ = 36 \qquad (6)$$

The inverse dependence in this expression of the amplitudes of Tilt° and Roll° on n_B means that the curvature of the molecule is proportional to the amplitude of local conformational changes. Since curvature is the inverse of the radius of a circle and the radius of a circular DNA molecule is proportional to the number of base pairs, curvature is also proportional to the inverse of n_B. Thus, when the amplitude of Tilt° and Roll° variations is larger, curvature is larger, and when the amplitude is smaller, curvature is smaller.

If, as considered here, the elastic constants f_{ii} impeding Tilt and Roll deformations are identical, the circular state remains a conformational energy minimum even if the constant −0.5 in eq 6 is replaced by other values. The restraint energy of such a molecule, however, differs from the ideal closed molecule in being greater than, albeit very close to, zero. Thus, the constant in the phase of the cosine functions in eq 6 is affected by the restraints in eq 2. For example, if we place restraints on the *y*-axes, rather than the *x*-axes, of terminal base pairs, the constant is −3. The effect of the starting structures on the optimized circular state, however, is limited.

Uniform Twisting and Uptake of Energy. Figure 3 shows the contributions of Tilt, Roll, and Twist to the dimer step energy at sequential positions along three 200 bp DNA circles in different minimum energy configurations: a torsionally relaxed molecule, an overtwisted chain, and an undertwisted chain with respective intrinsic Twist angles, θ_3^u, of 36°, 35.5°, and 36.5°. Because the energetic patterns repeat every 10 bp, data are plotted, as above, for only the first 20 bp steps. As reported in Table 2, the value of Twist° in the minimum energy configuration shows only limited dependence on the assumed intrinsic Twist. As a result, when the intrinsic Twist angle is decreased, Twist° exceeds the intrinsic value and the DNA is overtwisted. In the same way, when the intrinsic value of Twist is increased, the DNA becomes undertwisted. The change of intrinsic Twist θ_3^u from the value of 36° in the torsionally relaxed state introduces a constant contribution to the dimer energy at every base-pair

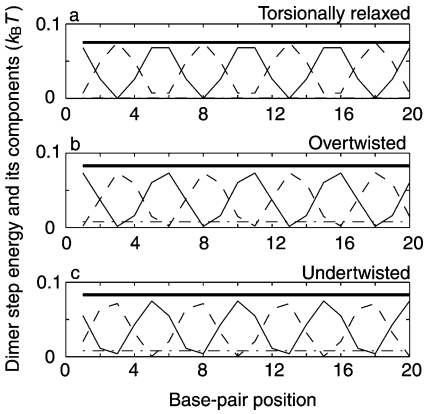


Figure 3. Dimer step energy (thick solid lines) and its bending and twisting components (Tilt: thin solid lines; Roll: dashed lines; Twist: dash−dotted lines) of (a) torsionally relaxed, (b) overtwisted, and (c) undertwisted DNA molecules, which are naturally straight at equilibrium, subject to an ideal elastic potential, and closed via energy minimization into 200 bp circles. The intrinsic Twist, θ_3^u, is fixed respectively at 36°, 35.5°, and 36.5° in the three structures. Because of the regular energetic patterns, data are shown for only the first 20 steps.

step, a well-known property of a naturally straight, homogeneous elastic rod.[24] Moreover, the energetic contributions associated with the deformations of Tilt and Roll are compensatory so that the dimer step energy is nearly constant along the chain contour. The uniform energy density along the chain is another well-known property of a naturally straight elastic rod with constant curvature.[24] This uniformity does not hold, however, if the chain is subject to local anisotropic bending, i.e., different force constants f_{ii} for Tilt and Roll.

Motions of Relaxed DNA Circles. The normal-mode analysis of energy-minimized DNA circles provides information on the frequencies and collective motions of the closed chain. We concentrate here on the very low frequency normal modes responsible for large-scale deformations of structure, studying both torsionally relaxed DNA and molecules with torsional stress. The normal modes of lowest frequency show one of three types of large-scale deformation, namely in-plane, out-of-plane, and torsional motions. These motions, which are consistent with the predictions of elastic rod theory,[14,15] are illustrated schematically in Figure 4 and reported as a function of frequency in the color-coded "spectrum" in Figure 5(a). The in-plane bending motion of the torsionally relaxed molecule involves changes in Tilt and Roll, whereas the out-of-plane bending motions always involve torsion. Local bending deformations also accompany the global torsional motions.

"Free" Torsional Motions. The DNA remains circular in the lowest (nearly zero) frequency mode, with the molecule rotating as a whole about its helical axis; see Figures 4(a) and 5(a), and the Supporting Information. Such behavior is predicted by the theory of elastic rod deformations, i.e., modes associated with *n* equal to 0 or 1 must be nonflexural and the rotation of the DNA in the $n = 0$ torsional mode is unimpeded.[30]

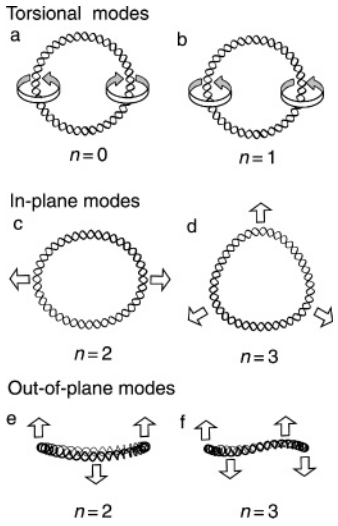


Figure 4. Schematic illustration of selected low-frequency normal modes of a closed circular double helical DNA molecule. The arrows denote the directions of bending or twisting in each mode. Images (a,c,e) depict the lowest frequency mode in each of the three types of global motions and images (b,d,f) the second lowest frequency modes. The values of *n* refer to indices used in elastic rod theory[14,15] to differentiate modes within the same class of motions. Double helical images generated with MolScript.[33]

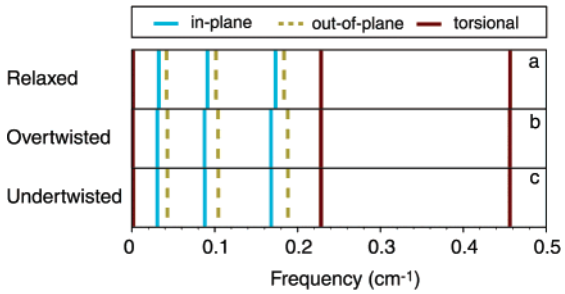


Figure 5. Color-coded spectra of the lowest frequency torsional (unbroken red line), in-plane (unbroken blue line), and out-of-plane (broken green line) modes of a 200 bp DNA molecule, which is naturally straight in its equilibrium rest state, subject to an ideal elastic potential, and closed into a circle with different values of intrinsic Twist: (top) torsionally relaxed ($\theta_3^u = 36°$); (middle) overtwisted ($\theta_3^u = 35.5°$); (bottom) undertwisted ($\theta_3^u = 36.5°$).

The sequential variation of base-pair step parameters which accompanies the "free" torsional motion is compared in Figure 2 with the corresponding values in the equilibrium reference state (filled-in vs corresponding open symbols). The plot captures the values of Roll and Tilt in a 20 bp fragment of the 200 bp circle at the instant when the energy of the DNA is raised by $k_BT/(2 \times 10^4)$. The very low-energy threshold is a consequence of the extremely low frequency, i.e., small energy change, of the mode and the limitation of normal-mode analysis to conformational fluctuations in the vicinity of the minimum energy state. (Step parameters lie far away from the reference state if the mode is assigned a higher energy.) Twist, by contrast, remains constant during the process. Thus, the nearly free rotation of the helix as a whole arises from concerted changes in local bending, rather than twisting, over a 10 bp helical repeat. The regular pattern of Roll and Tilt variations is emphasized by the cosine functions superimposed on the computed points in Figure 2.

As long as the variation in bending parameters is small compared to the range of Roll° and Tilt° in the minimum energy state (the conditions under which normal-mode analysis holds), the Roll and Tilt values will simply move back and forth along the chain contour, i.e., abscissa in Figure 2, as the amplitude and direction of fluctuations change over time. The resulting shift in the sites of maximum and minimum bending leads to the global rotational motion. For example, the minima of Roll° (thick solid lines) in Figure 2 occur at positions where the minor groove edges of base pairs 3, 13, 23, etc. face the center of the circle. The shift of phase associated with the normal mode moves base-pair steps such as these on the inside of the circle to the outer surface of the molecule, and vice versa, thereby producing rotational motion around the helical axis.

In-Plane Bending. The second and third lowest frequency modes correspond to in-plane motions, which deform the DNA circle to elliptical shapes; see Figure 4(c) and the Supporting Information. These modes with nearly equivalent frequencies differ only in the direction of global distortion, one mode generating a family of ellipses with major and minor axes rotated ~45° with respect to the corresponding axes of the second family of structures. Both the pure in-plane character of the modes and the equivalence of the frequencies are consistent with theoretical predictions. The directional differences in the in-plane bending modes are seen in the fluctuations of base-pair origins in the nonequilibrium states with respect to the (**n**°, **b**°, **t**°) Serret-Frenet frames embedded in each base pair of the minimum energy (perfectly circular) reference state. The plots in Figure 6-(a,b) show the base-pair displacements at the moment when the potential energy of the DNA in the two modes is raised by $k_BT/2$. The uniformly zero displacements along the **b**° axes of the planar circle (dashed lines in the figure) confirm the pure, in-plane nature of molecular deformation, and the larger sinusoidal displacements along the **n**° and **t**° axes (solid curves) the elliptical shapes. The major axes of the ellipses captured in Figure 6(a,b) run parallel to the vectors that join base pairs (20, 120) and (95, 195), respectively. These pairs of points correspond to the minima in the **n**°-axis displacements in the two plots. The 25 bp phase shift of corresponding curves of the two in-plane modes accounts for the different directions of global motion. The similar, yet out-of-phase pattern of conformational change observed here is reminiscent of the doubly degenerate pure bending modes of a linear DNA, with overall chain deflections in perpendicular directions.[1] Like the open chain, the closed DNA effects global distortions exclusively through Tilt and Roll, thereby qualifying as a bending mode; see Figure 7(a, b). As with the linear molecule, the degeneracy of the in-plane modes of a circular molecule is expected to break down for particular sequential repeats.

Examination of Figure 7(a,b) shows that the maximum or minimum fluctuations of Tilt and Roll appear at positions where Tilt° and Roll° assume a maximum or minimum in the equilibrium state and that the fluctuations are zero at

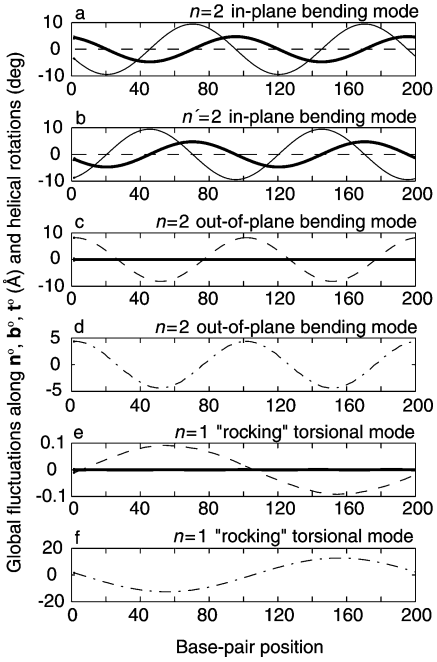


Figure 6. Displacement of base-pair origins and rotation of base-pair axes at the moment when the potential energy is raised by $k_BT/2$ in selected normal modes of a 200 bp torsionally relaxed DNA circle which is naturally straight in its equilibrium rest state. Displacements (thin solid, dashed, and thick solid lines) measured respectively along the **n**°, **b**°, **t**° axes of Serret-Frenet frames embedded in each base pair of the minimum energy DNA circle and found in (a,b) the pair of lowest frequency ($n = 2$) in-plane bending modes; (c) one of the lowest frequency ($n = 2$) out-of-plane bending modes; (e) one of the second lowest ($n = 1$) frequency "rocking" torsional modes. Rotational fluctuations (dash−dotted line) found in (d) one of the lowest frequency ($n = 2$) out-of-plane bending modes and (f) one of the second lowest ($n = 1$) frequency "rocking" torsional modes.

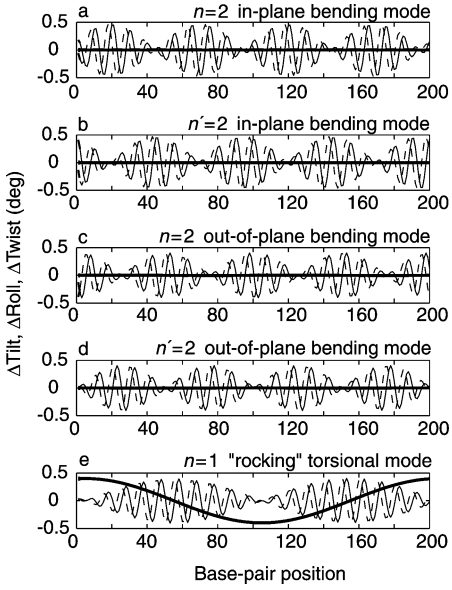


Figure 7. Fluctuations of local angular "step" parameters which are collectively responsible for selected normal modes of a 200 bp torsionally relaxed DNA circle which is naturally straight in its equilibrium rest state and subject to an ideal elastic potential: (a,b) the pair of lowest frequency ($n = 2$) in-plane bending modes; (c,d) the pair of lowest frequency ($n = 2$) out-of-plane bending modes; (e) one of the pair of second lowest frequency ($n = 1$) torsional modes. Plots illustrate the fluctuations of Tilt (thin solid lines), Roll (dashed lines), and Twist (thick solid lines) along the contour of the DNA molecule at the moment when the potential energy of the molecule is raised by $k_BT/2$; fluctuations are reversed a half cycle later of the mode.

steps where equilibrium values are zero. That is, the points of maximum or minimum Tilt and Roll do not change position in these two modes, a further indication that the bending motions are restricted to the plane of the energy-optimized circular structure. This pattern of dimer deformation in each of the modes is quite distinct from the traveling waves of local bending associated with the "free" torsional mode. The in-plane bending modes differ in phase such that the largest changes in curvature in one mode occur at positions where there are no deformations in the other mode and vice versa.

Addition of the two in-plane bending modes, as represented in Figure 7(a,b), creates a standing wave of chain deformation similar in character to each of its degenerate components, i.e., the closed molecule oscillates between circular and elliptical shapes via localized compensatory changes in Tilt and Roll. If the two modes are not in phase, the linear combination of modes produces a traveling elliptical wave with the sites of maximal conformational distortion moving over time along the chain contour (see the Supporting Information). The direction of motion and the magnitude of elliptical distortion of the traveling wave depend on the difference in phase angle $\Delta\delta$, i.e., the direction of the movement reverses if the sign of $\Delta\delta$ (defined over the range $-180° < \Delta\delta \leq 180°$) is changed. The proportions of the traveling ellipse are constant only when $\Delta\delta = \pm 90°$. Otherwise, both the proportions and the sites of elliptical deformation vary with time.

The local step parameters responsible for the one of the in-plane bending modes, obtained by adding the computed (ΔTilt, ΔRoll) deformations to the (Tilt°, Roll°) values that describe the minimum energy state, are presented in Figure 8. The agreement in phase of the (ΔTilt, ΔRoll) variation with the trigonometric equations (Table 2) that define the (Tilt°, Roll°) minimum energy state works positively at some positions, increasing the amplitude, and negatively at others, decreasing the amplitude. The amplitude assumes a maximum at the 20th and 120th base-pair steps, and minima at the 70th and 170th base-pair steps in the plot. As noted above, the magnitude of the amplitude is proportional to the curvature. Therefore, at the former pair of antipodal base-pair steps, the curvature is larger, and at the latter steps, the curvature is smaller than in the minimum energy structure. A half cycle later, the situation reverses, with lesser curvature found at the former positions and greater curvature at the latter ones.

Out-of-Plane Bending. The fourth and fifth lowest normal modes, also of close frequency, describe out-of-plane, cup-like bending deformations of the DNA circle; see Figure 4(e) and the Supporting Information. The null displacement, in

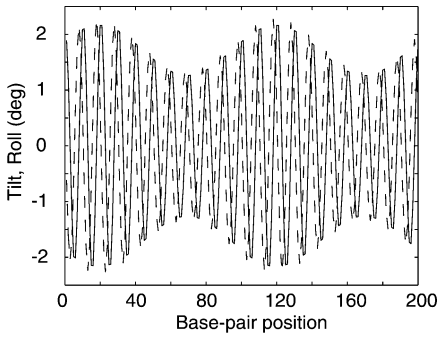


Figure 8. Values of Tilt and Roll (solid and dashed lines) along the contour of a 200 bp torsionally relaxed DNA circle at the moment when the potential energy is raised by $k_BT/2$ by lowest frequency ($n = 2$) in-plane bending motions. Data are obtained by adding the bending fluctuations in Figure 7(a) to the energy-optimized parameters of the covalently closed, molecule in Figure 2.

Figure 6(c), of base-pair origins in directions other than **b**° points to the pure out-of-plane character of the motions and the uptake of deformation through Tilt and Roll in Figure 7(c) to the role of bending. The major difference between the two out-of-plane modes is the direction of distortion, expressed at the local level by a 25 bp shift in the pattern of conformational fluctuations and at the global level by an orientational change of ~45°. Notably, the profiles of conformational change responsible for the out-of-plane modes are closely related to those associated with the in-plane deformations; see Figure 7(a−d). Values of ΔTilt and ΔRoll associated with in-plane deformations attain local maxima or minima at the base-pair steps that are unchanged during out-of-plane distortions and vice versa.

There is a rotational motion of the bases in the out-of-plane modes similar to that found in the lowest frequency torsional mode, the "free" torsional mode. The degree of rotation varies along the chain contour, and the direction of rotation is of the opposite sense at positions 50 bp apart. If adjacent segments of a perfect circular rod were to rotate in the opposite sense, torsional stress would build up between them, but as seen in Figure 7(c,d), ΔTwist is approximately zero in the out-of-plane modes. The torsional stress is relieved by the displacement of antipodal segments of the chain perpendicular to the plane of the circle. The fluctuations of the base-pair origin along the **b**° axis in Figure 6(c) are in perfect phase with the rotation of bases, monitored in Figure 6(d) by the positions of ($\mathbf{O}_{BP} - \mathbf{P}$) vectors on successive base pairs. Such deformations also agree with the theoretical expectation that the out-of-plane modes entail torsional motions.

Higher-Frequency Modes. The majority of very low frequency modes of the ideal DNA chain are either in- or out-of-plane bending motions (see the color-coded spectrum of frequencies in Figure 5(a)). Furthermore, the two types of modes occur as closely spaced pairs, with the 2-fold degeneracy of each mode leading to four independent motions of close frequency. The pure in-plane or out-of-plane character of the motions, however, disappears as the frequencies become larger. For example, the out-of-plane fluctuations (along the **b**°-axis) of the in-plane modes become increasingly larger, and it becomes impossible to tell whether the mode should be classified as an in-plane or an out-of-plane mode. That is, fluctuations along **b**° become comparable to those along **n**° and **t**°.

Like the bending modes of a naturally straight linear DNA molecule, the number of pure in-plane or pure out-of-plane bending modes of the circular chain is limited in number. For both in- and out-of-plane modes, ΔTilt and ΔRoll fluctuate in cycles of 10 bp steps, with null fluctuations of parameters occurring simultaneously at several steps (nodes). Reexamination of Figure 7(a−d) reveals four such nodes, separated by 50 bp increments, for the lowest frequency in-plane and out-of-plane bending modes of a 200 bp closed circle. The number of nodes increases, and the distance between adjacent nodes decreases as the frequency becomes higher. If adjacent nodes approach too closely, ΔTilt and ΔRoll can no longer fluctuate in cycles of 10 bp steps, thereby limiting the number of pure in-plane or out-of-plane modes. The 200 bp DNA circle has only eight pairs of fairly pure in-plane modes and eight pairs of fairly pure out-of-plane modes, with the highest frequencies of the in-plane modes being 0.869 cm^{-1} and 0.870 cm^{-1}, and those of the out-of-plane modes 0.877 cm^{-1} and 0.878 cm^{-1}. There are 18 nodes in these four higher frequency modes.

A second ($n = 1$) torsional mode, which preserves the circular structure of DNA and allows for the opposing rotation of base pairs around the helical axis, is found at higher frequency (Figures 4(b) and 5(a) and the Supporting Information). As is clear from Figure 6(e,f), the rotational motion (illustrated in the latter plot) is accompanied by small (±0.1 Å) out-of-plane base-pair displacements (dashed curves in the other plot). The latter features are consistent with the small oscillatory rocking motion predicted by elastic rod theory for this mode.[14] [The term rocking is characteristic of a standing mode; the motions of the traveling modes generated by the linear combination of degenerate $n = 1$ torsional modes resemble those of a precessing top.] The opposing rotations at antipodal sites on the circle introduce torsional stress on the intermediate regions, leading to the build-up of ΔTwist illustrated by heavy solid lines in Figure 7(e). The bending angles are unchanged at the sites of maximum over- and underwinding but are maximally perturbed when ΔTwist is zero. The patterns of ΔTilt and ΔRoll resemble those for the out-of-plane bending motions in terms of the ~10 bp phase of the computed fluctuations. The sites that bend more strongly or become more straightened around the closed structure are more widely spaced along the chain contour than the sites of strongest and weakest bending in the out-of-plane mode, i.e., repetition at ~100 bp vs ~50 bp intervals. A degenerate torsional mode of the same type introduces torsional stress and bends the DNA in the same fashion but is displaced 50 bp further along the chain contour. The degeneracy of the $n = 1$ torsional mode also matches theoretical expectations.

Motions of Supercoiled Molecules. *Normal-Mode Frequencies.* Slight over- or undertwisting of the DNA circle introduces subtle changes in the computed normal-mode frequencies. As evident from the color-coded spectra of lowest frequency modes in Figure 5(a−c), the torsional

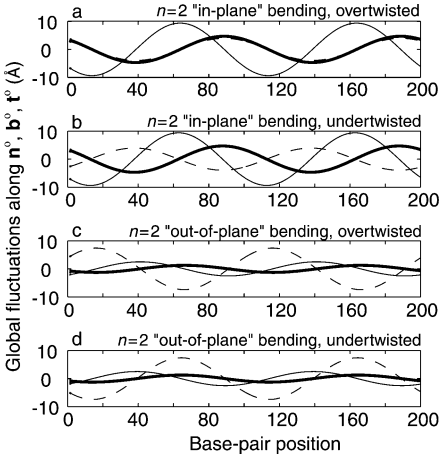


Figure 9. Displacement of base-pair origins of (a,c) overtwisted ($\theta_3^u = 35.5°$) and (b,d) undertwisted ($\theta_3^u = 36.5°$) DNA circles along the **n**°, **b**°, **t**° axes of Serret-Frenet frames embedded in each base pair of the minimum energy DNA circle. Data reported at the moment when the potential energy of the naturally straight, 200 bp molecule is raised by $k_BT/2$ in selected normal modes: (a,b) one of the lowest ($n = 2$) frequency, predominantly in-plane bending modes; (c,d) one of the lowest ($n = 2$) frequency, predominantly out-of-plane bending modes. See legend to Figure 6.

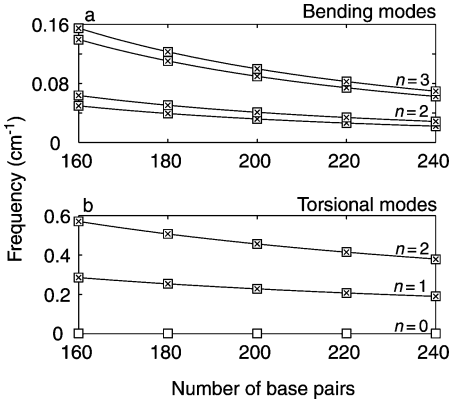


Figure 10. Normal-mode frequencies of (a) the two lowest ($n = 2, 3$) pairs of in-plane bending modes and the two lowest ($n = 2, 3$) pairs of out-of-plane bending modes and (b) the three lowest ($n = 0, 1, 2$) torsional modes, as a function of chain length, of a torsionally relaxed circular DNA, which is naturally straight in its equilibrium rest state and subject to an ideal elastic potential. The degeneracy of certain modes is evident from the computed frequencies, which are distinguished by the open boxes and cross symbols. The theoretically predicted chain length dependence of the frequencies, based on elastic rod theory,[14,15] is shown as solid lines. Note that the out-of-plane modes have higher frequencies than the in-plane modes of corresponding *n.*

modes (unbroken red lines) of the overtwisted ($\theta_3^u = 35.5°$) and undertwisted ($\theta_3^u = 36.5°$) molecules appear to coincide with those of the torsionally relaxed circle of the same length (200 bp), but the spacing between in-plane and out-of-plane frequencies (unbroken blue and broken green lines, respectively) differs. The in-plane bending motions of the supercoiled chains are consistently lower in frequency (energy) and the out-of-plane motions are consistently higher in value than the corresponding modes of the relaxed circle, with the differences becoming more pronounced as the frequencies of the modes increase. These trends match theoretical expectations as detailed below.

Mixed Bending Modes. Elastic rod theory predicts that pure in-plane or out-of-plane modes do not exist when the ideal rod is torsionally stressed. This is also true in our DNA model. Figure 9 illustrates the mixed nature of global structural fluctuations for two 200 bp chains with intrinsic Twist 35.5° and 36.5°. As is clear from Figure 9(a,b), the base-pair origins undergo small (±3 Å) out-of-plane displacements along the **b**° axes of the planar circle (dashed curves) as the molecule concomitantly deforms to elliptical shapes, and as evident from Figure 9(c,d), the chain fluctuates along the **n**° and **t**° axes in the plane of the ring (solid curves) as it also bends out-of-plane to cup-like shapes. Thus, the former mode is no longer a pure in-plane bending mode, and the latter is no longer a pure out-of-plane bending mode if the circular molecule is torsionally stressed.

Comparison of Figure 9 with the corresponding plots in Figure 6(a−c) for the torsionally relaxed circle shows that the normal modes of the supercoiled rings are combinations of pure in-plane and pure out-of-plane bending modes. For example, the low-frequency bending mode of the overtwisted DNA molecule in Figure 9(a), with intrinsic Twist 35.5°, is obtained by combining the in-plane mode of the relaxed circle from Figure 6(a), with the out-of-plane mode in Figure 6(c). Figure 9(b−d) are obtained in a similar fashion.

The variation of Tilt and Roll reflects this combination of pure in-plane and out-of-plane bending motions. The fluctuations of ΔTilt and ΔRoll in the lowest frequency bending modes of the over- and undertwisted DNA circles are a mixture of the change in step parameters that accompany pure in-plane and out-of-plane bending of the relaxed circle. The labels of the mixed bending modes are described by the more dominant of the two modes. Thus, the "in-plane" modes of the supercoiled circles have more pure in-plane than pure out-of-plane character, and the pure out-of-plane deformations outweigh the pure in-plane contributions to ΔTilt and ΔRoll in the out-of-plane modes.

Comparison with Theory. *Chain Length Dependence.* To study the dependence of the normal modes on chain length, we created five torsionally relaxed circular DNAs of different length—160, 180, 200, 220, 240 bp, and performed normal-mode analyses of each chain. Figure 10(a) shows the frequencies of the two lowest pairs of in-plane bending modes and the two lowest pairs of out-of-plane bending modes as a function of chain length. The degeneracy of the bending modes is evident from the superposition of computed frequencies, with one of each pair of modes denoted by an open box and the other by a cross symbol. The overlap of these values with the theoretically predicted frequencies[14,15] of an inextensible, naturally straight, homogeneous rod closed into a circle, shown by the solid curves in the figure, is remarkable (the numerical agreement is within 3×10^{-4} cm^{-1}).

As mentioned earlier, the theoretical frequencies are determined in units of $(4/r^3)(A/\pi\rho)^{1/2}$, where r is the radius of the elastic rod and ρ is its mass density. To make the comparison with the calculated values, we use a bending

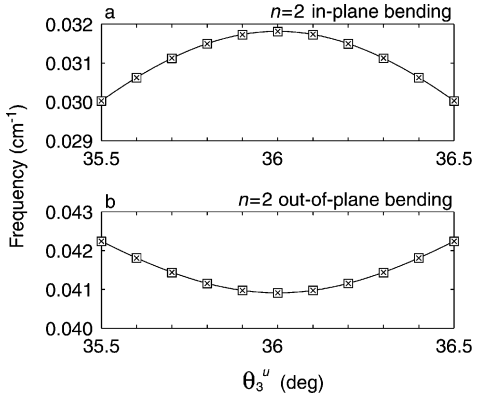


Figure 11. Normal-mode frequencies of the lowest ($n = 2$) pair of (a) in-plane and (b) out-of-plane bending modes, as a function of intrinsic Twist, θ_3^u, of a 200 bp circular DNA molecule, which is naturally straight in its equilibrium rest state. See legend to Figure 10.

rigidity A of 2.066×10^{-19} erg-cm deduced in Table 1 from the minimum energy circular state and set the radius r to 9.45 Å. This value of r is obtained by equating the inertial moment around the longitudinal axis of a straight rod of circular cross-section ($M r^2$) to the corresponding value for a B-DNA helix (calculated as a summation over all atoms of the DNA, $\Sigma m_i r_i^2$, where m_i is the mass of ith atom and r_i is the displacement of the atom from the helical axis). See ref 31 for the computational procedure used to locate the helical axis. A mass density of DNA of 0.065 g/mol-Å^3 is used. Finally, the circular frequency unit (1.17×10^{12} Hz) has been converted to wavenumbers, cm^{-1}.

The number n in Figure 10 is the index used in the theory to differentiate normal modes within the same class of motions (i.e., in-plane, out-of-plane, or torsional modes) and is equal to half the number of nodes found in the mode. As evident from the figure and from earlier discussion of the torsionally relaxed 200 bp circle, out-of-plane modes have higher frequencies than in-plane modes of corresponding n. The similar chain length dependence of the computed and theoretically predicted frequencies of the three lowest torsional modes is apparent in Figure 10(b). The agreement between numerical values is again quite good, despite the fact that the theoretically predicted frequency of the $n = 0$ torsional mode is zero and the computed values are small nonzero values.

Twist Dependence. As noted above, the DNA circle can be supercoiled by changing the value of the intrinsic Twist θ_3^u. For example, a change of the intrinsic Twist of a 200 bp DNA circle from the value of 36° found in the torsionally relaxed state to 35.5° introduces excess Twist in the molecule, $200 \times (36-35.5)° = 100°$. The latter value is one of the variables used in the analytical treatment of the normal-mode frequencies of a closed elastic ring.[14,15] The theoretically predicted frequencies of the lowest ($n = 2$) in-plane and out-of-plane bending modes are compared with values computed, respectively, for over- and undertwisted 200 bp circular DNA molecules in Figure 11(a,b). The agreement is again quite good, with the frequencies of the dominant in-plane mode reduced when the molecule is over- or undertwisted and the frequencies of the dominant out-of-plane mode increased under the same conditions.

In agreement with the theoretical method described in connection with eq 3 to determine the normal-mode frequencies, the torsional mode frequencies are nearly independent of excess twist. The respective frequencies of the $n = 1$ and $n = 2$ torsional modes of a 200 bp circle are predicted by the theory to be 0.2281 and 0.4561 cm^{-1}. The average values of the corresponding normal-mode frequencies of 11 different DNA circles, with intrinsic Twist varied from 35.5° to 36.5° at increments of 0.1°, are 0.2282 and 0.4562 cm^{-1}. Each of these modes is doubly degenerate, with some slight differences in the computed standard deviations of normal-mode frequencies: 2.82 vs 2.25×10^{-6} cm^{-1} and 4.60 vs 2.96×10^{-6} cm^{-1} for the respective modes. The very small deviations show that the computed frequencies are essentially independent of added Twist, in further agreement with predictions of elastic rod theory.

Conclusions

The motions of an ideal, closed circular DNA molecule obtained in the present normal-mode calculations show remarkable agreement with the theoretically predicted dynamic properties of an elastic rod. The energy-optimized closed circular states and the types of low frequency motions of a 200 bp molecule which is naturally straight in its relaxed linear state, inextensible, and capable of isotropic bending and twisting of adjacent base pairs, reproduce a wide variety of features, including (1) the uniformity of twist density and energy density in the minimum energy state; (2) the near zero frequency "free" torsional mode of the torsionally relaxed circle; (3) the higher-order torsional rocking modes and the pure in-plane and out-of-plane bending motions of the relaxed circle; (4) the mixed bending modes of the supercoiled chain; (5) the degeneracy of modes; and (6) the changes in normal-mode frequencies with variation of chain length and imposed supercoiling. The successful comparison validates extension of the computational procedure to study effects of specific base sequences on the supercoiled states and global motions of more realistically modeled DNA minicircles.

The curves fitted to the sequential variation of step parameters in both the minimum energy configuration and the different normal modes reveal the interplay between local conformational motions and global chain dynamics. The patterns of local conformational deformation responsible for the overall motions of the closed duplex persist in more realistically modeled DNA minicircles and, as illustrated in the companion article,[32] can be used to understand the effects of intrinsic curvature on large-scale global motions.

Acknowledgment. Support of this work through U.S.P.H.S. Grant GM34809 and the New Jersey Commission on Science and Technology (Center for Biomolecular Applications of Nanoscale Structures) is gratefully acknowledged. Computations were carried out at the Rutgers University Center for Computational Chemistry.

Supporting Information Available: Animation files of the normal modes of a 200 bp, torsionally relaxed, closed circular DNA molecule, which is naturally straight in its equilibrium rest state and subject to an ideal elastic potential. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Matsumoto, A.; Olson, W. K. Sequence-dependent motions of DNA: a normal mode analysis at the base-pair level. *Biophys. J.* **2002**, *83*, 22−41.

(2) Gō, N.; Noguti, T.; Nishikawa, T. Dynamics of a small protein in terms of low-frequency vibrational modes. *Proc. Natl. Acad. Sci., U.S.A.* **1983**, *80*, 3693−3700.

(3) Levitt, M.; Sander, C.; Stern, P. S. Protein normal-mode dynamics: trypsin inhibitor, crambin, ribonuclease and lysozyme. *J. Mol. Biol.* **1985**, *181*, 423−447.

(4) Nishikawa, T.; Gō, N. Normal modes of vibration in bovine pancreatic trypsin inhibitor and its mechanical property. *Proteins* **1987**, *2*, 308−329.

(5) Diekmann, S.; Dickerson, R. E.; Bansal, M.; Calladine, C. R.; Hunter, W. N.; Kennard, O.; Lavery, R.; Nelson, H. C. M.; Olson, W. K.; Saenger, W.; Shakked, Z.; Sklenar, H.; Soumpasis, D. M.; Tung, C.-S.; von Kitzing, E.; Wang, A. H.-J.; Zhurkin, V. B. Definitions and nomenclature of nucleic acid structure parameters. *J. Mol. Biol.* **1989**, *205*, 787−791.

(6) 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.

(7) Chandrasekaran, R.; Arnott, S. The structures of DNA and RNA helices in oriented fibers. In *Landolt-Börnstein Numerical Data and Functional Relationships in Science and Technology, Group VII/1b, Nucleic Acids*; Saenger, W., Ed.; Springer-Verlag: Berlin, 1989; pp 31−170.

(8) 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.

(9) 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.

(10) Kratky, O.; Porod, G. Röntgenuntersuchung Gelöster Fadenmoleküle. *Rec. Trav. Chim.* **1949**, *68*, 1106−1122.

(11) Shore, D.; Baldwin, R. L. Energetics of DNA twisting. I. Relation between twist and cyclization probability. *J. Mol. Biol.* **1983**, *170*, 957−981.

(12) 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.

(13) Frank-Kamenetskii, M. D.; Lukashin, A. V.; Anshelevich, V. V.; Vologodskii, A. V. Torsional and bending rigidity of the double helix from data on small DNA rings. *J. Biomol. Struct. Dynam.* **1985**, *2*, 1005−1012.

(14) Coleman, B. D.; Lembo, M.; Tobias, I. A new class of flexure-free torsional vibrations of annular rods. *Meccanica* **1996**, *31*, 565−575.

(15) Tobias, I. A theory of thermal fluctuations in DNA miniplasmids. *Biophys. J.* **1998**, *74*, 2545−2553.

(16) Song, L. U.; Schurr, J. M. Dynamic bending rigidity of DNA. *Biopolymers* **1990**, *30*, 229−237.

(17) Okonogi, T. M.; Alley, S. C.; Reese, A. W.; Hopkins, P. B.; Robinson, B. H. Sequence-dependent dynamics of duplex DNA: the applicability of a dinucleotide model. *Biophys. J.* **2002**, *83*, 3446−3459.

(18) Braun, W.; Yoshioki, S.; Gō, N. Formulation of static and dynamic conformational energy analysis of biopolymer systems consisting of two or more molecules. *J. Phys. Soc. Jpn.* **1984**, *53*, 3269−3275.

(19) Higo, J.; Seno, Y.; Gō, N. Formulation of static and dynamic conformational energy analysis of biopolymer systems consisting of two or more molecules—avoiding a singularity in the previous method. *J. Phys. Soc. Jpn.* **1985**, *54*, 4053−4058.

(20) Olson, W. K.; Bansal, M.; Burley, S. K.; Dickerson, R. E.; Gerstein, M.; Harvey, S. C.; Heinemann, U.; Lu, X.-J.; Neidle, S.; Shakked, Z.; Sklenar, H.; Suzuki, M.; Tung, C.-S.; Westhof, E.; Wolberger, C.; Berman, H. M. A standard reference frame for the description of nucleic acid base-pair geometry. *J. Mol. Biol.* **2001**, *313*, 229−237.

(21) Wang, J. C. Variation of the average rotation angle of the DNA helix and the superhelical turns of closed cyclic lambda DNA. *J. Mol. Biol.* **1969**, *43*, 25−39.

(22) Depew, R. E.; Wang, J. C. Conformational fluctuations of DNA helix. *Proc. Natl. Acad. Sci., U.S.A.* **1975**, *72*, 4275−4279.

(23) Anderson, P.; Bauer, W. Supercoiling in closed circular DNA: dependence upon ion type and concentration. *Biochemistry* **1978**, *17*, 594−600.

(24) Landau, L. D.; Lifshitz, L. M. *Theory of Elasticity*; Pergamon Press: Oxford, 1986; Chapter 2.

(25) Coleman, B. D.; Olson, W. K.; Swigon, D. Theory of sequence-dependent DNA elasticity. *J. Chem. Phys.* **2003**, *118*, 7127−7140.

(26) 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.

(27) Namoradze, N. Z.; Goryunov, A. N.; Birshtein, T. M. On conformations of the superhelix structure. *Biophys. Chem.* **1977**, *7*, 59−70.

(28) Levitt, M. How many base-pairs per turn does DNA have in solution and in chromatin? Some theoretical calculations. *Proc. Natl. Acad. Sci., U.S.A.* **1978**, *75*, 640−644.

(29) Sussman, J. L.; Trifonov, E. N. Possibility of nonkinked packing of DNA in chromatin. *Proc. Natl. Acad. Sci., U.S.A.* **1978**, *75*, 103−107.

(30) Tobias, I.; Coleman, B. D.; Lembo, M. A class of exact dynamical solutions in the elastic rod model of DNA with implications for the theory of fluctuations in the torsional

motion of plasmids. *J. Chem. Phys.* **1996**, *105*, 2517−2526.

(31) Matsumoto, A.; Gō, N. Dynamic properties of double-stranded DNA by normal mode analysis. *J. Chem. Phys.* **1999**, *110*, 11070−11075.

(32) Matsumoto, A.; Tobias, I.; Olson, W. K. Normal-mode analysis of circular DNA at the base-pair level. 2. Large-scale configurational transformation of a naturally curved molecule. **2005**, *1*, 130−142.

(33) Kraulis, P. J. MolScript: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **1991**, *24*, 946−950.

CT049950R

JCTC Journal of Chemical Theory and Computation

Normal-Mode Analysis of Circular DNA at the Base-Pair Level. 2. Large-Scale Configurational Transformation of a Naturally Curved Molecule

Atsushi Matsumoto,[†,‡] Irwin Tobias,[†] and Wilma K. Olson*[,†]

Department of Chemistry and Chemical Biology, Rutgers, The State University of New Jersey, Wright-Rieman Laboratories, 610 Taylor Road, Piscataway, New Jersey 08854-8087, and Quantum Bioinformatics Group, Center for Promotion of Computational Science and Engineering, Japan Atomic Energy Research Institute, 8-1 Umemidai, Kizu, Kyoto 619-0215, Japan

Received August 31, 2004

Abstract: Fine structural and energetic details embedded in the DNA base sequence, such as intrinsic curvature, are important to the packaging and processing of the genetic material. Here we investigate the internal dynamics of a 200 bp closed circular molecule with natural curvature using a newly developed normal-mode treatment of DNA in terms of neighboring base-pair "step" parameters. The intrinsic curvature of the DNA is described by a 10 bp repeating pattern of bending distortions at successive base-pair steps. We vary the degree of intrinsic curvature and the superhelical stress on the molecule and consider the normal-mode fluctuations of both the circle and the stable figure-8 configuration under conditions where the energies of the two states are similar. To extract the properties due solely to curvature, we ignore other important features of the double helix, such as the extensibility of the chain, the anisotropy of local bending, and the coupling of step parameters. We compare the computed normal modes of the curved DNA model with the corresponding dynamical features of a covalently closed duplex of the same chain length constructed from naturally straight DNA and with the theoretically predicted dynamical properties of a naturally circular, inextensible elastic rod, i.e., an O-ring. The cyclic molecules with intrinsic curvature are found to be more deformable under superhelical stress than rings formed from naturally straight DNA. As superhelical stress is accumulated in the DNA, the frequency, i.e., energy, of the dominant bending mode decreases in value, and if the imposed stress is sufficiently large, a global configurational rearrangement of the circle to the figure-8 form takes place. We combine energy minimization with normal-mode calculations of the two states to decipher the configurational pathway between the two states. We also describe and make use of a general analytical treatment of the thermal fluctuations of an elastic rod to characterize the motions of the minicircle as a whole from knowledge of the full set of normal modes. The remarkable agreement between computed and theoretically predicted values of the average deviation and dispersion of the writhe of the circular configuration adds to the reliability in the computational approach. Application of the new formalism to the computed modes of the figure-8 provides insights into macromolecular motions which are beyond the scope of current theoretical treatments.

Introduction

Although the average properties of polymeric DNA resemble those of an ideal elastic rod, the fine structure of the double helix carries a sequence-dependent structural and energetic code which helps to organize the overall folding of the long, threadlike molecule, and which also governs the susceptibility of DNA to interactions with other molecules. Individual base-pair steps adopt characteristic spatial forms and show different deformational tendencies in high-resolution structures.[1] These local turns and twists, if appropriately concat-

* Corresponding author phone: (732)445-3993; fax: (732)445-5958; e-mail: olson@rutchem.rutgers.edu.

[†] Rutgers, The State University of New Jersey.

[‡] Japan Atomic Energy Research Institute.

10.1021/ct049949s CCC: $30.25
Published on Web 12/21/2004

enated and then repeated in phase with the (~10 bp/turn) double helical repeat, introduce a natural "static" curvature or superhelicity in the DNA,[2−4] which in turn guides the spatial arrangements of longer molecules.[5−7]

Calculations that account for the natural curvature of DNA indicate that polymers with such features adopt completely different three-dimensional arrangements from an ideal, naturally straight elastic rod. For example, a naturally closed circular duplex is expected to take up ligand-induced superhelical stress through out-of-plane folding, gradually converting at natural levels of supercoiling into a 2-fold symmetric collapsed (clamshell-like) figure-8 configuration,[8−15] whereas a closed ideal rod retains a circular shape and snaps suddenly into a plectonemic structure at a characteristic level of supercoiling.[16−18]

The effect of curvature on the dynamical features of DNA, such as the retarded movement of naturally curved sequences on electrophoretic gels, is less well understood. Most modeling studies of the dynamics of naturally curved helices have focused to date on the variation of chemical fine structure extracted from all-atom simulations of the motions of short oligonucleotide duplexes.[19−24] Other work has addressed the bending, spinning, and tumbling of the molecule as a whole in the context of the physical manipulation of naturally curved elastic rods,[25] time-resolved electron microscopic images of single naturally curved molecules,[26] and selected spectroscopic properties of DNA chains containing curved fragments, e.g., electric dichroism decay and fluorescence depolarization of intercalated ethidium dyes in short, naturally curved sequences.[27,28]

Much less is known about the internal dynamics of supercoiled molecules with intrinsic curvature. The insertion of curved sequences in a naturally straight DNA is reported to reduce the internal motions that underlie the dynamic light scattering of supercoiled plasmids.[29] That is, the global configuration of the closed circular molecule is stiffened in the presence of curved DNA such that the likelihood of close approach between interacting fragments is increased and the slithering of individual residues past one another is decreased.[30] By contrast, sufficient increase in the intrinsic curvature of a closed circular molecule introduces a bimodality in the distribution of Monte Carlo simulated configurations of DNA.[31]

In this paper, as a next step in understanding the behavior of curved DNA, we investigate the internal dynamics of a covalently closed, naturally circular double helix. We compare the computed normal modes of such a molecule with the corresponding dynamical features of a cyclized duplex of the same chain length constructed from naturally straight DNA and with the theoretically predicted dynamical properties of a naturally circular, inextensible elastic rod. We present and make use of a general analytical treatment of the configurational fluctuations of an elastic rod. We examine mesoscopic pieces of DNA, fragments too long to investigate at the all-atom level and too complex to approximate as hinged objects, e.g., rigid rods connected by flexible joints. We vary the degree of intrinsic curvature and the superhelical stress on the DNA and consider the normal-mode fluctuations of both the circle and the stable figure-8 configuration under conditions where the energies of the two states are comparable. In this way we are able to decipher the low frequency modes and the enhancement in overall flexibility that underlie the large-scale rearrangement of the naturally curved molecule between the two configurations and gain new insight into the circle to figure-8 transition of supercoiled DNA.

Methods

Computational Treatment. We consider a chain which forms a closed minicircle in its equilibrium rest state. The rest state is defined by the base-pair step parameters identified in the companion paper[32] for the energy-minimized circular form of a DNA molecule which is naturally straight at equilibrium. The natural minicircle is thus described by a 10 bp repeating pattern of intrinsic local structure, with the bending components at each base-pair step, (θ_1^u, θ_2^u), equated to the values of Tilt° and Roll° along the contour of the minimized circular configuration. Values of Tilt and Roll used as references in the calculation of energy, i.e., θ_1^u and θ_2^u in eq *I*-1, where the *I* refers to the companion paper,[32] thus depend on chain length and imposed intrinsic Twist (see Table *I*-2). The treatment is applicable to lengths of DNA such that, when the molecule is constrained to be planar, the total twist of the ideal minicircle, $Tw = \Sigma\, \theta_3^\circ/360°$, is an integer. Here this constraint is satisfied by choosing θ_3° for the planar molecule to be 36° at all base-pair steps and n_B, the number of base-pair steps, to be 200. The normalized sum, which is equal to 20 is the linking number *Lk* of the closed ring, i.e., the number of times the two strands of the double helix wrap around one another. Values of the intrinsic Twist θ_3^u are assumed, however, to be independent of sequence and are assigned a range of values consistent with known environmentally induced changes, e.g., the dependence on temperature or ionic strength.[33−35] If there are no spatial constraints on the ends of the chain, the variation of θ_3^u to a value different from θ_3° converts the circular equilibrium structure to a helical configuration.[36−39] If the chain ends are covalently linked, the total increase or decrease of intrinsic Twist relative to the unligated structure, $\Delta Tw° = (\theta_3^\circ - \theta_3^u)n_B/360°$, imposes torsional stress on the naturally circular molecule. The excess twist in the closed circular configuration, $\Delta Tw°$, is equal to $-\tau^u L/2\pi$, where τ^u is the torsion of the helical pathway of the unlinked chain with intrinsic Twist θ_3^u and *L* is the length of the helical axis.[39−41] (The quantity $\Delta Tw°$, frequently called the linking number difference ΔLk, is the sum of the excess twist ΔTw and the writhing number *Wr*, or writhe for short, of any other configuration.) The molecule is assumed to be inextensible with the displacement of base pairs assigned values characteristic of B DNA, namely zero values of Slide and Shift and a Rise of 3.4 Å. The contour length *L* is therefore equal to 3.4 n_B Å.

The natural minicircle is subject to the same simplified elastic potential as the ideal, naturally straight, inextensible DNA molecule treated in the companion paper.[32] That is, the molecule bends isotropically at all base-pair steps, and the deformations of individual base-pair step parameters are independent of one another. Even though the equilibrium

structure of the DNA is a closed circle, restraints must be introduced in the normal-mode analysis to ensure that the chain termini are connected in the nonequilibrium states. Otherwise, the two ends of the double helix would fly apart as the chain undergoes conformational fluctuations. Thus, a restraint energy term like eq *I*-2 is included in the potential energy function, and an energy minimization step is carried out prior to normal-mode calculations.

Analytical Treatment. It was pointed out in paper *I*, in a discussion of some of the results of an analytical treatment of the normal modes of a circular DNA formed from a naturally straight elastic rod, that the normal-mode frequencies can be obtained by finding the roots of a polynomial cubic in the square of the frequencies. The same turns out to be the case for circular rings formed from intrinsically curved rods.[39] In this paper we compare the frequencies of some of the computed low-order modes of a naturally closed DNA minicircle with those determined from the analytical theory.

The ensemble average of various properties of a collection of identical elastic rods in thermal equilibrium at a temperature T can be extracted from the configuration integral

$$Z = \int e^{-E_\eta/k_B T}\, d\eta \tag{1}$$

an integral of $\exp[-E_\eta/k_B T]$ over all configurations, where E_η is the elastic energy of the rod in a configuration denoted by η and k_B is the Boltzmann constant. Since for the small rings we treat here, the only configurations which make a significant contribution to Z are those close to the equilibrium configuration, we first rewrite eq 1 in the form

$$Z = e^{-E^e/k_B T} z \tag{2}$$

where

$$z = \int e^{-(E_\eta - E^e)/k_B T} d\eta \tag{3}$$

and E^e is the elastic energy of the equilibrium configuration. To obtain information about the distribution of the writhe Wr for the ring-like molecules being considered here, it turns out, as we show below, that what is needed is the explicit dependence of $E_\eta - E^e$ on the topological invariant, the excess twist $\Delta Tw°$ characterizing the circular equilibrium configurations. Only the elastic twist energy contains a term that explicitly involves $\Delta Tw°$, namely, $(2\pi^2 C/L)(\Delta Tw° - Wr)^2$, where C is the twisting modulus. Therefore there is only one term in $E_\eta - E^e$ which contains $\Delta Tw°$ explicitly, $-(4\pi^2 C/L)\Delta Tw° \Delta Wr$, where ΔWr is the writhe associated with a configuration relative to that in the equilibrium configuration. Given the form of this expression, if the integration in eq 3 is now carried out over all configurations of a given writhe, z is of the form

$$z = \int_{-\infty}^{+\infty} e^{[4\pi^2 C/Lk_B T]\Delta Tw° \Delta Wr} F(Wr)\, dWr \tag{4}$$

where $F(Wr)$ is a function of the writhe alone.

The integrand in eq 4 represents the distribution function for the writhe. We see that $\langle \Delta Wr \rangle$, the ensemble average of the writhe relative to that in the equilibrium configuration is given by

$$\langle \Delta Wr \rangle = \left(\frac{Lk_B T}{4\pi^2 C}\right) \frac{\partial \ln z}{\partial \Delta Tw^o} \tag{5}$$

The variance of the writhe, $\langle Wr^2 \rangle - \langle Wr \rangle^2$, is obtained by differentiating ln z again, or, given eq 5, it can be written

$$\langle Wr^2 \rangle - \langle Wr \rangle^2 = \left(\frac{Lk_B T}{4\pi^2 C}\right) \frac{\partial \langle \Delta Wr \rangle}{\partial \Delta Tw^o} \tag{6}$$

For small elastic rings, the configuration integral is proportional to the high-temperature form of the partition function for a collection of harmonic oscillators having the frequencies $\omega_i(\Delta Tw°)$ of the normal modes of the elastic ring. That is, $z(\Delta Tw°)$ is proportional to the product

$$\prod_i \frac{k_B T}{\hbar \omega_i(\Delta Tw°)}$$

where $\hbar$ is Planck's constant divided by 2π. Knowing the dependence of the normal-mode frequencies on $\Delta Tw°$ is therefore sufficient for determining the average writhe and the variance of the writhe.

In a later section we also compare the average writhe, as given by eq 5, and the variance of the writhe, as given by eq 6, for the two approaches, the computational treatment and the analytical theory.

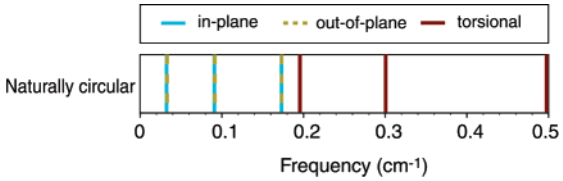


Figure 1. Color coded-spectrum of lowest frequency torsional (unbroken red line), in-plane (unbroken blue line), and out-of-plane (broken green line) modes of a 200 bp torsionally relaxed, inextensible, naturally circular DNA molecule subject to an ideal elastic force field.

Results and Discussion

Natural Minicircle. We start with a 200 bp DNA minicircle with an equilibrium Twist $\theta_3^u = \theta_3^\circ = 36°$, i.e., $Lk = 20$, and local intrinsic bending, given by the variation of Tilt° and Roll° in Table *I*-2, which naturally closes the chain into a circle. As evident from the color-coded spectrum of lowest frequency normal modes in Figure 1, the naturally circular molecule exhibits the same kinds of global motions as a straight chain with covalently linked ends in the torsionally relaxed state, namely in-plane and out-of-plane bending (unbroken blue and broken green lines, respectively) plus large-scale torsional movements of the polymer (red lines) about the circular helical axis. Unlike circles made up of naturally straight DNA, where the ease of in-plane and out-of-plane bending differs, the frequencies of in-plane and out-of-plane deformations of the closed naturally circular molecule are virtually identical. Moreover, these frequencies are roughly equivalent to the frequency of in-plane bending of a cyclized naturally straight chain (Figure *I*-5).

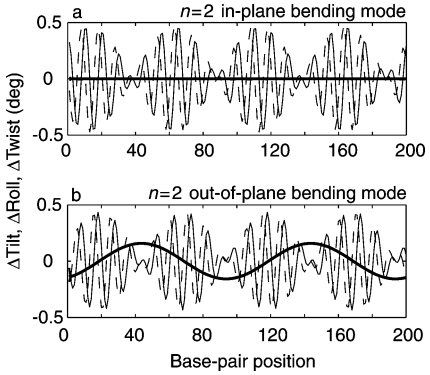


Figure 2. Fluctuations of local angular "step" parameters which are collectively responsible for selected normal modes of a 200 bp torsionally relaxed DNA which forms a natural minicircle in its equilibrium rest state and is subject to an ideal elastic potential: (a) one of the pair of lowest frequency ($n = 2$) in-plane bending modes and (b) one of the pair of lowest frequency ($n = 2$) out-of-plane bending modes. Plots illustrate the fluctuations of Tilt (thin solid lines), Roll (dashed lines), and Twist (thick solid lines) along the contour of the DNA molecule at the moment when the potential energy of the molecule is raised by $k_BT/2$; fluctuations are reversed a half cycle later of the mode.

The difference in the out-of-plane bending modes of naturally circular vs naturally straight DNA arises from a different pattern of local conformational motions. The fluctuations in base-pair step parameters (ΔTilt, ΔRoll, ΔTwist) which give rise to the lowest frequency in-plane and out-of-plane motions of a 200 bp covalently closed, naturally circular molecule, are reported in Figure 2. Comparison of these plots, which capture the local conformational distortions at the instant when the energy of the DNA is raised by $k_BT/2$, with those computed for the straight molecule closed into a circle (Figure *I*-7) reveals a notable difference in the twisting of the intrinsically curved chain. Whereas ΔTwist is close to zero for the out-of-plane deformations of straight DNA, it assumes nonzero values for the corresponding changes in the natural minicircle. By contrast, the patterns of fluctuations associated with the in-plane modes are similar for the two types of circular molecules.

The nonzero ΔTwist in the out-of-plane modes is a natural consequence of intrinsic curvature. Suppose we have a straight piece of DNA with unlinked ends and a planar, curved molecule, e.g., a fragment of naturally circular DNA, with the same contour length and we introduce the same amount of excess Twist at the central base-pair step in the two molecules. The straight DNA retains its original linear global shape, but the curved DNA responds to the imposed deformation through an out-of-plane configurational rearrangement. A DNA which is intrinsically more curved would undergo an even larger out-of-plane movement. The out-of-plane motions of intrinsically curved DNA can thus be effected by changes of Twist as well as by changes of Tilt and Roll, and a more curved piece of DNA can undergo out-of-plane motions more easily. Although the situation is complicated by the constraints of covalent bond formation in cyclic molecules, the involvement of nonzero ΔTwist persists in the out-of-plane modes of naturally circular DNA. Just as a linear molecule of greater curvature undergoes larger out-of-plane movements than a straighter fragment subject to the same amount of added Twist, a circular DNA molecule made up of highly curved pieces is expected to have lower out-of-plane bending frequencies and larger contributions from ΔTwist to the normal modes of bending than a cyclized molecule constructed from naturally straight DNA.

As is clear from comparison of Figures 1 and *I*-5, the torsional frequencies of the intrinsically curved molecule are much higher than those of a cyclized, naturally straight DNA molecule. Whereas the lowest frequency torsional mode of a naturally straight chain closed into a circle is very close to zero, the corresponding frequency of the natural minicircle is much higher. The same conformational mechanism, namely concerted changes in Tilt and Roll, which move base pairs from the inside to the outside of the circle and vice versa, effects global torsional movements in the two molecules. The deformations, however, place a greater conformational energy penalty on the natural minicircle than on cyclized straight DNA. The uniformity of equilibrium Tilt and Roll in the straight chain, $\theta_1^u = \theta_2^u = 0°$, gives rise to a residue-invariant contribution to the bending energy, that allows for all rotational orientations of base pairs and consequent "free rotation" of base pairs about the global helical axis. The corresponding shift of Roll and Tilt in the naturally circular molecule is energetically more costly than that in the straight chain. As a result, the naturally circular DNA has a higher barrier to large-scale helical rotation and higher ($n = 0$) torsional frequencies than a closed, intrinsically straight, ideal rod.

DNA Circles with Variable Intrinsic Curvature. We next consider a series of naturally curved molecules of varying intrinsic curvature κ^u, but all of a length corresponding to 200 bp and all planar ($\tau^u = 0$) in their undeformed open configuration. As discussed previously, when the condition of a uniform double helical repeat of 10 bp per turn is also satisfied, the closed, torsionally relaxed, circular molecule having a curvature $\kappa° = 2\pi/200\Delta s$, where (base-pair displacement) $\Delta s = 3.4$ Å, is in a minimum energy configuration. We report in Figure 3 the dependence of the frequencies of various kinds of global deformations on the value $q = \kappa^u/\kappa°$ for a series of such 200 bp DNA minicircles. (For a given value of the ratio *C*/*A* of the torsional and bending constants, there is a value of q, above which the circle is no longer stable.[39] In the present case, this occurs for a value of q somewhat greater than 2.) Note that, although the values of Tilt, Roll, and Twist of the minimum energy configurations of the minicircles are independent of q, the amplitudes of Roll and Tilt differ from that in the open undeformed configurations (see Table *I*-2) by an amount $(1 - q)\ 360°/200$. We can thus state that the elastic bending energy of the minicircles is proportional to $(1 - q)^2$. The ratio q can also be expressed in terms of the contour length $\tilde{L}$, measured in base pairs, for which the open molecule would form a complete circle, namely, $q = 200/\tilde{L}$.

The variation of the computed lowest frequency ($n = 0$) torsional mode in Figure 3(a) shows remarkable agreement with the theoretically predicted dependence on q.[41] The

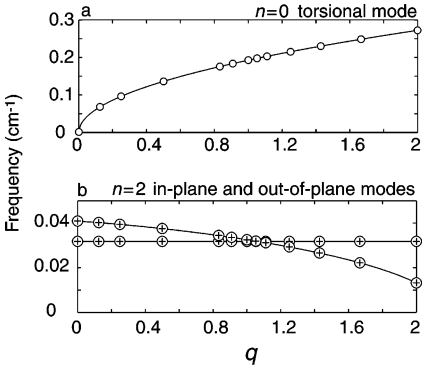


Figure 3. Normal-mode frequencies for (a) the $n = 0$ torsional and (b) the $n = 2$ in-plane and out-of-plane bending modes of naturally curved molecules which are closed into a chain of 200 bp. Data are reported as a function of the ratio $q = \kappa^u/\kappa^\circ$ of the intrinsic curvature κ^u to the curvature κ° of the minimum energy configuration of the natural minicircle. Computed values of the torsional mode frequencies (denoted by o) are compared with the theoretically predicted frequencies (shown by the smooth curve). The degeneracy of the in-plane and out-of-plane modes is evident from the computed frequencies, which are distinguished by o and + symbols and overlaid on the corresponding theoretically predicted curves.

numerical data (open circles) closely match the expected proportionality to $q^{1/2}$ (smooth curve). As observed in paper *I*, the lowest torsional frequency of a circle made from a naturally straight rod is zero. Here we see that this behavior follows from the null value of q. Moreover, the computed magnitude of the lowest torsional frequency of the naturally curved molecules is identical to the theoretically predicted value, e.g., a computed and predicted frequency of 0.19253 cm^{-1} for the natural $q = 1$ minicircle.

There is similar correspondence in Figure 3(b) between the computed and theoretically predicted frequencies of the lowest ($n = 2$) in-plane and out-of-plane bending modes of circular molecules with different degrees of intrinsic curvature. As noted above, the frequency of the out-of-plane bending mode is higher than that of the in-plane mode if the molecule is naturally straight ($q = 0$) but is of comparable magnitude if the DNA forms a natural minicircle ($q = 1$). The two modes are predicted by the theory to be identical in the present case when $q = 1.06$ and found by the calculations to be equivalent at approximately the same value. (An exact comparison is precluded by the limitations on chain length in the calculations, i.e., multiples of 10 bp.[32]) The ease of out-of-plane bending becomes greater than that of in-plane deformation, i.e., of lower frequency and lower energy, if q exceeds this threshold. That is, molecules which are more strongly bent, i.e., chains which cyclize into smaller rings than the natural minicircle, show a natural tendency to fluctuate out of the plane of the 200 bp circle. Indeed, when $q = 2$ and $C/A = 1$ and the molecule is closed into a circle two times the length of its equilibrium rest state, the barrier to out-of-plane deformations is removed, and the frequency of the mode is close to zero.[39,42] By contrast, the frequencies of the in-plane modes of the torsionally relaxed minicircle are predicted and found through computation to be constant over this range of q.

The difference in Twist fluctuations noted above for circles of naturally straight and naturally circular DNA also depends on the value of q. That is, when q is small, the amplitude of ΔTwist in the out-of-plane bending modes is small compared to that of either ΔTilt or ΔRoll, but as q increases in value, the amplitude of ΔTwist becomes comparable to the amplitudes of the local bending parameters (data not shown).

Properties of Supercoiled Molecules. *Figure-8 Minimum.* As with naturally straight DNA, superhelical stress can be introduced into the natural minicircle by changing the intrinsic Twist. If the change is sufficiently large, a global configurational rearrangement takes place, with the DNA adopting a figure-8 rather than a circular minimum energy state. This transition also occurs in closed molecules made up of naturally straight DNA,[17,18] but since there are no self-contact terms[43] in the present calculations, the figure-8 configuration of a naturally straight DNA is unstable and not found upon energy minimization. In the case of naturally circular DNA, energy minimization identifies a figure-8 minimum energy structure, which makes it possible to monitor details of the large-scale (circle to figure-8) spatial rearrangement.

Here we again study a 200 bp natural DNA minicircle subject to the same ideal elastic force field employed above, i.e., $A = 2.1 \times 10^{-19}$ erg-cm, $C = 2.9 \times 10^{-19}$ erg-cm. The equilibrium values of the base-pair step parameters are taken from the expressions for Tilt° and Roll° in Table *I*-2 or eq *I*-6, which close a naturally straight molecule of specified length and equilibrium Twist, θ_3°, into a circle. The transition to the figure-8 occurs when the intrinsic Twist θ_3^u differs by about $\pm 1.8°$ from θ_3°, changes which are equivalent to the introduction of $\pm 360°$ of additional twist into the DNA. By contrast, 1.25 additional helical turns are required to effect the interchange of stability between a closed circle and the figure-8 configuration of naturally straight DNA under the same elastic potential, i.e., a critical twist increment of $\pm\sqrt{3}\ A/C$ helical turns.[16−18] Figure 4 shows the minimum energy figure-8 structure obtained when the intrinsic Twist of the minicircle differs from θ_3° by $\pm 1.8°$, i.e., $\theta_3^u = 34.2°, 37.8°$. As is clear from the color coding in the figure, the Twist of individual base-pair steps is nonuniformly distributed along the two configurations. The uptake of Twist is concentrated in the center of the structures. The twisting of successive base pairs remains close to the 36° value characteristic of torsionally relaxed DNA in the 180° turns at the two (hairpin) ends of each structure. The slight difference in the respective writhes of the two figure-8's, +1.05 and −1.05, from the values (± 1) characteristic of the ideal planar forms reflect the finite radius of the DNA model (~10 Å). Details of the best-fit cosine functions, which describe the variation of base-pair step parameters along the minimum energy structures, are summarized in Table 1. The two dominant terms are presented. Comparison of these functions with those fitted to the minimum energy circular state adopted by the same chain (Table *I*-2) reveals an additional cosine term of wavelength of 11.1 bp or 9.1 bp, numbers corresponding respectively to 10/9 or 10/11 of the 10-fold helical repeat of the relaxed equilibrium state.

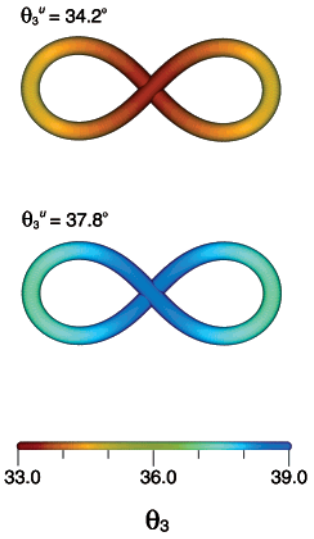


Figure 4. Computer-generated representation[52] of the minimum energy figure-8 configurations of a natural 200 bp DNA minicircle obtained by changing the intrinsic Twist θ_3^u by $\pm 1.8°$ from the equilibrium value θ_3^o in the torsionally relaxed state. The color coding depicts the value of Twist θ_3, in degrees, at consecutive base-pair steps along the two equilibrium structures.

Table 1. Base-Pair Step Parameters at the *m*th Dimer Step of an Ideal, Inextensible[a] Supercoiled DNA Circle of 200 bp in the Figure-8 Minimum Energy State

description	sequential conformational state
$\theta_3^u = 34.2°$ (figure-8 form)	
	Tilt° = 1.857 cos((360/11.1)(m + 0.553)) + 2.468 cos(36(m − 0.500))
	Roll° = 1.857 cos((360/11.1)(m + 0.553) + 90) + 2.468 cos(36(m − 0.500) + 90)
	Twist° = 34.108 + 0.671 cos((360/100)(m − 9.979))
$\theta_3^u = 34.25°$ (figure-8 form)	
	Tilt° = 1.864 cos((360/11.1)(m + 0.508)) + 2.460 cos(36(m − 0.500))
	Roll° = 1.864 cos((360/11.1)(m + 0.508) + 90) + 2.460 cos(36(m − 0.500) + 90)
	Twist° = 34.308 + 0.663 cos((360/100)(m − 9.571))
$\theta_3^u = 37.8°$ (figure-8 form)	
	Tilt° = 2.468 cos(36(m − 0.500)) + 1.857 cos((360/9.1)(m − 1.042))
	Roll° = 2.468 cos(36(m − 0.500) + 90) + 1.857 cos((360/9.1)(m − 1.042) + 90)
	Twist° = 37.891 − 0.670 cos((360/100)(m − 6.460))

[a] (Shift°, Slide°, Rise°) = (0 Å, 0 Å, 3.4 Å).

Bending Modes of the Torsionally Stressed Minicircle. Figure 5 reports the lowest bending frequencies of the natural minicircle as a function of intrinsic Twist θ_3^u. This figure is a counterpart to Figure *I*-11(a), obtained for a naturally straight elastic chain closed into a circle of the same length (200 bp). Here the range of imposed stress is extended beyond that presented in paper *I*. The computed bending modes are represented by discrete points and the predictions of theory by smooth curves. The superposition of symbols—open circles and plus signs for the natural minicircle, open boxes and cross symbols for the over- or undertwisted circle

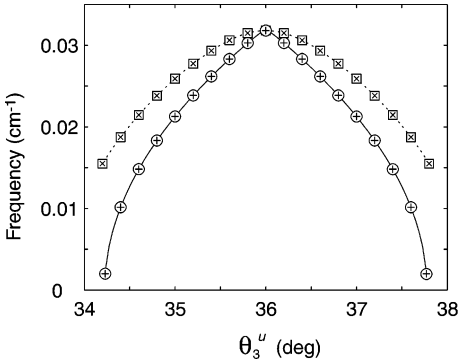


Figure 5. Lowest normal-mode frequencies of bending of a 200 bp natural DNA minicircle subject to an ideal elastic force field and the corresponding cyclic polymer made up of naturally straight DNA as a function of the intrinsic Twist θ_3^u. The degenerate frequencies obtained through computations are distinguished by symbols (o and + for the natural minicircle; open box and cross × for the closed, naturally straight DNA). The theoretically predicted values are represented by smooth curves (unbroken for the natural minicircle and broken for the circularized straight DNA).

made up of straight DNA—highlights the degeneracy of the configurational fluctuations, and the nearly perfect fit of these points to the smooth (respectively unbroken and broken) curves illustrates the remarkable agreement of computation and theory. The computed frequencies of the natural minicircle are limited to the range of θ_3^u within which and slightly beyond the limit where the circular form is lower in energy than the figure-8 configuration. The theoretical frequencies are reported for values of θ_3^u up to the point at which the predicted variation in $\langle \Delta Wr \rangle$, the average deviation of the writhe, becomes unphysical.

The primary difference in behavior between the two cyclized polymers lies in the much greater sensitivity of the normal-mode frequencies to changes in the intrinsic Twist in the natural minicircle compared to the cyclized polymer of naturally straight DNA. The decrease in the lowest frequency bending mode in the former chain is more than twice that found for a $\pm 0.5°$ increment of θ_3^u in the latter molecule. Furthermore, the frequency of deformation of the natural minicircle drops precipitously if the intrinsic Twist is changed slightly more, approaching a value of zero if θ_3^u is changed by $\pm 1.8°$, the same critical value associated with the interchange of the naturally circular and figure-8 minimum energy rest states. The very low frequency of these modes indicates that the energetic cost of deforming the over- or undertwisted circle into a different shape is negligible. The bending frequencies of the cyclized naturally straight DNA approach the same low-energy values when θ_3^u differs by $\pm 2.26°$ from the equilibrium state.

The nature of this large-scale configurational rearrangement is evident from the computed fluctuations in Figure 6 of individual base-pair origins with respect to (**n**°, **b**°, **t**°) Serret-Frenet coordinate frames embedded in each base pair of the energy-minimized, circular reference state. The plots show the displacements of individual residues in over- and undertwisted ($\theta_3^u = 34.25°$ and $37.75°$) natural minicircles at the moment when the potential energy of the DNA is

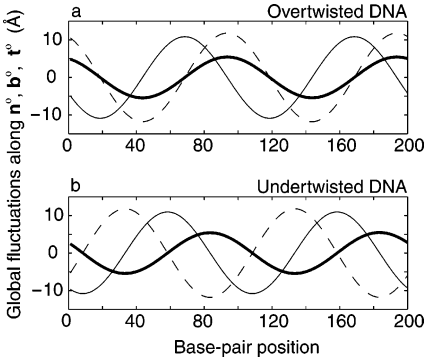


Figure 6. Displacement of the origins of base-pair axes at the moment when the potential energy is raised by $k_BT/50$ in the lowest frequency bending modes of (a) an overtwisted ($\theta_3^u = 34.25°$) or (b) an undertwisted ($\theta_3^u = 37.75°$) 200 bp natural DNA minicircle. Displacements (thin solid, dashed, and thick solid lines) measured respectively along the **n**°, **b**°, **t**° axes of Serret-Frenet frames embedded in each base pair of the minimum energy configuration.

raised by $k_BT/50$ along the lowest frequency mode. The very low-energy threshold in the example is a consequence of the very low frequency of the mode and the restriction of normal-mode analysis to conformational fluctuations in the vicinity of the minimum energy state. (Step parameters lie very far away from the reference state if the mode is assigned a higher energy.) As evident from the displacement along the **b**° axes (normals) of the planar circle (dashed curves), the global motion is no longer a pure in-plane bending mode upon supercoiling. The patterns of macromolecular displacement are quite similar, in terms of relative phase, to the deformations reported in Figure *I*-9 for over- and undertwisted closed circles made up of naturally straight DNA. For example, the largest moves along the **b**° and **t**° axes of the overtwisted natural minicircle in Figure 6(a) occur at the same positions as those of the overtwisted ideal DNA rod in Figure *I*-9(a), and the greatest changes along **n**° again appear 25 bp ahead of these points. The deformations of the natural minicircle, however, are much greater in magnitude than those of the naturally straight molecule under corresponding superhelical stress. The relative contribution of out-of-plane (**b**°-axis) motions also differs in the two molecule, i.e., greater displacements along **b**° and **n**° than along **t**° in the naturally closed molecule but more pronounced motions along **n**° than along either **b**° or **t**° in over- and undertwisted circles composed of naturally straight DNA.

Local Conformational Responses of the Stressed Minicircle. The fluctuations of local step parameters responsible for the lowest frequency mixed bending modes of the over- or undertwisted natural minicircle are summarized by a series of best-fit cosine functions in Table 2. The expressions describe the sequential variation of ΔTilt, ΔRoll, and ΔTwist of one of the degenerate modes for selected values of intrinsic Twist at the instant when the energy is raised by $k_BT/50$. The conformational patterns of the other of the degenerate modes are related by a phase shift of 90° in the fitted cosine functions. In all cases, the sequential variation in dimer bending is described by a sum of cosine functions, one with wavelength 11.1 bp and the other with wavelength 9.1 bp.

Table 2. Fluctuations of Base-Pair Step Parameters at the *m*th Dimer Step of an Ideal, Inextensible,[a] Naturally Circular, Supercoiled DNA Circle of 200 bp in the Lowest Frequency In-Plane Bending Mode

description	sequential conformational distortions
$\theta_3^u = 34.25°$ (overtwisted)	
	$\Delta\text{Tilt} = -0.522\cos((360/11.1)(m + 1.401)) - 0.0215\cos((360/9.1)(m - 2.055))$
	$\Delta\text{Roll} = -0.522\,(\cos((360/11.1)(m + 1.401) + 90) - 0.0215\cos((360/9.1)(m - 2.055) + 90)$
	$\Delta\text{Twist} = -0.179\cos((360/100)(m - 17.610))$
$\theta_3^u = 36°$ (torsionally relaxed)	
	$\Delta\text{Tilt} = -0.0472\cos((360/11.1)(m + 0.193)) - 0.0473\cos((360/9.1)(m + 1.652))$
	$\Delta\text{Roll} = -0.0472\cos((360/11.1)(m + 0.193) + 90) - 0.0473\cos((360/9.1)(m + 1.652) + 90)$
	$\Delta\text{Twist} = 0.000$
$\theta_3^u = 37.75°$ (undertwisted)	
	$\Delta\text{Tilt} = -0.0215\cos((360/11.1)(m + 0.268)) - 0.522\cos((360/9.1)(m - 1.128))$
	$\Delta\text{Roll} = -0.0215\cos((360/11.1)(m + 0.268) + 90) - 0.522\cos((360/9.1)(m - 1.128) + 90)$
	$\Delta\text{Twist} = +0.179\cos((360/100)(m - 7.412))$

[a] (Shift°, Slide°, Rise°) = (0 Å, 0 Å, 3.4 Å).

If the minicircle is torsionally relaxed, the amplitudes of the two terms are roughly equivalent. The function characterized by wavelength 11.1 bp dominates if the intrinsic Twist is decreased to 34.25°, and the function characterized by wavelength 9.1 dominates if the intrinsic Twist is increased to 37.75°. The amplitude of ΔTilt or ΔRoll is fairly large in both cases, increasing by more than 0.5° at some base-pair steps. If Tilt and Roll vary independently of one another, there is, by definition, a 90° phase shift in the terms used to describe the sequential variation of ΔTilt and ΔRoll.

The sequential variation of base-pair step parameters in the minimum energy figure-8 structures of the natural minicircles with intrinsic Twist varied by ±1.8° from the ($\theta_3^\circ = 36°$) equilibrium state (Table 1) bears a remarkable resemblance to the computed fluctuations of local variables in the lowest frequency mixed bending mode of the over- and undertwisted circles (Table 2). As noted above, an additional cosine term of wavelength 11.1 or 9.1 bp appears in the functions fitted to the sequential variation of Tilt and Roll along the optimized figure-8 structures and a term of the same period dominates the bending modes of the natural minicircle with $\theta_3^u = 34.25°$ or 37.75°. Given that no other normal modes of the circle show a comparable decrease in frequency and energy with imposed supercoiling, it is highly likely that these modes guide the transition pathway between the circular and figure-8 forms (see below).

Normal Modes of the Figure-8. Normal-mode analysis of the stable figure-8 minimum of the same ($\theta_3^u = 34.25°$) overtwisted DNA molecule yields a complementary picture of configurational deformation. The lowest (nearly zero) frequency motion of the figure-8 is a slithering motion of the duplex which has no effect on overall macromolecular shape, i.e., the point of chain self-contact simply translocates freely along the molecular contour. The fluctuations of local step parameters responsible for the slithering mode are described by the best-fit cosine functions in Table 3. The expressions, which contain only the dominant contribution

Table 3. Fluctuations of Base-Pair Step Parameters at the *m*th Dimer Step of an Ideal, Inextensible,[a] Naturally Circular, Supercoiled DNA of 200 bp about the Minimum Energy Figure-8 Configuration in the Two Lowest Frequency Bending Modes

description	sequential conformational distortions
$\theta_3^u = 34.25°$ (mode 1, slithering[b])	
	$\Delta\text{Tilt} = 0.185\cos((360/11.1)(m + 0.508) - 90)$
	$\Delta\text{Roll} = 0.185\cos((360/11.1)(m + 0.508))$
	$\Delta\text{Twist} = -0.066\cos((360/100)(m - 9.571) - 90)$
$\theta_3^u = 34.25°$ (mode 2, bending)	
	$\Delta\text{Tilt} = 0.076\cos(36(m - 0.500)) - 0.098\cos((360/9.1)(m - 1.325))$
	$\Delta\text{Roll} = 0.076\cos(36(m - 0.500) + 90) - 0.098\cos((360/9.1)(m - 1.325) + 90)$
	$\Delta\text{Twist} = -0.638 + 0.034\cos((360/100)(m - 9.671))$

[a] (Shift°, Slide°, Rise°) = (0 Å, 0 Å, 3.4 Å). [b] Parametric values when the energy is raised by $k_BT/(2 \times 10^4)$.

for each base-pair step parameter at the moment when the energy is raised by $k_BT/(2 \times 10^4)$, closely resemble the sequential variation of base-pair step parameters along the figure-8 minimum energy state (Table 1). Specifically, the expressions for the fluctuations in Roll and Tilt are obtained by shifting the phase by 90° and reducing the amplitude of the cosine functions with wavelength 11.1 bp. Such variation of parameters is reminiscent of the local conformational changes found in the lowest frequency, "free" torsional mode of a naturally straight DNA closed into a circle, where a corresponding change in phase results in the movement of base-pair steps from the inside to the outer surface of the molecule, and vice versa.[32] In the case of the lowest frequency slithering mode of the figure-8, the local conformational changes move the sites of maximum and minimum bending, located respectively at the two tips and the central crossing points of the figure-8, back and forth along the chain contour. The second lowest frequency motion of the figure-8 is a mixed bending mode which assists in opening the collapsed, self-contacted structure to the circular form (see below). As evident from Table 3, where the fluctuations of local step parameters needed to raise the energy by $k_BT/2$ are reported as fitted trigonometric functions, the residue-invariant increase (or decrease a half cycle later) of ΔTwist dominates this mode.

Thermal Fluctuations. The global motions associated with a collection of thermally fluctuating 200 bp minicircles are illustrated in Figure 7. Equation 5 was used to compute $\langle\Delta Wr\rangle$, the average deviation of the writhe from that in the equilibrium configuration, and eq 6 to compute the square root of the variance of the writhe $(\langle Wr^2\rangle - \langle Wr\rangle^2)^{1/2}$, both as a function of torsional stress measured by $\Delta Tw° = (\theta_3° - \theta_3^u) \times 200/360°$. The open circles and the ×'s give the results of the computations for naturally curved and intrinsically straight DNA, respectively. In both of these cases, the equilibrium configuration is circular, and the results closely match the predictions of the analytical theory given by the solid and broken curves. In the case of the filled-in circles, the equilibrium configuration is figure-8 like, a case to which the analytical theory has not been applied. The

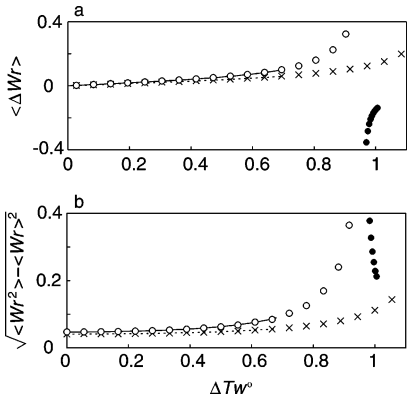


Figure 7. Variation of the average deviation of the writhe $\langle\Delta Wr\rangle$ from that in the circular equilibrium configuration and the square root of the variance of the writhe, $\sqrt{\langle Wr^2\rangle - \langle Wr\rangle^2}$ vs the total imposed twist, $\Delta Tw°$, of a closed 200 bp naturally circular DNA (open circles), a circular chain of the same length constructed of naturally straight DNA (× symbols), and the figure-8 configuration adopted by the naturally circular molecule (filled-in circles). The theoretically predicted behavior of the two circular configurations is represented respectively by broken and unbroken curves over the range in which the theory is valid.

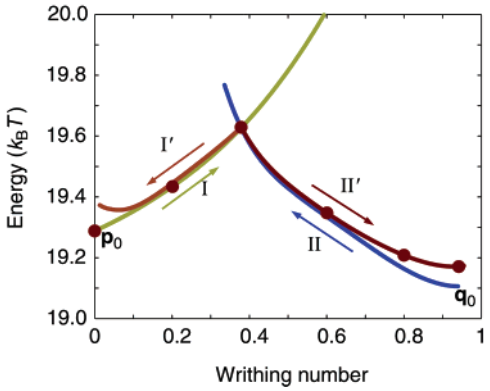


Figure 8. The variation of energy versus writhe of an overtwisted ($\theta_3^u = 34.25°$), naturally circular, 200 bp DNA molecule perturbed along its lowest frequency modes to transient configurational states intermediate between the minimum energy circular and figure-8 forms. The large dots correspond to states illustrated in Figure 10. (See text for details of transition pathways.)

excellent agreement between the present calculations and the analytical theory is evident in the graphs, as is the difference in properties of the rings formed from curved DNA as opposed to straight. The data show the increased flexibility of the natural minicircles as compared with circles formed from intrinsically straight chains for all values of $\Delta Tw°$. The greater sensitivity of the writhe-altering fluctuations of the natural minicircles to increasing torsional stress is also evident.

Pathways of Large-Scale Configurational Rearrangement. *Excursions of the Circle*. Figure 8 reports the variation in both energy and writhe of an overtwisted ($\theta_3^u = 34.25°$), naturally curved DNA minicircle perturbed along its lowest frequency normal modes to transient configurational states intermediate between the (circular and figure-8) minimum

energy configurations found to be stable under these conditions. Path I corresponds to deformations of the circle along the lowest frequency (mixed bending) mode detailed in Table 2 and Path II to the corresponding normal-mode bending distortions of the figure-8 (see below). A writhe of zero corresponds to the circle and a value of unity to an ideal, planar figure-8 configuration. The energy cost of large-scale configurational rearrangements (monitored by the writhe) between the circle and figure-8 is quite small.

The intermediate configurational states in Figure 8 are obtained by recursive introduction of small normal-mode distortions of base-pair step parameters followed by rapid energy minimization. Each configuration of DNA is described by 1200 parameters (6 rigid-body parameters per base-pair step × 200 base-pair steps). The initial (minimum energy) configuration is defined by a 1200 dimensional vector $\mathbf{p}_0$, with elements corresponding to the sequential variation of step parameters around the circle, and is deformed to $\mathbf{p}_0 + \alpha\Delta\mathbf{p}_0$, where α is a constant and the displacement vector $\Delta\mathbf{p}_0$ is the normalized lowest frequency normal-mode vector of the circle. A short run of (conjugate gradient) energy minimization is then carried out to avoid high energy states. The minimization is stopped when the decrease of energy per iteration is small ($<5 \times 10^{-5}\ k_BT$). More thorough energy minimization would return the configuration to the initial state $\mathbf{p}_0$. The constant α is chosen to be small enough so that the number of iterations per minimization cycle is at most 5. The new configuration $\mathbf{p}_1$ is then deformed to $\mathbf{p}_1 + \alpha\Delta\mathbf{p}_1$, where the normalized displacement vector $\Delta\mathbf{p}_1$ is defined by the configurational change from the initial structure, $(\mathbf{p}_1 - \mathbf{p}_0)/|\mathbf{p}_1 - \mathbf{p}_0|$. Energy minimization is performed as in the preceding step, and the new configuration $\mathbf{p}_2$ is obtained. This process is repeated, so that in the kth repetition, configuration $\mathbf{p}_k$ is deformed to $\mathbf{p}_k + \alpha\Delta\mathbf{p}_k$, where $\Delta\mathbf{p}_k$ is defined by the direction of the last configurational move, $(\mathbf{p}_k - \mathbf{p}_{k-1})/|\mathbf{p}_k - \mathbf{p}_{k-1}|$. Energy minimization follows, and the new configuration $\mathbf{p}_{k+1}$ is obtained. In this way, the series of configurations $\mathbf{p}_1 \ldots \mathbf{p}_K$ along Path I is generated, the energy of which is plotted versus the writhe in Figure 8. As should be clear from the above description, the displacement vector $\Delta\mathbf{p}_k$ changes as the configuration of the molecule changes. The correlation of the displacement vectors along Path I with the initial displacement vector $\Delta\mathbf{p}_0$ is reported in Figure 9, Curve *a*. The correlation is greatest at the start of the transformation and decreases approximately linearly with the increase in writhe, i.e., departure from the equilibrium reference state.

Miyashita et al.[44] have recently reported an analogous global transformation of a protein using an elastic network model of amino acid interactions. They use a scheme much like ours to generate intermediate conformational states along the transition pathway between open and closed forms of the molecule. The displacement vector $\Delta\mathbf{p}_k$ ($k > 0$) used to generate successive intermediate states, however, is a combination of the low-frequency normal modes of the current conformation $\mathbf{p}_k$. In contrast to the elastic treatment of DNA, where intermediate conformational states are not minimum energy structures, all molecular states can be regarded as minima in the elastic network model. Thus, we

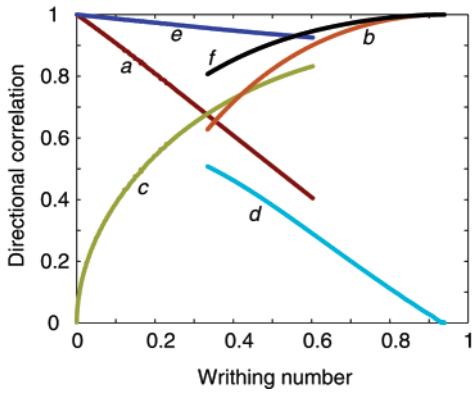


Figure 9. Correlations, plotted against the writhe, of the displacement vectors $\Delta\mathbf{p}_k$ and $\Delta\mathbf{q}_l$ of intermediate configurational states with the normal (bending) mode vectors $\mathbf{p}_0$ and $\mathbf{q}_0$ of the minimum energy circular and figure-8 forms of the overtwisted DNA molecule described in Figure 8: (a) $\Delta\mathbf{p}_k \cdot \Delta\mathbf{p}_0$; (b) $\Delta\mathbf{q}_l \cdot \Delta\mathbf{q}_0$; (c) $-\Delta\mathbf{p}_k \cdot \Delta\mathbf{q}_0$; (d) $-\Delta\mathbf{q}_l \cdot \Delta\mathbf{p}_0$, and the corresponding contributions from the two normal-mode vectors ($\Delta\mathbf{p}_0$ and $\Delta\mathbf{q}_0$) to the displacement vectors ($\Delta\mathbf{p}_k$ and $\Delta\mathbf{q}_l$); (e) $((\Delta\mathbf{p}_k \cdot \Delta\mathbf{p}_0)^2 + (\Delta\mathbf{p}_k \cdot \Delta\mathbf{q}_0)^2)^{1/2}$; (f) $((\Delta\mathbf{q}_l \cdot \Delta\mathbf{p}_0)^2 + (\Delta\mathbf{q}_l \cdot \Delta\mathbf{q}_0)^2)^{1/2}$.

cannot perform normal mode calculations at each stage of conformational transformation and, instead, move from state to state using the aforementioned iterative minimization procedure. The fact that the transition pathway can be described by a small number of normal modes suggests the possibility of identifying a smooth, realistic conformational pathway with a low-energy barrier using more sophisticated approaches, such as path integral techniques.[45]

Excursions of the Figure-8. The second lowest frequency mixed bending mode of the figure-8 is responsible for the large-scale configurational rearrangement needed to open the collapsed, self-contacted structure to the circular form. The series of configurations $\mathbf{q}_1 \ldots \mathbf{q}_L$ along Path II in Figure 8 is generated, starting from the energy-minimized figure-8 configuration $\mathbf{q}_0$ (writhe = 0.94) and the normalized initial displacement vector $\Delta\mathbf{q}_0$ associated with the second lowest frequency bending mode. As with the deformed states of the circle, the correlation of the displacement vector $\Delta\mathbf{q}_l$ with the initial displacement vector $\Delta\mathbf{q}_0$ decreases as the configuration of the molecule changes from its original state, Curve *b* in Figure 9.

Interestingly, the displacement vector $\Delta\mathbf{p}_k$ on Path I describing perturbations of the circle is correlated with the initial displacement vector $\Delta\mathbf{q}_0$ of the figure-8. The correlation $\Delta\mathbf{p}_k \cdot (-\Delta\mathbf{q}_0)$ is plotted as Curve *c* in Figure 9. Initially at the minimum energy (circular) configuration (where $k = 0$), the correlation is close to zero, indicating that the two vectors ($\Delta\mathbf{p}_0$ and $\Delta\mathbf{q}_0$) are almost perpendicular to each other. The correlation increases as the writhe increases. The increase of the correlation complements the decrease of the correlation $\Delta\mathbf{p}_k \cdot \Delta\mathbf{p}_0$ (Curve *a*). Indeed, the contribution from the two directions $((\Delta\mathbf{p}_k \cdot \Delta\mathbf{p}_0)^2 + (\Delta\mathbf{p}_k \cdot \Delta\mathbf{q}_0)^2)^{1/2}$, which is plotted as Curve *e* in Figure 9, is close to unity, indicating that these two normal-mode vectors play dominant roles in the circle to figure-8 transition (Path I). The series of displacement vectors $\Delta\mathbf{q}_l$ along Path II is similarly correlated with $\Delta\mathbf{p}_0$. The correlation $\Delta\mathbf{q}_l \cdot (-\Delta\mathbf{p}_0)$ is plotted

as Curve *d* in Figure 9. In this case also, the contribution from the two directions $((\Delta\mathbf{q}_l \cdot \Delta\mathbf{p}_0)^2 + (\Delta\mathbf{q}_l \cdot \Delta\mathbf{q}_0)^2)^{1/2}$, Curve *f* in Figure 9, is close to unity, indicating that the two normal-mode vectors play dominant roles in the reverse (figure-8 to circle) transformation (Path II).

Intermediate States. Although the writhe is very effectively changed if the DNA is deformed along Paths I and II in Figure 8, the minimum energy configurations of the figure-8 and circular forms ($\mathbf{q}_0$ and $\mathbf{p}_0$, respectively) cannot be reached or approached from the opposing minimum. The direction of the displacement vectors $\Delta\mathbf{p}_n$ and $\Delta\mathbf{q}_m$ must be changed discontinuously at some point to approach the opposite minimum energy states. Paths I′ and II′ in Figure 8 are obtained by such changes. Configurations $\mathbf{p}_{k'}$ and $\mathbf{q}_{l'}$ are the closest points on Paths I and II, respectively, to their point of intersection, differing from one another by a root-mean-square distance of 5.6 Å. Path II′ is obtained by using $\mathbf{p}_{k'}$ as the starting configuration and the $-\Delta\mathbf{q}_{l'}$ as the initial displacement vector. Path I′ is similarly obtained from $\mathbf{q}_{l'}$ and $-\Delta\mathbf{p}_{k'}$. The differences between Paths I and I′ near $Wr = 0$ and those between Paths II and II′ near $Wr = 1$ stem primarily from insufficient energy minimization in the generation of intermediate configurations. The minimum energy configurations $\mathbf{p}_0$ and $\mathbf{q}_0$, where normal-mode calculations are carried out, can be reached only by thorough energy optimization, e.g., Newton−Raphson minimization.

The continuous transformation of the circle to the figure-8 along Paths I and II′ is illustrated in Figure 10 and in the Supporting Information. Starting from the circular configuration, the deformation proceeds primarily via the change of Tilt and Roll (see the expressions in Table 2 for the variation of step parameters in the lowest frequency bending mode of the circle and the dominant role of the mode shown by Curve *a* in Figure 9). A residue-invariant decrease of Twist, which releases the excess Twist in the molecule, gradually comes into play (see Curve *c* in Figure 9 and the expressions in Table 3 for the variation of step parameters in the lowest frequency bending mode of the figure-8). The residue-invariant decrease of Twist becomes dominant when the configuration nears the intersection point in Figure 8, i.e., intermediate state, and further configurational rearrangement to the figure-8 proceeds almost exclusively through the release of excess Twist (noted by the color coding). Although the distribution of Twist along the contour of the intermediates is nonuniform, the incremental changes in overall twist, *Tw*, between successive configurational states is uniform, reflecting the regular increments of the writhe and the well-known invariance of the linking number ($Lk = Wr + Tw$).[46] The configurations adopted in the reverse transition from the figure-8 to the circle along Paths II and I′ are very similar to those shown in Figure 10. The energy barrier is slightly lower, ~0.3 k_BT or ~0.2 kcal/mol at 300 K, if the molecule follows Paths I and II′ from the circle to the figure-8, rather than the reverse Paths (II and I′) from the figure-8 to the circle, where the barrier is ~0.5 k_BT or ~0.3 kcal/mol at 300 K. It should be noted that explicit treatment of chain self-contact could change the normal-mode frequencies as well as the minimum energy of the figure-8 from the values computed here in the absence of such a correction.

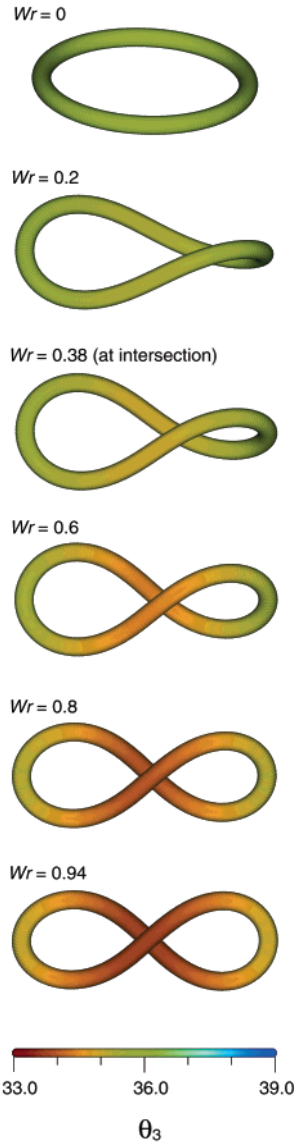


Figure 10. Computer-generated snapshots[52] of the configurational pathway between circular and figure-8 configurations obtained by deforming an overtwisted ($\theta_3^u = 34.25°$), naturally circular DNA minicircle along paths I and II′ in Figure 8. The color-coding of Twist is identical to that in Figure 4.

Discussion

The minicircles studied in this work are comparable in length and degree of supercoiling to the DNA loops which are formed by various regulatory proteins and enzymes that bind in tandem to sequentially distant parts of the long chain molecule.[47,48] The influence of natural curvature on the global motions of the minicircles found here can thus provide insight into how DNA loops of several hundred base pairs might respond to changes in nucleotide sequence. The sequence of base pairs in such loops determines the degree of intrinsic curvature of the spatially constrained molecule.[2−4]

Here we find that covalently closed DNA duplexes with natural curvature are torsionally stiffer but, when placed

under superhelical stress, are capable of greater bending deformations than minicircles which are made up of naturally straight DNA. The degree of curvature changes the character of global bending, i.e., the relative frequencies of in-plane vs out-of-plane deformations. Whereas a covalently closed, naturally straight duplex distorts more easily via in-plane than out-of-plane bending deformations, a natural minicircle is just as likely to bend via either route and a chain, which is curved more tightly than the natural minicircle, preferentially deforms out of the plane of the circle (Figure 3).

Whereas the naturally straight DNA rotates freely about its global helical axis, there is a barrier impeding large-scale helical twisting of curved DNA. In the absence of intrinsic curvature, no single orientation of the closed duplex is preferred over any other, and all sites are expected to be equally accessible to a ligand, such as DNase I, which preferentially contacts the (outer) convex surface of its DNA target.[49,50] By contrast, the introduction of natural curvature is predicted to restrict rotation of the DNA as a whole about its helical axis, thereby favoring the minimum energy configuration and limiting enzymatic access to residues located on the inside of the ring. The enzymatic cleavage pattern of a naturally closed DNA minicircle is thus expected to include regularly spaced sites of enhanced cutting alternating every half helical turn with sites of suppressed cutting. Other types of naturally closed molecules, such as curved DNA molecules generated from alternating fragments of naturally straight and naturally rolled base-pair steps,[14,15] are expected to exhibit the same global properties.

The barrier opposing global bending of the natural minicircle lowers significantly when the molecule is over- or undertwisted. The frequency, i.e., energy, of global bending decreases in value upon supercoiling (Figure 5), and if the imposed stress is sufficiently large, global configurational rearrangement of the circle to the figure-8 form takes place. Because the bending frequencies of the natural minicircle are much more sensitive to changes in intrinsic Twist than are those of the cyclized naturally straight polymer, the large-scale configurational interchange occurs more easily in the molecule with intrinsic curvature. The dominant ($n = 2$) modes of the two kinds of molecules, however, are of similar mixed bending character, i.e., the base pairs move out of the plane of the circle as the supercoiled molecule concomitantly deforms to elliptical shapes. The mechanism of conformational transformation between circle and figure-8 is thus expected to be similar in the two types of DNA.

Intermediate states constructed from the computed structures and dominant bending modes of the two minimum energy forms (Figures 8 and 10) suggest that the circle to figure-8 transformation involves two distinct types of conformational rearrangement. Localized changes in bending (Tilt and Roll) initially dominate as the circle deforms to an elongated, nonplanar intermediate state, and subsequent transformation to the figure-8 minimum proceeds via the uptake of twisting (Tables 2 and 3 and Figure 10). The twist density is nonuniform in both the intermediate states and the stable figure-8 minimum, with the imposed stress taken up preferentially at the sites of closest interstrand contact in the straighter central parts of the structures (Figure 4).

The range of low-energy states identified on the basis of the dominant normal modes of the circle and figure-8 is consistent with the mixture of spatial forms found in previous Monte Carlo simulations of a much longer (486 bp) DNA circle under superhelical stress.[51] The ($\sim$0.4 k_BT) potential barrier between the two states at the midpoint of the transition of the 200 bp natural minicircle is remarkably similar to the free energy (0.2 k_BT) reported previously for the longer, naturally straight DNA, with slightly different elastic constants and under the influence of a screened Coulombic potential. The present study tracks the lowest energy pathway of interconversion between circular and figure-8 configurations via the dominant thermal fluctuations of the two minimum energy states, whereas the Monte Carlo findings are based on the characteristics of a broad, random sample of configurational states. The treatment of normal modes provides mechanistic insights into configurational rearrangements which cannot be gleaned from Monte Carlo and other stochastic approaches.

Statistical mechanical considerations make it possible to characterize the large-scale motions in terms of the full set of normal-mode frequencies of covalently closed DNA molecules and can be applied to either computed or theoretically predicted modes (Figure 7). The thermal fluctuations in global structure are described in terms of the average deviation and variance of the writhe. The intrinsic, out-of-plane response of curved DNA to imposed torsional stress underlies its greater global deformability compared to a naturally straight molecule. The writhe, a measure of the chiral distortions from planarity of a closed curve, is sensitive to the intrinsic conformational response of curved DNA to imposed twist. The uniform twisting of base pairs along a curved, unligated molecule results in a helical configuration, the handedness and proportions of which depend respectively on the sign and magnitude of imposed twist. The same type of local deformations of an open piece of straight DNA merely reorients the bases at either end of the molecule without change of global shape. The covalent closure of the ends of the naturally curved duplex suppresses the torsionally induced configurational response of the linear molecule and converts the preferred helical configuration to the out-of-plane bending modes which dominate the global fluctuations of the closed polymer. The localized twisting of adjacent residues at a single site along a curved DNA similarly produces a chiral arc. The binding of an untwisting agent to a natural minicircle is therefore expected to enhance the global motions of a DNA minicircle by a similar mechanism, converting the end-to-end separation of the bound linear form into an out-of-plane bending mode in the closed molecule.

Finally, the remarkable agreement between the computed and theoretically predicted dependence of the normal modes of naturally curved DNA on the degree of curvature and torsional stress and the identical descriptions of the global motions of circular molecules add to the reliability of the normal-mode analysis of DNA at the base-pair level and increase confidence in the computed dynamic properties of

configurations such as the figure-8 which are beyond the scope of current theory.

Acknowledgment. Support of this work through U.S.P.H.S. Grant GM34809 and the New Jersey Commission on Science and Technology (Center for Biomolecular Applications of Nanoscale Structures) is gratefully acknowledged. Computations were carried out at the Rutgers University Center for Computational Chemistry.

Supporting Information Available: Animation files of the normal modes of a 200 bp DNA minicircle, which is naturally circular in its equilibrium rest state, governed by an ideal elastic potential, and subjected to torsional stress. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) 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.

(2) Trifonov, E. N. DNA in profile. *Trends Biochem. Sci.* **1991**, *16*, 467−470.

(3) 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.

(4) Hagerman, P. J. Straightening out the bends in curved DNA. *Biochim. Biophys. Acta* **1992**, *1131*, 125−132.

(5) Laundon, C. H.; Griffith, J. D. Curved helix segments can uniquely orient the topology of supertwisted DNA. *Cell* **1988**, *52*, 545−549.

(6) Yang, Y.; Westcott, T. P.; Pedersen, S. C.; Tobias, I.; Olson, W. K. The effect of sequence-directed bending on DNA supercoiling. *Trends Biochem. Sci.* **1995**, *20*, 313−319.

(7) Chirico, G.; Langowski, J. Brownian dynamics simulations of supercoiled DNA with bent sequences. *Biophys. J.* **1996**, *71*, 955−971.

(8) Bauer, W. R.; Lund, R. A.; White, J. H. Twist and writhe of a DNA loop containing intrinsic bends. *Proc. Natl. Acad. Sci., U.S.A.* **1993**, *90*, 833−837.

(9) White, J. H.; Lund, R. A.; Bauer, W. R. Twist, writhe, and geometry of a DNA loop containing equally spaced coplanar bends. *Biopolymers* **1996**, *38*, 235−250.

(10) Charitat, T.; Fourcade, B. Metastability of a circular O-ring due to intrinsic curvature. *Eur. Phys. J. B* **1998**, *1*, 333−336.

(11) Olson, W. K. DNA higher-order structures. In *Oxford Handbook of Nucleic Acid Structure*; Neidle, S., Ed.; Oxford University Press: Oxford, U.K., 1999; pp 499−531.

(12) White, J. H.; Lund, R. A.; Bauer, W. R. Effect of salt-dependent stiffness on the conformation of a stressed DNA loop containing initially coplanar bends. *Biopolymers* **1999**, *49*, 605−619.

(13) Garrivier, D.; Fourcade, B. Twisting DNA with variable intrinsic curvature. *Europhys. Lett.* **2000**, *49*, 390−395.

(14) Coleman, B. D.; Olson, W. K.; Swigon, D. Theory of sequence-dependent DNA elasticity. *J. Chem. Phys.* **2003**, *118*, 7127−7140.

(15) 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.

(16) Zajac, E. E. Stability of two planar loop elasticas. *J. Appl. Mech. Trans. ASME Ser. E* **1962**, *29*, 136−142.

(17) Le Bret, M. Catastrophic variation of twist and writhing of circular DNAs with constraint? *Biopolymers* **1979**, *18*, 1709−1725.

(18) Benham, C. J. The onset of writhing in circular elastic polymers. *Phys. Rev. A* **1989**, *39*, 2582−2586.

(19) Young, M. A.; Srinivasan, J.; Goljer, I.; Kumar, S.; Beveridge, D. L.; Bolton, P. H. Structure determination and analysis of local bending in an A-tract DNA duplex: comparison of results from crystallography, nuclear magnetic resonance, and molecular dynamics simulation on d(CGCAAAAATGCG). *Methods Enzymol.* **1995**, *261*, 121−144.

(20) Sherer, E. C.; Harris, S. A.; Soliva, R.; Orozco, M.; Laughton, C. A. Molecular dynamics studies of DNA A-tract structure and flexibility. *J. Am. Chem. Soc.* **1999**, *121*, 5981−5991.

(21) Sprous, D.; Young, M. A.; Beveridge, D. L. Molecular dynamics studies of axis bending in $d(G_5\text{-}(GA_4T_4C)_2\text{-}C_5)$ and $d(G_5\text{-}(GT_4A_4C)_2C_5)$: effects of sequence polarity on DNA curvature. *J. Mol. Biol.* **1999**, *285*, 1623−1632.

(22) Strahs, D.; Schlick, T. A-tract bending: insights into experimental structures by computational models. *J. Mol. Biol.* **2000**, *301*, 643−663.

(23) McConnell, K. J.; Beveridge, D. L. Molecular dynamics simulations of B′-DNA: sequence effects on A-tract-induced bending and flexibility. *J. Mol. Biol.* **2001**, *314*, 23−40.

(24) Mazur, A. K.; Kamashev, D. E. Comparative bending dynamics in DNA with and without regularly repeated adenine tracts. *Phys. Rev. E* **2002**, *66*, art. no. 011917.

(25) Koehler, S. A.; Powers, T. R. Twirling elastica: kinks, viscous drag, and torsional stress. *Phys. Rev. Lett.* **2000**, *85*, 4827−4830.

(26) Scipioni, A.; Zuccheri, G.; Anselmi, C.; Bergia, A.; Samori, B.; DeSantis, P. Sequence-dependent DNA dynamics by scanning force microscopy time-resolved imaging. *Chem. Biol.* **2002**, *9*, 1315−1321.

(27) Porschke, D.; Schmidt, E. R.; Hankeln, T.; Nolte, G.; Antosiewicz, J. Structure and dynamics of curved DNA fragments in solution: evidence for slow modes of configurational transitions. *Biophys. Chem.* **1993**, *47*, 179−191.

(28) Chirico, G.; Collini, M.; Toth, K.; Brun, N.; Langowski, J. Rotational dynamics of curved DNA fragments studied by fluorescence polarization anisotropy. *Eur. Biophys. J.* **2001**, *29*, 597−606.

(29) Kremer, W.; Klenin, K.; Diekmann, S.; Langowski, J. DNA curvature influences the internal motions of supercoiled DNA. *EMBO J.* **1993**, *12*, 4407−4412.

(30) Klenin, K. V.; Frank-Kamenetskii, M. D.; Langowski, J. Modulation of intramolecular interactions in superhelical DNA by curved sequences. A Monte Carlo simulation study. *Biophys. J.* **1995**, *68*, 81−88.

(31) Katritch, V.; Vologodskii, A. The effect of intrinsic curvature on conformational properties of circular DNA. *Biophys. J.* **1997**, *72*, 1070−1079.

(32) Matsumoto, A.; Tobias, I.; Olson, W. K. Normal-mode analysis of circular DNA at the base-pair level. 1. Comparison of computed motions with the predicted behavior of an ideal elastic rod. **2005**, *1*, 117−129.

(33) Wang, J. C. Variation of the average rotation angle of the DNA helix and the superhelical turns of closed cyclic lambda DNA. *J. Mol. Biol.* **1969**, *43*, 25−39.

(34) Depew, R. E.; Wang, J. C. Conformational fluctuations of DNA helix. *Proc. Natl. Acad. Sci., U.S.A.* **1975**, *72*, 4275−4279.

(35) Anderson, P.; Bauer, W. Supercoiling in closed circular DNA: dependence upon ion type and concentration. *Biochemistry* **1978**, *17*, 594−600.

(36) 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.

(37) Tobias, I.; Olson, W. K. The effect of intrinsic curvature on supercoiling − predictions of elasticity theory. *Biopolymers* **1993**, *33*, 639−646.

(38) Dubochet, H.; Bednar, J.; Furrer, P.; Stasiak, A. Z.; Stasiak, A. Determination of the DNA helical repeat by cryo-electron microscopy. *Struct. Biol.* **1994**, *1*, 361−363.

(39) Tobias, I. Thermal fluctuations of small rings of intrinsically helical DNA treated like an elastic rod. *Philos. Trans. R. Soc.* **2004**, *362*, 1387−1402.

(40) Tobias, I.; Coleman, B. D.; Lembo, M. A class of exact dynamical solutions in the elastic rod model of DNA with implications for the theory of fluctuations in the torsional motion of plasmids. *J. Chem. Phys.* **1996**, *105*, 2517−2526.

(41) Tobias, I. A theory of thermal fluctuations in DNA miniplasmids. *Biophys. J.* **1998**, *74*, 2545−2553.

(42) Manning, R. S.; Hoffman, K. A. Stability of *n*-covered circles for elastic rods with constant planar intrinsic curvature. *J. Elasticity* **2001**, *62*, 1−23.

(43) Westcott, T. P.; Tobias, I.; Olson, W. K. Modeling self-contact forces in the elastic theory of DNA supercoiling. *J. Chem. Phys.* **1997**, *107*, 3967−3980.

(44) 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.

(45) Tomimoto, M.; Gō, N. Analytic theory of pseudorotation in five-membered rings. Cyclopentane, tetrahydrofuran, ribose, and deoxyribose. *J. Phys. Chem.* **1995**, *99*, 563−577.

(46) White, J. H. Self-linking and the Gauss integral in higher dimensions. *Am. J. Math.* **1969**, *91*, 693−728.

(47) Adhya, S. Multipartite genetic control elements: communication by DNA loop. *Annu. Rev. Genet.* **1989**, *23*, 227−250.

(48) Schleif, R. DNA looping. *Annu. Rev. Biochem.* **1992**, *61*, 199−223.

(49) Suck, D.; Oefner, C. Structure of DNase I at 2.0 Å resolution suggests a mechanism for binding to and cutting DNA. *Nature* **1986**, *321*, 620−625.

(50) Lahm, A.; Suck, D. DNase I-induced DNA conformation. 2 Å structure of a DNase I-octamer complex. *J. Mol. Biol.* **1991**, *221*, 645−667.

(51) Gebe, J. A.; Schurr, J. M. Thermodynamics of the first transition in writhe of a small circular DNA by Monte Carlo simulation. *Biopolymers* **1995**, *38*, 493−503.

(52) Kraulis, P. J. MolScript: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **1991**, *24*, 946−950.

CT049949S

JCTC Journal of Chemical Theory and Computation

Rapid Estimation of Solvation Energy for Simulations of Protein−Protein Association

David S. Cerutti,*[,§] Lynn F. Ten Eyck,[‡,§,||] and J. Andrew McCammon[†,§,||]

Department of Chemistry and Biochemistry, Howard Hughes Medical Institute, Department of Pharmacology, and San Diego Supercomputer Center, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0365

Received September 3, 2004

Abstract: We have formulated the Energy by Linear Superposition of Corrections Approximation (ELSCA) for estimating the electrostatic and apolar solvation energy of bringing two proteins into close proximity or into contact as defined by the linearized Poisson−Boltzmann model and a linear function of the solvent-accessible surface area. ELSCA utilizes potentials of mean force between atom types found in the AMBER *ff*99 force field, a uniform distance-dependent dielectric, and a potential that mimics the change in solvent accessible surface area for bringing two solvated spheres into contact. ELSCA was trained by a linear least-squares fit on more than 39 000 putative complexes, each formed from pairs of nonhomologous proteins with a range of shapes, sizes, and charges. The training set was also designed to capture various stages of complex formation. ELSCA was tested against over 8000 non-native complexes of 45 enzyme/inhibitor, antibody/antigen, and other systems that are known to form complexes and gives an overall correlation of 0.962 with PBSA-derived energies for these complexes. The predictions have a slope of 0.89 on the actual values with a bias of 11.1 kcal/mol. When applied to native complexes of these 45 protein systems, ELSCA reproduces PBSA results with a correlation of 0.787, a slope of 1.13, and a bias of 13.0 kcal/mol. We report parameters for ELSCA in the context of the AMBER *ff*99 parameter set. Our model is most useful in macromolecular docking and protein association simulations, where large portions of each molecule may be considered rigid.

Introduction

The protein docking problem, most generally the challenge of finding the structure of a complex of two proteins given the three-dimensional structures of the isolated components, is difficult due to the degree of sampling required and the limitations of scoring functions. In particular, while gas-phase molecular mechanics energies are readily computed by docking programs, the change in solvation energy involved in forming a particular complex is not rigorously treated by most scoring functions presently in use. Although implicit-solvent methods such as Poisson−Boltzmann/Surface Area (PBSA) and Generalized Born/Surface Area (GBSA) are widely used for calculating solvation energies during complex formation,[3,24,37] they are applicable for screening at most a few thousand putative complexes after lower-resolution searches. If the solvation energy involved in forming a complex between two proteins were formulated as a sum of pairwise additive interactions between atoms of the ligand and receptor, this quantity could be readily computed along with the gas-phase molecular mechanics energies that most real-space docking programs already employ, abrogating the need for refinement of the solutions produced and extending implicit-solvent methodology for protein docking and efficient simulations of macromolecular encounters.

* Corresponding author phone: (858)534-2798; e-mail: dcerutti@mccammon.ucsd.edu.
† Howard Hughes Medical Institute.
‡ San Diego Supercomputer Center.
§ Department of Chemistry and Biochemistry.
|| Department of Pharmacology.

10.1021/ct049946f CCC: $30.25
Published on Web 11/24/2004

In implicit solvent methodology, the solvation free energy is typically approximated as the sum of electrostatic and nonpolar contributions.[38] Numerous studies have demonstrated the importance of electrostatics in biomolecular recognition.[2,6,7,21,43,49] In solvated systems, electrostatic forces may act over long ranges to steer protein partners and accelerate the rate of complex formation,[20,22] affect the structure of highly charged nucleic acids,[1] and order the ions around those structures.[11] Favorable electrostatic interactions exist between polar or charged regions of biomolecules and polar water molecules. During the final stages of the association of biomolecules, stripping away nearby water molecules makes an unfavorable contribution to the binding free energy and may even outweigh salt bridges and favorable dipole interactions formed by the association.[40] To whatever degree hydrophobic and apolar forces drive the final stages of complex formation, electrostatic effects may still help to specify interactions. In contrast, nonpolar interactions comprise the work of cavity formation and the dispersion forces acting between solute and solvent.[38] Breaking solute−solvent dispersion interactions during complex formation is energetically unfavorable, but this is compensated by burial of solvent-accessible surface area which reduces the work required to form a solvent cavity for the proteins.

Accurate estimates of the electrostatic interaction energy of two solvated biomolecules at long range can be calculated accurately and cheaply so as to predict parameters such as relative association kinetics through Brownian dynamics studies.[20,23] However, at close range, accurate electrostatic energies are much more expensive to calculate due to desolvation effects and the disparity between the polarity of water and the apolar residues in the cores of proteins. This is a major obstacle to accurate prediction of absolute association rates of biomolecules and the structures of their complexes. The Poisson−Boltzmann approximation, a merger of Poisson's equation for macroscopic electrostatics and a mean-field treatment of the spectator ions, has been used extensively for analyzing the electrostatics of macromolecules, particularly the association of proteins,[18,20,40] but this computation typically requires several minutes on a modern workstation. Approximations to Poisson-derived hydration energies such as Generalized Born parametrizations[17,28] are faster, but are still expensive for lengthy macromolecular simulations and far too expensive for general searches in protein docking applications.

The nonpolar solvation energy of proteins is commonly formulated as a linear function of the solvent-accessible surface area, based on observations of the solvation energy of linear alkanes.[42] While this rule does not generalize perfectly to small molecules of other shapes,[52] nor to proteins,[29] most Poisson−Boltzmann electrostatic models are benchmarked alongside such models (see ref 41), for example) and so require this term for computation of total solvation energy. While calculating the solvent-accessible surface area of a molecule with N atoms can be written as an O(Nlog$N$) problem,[39] the number of computations per atom is still large, and the amount of coding needed to rigorously implement this calculation in a docking or simulation program is considerable. Square-well potentials[33] and other rough measures of shape complementarity[12] are therefore used.

For computational convenience, scoring functions presently used in protein docking applications neglect the most physically meaningful aspects of solvation: the fact that buried atoms in the ligand oppose electrostatic fields much less effectively than polar water molecules, the solute−solvent dispersion interactions, and the change in the work of cavity formation upon binding. At long ranges, these effects are negligible, but protein docking requires millions to billions of accurate estimates for closely interacting proteins, and simulations of macromolecular encounters require accurate energy estimates for all degrees of separation.

In the present work we introduce the Energy by Linear Superposition of Corrections Approximation (ELSCA), a correction scheme based on a distance-dependent dielectric (DDD) that is uniform (e.g. its form is the same regardless of the local environment), a scalable function describing buried surface area between two interacting spheres, and a set of potentials of mean force (PMFs) between distinct types of atoms. We compute solvation energies using PBSA parameters for modeling proteins with the AMBER *ff*99 parameter set[14,49] for more than 39 000 putative complexes and close associations of 21 distinct proteins with differing charges, shapes, and sizes. For convenience of implementation, ELSCA is fit to reproduce the change in solvation energy plus the gas-phase Coulombic association energy of two proteins. To demonstrate ELSCA's transferability, we apply it to an independent set of 45 native and over 8000 non-native protein complexes. All functions used in this model can be conveniently superimposed on grid-based potentials or lookup tables, implying no additional computational cost during docking studies and simulations. The purpose of ELSCA is most like that of a set of potentials of mean force developed by Jiang and co-workers[26] for the total free energy change of bringing two proteins together in a particular conformation. However, ELSCA should also be considered in the context of other pairwise potentials for calculating the solvation energy of individual species such as EEF1[27] and atomic contact energies.[54]

Methods

Approximation of the Total Electrostatic and Nonpolar Solvation Energy. To break Poisson−Boltzmann electrostatics and change in solvent-accessible surface area into the pairwise interactions of ELSCA, we make use of four approximations. The first and most basic approximation, a screened Coulombic interaction, stems directly from Debye−Hückel theory[34] and is the dominant contribution for interactions at distances greater than 60 Å. The second approximation, a screened Coulombic interaction attenuated by a uniform distance-dependent dielectric (DDD), affects intermediate and short-range interactions. The third, a set of pair potentials (PMFs) between sixteen atom types, is effective at distances less than 15 Å. The fourth, a function that roughly reproduces the change in solvent accessible surface area as two spheres of given radii approach one another, is

effective at distances on the order of the diameter of the solvent probe. The first approximation has no adjustable parameters, but the second and third have a total of 413 (see below) and the fourth is scaled by a single parameter. All parameters are fit simultaneously by solving a linear least-squares problem. The total electrostatic and solvation energy is the sum of energies from all four approximations.

At large distances, the electrostatic interaction of two charges in a neutral solvent is given by Coulomb's law scaled by the appropriate solvent dielectric constant. In dilute ionic solutions, however, the long-range electrostatic potential is screened exponentially as described by Debye−Hückel theory. We account for this explicitly using [1], where i and j run over the atoms in the ligand and receptor, k runs over the two ionic species in the continuum solvent, q represents an atomic partial charge, β is the inverse product of Boltzmann's constant and the absolute temperature, and N is the number density of an ionic species in bulk solvent.

$$E_{\text{basic}} = \sum_{i,j} \frac{q_i q_j}{4\pi\epsilon_0\epsilon_{\text{water}} r_{ij}} \exp\left[- \sqrt{\sum_k \frac{\beta q_k^2 N_k}{\epsilon_0\epsilon_{\text{water}}}} r_{ij} \right] \quad (1)$$

Distance-dependent dielectrics are often computationally convenient, whether realistic or not. In the docking problem, a uniform DDD is particularly convenient as the energy remains a sum of pairwise interactions. We formulate the second part of our method using a set of Gaussians to attenuate the screened Coulombic potential as shown in [2].

$$E_{\text{DDD}} = E_{\text{basic}} \left[1 + \sum_{i,j} \left(\sum_{\alpha=1}^{5} \exp\left(- \frac{0.02}{\alpha^2} r_{ij}^2 \right) S_{\text{DDD}}^{(\alpha)} \right) \right] \quad (2)$$

The value of 0.02 in the exponential argument was chosen along with five Gaussian terms to provide a basis for finely tuning the dielectric constant at distances between 0 and 60 Å, where the interacting ligand and receptor are no longer adequately described as collections of Debye−Hückel spheres.

The electrostatic energy of closely associated proteins is not adequately described by pairwise charge−charge interactions between the solutes due to desolvation effects and the fact that protein interiors dampen electrostatic fields much less effectively than bulk water. To account for these effects while still using pairwise interactions, we parametrize a set of $M(M+1)/2$ PMFs for M atom types. Each PMF is in turn a linear combination of k_{PMF} basis functions $g(r)$ [3], giving $k_{\text{PMF}}M(M+1)/2$ scaling parameters that must be solved for. The correction to the electrostatic energy E_{corr} is given by [3 and 4]:

$$f_{np}(r) = \sum_{\alpha=1}^{k_{\text{PMF}}} g_\alpha(r) S_{\text{np}}^{(\alpha)} \quad (3)$$

$$E_{\text{corr}} = \sum_{i=1}^{L} \sum_{j=1}^{R} \sum_{n=1}^{M-1} \sum_{p=n+1}^{M} \delta_{T_i n} \delta_{T_j p} f_{np}(r_{ij}) \quad (4)$$

Above, n and p are involved in summations over all atom types, T_i and T_j represent the types of the ith and jth atoms, respectively, $S_{np}^{(\alpha)}$ represents the αth scaling parameter of the PMF between atoms of types n and p to be solved for, and δ is the Kronecker delta. The total electrostatic estimate is thus a linear function of the scaling parameters $S_{np}^{(\alpha)}$ and $S_{\text{DDD}}^{(\alpha)}$. With a number of putative complexes generating at least $(k_{\text{PMF}}M(M+1)/2) + 6$ sets of distinct coefficients, the scaling parameters may be obtained by solving a least squares problem $\mathbf{A}\boldsymbol{s} = \boldsymbol{b}$. Each row of the matrix $\mathbf{A}$ is a set of coefficients for each $S_{\text{DDD}}^{(\alpha)}$ and $S_{np}^{(\alpha)}$ obtained from one putative complex. $\boldsymbol{s}$ represents the vector of scaling parameters for the DDD and PMFs, and $\boldsymbol{b}$ represents the vector of total electrostatic and nonpolar solvation energies. Operationally, each row of matrix $\mathbf{A}$ is filled by first initializing it to zero and then looping over atoms j of one molecule nested within a loop over atoms i of the other. In the inner loop, the value of each relevant basis function given the distance r_{ij} is added to the appropriate column of the row of matrix $\mathbf{A}$.

The PMF basis functions $g(r)$ must be well-behaved in that $g(r)$ has finite values at $r = r_0$ and at $r = r_{\text{cut}}$, where r_{cut} is the cutoff distance for applying the PMF (set equal to the cutoff for vdW interactions in our model) and $r_0 < r_{\text{cut}}$. Because the electrostatic interactions between receptor and ligand at long ranges are approximated well by treating the ligand as a set of point charges moving within the electrostatic field of the receptor, it is preferable to have PMFs, and therefore to select basis functions, that diminish as r approaches r_{cut}. Piecewise Gaussians of the form 5 were chosen.

$$g_\alpha(r) = \begin{cases} \exp\left(- \dfrac{\beta_{\text{PMF}}}{(\alpha + \lambda_{\text{PMF}})^2} (r - \sigma_{np})^2 \right), & r \geq \sigma_{np} \\ 1, & r < \sigma_{np} \end{cases} \quad (5)$$

In [5], σ_{np} is the Lennard-Jones parameter for atom types n and p except in cases where n or p is HO (see Table 1), in which case σ_{np} is zero.

For constructing the PMFs, three criteria were considered in determining the number of basis functions k_{PMF} in [4] and

Table 1. Atom Types for ELSCA Parameters[a]

ELSCA name	AMBER *ff*99 names
CT	CT
C	C
CR	CA, CC, CV, CW, CR, CB, C*, CN
N	N, NA, NB, N2, N3
O	O
O2	O2
OH	OH
S	S, SH
H	H, HS
HA	HA
HP	HP
HC	HC
HO	HO
H1	H1
H4	H4
H5	H5

[a] The atom types defined for ELSCA comprise all those found in amino acids in AMBER *ff*99. Distinct Lennard-Jones parameters as well as considerations to broad classes of chemical groups led to these choices.

spread of basis functions as dictated by the β_{PMF} and λ_{PMF} in [5]: the RMSD of estimated energies compared to those derived by PBSA, the similarity in correlations obtained in the training set versus those in the test set, and the range of values taken by the PMFs as r approached σ_{np}. The third criterion was included to ensure that no particular pairwise interaction could drastically affect the estimated score, based on our observation that ELSCA could become particularly sensitive to atoms in close contact if given large values of β_{PMF} in conjunction with a small λ_{PMF}.

Distinct atom types for defining the PMFs were first identified by distinct Lennard-Jones parameters within AMBER *ff*99 (see *Methods*, *Protein Docking*). Distinctions were also made between *sp2* hybridized carbon atoms in aromatic groups, *sp2* hybridized carbon atoms in carbonyl groups, oxygen atoms in amides or carbonyl groups, and oxygen atoms in carboxylic acids. These distinctions help to distinguish between grossly different electronic environments but do not permit ELSCA to parametrize extremely rare atom types, such as carbons in imidazole rings, which would lead to overfitting. The AMBER *ff*99 atom types, and the corresponding names we give them, are found in Table 1. In total, these 16 atom types provided 136 PMFs for fitting.

Finally, at very short distances, the change in solvent-accessible surface area was approximated by a piecewise function that closely reproduces the change in solvent-accessible surface area when two spheres of radii R_i and R_j ($R_i > R_j$) come into contact with each other considering solvent probe radius R_s. This function is $4\pi(Rj + R_s)^2$ for $r < (R_i - R_j)$, a cubic spline connecting the points $(R_i - R_j, 4\pi(R_j + R_s)^2)$ and $(R_i + R_j + 2R_s, 0)$ with zero derivative at its extrema, and zero elsewhere. This function is scaled by a constant fit in tandem with the PMFs and DDD. This is a first-order correction to the SA term, in analogy to the DDD as a first-order correction to the PB term. We set R_s equal to the solvent probe radius used in PBSA calculations (see below). However, for convenience, the atomic radii for this calculation were taken from the Lennard-Jones parameters in *ff*99. While these radii do not perfectly correspond to the AMBER *ff*99 PBSA radii, they are adequate for estimating the change in surface area in our model.

Protein Modeling. For training ELSCA, a set of high-resolution, nonhomologous proteins was chosen from the Protein Data Bank[5] using the PISCES server[48] to have R-factor less than 0.18 and less than 5% sequence similarity. From the resulting list of 441 proteins, 21 were chosen to represent a range of sizes and charges, with the additional criteria that there be no cofactors present in the selected proteins and that their structures be fully resolved. PDB ID numbers and summary data for the systems are given in Table 2. Hydrogens found in AMBER *ff*99 were added using Leap in the AMBER8[55] package, but to give some consideration to the protonation state of histidines, the pdb2gmx utility from the GROMACS software package[4,15,30,45,50] was used to dictate protonation on N^{ϵ} (the default state for pdb2gmx), N^{δ}, or both. Because the initial structure fed to our docking program is taken as a reference state, each structure was energy-minimized in vacuo using the Sander program in AMBER8 through 50 cycles of steepest descent and up to 950 cycles of conjugate gradient minimization or until the potential energy converged to 1.0×10^{-4} kcal/mol. Backbone heavy atoms were restrained to their initial coordinates by 1.0 kcal/mol-Å^2 harmonic potentials during the minimization.

Table 2. Proteins Used in Training

PDB ID[b]	resolution[a]	residues	charge (e)
1jcd	1.30	50	−2
1gvd	1.45	52	+8
1mof	1.70	53	−3
1i2t	1.04	61	−1
3ebx	1.40	62	+2
1ok0	0.93	74	−5
1c5e	1.10	95	−2
1lni	1.00	96	−7
1jo0	1.37	97	+3
1eaj	1.35	127	−3
2lis	1.35	131	+11
1cxq	1.02	162	+10
1nwa	1.50	168	0
1mf7	1.25	194	+3
1qhv	1.51	195	−1
1g61	1.30	225	−14
1dj0	1.50	264	+4
1gxm	1.32	324	+4
1qcx	1.70	359	−7
1ug6	0.99	426	−6
1iat	1.62	556	+3

[a] Crystal diffraction resolution, Å. [b] Of the cases listed, where the PDB file contained multiple chains, the A chain was the one selected by PISCES.

Protein Docking. To parametrize ELSCA as described in *Approximation to the Total Electrostatic Energy*, we required a diverse set of pairs of proteins in various hypothetical complexes that are packed together to varying degrees. We gave consideration to side-chain flexibility in our docking by allowing all side-chains to take on configurations found in a detailed rotamer library,[19] but in order to avoid biasing the training set toward complexes with favorable electrostatics, only steric factors were included in the scoring function. Each of the 21 proteins in the homology-culled set was first docked to every other along 10 000 random approach vectors. This set was pruned of any docked solutions that put the two proteins in very similar orientations. The top-scoring 100 complexes generated in the pruned set (the tightly packed set) were subjected to additional refinement by allowing additional rigid backbone moves with side-chain flexibility as described above. The next 60 high-scoring complexes (the loosely packed set) were refined with flexible side-chains but no additional backbone moves. Every 50th complex from the rest of the pruned set, up to a total of 50 (the perturbed set), was perturbed by pulling the proteins apart by one to four solvent probe diameters and then performing flexible side-chain refinement. Although we intend the software that performed these manipulations to be used in protein docking, this cursory scheme was intended to generate random juxtapositions of proteins with varying degrees of tightness, not correct docked answers.

PBSA Calculations. The thermodynamic cycle used to calculate the total electrostatic energy of complex formation by solving the linearized Poisson−Boltzmann equation [6] has been extensively used in the literature and is clearly illustrated in the work of Elcock and coworkers.[20]

$$\nabla\cdot(\epsilon(\mathbf{r})\nabla\phi(\mathbf{r})) = -4\pi[\rho_{\text{pro}} - \sum_k \beta q_k^2 N_k \phi(\mathbf{r})] \qquad (6)$$

In [6], ϵ represents the dielectric scalar field, ϕ is the electrostatic potential, ρ_{pro} is the macroscopic charge density due to the solute, β is the reciprocal of temperature times Boltzmann's constant, and q_k and N_k are the charge and number density of the kth ion in solution. The electrostatic energy of a solute is given by [7]:

$$E_{\text{elec}} = \frac{1}{2}\int\int_V\int \phi(\mathbf{r})\rho_{\text{pro}}(\mathbf{r})\mathrm{d}\mathbf{r} \qquad (7)$$

The electrostatic free energy of forming a complex between two proteins with low internal dielectric ϵ_{pro} in continuum water with high dielectric ϵ_{water} is equal to the sum of the energies required to bring each isolated protein out of water into a phase in which the dielectric of the continuum solvent is ϵ_{pro}, the energy of bringing the two proteins together as calculated by Coulombic interactions with dielectric ϵ_{pro}, and the energy of returning the complex to the continuum water. This cycle is equivalent to calculating the electrostatic energies E_{elec} of each separate protein and the complex using identical grids, subtracting the self-energy artifacts created by mapping the point charges to the grid points, and taking the difference in the energy of the complex and its components.

We used UHBD[32] to solve the LPBE for each putative complex produced by our docking studies on each protein system. Grids with identical dimensions and alignment were calculated for each separate protein and the complex. To minimize boundary artifacts, a coarse grid was first calculated, large enough that solvent filled roughly 95% of the volume, using Coulombic potentials with Debye−Hückel screening for boundary conditions. A finer, focused grid was used for the thermodynamic calculations, generated using the coarse grid for boundary conditions and large enough to encompass the complex by at least 5 Å on all sides.

In our PB formulation, we conformed to the methods used for PBSA in AMBER *ff*99. For all proteins, ϵ_{pro} was set to 1 and ϵ_{water} was set to 80. Dielectric smoothing[16] was applied in all cases. A monovalent implicit ion concentration of 0.100 M was used throughout the studies with an implicit ion radius of 2.1 Å. A grid spacing of 0.5 Å, or the minimum spacing needed to satisfy the aforementioned size requirements with a grid of 240 points on a side, was used for the focused grids. In all calculations, the low-dielectric region was defined by the solvent-excluded volume with a solvent probe of radius 1.6 Å and solute atom radii optimized for PBSA calculations in the context of AMBER *ff*99 (readers should consult the AMBER 8 source code file pb_init.f for details). To ensure convergence in each result, the UHBD finite-difference solver was permitted to run for up to 200 iterations, twice the default number.

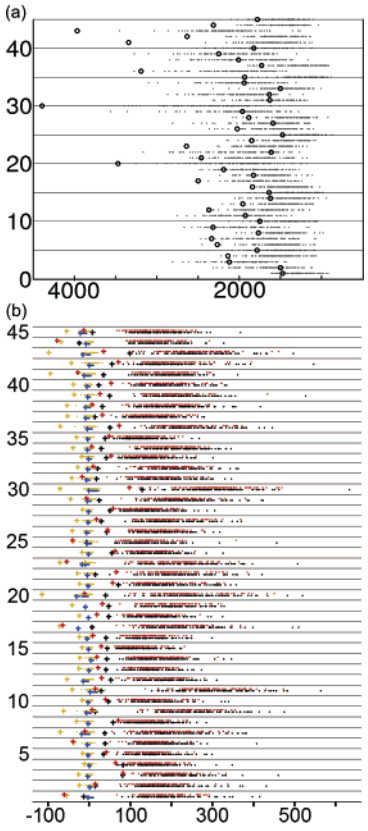


Figure 1. Properties of native vs non-native protein complexes. In (a), the solvent-accessible surface area buried by non-native complexes is shown by gray points, that of native complexes by black donuts. The *x*-axis has units of Å^2. In (b), the electrostatic energy of the complexes according to numerous methods is shown. Plus signs represent native complexes and points represent non-native ones. PB-derived energies (without consideration to the apolar solvation energy given by a change in SASA) are shown in black, analytic Coulombic ($\epsilon = 78$) just below the PB results in blue, linear DDD ($\epsilon = 5r$) just above the PB results in yellow, and ELSCA results (including apolar solvation energy) above those of the linear DDD in red. The *x*-axis has units of kcal/mol. Each pair of proteins is represented on a different level in each chart. In ascending order on both charts, the pairs are the separated components of crystal structures 1A0O, 1BVK, 1KXV, 1STF, 1ACB, 1CGI, 1FSS, 1TAB, 1AHW, 1CHO (10), 1GLA, 1MAH, 1TGS, 2PTC, 1CSE, 1MEL, 1UDI, 2SIC, 1AVW, 1DFJ (20), 1IAI, 1MLC, 1UGH, 2SNI, 1AVZ, 1DQJ, 1IGC, 2TEC, 1BQL, 1EFU (30), 1JHL, 2VIR, 1BRC, 1EO8, 1PPE, 3HHR, 1BRS, 1FBI, 1QFU, 2JEL (40), 4HTC, 1BTH, 1FIN, 1SPB, and 2KAI.

Test Cases for ELSCA. To validate ELSCA, we selected 45 protein complexes from a protein:protein docking benchmark[13] of nonredundant test cases. PDB identifiers for these complexes are given in Figures 1 and 2. Because protein complex structures are considerably rarer than those of individual proteins, we accepted members of the benchmark into our test set so long as they did not have deeply buried cofactors or cofactors near the binding site. Residues immediately before or after an unresolved portion of the backbone were treated as C- and N-terminal residues, respectively. Unresolved side-chains were rebuilt using SCWRL3.0.[8,10] Each protein complex was energy-minimized

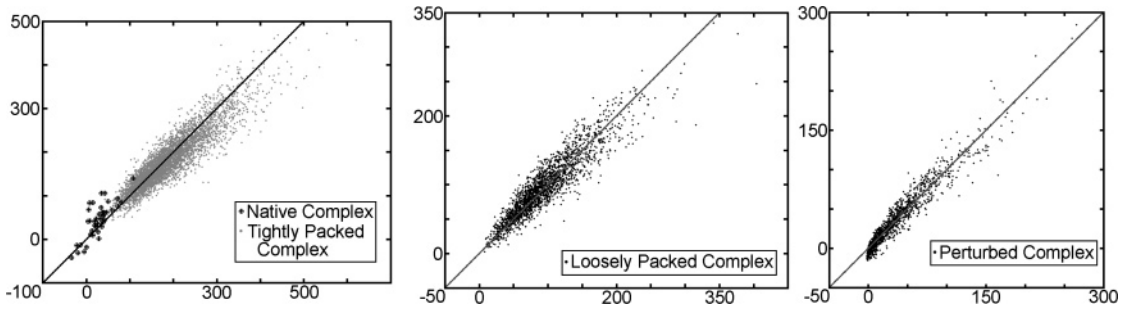


Figure 2. ELSCA predictions for native, non-native, and separated complexes. ELSCA was trained on 39 153 non-native complexes of proteins that are not known to form complexes and then tested on 8161 non-native and 45 native complexes of proteins that are known to associate. The *x*-axis gives the energy, in kcal/mol, of the complex according to the AMBER *ff*99 PBSA formulation, the *y*-axis the ELSCA-predicted value. Predictions for each of the three complex types described in *Methods: Protein Docking* are displayed separately.

according to the protocol used for the training set proteins. The complexes were then separated and docked to their partners in the manner described in *Methods: Protein Docking*. No docking between noninteracting proteins was attempted in the test set. None of the complexes produced by the cursory manipulation scheme came within 10 Å RMSD of the native complex.

Results

Protein Docking and Complex Selection. To parametrize a set of pairwise interactions that emulate PBSA results for general interacting proteins, we created a set of putative protein complexes from a nonhomologous set of peptide chains, all with different orientations and degrees of association. In docking the proteins, we permitted limited flexibility by allowing all side-chains in either molecule to take on discrete rotamer states (see *Methods*, *Protein Docking*), which permitted closer associations and helped to ensure that our training set included some diversity in the conformation of each molecule.

To test whether our methods of producing the training set can generate complexes that appear native by geometric and electrostatic standards, we compared the amount of SASA buried by native complexes in the test set to that buried by tightly packed configurations of the same proteins (Figure 1a). For some complexes our docking method produced non-native structures with comparable buried SASA, although native complexes often bury more surface area than any of the decoys. Figure 1b reveals that our method only rarely creates tightly packed complexes with comparable electrostatics to the native configuration. According to our Poisson−Boltzmann model, the electrostatic energy of the native complexes is optimal for nearly every pair of proteins in the test set, and there is a considerable variation in the energy of non-native complexes.

We are currently unable to test whether PB electrostatics are truly discriminatory of native complexes because we have not yet used PB directly to energy-minimize putative complexes of interacting proteins. Even while we are working to overcome this limitation, the data in Figure 1b suggest that the PB electrostatic model is an outstanding indicator, relative to other simple electrostatic models, of a native complex among other members of the tightly packed set. In contrast to PB, a homogeneous dielectric model ($\epsilon = 78$) typically suggests a "golf-course" electrostatic energy landscape in protein association. The linear DDD ($\epsilon = 5r$) appears to be much better at distinguishing native complexes from non-native, a result that may help to explain the success of linear DDDs in other docking applications.[53,25] However, such a model ultimately has no physical basis, which would hinder attempts to develop more complete potentials that take into account nonpolar solvation effects.

Optimizing Basis Functions. A large set of PMFs was the primary means of accurately reproducing Poisson−Boltzmann electrostatics using pairwise interactions. However, a uniform DDD and a crude surface area function helped make gross corrections to the electrostatic and nonpolar solvation energy estimates so that the PMFs could make finer corrections. A screened Coulombic potential ensured that long-range electrostatic interactions were realistically captured. As stated in *Methods*, the basis set for the DDD was chosen so that the dielectric constant could be tuned at short-range and intermediate distances, where the PMFs taper and the screened Coulombic approximation begins break down. We found $k_{PMF} = 3$, $\beta_{PMF} = 0.6$, and $\lambda_{PMF} = 1$ to yield good results according to the criteria set forth in *Methods*; addition of more basis functions could not significantly improve the quality of the fit nor could changing the broadness of the set of basis functions (data not shown).

A successful model must correctly predict the energy of both native and non-native complexes. As shown in Figure 2, ELSCA is very good at predicting the energies of non-native complexes of proteins in the test-set, yielding an overall correlation coefficient of 0.962 (if ELSCA were applied to its own training set, the correlation is 0.964) and a RMSD of 26.8 kcal/mol between the predicted and calculated energies (Figure 2). The regression line for predicted on calculated values of all non-native complexes has a slope of 0.89 and bias 11.1 kcal/mol. For predicting native complexes, the correlation is somewhat lower (0.787), the RMSD is 29.2 kcal/mol, and the regression line shows a slope of 1.14 with a bias of 13.0 kcal/mol. For comparison, a homogeneous dielectric model yields a poor correlation of 0.229 with our PB results for native complexes, and a linear distance-dependent dielectric model yields a correlation of only 0.083.

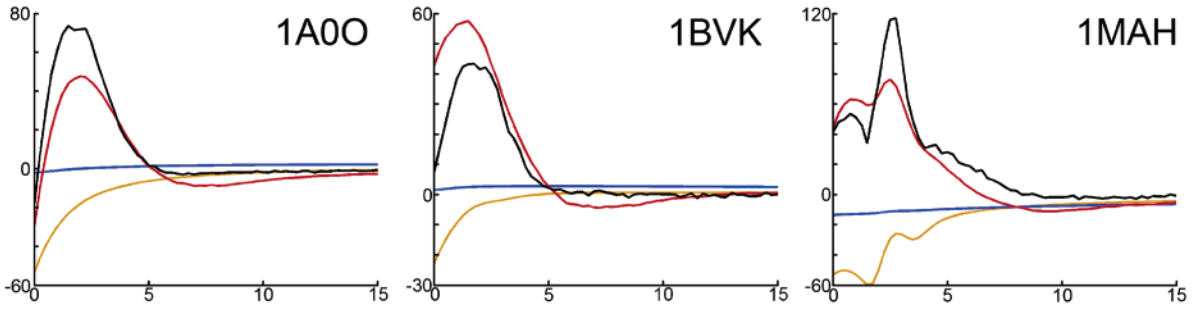


Figure 3. Reproduction of energy landscapes for three complexes. ELSCA was used to reproduce electrostatic plus apolar solvation energy (*y*-axis, units of kcal/mol) along a hypothetical, linear reaction pathway for three different protein systems (*x*-axis denotes distance from the native state, in Å). Similar to Figure 1b, black represents PBSA energy, red ELSCA, blue a homogeneous dielectric ($\epsilon = 78$), and yellow a linear DDD ($\epsilon = 5r$).

Energy along Reaction Pathways. To demonstrate ELSCA's ability to predict the features of the energy landscape, we selected three protein complexes with different characteristics from the test set and computed the total electrostatic and solvation energy using PBSA, ELSCA, $\epsilon = 78$, and $\epsilon = 5r$ as the ligand was pulled directly away from the receptor along a putative reaction pathway. Proteins were held rigid during these tests to eliminate energy contributions from conformational change in either protein. Measurements were taken at 0.25 Å intervals, up to 15.0 Å from the native state. We selected systems 1MAH (acetylcholinesterase and fasciculin-2), 1BVK (lysozyme bound to a lysozyme-specific antibody), and 1A0O (chea in complex with chey). Despite quantitative errors in the energy estimates noted above and in Figure 2, Figure 3 demonstrates that ELSCA creates a smooth potential energy landscape that reproduces the features of the benchmark PBSA model, and thus the critical effects of solvation in protein interactions, in ways that simpler electrostatic models cannot. ELSCA can produce energies that are wrong in the native state, but as shown by the plot of 1BVK approach, ELSCA at least reproduces the qualitative features of the energy landscape. Comparison of the parametrized ELSCA model to the other electrostatic models (Figure 1b) reveals that the native state is again predicted to be the minimum or (in one case only) within 3 kcal/mol of the minimum energy of the tightly packed set. Figure 1b compares ELSCA to PB electrostatic energies; we remind the reader that ELSCA includes a SASA term in its formulation (see *Methods: Approximation to the Total Electrostatic and Nonpolar Solvation Energy*) that cannot be simply dissected from the electrostatic contributions. However, the effect of the SASA term is small (5 kcal/mol per 1000 Å^2 buried SASA) compared to the magnitudes of the PB electrostatic energies.

The parameters for the surface area term, PMFs, and uniform DDD obtained by the least-squares fitting are given in the Supporting Information.

Discussion

Reproduction of PBSA Results. ELSCA's estimates compare to PBSA-derived energies with very high correlations. Furthermore, ELSCA is equally predictive of two very large, nonredundant sets of proteins, indicating that it is applicable to proteins in general. The parameters k_{PMF}, β_{PMF}, and λ_{PMF}, which are most important for correctly determining the energy of closely packed complexes, are somewhat arbitrary; a comparable quality of fit can be obtained with numerous PMF and DDD bases, but significant improvements over our formulation were not found. These results suggest that ELSCA is nearly the best possible rendering of the PBSA free energy model in the form of pairwise interactions between the atoms of each solute; additional improvements require consideration of the local atomic environments.

In this study, we have parametrized ELSCA over a set of non-native complexes that were not optimized in terms of electrostatics because the vast majority of structures that will be sampled in protein docking or association simulations will be non-native. Despite the fact that electrostatic complementarity typifies protein interfaces, the parameters transfer well to native complexes. Indeed, the data points for native complexes in Figure 2 can be seen as an extension of the data for the most tightly packed non-native complexes. The lower correlation is largely due to the fact that native complexes have consistently low energy in the PBSA model, whereas non-native, tightly packed complexes can have a broad range of energies. Further improvements in this regime may be possible by including hundreds to thousands of natural interfaces in the training set.

ELSCA's potentials of mean force and uniform distance-dependent dielectric are meant to be superimposed on van der Waals and electrostatic grids and lookup tables so that ELSCA may be as computationally tractable as any gas-phase energy calculation or grid-based calculation used in simulations and docking studies. Implicit solvent simulations of protein association can be significantly accelerated using ELSCA rather than a model such as GBSA which depends on frequently recalculating atomic polarization radii. Significant enhancements are also possible for methods based on rotamer libraries which require pairwise additive interactions. Once the ELSCA functions are stored in lookup tables (and, ideally, mapped to grids), the implementation is trivial. Because the PMFs used in ELSCA are smoothly varying, grid-based artifacts should be minimal when ELSCA is applied in grid-based calculations.

Alone, none of the PMFs has physical significance and the DDD is realistic only at distances greater than 60 Å, where it reproduces Debye−Hückel screening effects of the ionic solvent. Together, however, all these functions repro-

duce a model that can realistically account for desolvation and solvent ion effects. For docking studies, ELSCA is expected be nearly as useful as the PBSA model because it reproduces the features of the PBSA energy landscape (Figure 3) and obtains few competing minima that PBSA would not have (Figure 1a,b). However, there is ongoing debate about the validity of implicit solvent models. ELSCA has promise as an investigative tool, but like Generalized Born and other theoretically benchmarked approximations it is ultimately limited by the model it emulates.

ELSCA in Context. A number of studies have used potentials of mean force between multiple atom types to approximate nonlinear functions in molecular modeling, many of them stemming from the work of Miyzawa and Jernigan.[35,36] Such models often rely on short-ranged potentials parametrized by statistical distributions found in a homology-culled database of protein structures such as the Protein Data Bank.[5] The set of atomic contact energies (ACE)[54] is a very detailed example of such "knowledge-based" potentials, using step-functions to describe the transfer of various atoms in the selected force field between hydrophobic and aqueous environments. ACE accounts for solvation energy contributions to protein stability but must be augmented by an approximation for solvent screening electrostatic interactions. In contrast to other PMF models that are only able to assess the energy of a bound state, ELSCA is intended specifically for estimating the electrostatic energy of protein:protein association as the partners come within a few solvent layers of one another and as they bind. The basis functions used in ELSCA are continuously differentiable and longer-ranged than those used in most other PMFs, essential features of a model that is intended to predict interactions over a range of intermolecular separations.

By assuming a Hamiltonian that separates into solvation free energy and molecular-mechanics interactions and the existence of pairwise additive potentials of mean force for various chemical groups, Lazaridis and Karplus built the EEF1 model[27] for solvation energy of proteins. These are the same basic assumptions that were used to construct ELSCA. Again like ELSCA, EEF1 uses Gaussians as the basis functions for its PMFs and works in tandem with a uniform distance-dependent dielectric ($\epsilon = r$ in this case) and can be superimposed on the gas-phase molecular mechanics energy for a modest computational cost. EEF1 is intended for use in protein folding and single-protein dynamics problems, where the internal energy of the molecule varies as its conformation changes. Such a model could be a powerful complement to ELSCA, which addresses the change in solvation and electrostatic association energy in complex formation, not the energy of changing each molecule's configuration from a particular reference state.

One other pairwise additive PMF model should be mentioned because of its notable similarity to ELSCA in design and purpose. Jiang and co-workers[26] built a model for the total free energy of protein:protein association using four general heavy atom types, distinguishing between atoms of each type on the ligand and receptor to construct sixteen PMFs. This model achieved a correlation coefficient of 0.75 between its scores and the experimental binding energies of 28 protein complexes and recovered important features such as a potential well between hydrogen-bond donors and acceptors at the correct hydrogen-bonding distance. This is nearly the consistency by which ELSCA, given its current training set, reproduces native complex energies, but there is no way to judge the model's accuracy in reproducing nonisolable protein complexes and intermediates. ELSCA's primary advantages in reproducing its training data are likely the consistency by which the data is generated and the limitless number of data points available, which permits distinctions between many more atom types. Again, this is an advantage only if the model ELSCA emulates is realistic.

Future Directions. We plan to extend ELSCA's PMFs to estimate the solute−solvent dispersion energy as derived by a continuum model based on volume integration while retaining the linear function of surface area for the work of cavity formation.[29,52] ELSCA is implemented in a developmental docking program, GRAPPLE, which is ultimately intended to emulate state-of-the-art binding energy calculation techniques[24,31,37] to dock proteins and determine transition paths in binding by computing the free-energy landscape of the interaction at high resolution.

Acknowledgment. David Cerutti thanks Drs. Stewart Adcock and Steven Bond for advice on numerical methods and Dr. Robert Konecny and the Center for Theoretical Biophysics (NSF grants MCB-0084797, PHY-0216576, and PHY-0225630, with additional support from NIH grant NBCR-RR08605i) for assistance with large-scale computing resources. This research was funded in part by NIH grant 2 T32 GM07240-27, the Achievement Rewards for Collegiate Scholars San Diego chapter, and grants from Howard Hughes Medical Institute, the National Science Foundation, the National Institutes of Health, the National Biomedical Computing Resource, the San Deigo Supercomputer Center, the W.M. Keck Foundation, and Accelrys, Inc.

Supporting Information Available: Coefficients for ELSCA to simulate protein encounters and notes on implementation of ELSCA-based potentials. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Angelini, T. E.; Liang, H.; Wriggers W.; Wong, G. C. Like-Charge Attraction Between Polyelectrolytes Induced by Counterion Charge Density Waves. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*(*15*), 8634−8637.

(2) Bahadur, R. P.; Chakrabarti, P.; Rodier, F.; Janin, J. Dissecting Subunit Interfaces in Homodimeric Proteins. *Proteins: Struct., Funct., Genet.* **2003**, *53*, 708−719.

(3) Baker, N. A.; Sept, D.; Joseph, S.; Holst, M. J.; McCammon, J. A. Electrostatics of Nanosystems: Application to Microtubules and the Ribosome. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10037−10041.

(4) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. GROM American Chemical Society: A Message-Passing Molecular Dynamics Implementation. *Comput. Phys. Commun*. **1995**, *91*, 43−45.

(5) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235−242.

(6) Bogan, A. A.; Thorn, K. S. Anatomy of Hot Spots in Protein−Protein Interfaces. *J. Mol. Biol*. **1998**, *280*, 1−9.

(7) Bourne, Y.; Taylor, P.; Marchot, P. Acetylcholinesterase Inhibition by Fasciculin: Crystal Structure of the Complex. *Cell* **1995**, *83(3)*, 503−12.

(8) Bower, M. J.; Cohen, F. E.; Dunbrack, R. L. Prediction of Protein Side-chain Rotamers from a Backbone-dependent Rotamer Library: A New Homology Modeling Tool. *J. Mol. Biol.* **1997**, *267*, 1268−1282.

(9) Caflisch, A.; Karplus, M. Acid and Thermal Denaturation of Barnase Investigated by Molecular Dynamics Simulations. *J. Mol. Biol.* **1995**, *252*, 672−708.

(10) Canutescu, A. A.; Shelenkov, A. A.; Dunbrack, R. L., Jr. A Graph Theory Algorithm for Protein Side-Chain Prediction. *Protein Sci.* **2003**, *12*, 2001−2014.

(11) Cerutti, D. S.; Wong, C. F.; McCammon, J. A. Brownian Dynamics Simulations of Ion Atmospheres around Polyalanine and B-DNA: Effects of biomolecular dielectric. *Biopolymers* **2003**, *70(3)*, 391−402.

(12) Chen, R.; Weng, Z. A Novel Shape-Complementarity Scoring Function for Protein−Protein Docking. *Proteins: Struct., Funct., Genet.* **2003**, *51*, 397−408.

(13) Chen, R.; Mintseris, J.; Janin, J.; Weng, Z. A Protein−Protein Docking Benchmark. *Proteins: Struct., Funct., Genet.* **2003**, *52(1)*, 88−91.

(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.; Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.

(15) Daura, X.; Oliva, B.; Querol, E.; Avilés, F. X.; Tapia, O. On the Sensitivity of MD Trajectories to Changes in Water−Protein Interaction Parameters: The Potato Carboxypeptidase Inhibitor in Water as a Test Case for the GROMOS Force Field. *Proteins: Struct., Funct. Genet.* **1996**, *25(1)*, 89−103.

(16) Davis, M. E.; McCammon, J. A. Dielectric Boundary Smoothing in Finite Difference Solutions of the Poisson Equation: An Approach to Improve Accuracy and Convergence. *J. Comput. Chem.* **1991**, *12(7)*, 909−912.

(17) Dominy, B. N.; Brooks, C. L., III Development of a Generalized Born Model Parameterization for Proteins and Nucleic Acids. *J. Phys. Chem. B* **1999**, *103*, 3765−3773.

(18) Dong, F.; Vijayakumar, M.; Zhou, H. Comparison of Calculation and Experiment Indicates Significant Electrostatic Contributions to the Binding Stability of barnase and Barstar. *Biophys. J.* **2003**, *85*, 49−60.

(19) Dunbrack, R. L., Jr.; Karplus, M. Backbone-Dependent Rotamer Library for Proteins: Application to Side-Chain Prediction. *J. Mol. Biol.* **1993**, *230*, 543−574.

(20) Elcock, A. H.; Gabdoulline, R. R.; Wade, R. C.; McCammon, J. A. Computer Simulation of Protein−Protein Association Kinetics: Acetylcholinesterase-Fasciculin. *J. Mol. Biol.* **1999**, *291*, 149−162.

(21) Frisch, C.; Schreiber, G.; Johnson, C. M.; Fersht, A. R. Thermodynamics of the Interaction of barnase and Barstar: Changes in Free Energy Versus Changes in Enthalpy upon Mutation. *J. Mol. Biol.* **1997**, *267(3)*, 696−706.

(22) Gabdoulline, R. R.; Wade, R. C. Protein−Protein Association: Investigation of Factors Influencing Association Rates by Brownian Dynamics Simulations. *J. Mol. Biol.* **2001**, *306*, 1139−1155.

(23) Gabdoulline, R. R.; Wade, R. C. Effective Charges for Macromolecules in Solvent. *J. Phys. Chem.* **1996**, *100*, 3868−3878.

(24) Gohlke, H.; Kiel, C.; Case, D. A. Insights into Protein−Protein Binding by Binding Free Energy Calculation and Free Energy Decomposition for the Ras-Raf and Ras-RalGDS Complexes. *J. Mol. Biol.* **2003**, *330*, 891−913.

(25) Gray, J. J.; Moughon, S.; Wang, C.; Schueler-Furman, O.; Kuhlman, B.; Rohl, C. A.; Baker, D. Protein−Protein Docking with Simultaneous Optimization of Rigid-Body Displacement and Side-Chain Conformations. *J. Mol. Biol.* **2003**, *331(1)*, 281−299.

(26) Jiang, L.; Gao, Y.; Fenglou, M.; Liu, Z.; Lai, L. Potential of Mean Force for Protein−Protein Interaction Studies. *Proteins: Struct., Funct., Genet.* **2002**, *46(2)*, 190−196.

(27) Lazardis, T.; Karplus, M. Effective Energy Functions for Proteins in Solution. *Proteins: Struct., Funct., Genet.* **1999**, *35*, 133−152.

(28) Lee, M. S.; Feig, M.; Salsbury, F. R., Jr.; Brooks, C. L., III. New Analytic Approximation to the Standard Molecular Volume Definition and Is Application to Generalized Born Calculations. *J. Comput. Chem.* **2003**, *24(11)*, 1348−1356.

(29) Levy, R.; Zhang, L. Y.; Gallicchio, E.; Felts, A. K. On the Nonpolar Hydration Free Energy of Proteins: Surface Area and Continuum Solvent Models for the Solute-Solvent Interaction Energy. *J. Am. Chem. Soc.* **2003**, *125(31)*, 9523−9530.

(30) Lindahl, E.; Hess, B.; ans van der Spoel, D. GROMACS 3.0: A Package for Molecular Simulations and Trajectory Analysis. *J. Mol. Model*. **2001**, *7*, 306−317.

(31) Luo, H.; Sharp, K. On the Calculation of Absolute Macromolecular Binding Free Energies. *Proc. Natl. Acad. Sci.* **2002**, *99(16)*, 10399−10404.

(32) 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. Electrostatics and Diffusion of Molecules in Solution: Simulations with the University of Houston Brownian Dynamics Program. *Comput. Phys. Commun.* **1995**, *91*, 57−95.

(33) Mandell, J. G.; Roberts, V. A.; Pique, M. E.; Kotlovyi, V.; Mitchell, J. C.; Nelson, E.; Tsigelny, I.; Ten Eyck, L. F. Protein Docking Using Continuum Electrostatics and Geometric Fit. *Protein Eng.* **2001**, *14(2)*, 105−113.

(34) McQuarrie, D. A. *Statistical Mechanics*; University Science Books: Sausalito, California, 2000.

(35) Miyazawa, S.; Jernigan, R. L. Residue-Residue Potentials with a Favorable Contact Pair Term and an Unfavorable High Packing Density Term for Simulation and Threading. *J. Mol. Biol.* **1996**, *256*, 632−644.

(36) Miyazawa, S.; Jernigan, R. L. Estimation of Effective Interresidue Contact Energies from Protein Crystal Structures: Quasi-Chemical Approximation. *Macromolecules* **1985**, *18*, 534−552.

(37) Noskov, S. Y.; Lim, C. Free Energy Decomposition of Protein−Protein Interactions. *Biophys. J.* **2001**, *81*, 737−750.

(38) Roux, B.; Simonson, T. Implicit Solvent Models. *Biophys. Chem.* **1999**, *78*, 1−20.

(39) Sanner, M. F.; Olson, A. J.; Spehner, J. C. Fast and Robust Computation of Molecular Surfaces. *Proceedings of the 11th ACM Symposium on Computational Geometry* C6−C7, 1995.

(40) Sheinerman, F. B.; Norel, R.; Honig, B. Electrostatic Aspects of Protein−Protein Interactions. *Curr. Opin. Struct. Biol.* **2000**, *10*, 153−159.

(41) Sitkoff, D.; Sharp, K. A.; Honig, B. Accurate Calculation of Hydration Free Energies Using Macroscopic Solvent Models. *J. Phys. Chem.* **1994**, *98*, 1978−1988.

(42) Sitkoff, D.; Sharp, K. A.; Honig, B. Correlating Solvation Free Energies and Surface Tensions of Hydrocarbon Solutes. *Biophys. Chem.* **1994**, *51(3)*, 397−409.

(43) Tai, K.; Shen, T.; Henchman, R. H.; Bourne, Y.; Marchot, P.; McCammon, J. A. Mechanism of Acetylcholinesterase Inhibition by Fasciculin: A 5-ns Molecular Dynamics Simulation. *J. Am. Chem. Soc.* **2002**, *124(21)*, 6153−6161.

(44) van Buuren, A. R.; Marrink, S. J.; Berendsen, H. J. C. A Molecular Dynamics Simulation of the Decane/Water Interface. *J. Phys. Chem.* **1993**, *97(36)*, 9206−9212.

(45) van der Spoel, D.; van Buuren, A. R.; Tieleman, P.; Berendsen, H. J. C. Molecular Dynamics Simulations of Peptides from BPTI: A Closer Look at Amide-Aromatic Interactions. *J. Biomolecular NMR* **1996**, *8*, 229−238.

(46) van Gunsteren, W. F.; Berendsen, H. J. C. *Gromos-87 Manual*; Biomos BV: Groningen, The Netherlands, 1987.

(47) Wang, T.; Wade, R. C. Implicit-Solvent Models for Flexible Protein−Protein Docking by Molecular Dynamics Simulation. *Proteins: Struct., Funct., Genet.* **2003**, *50*, 158−169.

(48) Wang, G.; Dunbrack, R. L. PISCES: a protein sequence culling server. *Bioinformatics* **2003**, *19*, 1589−1591.

(49) Wang, J.; Cieplak, P.; Kollman, P. A. How Well Does a Restrained Electrostatic Potential (RESP) Model Perform in Calculating Conformational Energies of Organic and Biological Molecules? *J. Comput. Chem.* **2000**, *21*, 1049−1074.

(50) Warshel, A.; Papazyan, A. Electrostatic Effects in Macromolecules: Fundamental Concepts and Practical Modeling. *Curr. Opin. Struct. Biol.* **1998**, *8*, 211−217.

(51) Yu, B.; Blaber, M.; Gronenborn, A. M.; Clore, G. M.; Caspar, D. L. Disordered Water within a Hydrophobic Protein Cavity Visualized by X-ray Crystallography. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 103−108.

(52) Zacharias, M. Continuum Solvent Modeling of Nonpolar Solvation: Improvement by Separating Surface Area Dependent Cavity and Dispersion Contributions. *J. Phys. Chem. A* **2003**, *107*, 3000−3004.

(53) Zacharias, M. Protein−Protein Docking with a Reduced Model Accounting for Side-Chain Flexibility. *Protein Sci.* **2003**, *12*, 1271−1282.

(54) Zhang, C.; Vasmatzis, G.; Cornette, J. L.; DeLisi, C. Determination of Atomic Desolvation Energies from the Structures of Crystallized Proteins. *J. Mol. Biol.* **1997**, *267-(3)*, 707−725.

(55) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, III, T. E.; DeBolt, S.; Ferguson, D.; Seibel, G.; 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.

CT049946F

JCTC Journal of Chemical Theory and Computation

Determination of Electrostatic Parameters for a Polarizable Force Field Based on the Classical Drude Oscillator

Victor M. Anisimov,[†] Guillaume Lamoureux,[‡] Igor V. Vorobyov,[†] Niu Huang,[†,||] Benoît Roux,[§] and Alexander D. MacKerell, Jr.*[,†]

Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, Maryland, Département de physique, Université de Montréal, C.P. 6128, succ. centre-ville, Montréal, Québec, Canada H3C 3J7, and Department of Biochemistry, Weill Medical College of Cornell University, New York, New York

Received September 28, 2004

Abstract: A procedure to determine the electrostatic parameters has been developed for a polarizable empirical force field based on the classical Drude oscillator model. Atomic charges and polarizabilities for a given molecule of interest were derived from restrained fitting to quantum-mechanical electrostatic potentials (ESP) calculated at the B3LYP/ cc-pVDZ or B3LYP/aug-cc-pVDZ levels on grid points located on concentric Connolly surfaces. The determination of the atomic polarizabilities requires a series of perturbed ESP maps, each one representing the electronic response of the molecule in the presence of a background charge placed on Connolly surfaces primarily along chemical bonds and lone pairs. Reference values for the partial atomic charges were taken from the CHARMM27 additive all-atom force field, and those for the polarizabilities were based on adjusted Miller's ahp atomic polarizability values. The fitted values of atomic polarizabilities were scaled to reflect the reduced polarization expected for the condensed media and/or to correct for the systematic underestimation of experimental molecular polarizabilities by B3LYP calculations. Following correction of the polarizabilities, the atomic charges were adjusted to reproduce gas-phase dipole moments. The developed scheme has been tested on a set of small molecules representing functional moieties of nucleic acids. The derived electrostatic parameters have been successfully applied in a preliminary polarizable molecular dynamics simulation of a DNA octamer in a box of water with sodium counterions. Thus, this study confirms the feasibility of the use of a polarizable force field based on a classical Drude model for simulations of biomolecules in the condensed phase.

1. Introduction

Computer simulations based on empirical force fields are now a standard procedure to investigate biological phenomena.[1] Empirical force field calculations, due to the simplicity of the potential energy function, allow for atomic detail studies of biomolecules with explicit representation of the condensed phase environment to be performed. However, it is essential that the force field accurately reproduces the experimental regimen to ensure the quality of results of such calculations.

The majority of force fields consist of electrostatic, van der Waals, and bonding energy terms calculated in a pairwise additive fashion.[2] The induced polarization, which arises from a perturbation of the electronic structure of the molecular

* Corresponding author phone: (410)706-7442; fax: (410)706-5017; e-mail: amackere@rx.umaryland.edu. Corresponding author address: 20 Penn Street, Baltimore, MD 21201.
[†] University of Maryland.
[‡] Université de Montréal.
[§] Weill Medical College of Cornell University.
[||] Current address: Department of Biopharmaceutical Sciences, UCSF, San Francisco, CA.

10.1021/ct049930p CCC: $30.25
Published on Web 12/03/2004

species in response to the external electric field, is typically incorporated implicitly by using enhanced fixed partial atomic charges reflecting the average polarization taking place in the condensed phase. Despite the apparent simplification, additive empirical force fields have been remarkably successful in modeling complex molecular systems for the last two decades.[3] However, there are shortcomings in the additive model,[4] emphasizing the need to account for many-body induced polarization effects in an explicit way, motivating the development of polarizable force fields.

Current polarizable models can be classified into three major categories: point dipole, charge transfer, and classical Drude oscillator (or shell models); information about each of these models can be found elsewhere.[4,5] While a variety of efforts are ongoing to apply the point dipole and charge-transfer methods to biological systems,[6−27] the classical Drude oscillator approach has only seen minimal attention.[28−30] The Drude oscillator polarizability model was first introduced by Paul Drude in 1900 as a simple approach to describe the dispersive properties of materials.[31] The classical version of this model has been successfully used in statistical mechanical studies of condensed systems to treat electronic polarizability[5] and has been recently implemented[32] into the CHARMM program.[33,34]

Creation of a force field for biomolecular systems traditionally starts with the development of a water model. The polarizable SWM4-DP water model based on the classical Drude oscillator formalism has recently been presented.[30] It was parametrized to reproduce properties of liquid water under ambient conditions as well as some gas-phase properties such as the dipole moment and the interaction energy of the water dimer. In the SWM4-DP water model the polarizability of the oxygen atom, which in this case is equivalent to the molecular polarizability (since no Drude particles for hydrogen atoms were considered), was found to be 1.04 Å^3, which is 0.724 of the experimental gas-phase molecular polarizability of water, 1.44 Å^3.[35] Such reduced polarizability, which is essential to reproduce liquid-phase properties of water, including the dielectric constant, has been attributed to the energy cost of overlapping electron clouds in the condensed phase opposing induction.[30,36−41]

In the present work, steps toward systematic development of a polarizable force field based on the classical Drude oscillator model are presented. The determination of the electrostatic parameters is considered as the first step toward this goal. This effort, which includes fitting of partial atomic charges and atomic polarizabilities to a series of electrostatic potentials (ESP) around a molecule, each in the presence of an individual background charge, is the main focus of the present study. The theoretical background for the classical Drude oscillator model and the methodological details of the electrostatic parameter fitting are discussed in the next section. Validation of the theoretical level used in the derivation procedure is presented followed by a discussion of the selection of the reference values for charges and polarizabilities. Then, an example is given for the application of that procedure to the cytosine base, a model compound for nucleic acids. Finally, the validity of the proposed approach is illustrated via a condensed phase simulation of a DNA octamer using a preliminary classical Drude polarizable force field.

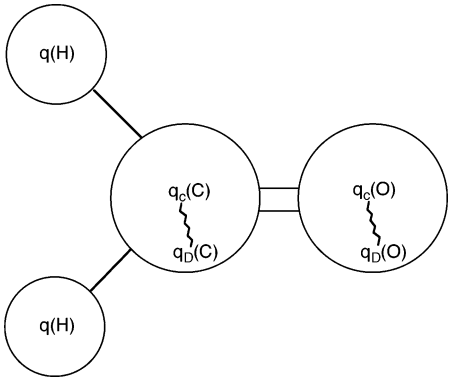


Figure 1. Classical Drude oscillator model using a formaldehyde molecule as an example. The displacements of the Drude particles, q_D, attached to the non-hydrogen atomic centers are exaggerated for clarity.

2. Theory and Methods

2.1. Classical Drude Oscillator Model. According to the classical Drude oscillator model, the polarizability is introduced by adding massless charged particles attached to each polarizable atom by a harmonic spring (Figure 1). Thus, a finite induced dipole is created and the partial atomic charge of atom A, $q(A)$, is redistributed between the Drude particle, $q_D(A)$, and the atomic core, $q_c(A)$. The positions of the Drude particles relative to the corresponding atomic centers are determined self-consistently by seeking the minimum energy consistent with the Born−Oppenheimer approximation. For the equilibrium position of the Drude particles, the atomic polarizability of atom A, $\alpha(A)$, is related to the charge $q_D(A)$ via the equation

$$\alpha(A) = q_D^2(A)/k_D \qquad (1)$$

where k_D is the force constant of the harmonic spring connecting a Drude particle to its corresponding atomic core. The magnitude of k_D is chosen to achieve small displacements of Drude particles from their corresponding atomic positions, $\mathbf{r}_D$, as required to remain close to the point-dipole approximation for the induced dipole associated with the atom−Drude pair.[32] Consequently, the atomic polarizability is determined by the amount of charge assigned to the Drude particle. Thus, the only adjustable parameters to be determined during parametrization of a polarizable atom A in the Drude model are the partial atomic charges $q(A) = q_c(A) + q_D(A)$. The development of a consistent protocol for the determination of $q_c(A)$ and $q_D(A)$ (or $\alpha(A)$) for a series of model compounds is, in part, the subject of the present study.

One of the major advantages of the classical Drude oscillator model is that it preserves the simple functional form of the pairwise additive force field and yet explicitly accounts for the electronic polarizability.[32] The electrostatic energy from the additive force field is substituted in the polarizable model by the Coulombic energy terms describing interactions between atomic cores and Drude particles and the self-energy of a polarizable atom treated via a harmonic term

$$U_{elec} = \sum_{A<B}^{N} \frac{q_c(A) \cdot q_c(B)}{|\mathbf{r}(A) - \mathbf{r}(B)|} + \sum_{A<B}^{N,N_D} \frac{q_D(A) \cdot q_c(B)}{|\mathbf{r}_D(A) - \mathbf{r}(B)|} + \sum_{A<B}^{N_D} \frac{q_D(A) \cdot q_D(B)}{|\mathbf{r}_D(A) - \mathbf{r}_D(B)|} + \frac{1}{2}\sum_{A}^{N_D} k_D |\mathbf{r}_D(A) - \mathbf{r}(A)|^2 \quad (2)$$

where N and N_D are the number of real atoms and Drude particles, respectively, q_c and q_D are the atomic core and Drude particle charges, respectively, and $\mathbf{r}$ and $\mathbf{r}_D$ are the positions of the real atoms and Drude particles, respectively. Modifying the classical force field energy function using eq 2 facilitates performing dynamical simulations with only minor modifications to existing programs, because the original energy functional form remains essentially unchanged. Computationally this means that the Drude particle positions for a given atomic configuration have to be self-consistently adjusted for each step of the dynamics simulation; however, such calculations are inefficient and rather expensive.[32] Therefore a molecular dynamics simulation algorithm based on the extended Lagrangian formalism[42,43] has been implemented, in which a small mass is attributed to the Drude particles, and the amplitude of the oscillators is controlled with a low temperature thermostat.[32] This technique allows the computationally expensive self-consistent field (SCF) regimen of molecular dynamics simulations to be avoided. Tests of the extended Lagrangian algorithm have shown that stable and accurate molecular dynamics trajectories can be generated, yielding liquid properties equivalent to the SCF regimen of molecular dynamics at a fraction of the computational cost.[32] Therefore, the classical Drude oscillator model for simulating atomic polarizability can be applied for molecular modeling in condensed media including macromolecular systems such as fully solvated nucleic acids, proteins, and lipid aggregates. However, the existing parameters from the nonpolarizable force fields need to be adjusted to take into account the presence of Drude particles.

2.2. Charge Fitting Scheme. The electrostatic properties of a molecular mechanics model with Drude polarizabilities are represented by atomic core charges $q_c(A)$ and Drude charges $q_D(A)$ producing the effective atomic charge $q(A)$ as their sum. Indeed, in the classical Drude oscillator polarizable model, the determination of atomic polarizabilities $\alpha(A)$ can be reduced to the determination of the partial charges of Drude particles, $q_D(A)$. Both $q(A)$ and $q_D(A)$ can be determined simultaneously, in a single fitting step.

Partial atomic charges are often obtained by optimizing the fit of an electrostatic potential ϕ^{MM} derived from the molecular mechanics (MM) model to a potential map ϕ^{QM} generated by quantum-mechanical (QM) calculations on a set of grid points $\{\mathbf{r}_g\}$ placed around the molecule. Although partial atomic charges of a nonpolarizable model can be extracted from a single QM potential map, adjusting the polarizabilities requires a series of response potential maps ϕ_p^{QM}, each one representing the altered charge distribution for the molecule in the presence of a small perturbing point charge z_p at a given position $\mathbf{r}_p$. A similar approach was used by Friesner et al. to derive parameters for the fluctuating charge and polarizable dipole models.[17,37,44] In our calculations the value of the perturbing charge z_p was arbitrarily chosen to be +0.5e. The MM potential at the g-th grid point (at the coordinate $\mathbf{r}_g$) for the molecule under the influence of a point-charge perturbation at position $\mathbf{r}_p$ is

$$\phi_{pg}^{MM} = \sum_{A}^{N}\left(\frac{q_c(A)}{|\mathbf{r}(A) - \mathbf{r}_g|} + \frac{q_D(A)}{|\mathbf{r}(A) + \mathbf{d}_p(A) - \mathbf{r}_g|}\right) + \frac{z_p}{|\mathbf{r}_p - \mathbf{r}_g|} \quad (3)$$

where $\mathbf{d}_p(A)$ is the Drude particle displacement from the corresponding atomic center position, $\mathbf{r}(A)$, in response to the perturbation p. The last term is the contribution from the perturbation charge itself.

During the fitting procedure, all core atomic and Drude charges have to be adjusted to minimize the discrepancy between the QM and MM potential maps, i.e., to minimize the following function:

$$\chi_\phi^2[q_c, q_D] = \sum_{p,g}(\phi_{pg}^{QM} - \phi_{pg}^{MM})^2 \quad (4)$$

The function ϕ_{pg}^{MM} has three unknown parameters: q_c, q_D, and $\mathbf{d}_p$. The first two are the subject of the standard least-squares fitting procedure, but the Drude particle displacement, $\mathbf{d}_p$, requires special consideration. Because of the implicit charge-dependence of the displacements $\mathbf{d}_p(A)$, the system of equations

$$\frac{\partial \chi_\phi^2}{\partial q(A)} = 0 \quad (5)$$

where $q(A)$ designates either q_c or q_D assigned to an atom A has to be solved iteratively. We use the Levenberg−Marquardt algorithm,[45] specially designed to minimize χ^2 functions (see below). First, Drude displacements, $\mathbf{d}_p$, are optimized to minimize total energy of the molecular system using atomic charges from the initial guess. This is followed by an ESP fitting step using the current positions of the atoms and Drude particles. The new set of fitted charges is again used to optimize the coordinates of the Drude particles. The iterative procedure is continued until eq 5 is satisfied.

Because the charge fitting problem is underdetermined, directly solving eq 5 usually leads to partial charges having poor chemical significance.[46] This is mainly due to the small contribution of some charges to the overall electrostatic potential associated with the screening of the charge on buried atoms by atoms located on the periphery of the molecule. To optimize individual charge contribution to the minimization function, it is necessary to penalize charge deviations from "chemically intuitive" reference values, as long as the penalty does not significantly deteriorate the quality of the fit. This requirement motivated the inclusion of restraints during charge fitting, referred to as restrained electrostatic potential (RESP) fitting. The original RESP scheme of Bayly et al.[46] minimizes

$$\chi^2 = \chi_f^2 + \chi_r^2 \quad (6)$$

through adding a penalty term in one of the two following forms

$$\chi_r^2 = w\sum_A^N [q(A) - \bar{q}(A)]^2 \quad (7)$$

$$\chi_r^2 = w\sum_A^N [\sqrt{q^2(A) + b^2} - b]^2 \quad (8)$$

where w is a weighting constant. The first restraint (7) is forcing the charges $q(A)$ to their "reference" values $\bar{q}(A)$, and the second restraint (8) favors smaller magnitude charges, where b is a hyperbolic stiffness parameter.

The RESP scheme can be generalized to the presence of Drude particles yielding the following parabolic (9) and hyperbolic (10) equations, respectively,

$$\chi_r^2 = \sum_A^N \{w[q(A) - \bar{q}(A)]^2 + w_D[q_D(A) - \bar{q}_D(A)]^2\} \quad (9)$$

$$\chi_r^2 = \sum_A^N \{w[\sqrt{q^2(A) + b^2} - b]^2 + w'_D[\sqrt{q_D^2(A) + b_D^2} - b_D]^2\} \quad (10)$$

where w and w_D are weighting factors for real atoms and Drude particles, respectively; b and b_D are the respective stiffness constants; $q(A)$ is the atomic charge representing the sum of atomic core and Drude particle charges, and $\bar{q}(A)$ is the reference charge. Due to the Drude charge − atomic polarizability formal equality (1) postulated by the classical Drude oscillator model, eqs 9 and 10 effectively lead to restraining the atomic polarizabilities.

To allow for additional flexibility of the fitted charges and polarizabilities, flat well potentials can be introduced into the fitting procedure. The parabolic restraint can be used only on the amount of charge deviating from $\bar{q}$ by more than a fixed charge tolerance q_{flat}. This allows the charge q to vary at no cost within the interval $[\bar{q} - q_{flat}, \bar{q} + q_{flat}]$ and creates a restraint only when the deviation is larger than q_{flat}. The charge fitting algorithm outlined above is implemented in the module FITCHARGE in the latest release of the CHARMM program.[33,34]

2.3. Grid Generation and Placement of Perturbation Charges. Electrostatic charge fitting procedures traditionally use QM electrostatic potentials determined on a cube based grid, with the grid points placed at an equidistant separation from each other. A limitation of this approach is the nonunique definition for the selection of the axes for the cube and lack of control over significance of each grid point. One approach to eliminate the orientation dependence is to use random point generation within the defined cube. However, this solution is not ideal due to reproducibility issues. In addition, for computational efficiency it seems reasonable to avoid placing grid points in regions having minimal chemical relevance. This is especially important when the molecular shape is significantly different from a spherical form. The well-known surface reflecting molecular shape, the Connolly surface,[47] is ideal for grid point placement. Such a structurally aware grid reduces the number of required grid points, facilitating the least-squares fitting by reducing the number of linear equations to solve, as well as maximizing the information content in selected regions around a molecule with the most chemical significance. The methodology is expected to be particularly advantageous for large molecules.

In practice, a Connolly surface is generated by overlapping atomic spheres and preserving only those regions of the spheres located on the periphery of a molecule. Surface building is based on using the van der Waals radii for the atomic spheres. Multiplication of all the atomic radii of the molecule by a constant creates a Connolly surface at the desired distance from the atomic centers. Choosing several different multiplication constants allows for creation of a set of nonintersecting Connolly surfaces. Similar considerations, as discussed above regarding the grid point placement, also apply to the placement of perturbation charges around the molecule. Multiple Connolly surfaces carrying the perturbation charges are required to probe molecular polarizability at different distances from the atomic centers. The chosen solution provides a simple mechanism to control the locations and number of point charges to place.

The simplest technique for point placement on a particular Connolly surface is to generate equidistant points on the atomic spheres at a selected density and to delete all the points which are within the overlapping spheres. The algorithm for placing grid points on the atomic sphere moves an atom to the center of the coordinate system, assigns the atom a unit radius, and generates points according to the sphere equation $x^2 + y^2 + z^2 = 1$. Next, the atomic sphere is scaled to the desired radius, and coordinates of the grid points are translated accordingly. Then a back translation of the atomic center to its original position is performed with simultaneous translation of coordinates of the generated grid points. These steps are applied to each atom of the molecule, and grid points in overlapping atomic spheres are deleted. This creates the final set of points situated on the Connolly surface.

By changing the number of grid points on a particular Connolly surface one can increase or decrease the contribution of that surface to the fitted molecular properties. Typically, large contributions from the nearest and most distant Connolly surfaces should be avoided; the nearest surface is approaching distances where the deviation from the atomic point charge approximation of the electronic distribution is still nonnegligible, while the contribution of the most distant surface may shift the accuracy of charge fitting to larger distances than those at which hydrogen bonding occurs. The following rationale may be used to identify the minimal number of Connolly surfaces for the placement of perturbation charges and grid points. The first layer of perturbation charges may be placed at distances typical for hydrogen bonds followed by a layer of grid points to detect an immediate ESP change caused by a perturbation charge. Several layers of perturbation charges are needed to adequately capture the orientational dependence of molecular polarizability. Therefore at least one more layer of perturbation charges is necessary followed by one more layer of grid points. Finally one additional distant layer of grid points is necessary to resolve the molecular dipole moment, which is a far-field molecular property. These operations yield five

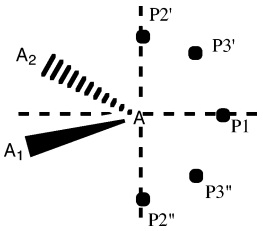


Figure 2. The placement of perturbation charges (P) for polarizability probing the region around lone pairs of sp^3 hybridized oxygen or sp^2 hybridized nitrogen atoms (A).

nonintersecting Connolly surfaces in total, two for perturbation charges and three for grid points, illustrated as (1) charges, (2) grid, (3) charges, (4) grid, and (5) grid according to increasing distance from the atomic centers. This number can be expanded as required based on computational tests.

The charge placement method employs a set of additional rules to the Connolly surface construction that further reduces the number of the charges. These additional rules are designed to assign perturbation charges to places of chemical significance, i.e., along covalently connected atoms and lone pairs where polarizability is expected to be largest. The perturbation points are prioritized into three groups according to the order of their generation. Charges placed along chemical bonds are first generated. If two atoms, A and B, are linked by a covalent bond, two perturbation points along the line A−B are generated that intersect with the corresponding Connolly surfaces on opposite sides of the bond. A newly generated charge will not be saved if it is too close to a previously generated charge. The distance criteria imposed in this study is 1.5 Å. The second group of perturbation points is created to sample regions around lone pairs on sp^3 hybridized oxygen or sp^2 hybridized nitrogen atoms, A, when these atoms have just two covalently bound neighbors (A1 and A2). In this case a bisector line is drawn in the plane of the covalent bonds of atom A, dividing the valence angle (A1-A-A2) in half (Figure 2). On the side of the lone pairs a perturbation point, P1, is generated along this line where it intersects with the Connolly surface. Two more points, P2′ and P2″, are generated on the line coming through point A, perpendicular to the plane defined by atoms A1, A, and A2 on the two opposite sides of the plane. Two other points, P3′ and P3″, are generated on the bisector line of angles P1-A-P2′ and P1-A-P2″. All these points are placed on the intersection of the Connolly surface with the corresponding line.

The last group of charges is created to close gaps between the previously placed perturbation points providing nearly equivalent coverage of the molecular shape by perturbation charges. The generation of the grids and the placement of perturbation charges are performed using the stand-alone program CGRID developed in our laboratory.

2.4. Molecular Dipole and Polarizability. The dipole moment and polarizability are important molecular properties which may be used to validate the quality of derived electrostatic parameters through comparison with experimental and/or QM values. To calculate molecular dipole moments, μ, the position of Drude particles must be optimized self-consistently. Then μ can be calculated from the sum over all charges using the following equation

$$\mu = \sum_{A}^{N} q_c(A)\mathbf{r}(A) + \sum_{B}^{N_D} q_D(A)\mathbf{r}_D(A) \tag{11}$$

where N and N_D are the number of real atoms and Drude particles, respectively, q_c and q_D are the atomic core and Drude particle charges, respectively, and $\mathbf{r}$ and $\mathbf{r}_D$ are the positions of the real atoms and Drude particles, respectively.

Polarizability is a measure of the response of a molecular system to an external electric field. Experimental measurement of isotropic molecular polarizability is conventionally conducted by measuring the refractive index η[48]

$$\alpha = \frac{3}{4\pi n}\frac{\eta^2 - 1}{\eta^2 + 2} \tag{12}$$

where n is the number density of the gas or liquid. In the present MM picture, the fast linear response of the electronic density to the excitation of a beam of light is modeled by the quasi-instantaneous readjustment of the Drude particles. The total molecular polarizability α of the MM model can be calculated in analogy with its standard QM definition, as the sum over all excitation modes of the square of the response dipole divided by the excitation energy. Being a linear-response coefficient, it can be summed over the vibration modes ν of the molecule obtained from normal-mode analysis of the polarizable MM model, complete with atomic cores and Drude particles. In terms of the Cartesian normal mode vectors $\mathbf{A}_\nu$, the components of the tensor α are

$$\alpha_{ij} = \sum_{\nu > \nu_0} \frac{\partial \mu_i}{\partial \mathbf{A}_\nu} \cdot \frac{\partial \mu_j}{\partial \mathbf{A}_\nu} \left(\frac{\partial U}{\partial \mathbf{A}_\nu} \right)^{-1} \tag{13}$$

where μ_i are components of the molecular dipole moment and U is the potential energy of the model. The isotropic polarizability values α_{iso} can be calculated as a trace of the molecular polarizability tensor, i.e., $\alpha_{iso} = 1/3(\alpha_{xx} + \alpha_{yy} + \alpha_{zz})$. This formula was implemented in the VIBRAN module of the CHARMM program. Only the very-high-frequency normal modes, attributed to Drude particle excitations, are summed, since the polarizability is modeled through the movement of these auxiliary particles. Given the much lighter mass of the Drude particles, a frequency cutoff at $\nu_0 = 5000$ cm^{-1} ensures that the normal modes associated with the vibrations of nuclei are excluded. The zero-frequency rotation and translation modes are ignored as well. It should be noted that the calculation of molecular polarizabilities requires full optimization of the molecular structure, since the harmonic approximation used in the vibrational analysis is valid only for equilibrium geometries.

2.5. Computational Details. The determination of electrostatic parameters for the polarizable force field is a multistep iterative procedure and employs both QM and empirical force field calculations. QM calculations were performed using the Gaussian 98 suite of programs.[49] Geometry optimizations were performed at the MP2(fc)/6-31G(d) level of theory for neutral species and at the MP2(fc)/6-31+G(d,p)

level for ions.[50−53] This level of theory provides molecular geometries consistent with available gas-phase experimental data, and it has been previously utilized during optimization of the CHARMM27 all-atom empirical force field for nucleic acids.[54] Geometry optimizations of adenine, cytosine, and guanine bases were conducted with the amino group planarity enforced, since it is believed that an amino group will acquire an approximately planar geometry due to hydrogen bonding with complementary bases and solvent molecules. No constraints were imposed during optimizations of other test molecules.

Selection of the appropriate level of theory for determination of the ESPs required for calculation of the charges and polarizabilities is important. Ideally, to ensure high accuracy in the ESP, a highly correlated method in conjunction with a very large basis set should be used, since computed dipole moments and molecular polarizabilities strongly depend on the level of theory and size of the basis set employed. However, when developing a force field for biomolecular systems it is necessary to have a large number of model compounds in the training set, with many of those compounds being relatively large, i.e., greater than 20 non-hydrogen atoms. All these requirements are unlikely to be satisfied in full and therefore a compromise between accuracy of the theoretical model and its computational performance is necessary.

QM calculations of the molecular electrostatic potentials were performed on MP2 optimized geometries using the B3LYP hybrid functional[55−57] and the correlation-consistent double-ζ Dunning cc-pVDZ and aug-cc-pVDZ basis sets.[58] Single-point energy B3LYP calculations were performed with the tight convergence criteria producing the target QM ESP maps. Cartesian coordinates for grid points and perturbation charges for the ESP calculation were generated by the program CGRID described above and read by Gaussian from external files. The generated QM ESP maps were extracted from the Gaussian output to use as input to the electrostatic parameter fitting by the FITCHARGE module of the CHARMM program. Restrained fitting using the RESP algorithm employed a penalty function with 10^{-5} Å^{-2} for the weighting factor. The flat well potential described above was applied allowing penalty-free deviation of charges by 0.1e in both directions from the corresponding reference values. Coordinates of the Drude particles were self-consistently adjusted after each change of the optimized charges during the charge fitting procedure to minimize the potential energy of the system, whereas coordinates of real atoms were fixed to the corresponding MP2 geometry.

Following the fitting procedure the atomic charges and polarizabilities were scaled. Polarizability scaling is necessary to reflect the reduced polarization expected for the condensed media and/or to correct for systematic underestimation of the experimental values by B3LYP/cc-pVDZ calculations (see below). Atomic charge scaling for neutral compounds can be performed to reproduce experimental or high-level QM gas-phase target molecular dipole moments. In addition, rounding of fitted and scaled charges and polarizabities to 3 decimal places was performed to facilitate their transferability while preserving the value of the net molecular charge. Technical details of the scaling and rounding procedure are described in the Supporting Information. In this study, similarly to the SWM4-DP water model, the polarization of only "heavy" (i.e. non-hydrogen) atoms is considered although the method can be easily extended to all atoms at the increase of the computational expense.

In all Drude polarizable CHARMM calculations, the Drude particles were attached to the real atoms via a harmonic spring with a force constant of a 1000 kcal/(mol·Å^2). This force constant is of sufficient magnitude to prevent large displacement of the Drude particle from its atom and thus ensure the validity of the point dipole approximation.[30] In this scenario, the atomic polarizabilities unequivocally determine the magnitude of Drude charges from eq 1. In addition, the sign of the charges on Drude particles is irrelevant due to the point dipole approximation. We chose q_D to be negative by analogy with the electron charge.

Molecular dynamics (MD) simulations were performed at 300 K and 1 atm pressure using the new velocity Verlet integrator[32] implemented in CHARMM.[33,34] The extended Lagrangian double-thermostat formalism[32] was used in all polarizable MD simulations where a mass of 0.1 amu was transferred from real atoms to the corresponding Drude particles. The amplitude of their oscillation was controlled with a separate low-temperature thermostat (at a $T = 1$ K) to ensure that their time course stays close to the SCF regimen.[32] A Nosé-Hoover thermostat with a relaxation time of 0.1 ps was applied to all real atoms to control the global temperature of the system. A modified Andersen-Hoover barostat with a relaxation time of 0.1 ps was used to maintain the system at constant pressure. Condensed-phase MD simulations were performed using periodic boundary conditions and SHAKE to constrain covalent bonds involving hydrogens.[59] A 1 fs time step was used for the extended Lagrangian polarizable MD simulation. The electrostatic interactions were treated using particle-mesh Ewald (PME) summation[60] with a coupling parameter 0.34 and 6th order spline for mesh interpolation. Nonbond pair lists were maintained out to 14 Å, and a real space cutoff of 12 Å was also used for the Lennard-Jones parameters within the atom-based force switch algorithm.[61] The long-range van der Waals correction recently implemented into CHARMM was also applied.[62,63]

Validation of the electrostatic parameters optimization method, along with proof of concept that the classical Drude oscillator model is applicable to biomolecular condensed phase simulations, was obtained via a MD simulation of DNA in solution. The preequilibrated GAGTACTC duplex DNA structure solvated in a box of water with sodium ions was taken from our previous study.[64] 1746 water molecules and 14 sodium ions were used. The solvated molecular system contains 9586 atoms including Drude particles. Starting from the CHARMM27 additive force field equilibrated system,[64] the Drude particles and then solvent molecules were minimized for 200 steps using the steepest descent (SD) algorithm with all DNA real atoms fixed. Then the minimized structure was subjected to a 20 ps NPT MD simulation with all DNA atoms, excluding the Drude particles, harmonically constrained with a mass weighted

Table 1. Summary of Experimental and QM Calculated Dipole Moments (Debye)[a]

			B3LYP/cc-pVDZ		B3LYP/aug-cc-pVDZ	
molecule		experimental		ratio		ratio
water	H_2O	1.855 ± 0.004	1.910	1.03	1.854	1.00
propane	C_3H_8	0.084 ± 0.001	0.070	0.83	0.096	1.15
isobutene	C_4H_{10}	0.132 ± 0.002	0.110	0.83	0.147	1.12
pentene-1	C_5H_{10}	0.500[b]	0.400	0.80	0.447	0.89
toluene	C_7H_8	0.375 ± 0.01	0.380	1.01	0.405	1.08
fluoromethane	CH_3F	1.858 ± 0.002	1.757	0.95	1.873	1.01
fluorobenzene	C_6H_5F	1.600 ± 0.08	1.394	0.87	1.603	1.00
chlorobenzene	C_6H_5Cl	1.690 ± 0.03	1.667	0.99	1.749	1.03
methanol	CH_4O	1.700 ± 0.02	1.576	0.93	1.680	0.99
ethanol (trans)	C_2H_6O	1.440 ± 0.03	1.483	1.03	1.598	1.11
ethanol (gauche)	C_2H_6O	1.680 ± 0.03	1.538	0.92	1.727	1.03
dimethyl ether	C_2H_6O	1.300 ± 0.01	1.200	0.92	1.306	1.00
tetrahydrofuran	C_4H_8O	1.750 ± 0.04	1.695	0.97	1.865	1.07
trimethylamine	C_3H_9N	0.612 ± 0.003	0.450	0.74	0.591	0.97
dimethyl sulfide	C_2H_6S	1.554 ± 0.004	1.471	0.95	1.620	1.04
ethanethiol (trans)	C_2H_6S	1.580 ± 0.08	1.557	0.99	1.669	1.06
acetaldehyde	C_2H_4O	2.750 ± 0.006	2.600	0.95	2.965	1.08
acetone	C_3H_6O	2.880 ± 0.03	2.789	0.97	3.167	1.10
acetic acid	$C_2H_4O_2$	1.700 ± 0.03	1.613	0.95	1.826	1.07
methylformate	$C_2H_4O_2$	1.770 ± 0.04	1.778	1.00	1.931	1.09
dimethylamine	C_2H_7N	1.010 ± 0.02	0.890	0.88	1.043	1.03
imidazole	$C_3H_4N_2$	3.800 ± 0.4	3.668	0.97	3.773	0.99
pyrazole	$C_3H_4N_2$	2.200 ± 0.01	2.217	1.01	2.316	1.05
pyridine	C_5H_5N	2.215 ± 0.01	2.050	0.93	2.299	1.04
trimethyl phosphate	$C_3H_9O_4P$	3.18[c]	3.475	1.09	3.723	1.17
			AVER	0.94		1.05
			SD	0.08		0.06

[a] Ratios are calculated with respect to experimental values. Data on water were not included in the calculation of the average ratio (AVER) and the standard deviation (SD). Calculations were performed using MP2(fc)/6-31G(d) geometries except for water, for which experimental gas-phase geometry was utilized. Experimental dipole moments are from ref 35. The experimental uncertainties are indicated where available. [b] Questionable results because of undetermined error sources. [c] Liquid-phase measurements, which may have large errors because of association effects.

force constant of 2 kcal/(mol·Å^2) to equilibrate the solvent around the DNA. The final structure from that simulation was then subjected to two 200 step SD minimizations, first for Drude particles and then for all atoms prior to initialization of the production trajectory. The simulation was run for 1000 ps with coordinates saved every 2 ps for analysis.

3. Results and Discussion

3.1. QM Calculations of Molecular Dipoles and Polarizabilities. The performance of B3LYP/cc-pVDZ and B3LYP/aug-cc-pVDZ calculations on MP2 optimized geometries in reproducing experimental gas-phase dipole moments and polarizabilities was first verified. Model compounds for which experimental values of molecular polarizabilities and/or dipole moments are available were chosen to represent different chemical classes as well as building blocks of biomolecules. The experimental and B3LYP calculated dipole moments are summarized in Table 1. The data indicate that both B3LYP/cc-pVDZ and B3LYP/aug-cc-pVDZ single-point energy calculations generally provide reasonable estimates of the dipole moment magnitudes. In many cases some underestimation of the experimental values can be noted for the B3LYP/cc-pVDZ calculations, whereas augmenting this basis set with diffuse functions in general results in a slight overestimation of the gas-phase molecular dipole moments. The average ratio of calculated to experimental dipole moments for compounds listed in Table 1, excluding water, was 0.94 and 1.05 for the cc-pVDZ and aug-cc-pVDZ basis sets, respectively.

Experimental and calculated molecular polarizabilities for selected compounds are summarized in Table 2. In all cases the experimental gas-phase molecular polarizabilities are underestimated by the B3LYP/cc-pVDZ calculations. The degree of the discrepancy is the largest for water (almost 50%) and substantially smaller for the remaining compounds. Importantly, for the majority of compounds the ratio of B3LYP to experimental polarizabilities is quite uniform, with an average value of 0.83 ± 0.06. As data in Table 2 demonstrate the substantial deviation of calculated from experimental values of polarizabilities can be corrected by augmenting the cc-pVDZ basis set by diffuse functions, which provides better quantitative agreement with the experimental data. Thus, B3LYP/ aug-cc-pVDZ QM calculations may be considered the method of choice for calculation of the ESP data for electrostatic parameter fitting. However, QM calculations with the aug-cc-pVDZ basis set are much more computationally expensive and become impractical for large and flexible molecules such as nucleosides. For example, for the guanine nucleoside cc-pVDZ and aug-cc-pVDZ basis sets consist of 331 and 554 basis functions,

Table 2. Summary of Experimental and Calculated Molecular Polarizabilities (Å^3)[a]

				B3LYP/cc-pVdz		B3LYP/aug-cc-pVdz	
molecule		experimental	ahp Miller		ratio		ratio
water	H_2O	1.45	1.41	0.78	0.54	1.39	0.96
ethane	C_2H_6	4.47	4.44	3.61	0.81	4.31	0.96
propane	C_3H_8	6.29	6.28	5.27	0.84	6.14	0.98
isobutane	C_4H_{10}	8.14	8.11	6.94	0.85	7.98	0.98
butane	C_4H_{10}	8.20	8.11	6.94	0.85	7.99	0.97
pentene-1	C_5H_{10}	9.65	9.76	8.42	0.87	9.88	1.02
pentene-2	C_5H_{10}	9.84	9.76	8.61	0.87	10.00	1.02
cyclohexane	C_6H_{12}	11.00	11.01	9.60	0.87	10.68	0.97
benzene	C_6H_6	10.00	10.43	8.57	0.86	10.35	1.04
toluene	C_7H_8	11.80	12.27	10.48	0.89	12.39	1.05
fluoromethane	CH_3F	2.97	2.52	1.92	0.64	2.53	0.85
fluorobenzene	C_6H_5F	10.30	10.34	8.57	0.83	10.35	1.00
chlorobenzene	C_6H_5Cl	12.25	12.36	10.20	0.83	12.51	1.02
methanol	CH_4O	3.32	3.25	2.41	0.73	3.19	0.96
ethanol	C_2H_6O	5.11	5.08	4.11	0.80	5.04	0.99
dimethyl ether	C_2H_6O	5.29	5.08	4.14	0.78	5.09	0.96
ethanethiol	C_2H_6S	7.41	7.44	5.70	0.77	7.30	0.99
acetaldehyde	C_2H_4O	4.59	4.53	3.69	0.80	4.58	1.00
acetone	C_3H_6O	6.39	6.37	5.28	0.83	6.37	1.00
acetic acid	C_2H4O_2	5.10	5.17	4.05	0.79	5.14	1.01
methylformate	$C_2H_4O_2$	5.05	5.17	4.06	0.80	5.12	1.01
pyrazole	$C_3H_4N_2$	7.23	7.72	5.76	0.80	7.32	1.01
pyridine	C_5H_5N	9.18	9.73	7.87	0.86	9.55	1.04
adenine	$C_5H_5N_5$	13.10	15.05	11.82	0.90	14.44	1.10
cytosine	$C_4H_5N_3O$	10.30	11.12	9.32	0.90	11.60	1.13
guanine	$C_5H_5N_5O$	13.60	15.68	12.56	0.92	15.40	1.13
thymine	$C_5H_6N_2O_2$	11.23	12.11	10.34	0.92	12.44	1.11
trimethyl phosphate	$C_3H_9O_4P$	10.86	10.86	9.27	0.85	11.23	1.03
				AVER	0.83		1.01
				SD	0.06		0.06

[a] Ratios are calculated with respect to experimental values. Data on water were not included in the calculation of the average ratio (AVER) and the standard deviation (SD). Calculations were performed using MP2(fc)/6-31G(d) geometries except for water, for which experimental gas-phase geometry was utilized. Experimental polarizabilities are from ref 35. When several experimental estimates were available, the most recent data were used. Calculated empirical polarizabilities are those obtained from atomic component values using the additive ahp scheme from ref 65.

respectively, and single-point energy B3LYP calculations using 1.5 GHz CPU take around 2 and 12 h, respectively. Approximately 100 such calculations are required for each conformation of the model compound to generate the perturbed maps of the ESP. Thus, for larger systems, B3LYP/cc-pVDZ calculations combined with the appropriate scaling, can be used for determination of ESP target data.

Based on the data in Tables 1 and 2, the following QM approach is suggested for determination of the electrostatic parameters. B3LYP/aug-cc-pVDZ calculations will be used for the calculation of QM response electrostatic potential maps. The scale factor 0.724 is applied to the fitted values of atomic polarizabilities to reflect the reduced polarizability required for the condensed phase simulations.[30] For larger molecules B3LYP/cc-pVDZ calculations are recommended. The underestimation of experimental polarizabilities by this level of theory must be corrected by applying the inverse of the average ratio of calculated to experimental gas-phase molecular polarizabilities 1/0.83 to the empirical values of $\alpha(A)$ as an additional scale factor. Combining this scale factor with the 0.724 factor yields an overall scaling factor of 0.87 that should be applied to the atomic polarizabilities after the fitting procedure. In addition, care must be taken when applying this level of theory due to the presence of outliers with respect to the calculated molecular polarizabilities (e.g. fluoromethane, Table 2).

3.2. Reference Values for Atomic Charges and Polarizabilities. Ideally, charges and polarizabilities can be determined via free fitting, as the resultant charges and polarizabilities are theoretically invariant to the initial guess. However, free fitting often results in physically unrealistic values of the charges[46] and polarizabilities. For instance, the polarizability tends to be attracted to one or a few atoms typically in the central region of a molecule and negative charges are often obtained for hydrogen atoms attached to aliphatic carbons from such fitting (see Table 3S of the Supporting Information). Therefore, the use of restraints is necessary to obtain chemically meaningful charges and polarizabilities. The task of careful selection of initial values for charges and polarizabilities becomes an important step since in restrained fitting these charges and polarizabilities are used as reference values in eq 9.

The classical Drude oscillator formalism employs the concept of atomic polarizability and assigns a corresponding

Table 3. Initial Values of Atomic Polarizabilities

atom (group)		polarizabilities (Å^3)	
symbol	Miller[a]	unscaled[b]	scaled[c]
$CH_3(sp^3)$	CTE+3H	2.222	1.844
$CH_2(sp^3)$	CTE+2H	1.835	1.523
$CH(sp^3)$	CTE+H	1.448	1.202
$C(sp^2)$	CTR	1.352	1.122
$CH(sp^2)$	CTR+H	1.739	1.443
$CH_2(sp^2)$	CTR+2H	2.126	1.765
$C(sp^2$, br)	CBR	1.896	1.574
−OH	OTE+H	1.024	0.850
−O−	OTE	0.637	0.529
=O	OTR4	0.569	0.472
$-O^{(-)}$	OTA[d]	0.858	0.712
$NH_2(sp^3)$	NTE+2H	1.738	1.443
$NH(sp^3)$	NTE+H	1.351	1.121
$N(sp^3)$	NTE	0.964	0.800
$NH_2(sp^2)$	NPI2+2H	1.864	1.547
$NH(sp^2)$	NPI2+H	1.477	1.226
$N:(sp^2)$	NTR2	1.030	0.855
P	PTE	2.063[e]	1.712

[a] Hybrid names used by Miller[65] are given in the second column along with the number of attached hydrogen atoms. [b] Unscaled united atom values of ahp polarizabilities can be used as reference values in the charge fitting to the B3LYP/aug-cc-pVDZ ESP potentials. [c] United atom values of ahp polarizabilities were scaled by the factor 0.83 reflecting the underestimation of experimental gas-phase polarizabilities by B3LYP/cc-pVDZ calculations. [d] The anionic oxygen atom type (OTA) was not present in the original Miller's ahp scheme and was added to reflect its higher polarizability compared to the ester oxygen (OTE). [e] The atomic polarizability of phosphorus atom was substantially changed from its original value in Miller's ahp scheme (1.538) to obtain a better estimate of the B3LYP molecular polarizabilities for phosphates. See details in the text.

unique parameter to individual atoms. Unlike molecular dipoles and polarizabilities, the atomic charges and polarizabilities of atoms in molecules are not well-defined quantities that can be unambiguously determined. A variety of different schemes have been proposed to obtain values of atomic polarizabilities using both QM and experimental molecular polarizabilities.[65−71] They can be categorized into element, group, bond, and hybrid polarizability schemes.[65] These methods can be also classified as additive, where the molecular polarizability is considered as a sum of atomic contributions, and nonadditive, which usually rely on the iterative solution of nonlinear equations to obtain molecular polarizability values. Nonadditive atomic polarizability schemes of Applequist[66] and Thole[67] have been commonly used to obtain polarizabilities in polarizable force fields employing point dipole induction models.[21,25,72] However, the atomic polarizability values from these as well as the majority of other nonadditive schemes can be considered as parameters of that model and, therefore, may be inappropriate for use in other polarization models. To the best of our knowledge, no atomic polarizability values for molecules other than water have been suggested for the classical Drude oscillator model.

Based on the classical Drude oscillator formalism, the atomic polarizability, $\alpha(A)$, of atom A is theoretically independent of the polarizability of the other atoms. However, in practice, atomic polarizabilities will be interdependent. Such interdependence is systematically minimized in force fields through assignment of atom types, which are expected to show transferability within a given class of chemical compounds. To facilitate transferability, the initial guess for atomic polarizability should also be taken from an additive polarizability scheme. The atomic hybrid polarizability (ahp) scheme of Miller[65] is a good example of such a scheme, and, after some modifications, it has been used in this work to provide an initial guess of atomic polarizabilities.

In the original Miller scheme the atomic polarizability contributions were obtained from least-squares fitting to experimental gas-phase molecular polarizabilities for approximately 400 organic compounds.[65] In addition, the Miller atomic hybrid polarizabilities α_A(ahp) depend not only on the identity but also the hybridization state of a particular atom, which is similar to an empirical force field atom type concept. According to the Miller scheme, molecular polarizabilities are obtained by summing up atomic hybrid polarizabilities α_A(ahp).[65] In most cases the sum of ahp polarizabilities is very close to experimental values of molecular polarizabilities (Table 2). Values of α_A(ahp) are available for most atomic hybrids which are encountered in biological compounds. Besides, the additivity of the model allows summing up atomic polarizabilities for functional groups and easy adjustment of α_A(ahp) values by applying scale factors (see below). Thus, the α_A(ahp) values were used as the initial estimates for atomic polarizabilities in the electrostatic parameter fitting procedure. It should be noted that molecular polarizability in the classical Drude oscillator model is not the additive sum of the atomic polarizabilities. However, α_A(ahp) values can still be used as an initial guess and then adjusted during the fitting procedure.

Since no polarizability is assigned to hydrogens in the current model, the initial guess for atomic polarizabilities was constructed by adding ahp polarizabilities of the heavy atoms and their covalently bound hydrogen atoms, thereby constituting the "united atom" approximation for atomic polarizability. If B3LYP/cc-pVDZ calculations are used for the calculation of the response ESP maps, these "united atom" atomic polarizabilities have to be scaled by the factor of 0.83 introduced above. Thus, their sum will reproduce the molecular polarizability at the B3LYP/cc-pVDZ level of theory, which will avoid biasing during the restrained fitting. Following the fitting procedure the scale factors described above are applied. No initial atomic polarizability scaling is required if the B3LYP/aug-cc-pVDZ level of theory is employed. Both unscaled and scaled atomic polarizabilities for biologically important atom types are given in Table 3.

A few small adjustments to Miller's polarizabilities were introduced to improve agreement with the QM results. A new hybrid atom type for anionic oxygen (OTA) was introduced to reflect the substantially higher electronegativity of anions and was derived based on the polarizability of the acetate anion. The atomic polarizability of the P atom was substantially increased to obtain a better estimate for different phosphate species, especially those with methyl groups attached to ester oxygen atoms. The comparison between B3LYP/aug-cc-pVDZ calculated and the sum of corrected ahp values from Table 3 for a few relevant compounds are

given in Table 1S of the Supporting Information. Overall, the satisfactory agreement between QM and empirical values of molecular polarizabilities was obtained for most compounds. It should be emphasized that these values represent initial estimates of atomic polarizability values, which are then adjusted in the fitting procedure. Since in the classical Drude oscillator model the atomic polarizability is directly related to a partial charge assigned to the Drude particle, the atomic and Drude charges are determined in one step through charge fitting to the series of perturbed ESP maps obtained from QM calculations.

It is also necessary to have a good choice of reference atomic charges used in eq 9. Three different initial guesses were considered, namely CHARMM27,[54] NBO,[73] and Mulliken[74] charges. Initial guesses for the polarizabilities were kept the same in all three cases. In general, it was found that the reference charge selection affects the final atomic charges but not the polarizability values. An example for such calculations for the cytosine base is given in Table 2S of the Supporting Information. It was also found that some molecular properties, such as residue-water interaction energies, were not very sensitive to the reference atomic charge scheme chosen (see also Table 2S). In addition, NBO and Mulliken atomic charges as initial guesses have some drawbacks. For instance, NBO charges on aliphatic hydrogens are often very large (0.20−0.30) even in nonpolar compounds such as alkanes. This results in the undesirable overpolarization of the C−H bonds. Mulliken charge assignment has a variety of shortcomings including a strong dependence on the basis set and unrealistic population assignment in the presence of diffuse and polarization functions, among others.[75] Moreover, charges obtained from QM models are conformation dependent, thereby complicating selection of the initial guesses. Therefore, the CHARMM27 charges were selected as the initial guess. This choice facilitates transfer of already existing internal and Lennard-Jones CHARMM27 force field parameters to the polarizable model, although optimization of these parameters will be necessary due to the new charge distribution and inclusion of explicit polarization. A disadvantage of using CHARMM27 charges as the initial guess is that they are limited to molecules for which the force field parametrization has been performed. Currently there is no simple scheme to derive these charges for new molecular residues although fitting procedures based on electronegativity equalization schemes[5] may be useful.

3.3. Atomic Charge and Polarizability Derivation for Model Compounds. The methodology described above has been applied for the derivation of the electrostatic parameters for the model compounds shown in Figure 3, which comprise the most important functional moieties of nucleic acids. Details of the fitting procedure will be discussed for the cytosine base. The geometry was optimized at the MP2/6-31G(d) level of theory with the amino group geometry constrained to be planar. Three Connolly surfaces for grid points and two surfaces for perturbation charges were generated (see Figure 4). The total number of grid points and perturbation ions was 1327 and 57, respectively. Surface parameters and the number of points placed on each surface

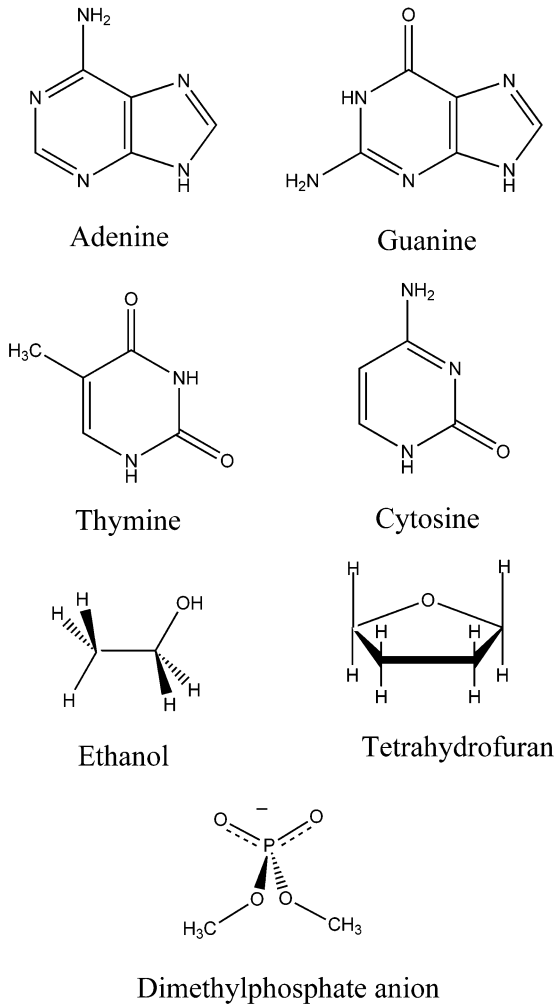


Figure 3. Model compounds used for the preliminary parameter development of the Drude polarizable CHARMM force field for nucleic acids.

are shown in Table 4. In total, 57 QM ESP maps were calculated, one for each placement of the perturbation charge, at the B3LYP/aug-cc-pVDZ level of theory. For other model compounds a similar strategy for the grid generation has been used. Reference partial atomic charges and atomic polarizabilities for the fitting procedure are shown in Table 4S of the Supporting Information. Fitting was performed under the RESP parabolic restraint with a 10^{-5} Å^{-2} weighting factor. Charges and polarizabilities were restrained separately to their corresponding initial values. The restraint was combined with a flat well potential with a half-width of 0.1e centered at the initial charge value. The restrained ESP fitting produced a final RMS error of $6.8 \cdot 10^{-4}$ e/Å with respect to the B3LYP potential, which is close to the RMS error of $3.9 \cdot 10^{-4}$ e/Å for the unrestrained fitting. The resultant atomic polarizabilities were then scaled by factor 0.724 to reflect reduced polarizability expected for the condensed phase (see above). This was followed by scaling the atomic charges to reproduce the B3LYP/aug-cc-pVZD gas-phase dipole moment. Final optimized partial atomic charges and atomic polarizabilities are shown in Table 4S of the Supporting Information for cytosine as well as other model compounds. In general, the fitted values of atomic charges decreased in

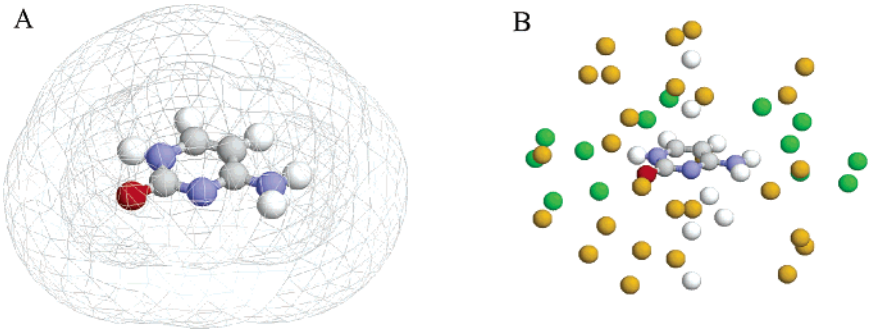


Figure 4. Electrostatic potential grid created based on Connolly surfaces around the cytosine base (A) and placement of perturbation charges around this molecule (B). For visualization purposes only grid 1 and grid 2 Connolly surfaces are displayed (see Table 4 for surface parameters). Perturbation charges are placed along chemical bonds (green spheres), around the N lone pair (white spheres), and in gaps between other charges (brown spheres) on the corresponding Connolly surfaces. The surfaces are not displayed in part B.

Table 4. Parameters Used for the Grid Point and Perturbation Charge Generation for the Cytosine Base

Connolly surface	size factor, f[a]	density factor, d	# of points
perturbation charges 1	2.2	1.1	46
grid 1	3.0	1.3	505
perturbation charges 2	4.0	0.1	11
grid 2	5.0	0.6	564
grid 3	6.0	0.2	258

[a] Size factors (multiplication constants) multiplied by the van der Waals radii of corresponding atoms (1.2 Å for H, 1.5 Å for C, 1.4 Å for N, 1.35 Å for O) define the distance from that atom to the generated Connolly surface. Density factors determine the relative density of points on a particular Connolly surface.

magnitude with respect to initial values. This is expected as polarization was taken into account implicitly in the CHARMM27 charges, whereas explicit polarization in the Drude polarizable CHARMM force field naturally results in decreased values of the atomic charge values in most cases (see Table 4S). In the fitting procedure redistribution of the polarizability also takes place. As a rule, the polarizability tends to accumulate on atoms that are located near the center of the molecules, which is mainly an artifact of the fitting procedure resulting from the inadequate sampling of the electrostatic potential for the "buried" central atom(s). The use of restraints allows a substantial reduction of this effect (see Table 3S).

In situations where model compounds have several stable conformers (e.g. dimethyl phosphate or ethanol), the grid generation and QM ESP calculations has been performed for the most relevant (i.e. the lowest-energy) nonequivalent conformers. For instance, gg and tg conformers have been used for the dimethyl phosphate electrostatic parameter derivation. The tt conformer of the DMP anion has not been taken into consideration since it lies substantially higher in energy (~3 kcal/mol compared to gg), and the corresponding conformation is not populated in nucleic acids. Then the fitting procedure was performed by simultaneously including the target ESP data for both conformers resulting in one set of charges and polarizabilities. The procedure tends to decrease the quality of the fit (i.e. increases the RMSE) but minimizes bias in electrostatic parameters toward one conformer.

The quality of the derived set of atomic charges and polarizabilities can be evaluated by the ability of the Drude polarizable model to reproduce the gas-phase dipole and polarizability. The corresponding data are presented in Table 5. The atomic charges and polarizabilities derived from the fitting procedure reproduce the reference experimental (if available) or QM gas-phase dipole moments within 3% for nucleic bases and tetrahydrofuran (THF) and within 10% for the stable conformers of the dimethyl phosphate anion (DMP) and ethanol, for which multiple conformers were used for fitting. The scaling of atomic polarizabilities results in the increase in the molecular dipole moments for the model compounds, which is corrected by the adjustment of partial atomic charges for all compounds except DMP. Since DMP is an ion the direct uniform charge scaling is not applicable. However the reproduction of the gas-phase dipole moment for DMP was considered to be not essential as the dipole moment of ionic species is undetermined, i.e., depends on the molecular orientation.

Ideally, the experimental and/or high-level QM gas-phase dipole moment values should be reproduced in Drude calculations for all neutral model compounds through the application of the atomic charge fitting. However, since the molecular dipole moment is also a function of the molecular geometry it may be necessary to perform fitting in an iterative fashion following adjustment of Lennard-Jones and internal parameters. In general, it should be noticed that the polarizable Drude model better reproduces gas-phase molecular dipole moments than additive CHARMM, which often overestimate them (Table 5) in order to reproduce condensed-phase properties of model compounds.

Concerning the molecular polarizabilities, those from the Drude model are underestimated by ca. 1−4% for nucleic bases, ~5% for ethanol and THF, and ca. ~14% for DMP compared to reference QM B3LYP/aug-cc-pVDZ data (Table 5). These differences may be a consequence of the different approaches used for determination of the QM and classical molecular polarizabilities as well as an artifact of the fitting procedure, which does not allow proper sampling for "buried" atoms. Future work will address the causes of these differences. However, the relative magnitudes of experimental and QM molecular polarizabilities are well reproduced by Drude polarizable CHARMM calculations.

Table 5. Summary of Calculated Molecular Dipole Moments and Polarizabilities for a Set of Model Compounds

					DRUDE	
model[a]	experiment	B3LYP/aug-cc-pVDZ	B3LYP/cc-pVDZ	CHARMM27[b]	unscaled[c]	scaled[d]
			Molecular Dipole Moments (Debye)			
ADE		2.43	2.35	2.94	2.43	2.25
CYT		6.72	6.24	7.88	6.70	6.92
GUA		6.97	6.84	7.61	6.96	6.81
THY		4.54	4.11	4.51	4.51	4.65
THF	1.75	1.87	1.70	2.34	1.91	1.72
DMP gg		5.43	4.88	*e*	5.42	*e*
DMP gt		4.82	4.31	*e*	4.75	*e*
DMP tt		2.70	2.35	*e*	2.43	*e*
ETOH t	1.44	1.60	1.48	2.36	1.70	1.68
ETOH g	1.68	1.73	1.54	2.40	1.81	1.73
			Isotropic Molecular Polarizability (Å^3)			
ADE	13.10	14.44	11.82		13.98	10.36
CYT	10.30	11.60	9.32		11.19	8.16
GUA	13.60	15.40	12.56		15.29	11.09
THY	11.23	12.44	10.34		12.09	8.72
THF		7.81	6.76		7.44	5.70
DMP gg		10.83	8.03		9.26	6.63
DMP gt		10.86	8.07		9.27	6.64
DMP tt		10.91	8.13		9.38	6.68
ETOH t	5.11	5.04	4.11		4.76	3.48
ETOH g		5.04	4.10		4.79	3.48

[a] ADE − adenine, CYT − cytosine, GUA − guanine, THY − thymine, THF − tetrahydrofuran, DMP − dimethyl phosphate anion, ETOH − ethanol. [b] CHARMM27 calculations were performed using the standard additive CHARMM27 force field for nucleic acids[54] and fully optimized geometry of the model compounds. [c] The fitted partial atomic charges and atomic polarizabilities were not scaled. MP2 optimized geometry was used, and only positions of Drude particles were optimized to make a direct assessment of the quality of the fit to the QM ESP. [d] The fitted atomic polarizabilities were scaled by a factor of 0.724. The full geometry optimization of model compounds was performed for molecular dipole and polarizability calculations. [e] Molecular dipole of ionic species depends on the molecular orientation and therefore cannot be compared directly for DMP using the QM and CHARMM optimized geometries.

3.4. Polarizable Condensed-Phase Molecular Dynamics Simulation of DNA. The ability to perform condensed-phase simulations of biologically relevant molecular systems is the ultimate goal of any macromolecular empirical force field. As a proof of concept of the Drude model and the presented electrostatic parameter derivation procedure, a MD simulation of the DNA duplex GAGTACTC in a SWM4-DP water box, including sodium counterions, was performed. Atomic charges and polarizabilities (Table 4S of the Supporting Information) of the model compounds shown in Figure 3 have been used. The values of atomic charges for terminal atoms had to be adjusted upon creating the covalent bonds in an oligonucleotide. For example, in pyrimidine bases the charge of the hydrogen H1 was summed into the N1 charge. The experimental polarizability of the Na^+ ion (0.157 Å^3)[76] scaled by 0.724 (see above) has been used.

Lennard-Jones (LJ) and bonded parameters for model compounds were also required for the simulation. Preliminary optimization of LJ parameters was performed to reproduce minimum interaction energies and geometries of model compounds with water and for base pairs. High-level ab initio LMP2/cc-pVQZ//MP2/6-31G(d) calculations for nucleic bases as well as THF and LMP2/aug-cc-pVTZ//MP2/6-31+G(d) calculations for the DMP were used as reference data.[77] In both the QM and empirical calculations the respective monomer geometries were fixed. The QM interaction distances were determined from the constrained MP2 optimization, and then interaction energies were calculated at the LMP2 level with a larger and more flexible basis set. In CHARMM the corresponding hydrogen bonding distances were sampled at a resolution of 0.01 Å, and the minimum interaction energies were obtained. The positions of Drude particles were self-consistently adjusted at every step of the potential energy scan. The hydrogen bond angles have been also sampled for some orientations of nucleic acid bases with water molecule, where it acts as a hydrogen bond donor to the carbonyl oxygen of the base.[54] The LJ parameters for model compounds were adjusted to minimize the difference between QM and CHARMM interaction energies and to reduce RMS deviation for both distances and energies across different complex orientations. Comparison of the present results with available CHARMM27 additive force field data[54] was also performed. The final values of the minimized base pair interaction energies are summarized in Table 5S of the Supporting Information, and those for the nucleic acid bases, DMP and THF with water are given in Tables 6S and 7S of the Supporting Information. For the nucleic acid bases the LJ parameter optimization was performed mainly based on the ab initio LMP2/cc-pVQZ//MP2/6-31G(d) base pairing interaction energies. The optimization of Lennard-Jones parameters for the DMP and THF was performed based on the LMP2//MP2 data on interactions with water.

Overall, the Drude model reproduces the QM base-base and model compound−water interactions at a level similar to that of CHARMM27. Base−base interactions are generally

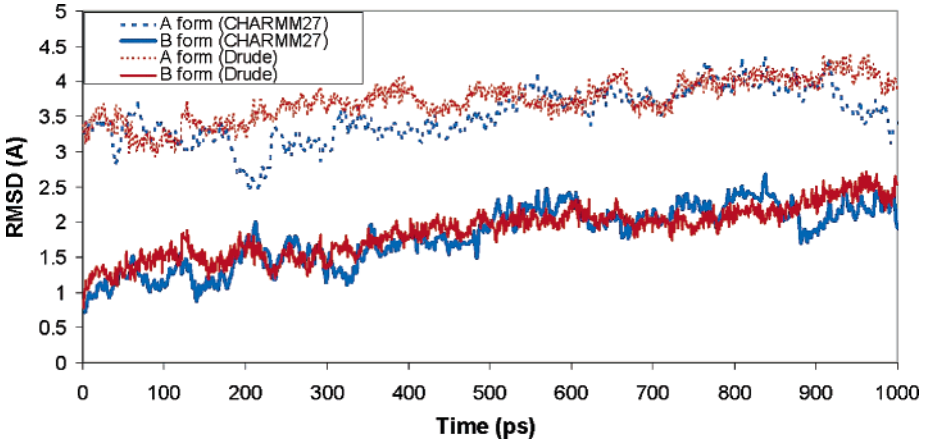


Figure 5. Root-mean-square deviation (RMSD) of heavy atom for six central residues of the GAGTACTC DNA duplex with respect to crystal structures of the A and B forms of DNA (dashed and solid lines) during the course of the MD simulation. RMSD from simulations using CHARMM27 pairwise additive force field and TIP3P water model (blue lines) and Drude polarizable CHARMM force field with SWM4-DP water model (red lines) are presented.

improved in the new model (Table 5S of the Supporting Information), although significant differences (e.g. the GG1 and GG2 interactions) that occur with CHARMM27 are still present. With the base−water interaction energies (Table 6S) the Drude model yields improved results for cytosine, while CHARMM27 is better for adenine, guanine, and thymine. Notably, the most significant improvement by the Drude model versus CHARMM27 is seen with dimethyl phosphate (Table 7S). This improvement includes the relative energies of the different interaction orientations and suggests that some of the largest gains in accuracy associated with the use of polarizable models may be expected with charged species. Thus, preliminary optimization of LJ parameters combined with electrostatic parameters determined using the methodology presented here yield a model that reproduces interaction energies to a level similar to CHARMM27 as judged by the reproduction of ab initio QM data. It is anticipated that more rigorous optimization of the LJ parameters, including the application of parameters obtained from condensed phase studies of small model compounds, will yield improved results for the studied interactions.

Selected internal parameters were also optimized. Equilibrium bond length and angle values were adjusted to reproduce average values of the bases targeting a survey of nucleic acid crystal structures from the NDB.[78] Force constants were adjusted to reproduce MP2/6-31G(d) frequencies scaled by 0.9434.[79] Torsional parameters for the C−O−P−O dihedral angles were adjusted to reproduce MP2/6-31+G(d) conformational energies for the DMP^- ion.[80]

Using this zero-generation polarizable model a MD simulation of the GAGTACTC DNA duplex in a box of SWM4-DP water box with sodium counterions was run for 1 ns. The results of this simulation were compared with a previously published MD simulation of this DNA sequence using the additive CHARMM27 force field and the TIP3P water model.[64] RMSD values for non-hydrogen atoms of the six central residues of the DNA duplex were calculated with respect to canonical A and B DNA structures for this duplex over the course of the simulation. These data are shown in Figure 5 and demonstrate that DNA structures from both Drude polarizable and CHARMM27 additive force field simulations remain closer to the B form versus the A form of DNA. More specific information may be obtained from the analysis of the base pairing interactions e.g. through monitoring of the N1···N3 distance between Watson−Crick bases. The average values over the course of the simulation are presented in Table 8S of the Supporting Information. These data indicate that the current Drude polarizable model gives reasonable agreement with the additive CHARMM27 force field. However, larger fluctuations of the N1···N3 distances for most base pairs from the Drude polarizable DNA simulation occur. Another test of the validity of the performed simulation is the analysis of the backbone dihedral parameters. Probability distributions of the backbone dihedral angles are presented in Figure 1S of the Supporting Information. The plots indicate that reasonable backbone dihedral angle probability distributions are obtained, being similar to those from the CHARMM27 MD simulation as well as data from a survey of the nucleic acid data bank (NDB)[78] in most cases. However, for some dihedrals there are noticeable differences in the relative conformer populations. The observed differences in the N1···N3 distances and in the dihedral distributions emphasize potential differences in molecular properties associated with the polarizable versus the additive CHARMM27 model. However, such differences also indicate the need for careful force field parametrization to be performed in order to properly implement a classical Drude based polarizable force field; such efforts are ongoing in our laboratory.

Importantly, the MD simulation was run in a reasonable amount of time. For instance, 50 ps of the extended Lagrangian DNA simulation took 23 CPU h to run on a 3 GHz Pentium IV computer. For comparison, the additive CHARMM27 force field simulation for the same system consumed approximately 5 h of CPU time. It should be noted that a larger time step of 2 fs was used in the CHARMM27 simulation, whereas a 1 fs time step was used for the extended Lagrangian simulation and that the TIP3P water model contains only 3 particles versus 5 particles in the SWM4-DP model.

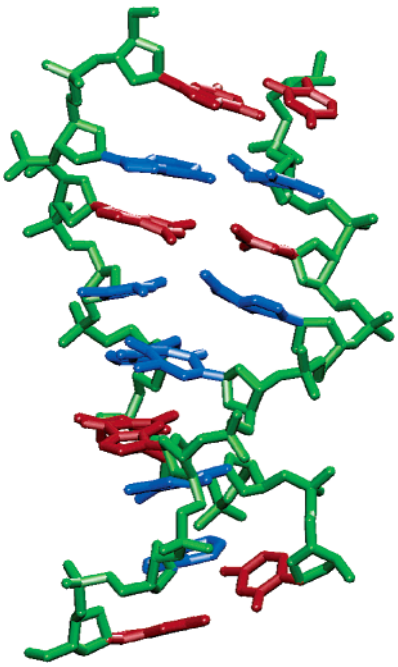


Figure 6. GAGTACTC duplex DNA molecule. Final structure from the 1000 ps Drude polarizable extended Lagrangian molecular dynamics simulation in a box of SWM4-DP water with sodium counterions.

Thus, applying the presented electrostatic parameter optimization methodology along with preliminary optimization of selected internal and LJ parameters produces a polarizable model of DNA that yields results comparable to those from the additive CHARMM27 nucleic acid force field. The final solvated DNA structure after 1000 ps is shown in Figure 6. In addition, the simulation system remained stable over the course of the trajectory as evidenced by a lack of significant drift in the molecular volume and potential energy of the system (not shown). These results indicate the applicability of the developed method for derivation of the electrostatic parameters and the resulting classical Drude based model of electronic polarizability for MD simulation studies of biological macromolecules in the condensed phase.

4. Conclusions

A general procedure for the determination of the electrostatic parameters for the classical Drude oscillator polarizable model, the partial atomic charges and atomic polarizabilities, is presented. This task is performed through fitting to a series of QM electrostatic potentials for a test molecule obtained in the presence of perturbation charges. Consequently, the partial atomic and Drude particle charges, where the latter are responsible for the atomic polarizabilities, are determined simultaneously in a single step. The QM level of theory for the electrostatic parameter fitting was determined via comparison of dipole moments and molecular polarizabilities for a variety of small organic molecules. The B3LYP/aug-cc-pVDZ level was selected based on this comparison, although the computationally less expensive B3LYP/cc-pVDZ level may be used for larger molecules with the appropriate scaling factors. Special emphasis was placed on the creation of the grid required for the electrostatic potential and placement of the perturbation charges. Versus the commonly used cubic grid, an approach based on the placement of grid points on a predetermined series of nonintersecting Connolly surfaces was developed. This approach reduces of the number of grid points by optimizing their placement around the test molecules. Perturbation charges were placed along chemical bonds, lone pairs, and in gaps between previous placed charges to provide equal coverage of the corresponding Connolly surfaces.

Consistent with previous work,[46] it was found that restraints were needed during the fitting procedure to avoid unphysical atomic charges and polarizabilities. Thus, generic reference values for the atomic charges and polarizabilities become important. Atomic charges from the CHARMM27 additive force field and atomic polarizabilities obtained using the atomic hybrid polarizability scheme of Miller[65] were identified as suitable reference values. However, it was necessary to adjust Miller's ahp atomic polarizability values to take into account the united-atom polarizability model used in this study, i.e., polarizabilities of H atoms are added to that of the corresponding heavy atom to which they are bonded. Furthermore, fitted values of atomic polarizabilities were scaled to reflect the reduced polarization which appears to take place in the condensed phase.[30,36−41] Studies on the SWM4-DP[30] and other polarizable models[36,37] have shown such scaling to be necessary to reproduce condensed-phase properties. In this study we used the same scaling factor as was used for the SWM4-DP water model. Studies are underway in our laboratory to determine if the scale factor based on the water molecule (0.724) is appropriate for other small organic molecules. These tests will involve the ability of the scaling approach to reproduce pure solvent properties, including dielectric constants and free energies of solvation. Use of scaling of the polarizabilities required readjustment of the partial atomic charges to ensure that gas-phase dipole moments were reproduced. The quality of the fits was evaluated via comparison of calculated molecular dipole moments and polarizabilities with the available gas-phase experimental and QM values.

The developed scheme for the determination of atomic charges and polarizabilities has been tested on a set of small molecules representing functional moieties of nucleic acids. All other parameters have been taken from the all-atom additive CHARMM27 force field for nucleic acids. Selected Lennard-Jones parameters were adjusted to reproduce QM data on interactions of model compounds with water as well as nucleic acid base pairing interactions. Selected internal parameters were optimized to reproduce experimental and QM data on molecular geometries, vibrational frequencies, and rotational barriers. The resulting zero-generation force field has been successfully applied in a 1 ns polarizable MD simulation of a DNA octamer in aqueous solution. This simulation validates the feasibility of the developed methodology for the determination of partial atomic charges and polarizabilities as well as the use of the Drude oscillator model to include electronic polarizability in biomolecular systems. Future efforts will apply the methodology developed in this work, along with an iterative parameter optimization scheme that includes the internal and LJ parameters, to

develop a nonadditive empirical force field for molecules of biological and pharmacological interest.

Acknowledgment. We gratefully acknowledge financial support from the NIH (GM51501) to A.D.M. and from the NSF (NSF 0415784) to G.L. and B.R. Computer time allocations were received from the Pittsburgh Supercomputing Center and the DOD ASC Major Shared Resource Computing and High Performance Computing.

Supporting Information Available: The comparison of QM and empirical molecular polarizabilities; initial and fitted values of atomic charges and polarizabilities for model compounds; the comparison of QM, CHARMM27, and Drude polarizable CHARMM calculations for nucleic base pairing interactions and interactions of model compounds with water; the comparison of CHARMM27 and Drude polarizable CHARMM average base pairing distances and backbone dihedral angle distributions from the DNA simulation; and the description of the technical details of the charge and polarizability scaling and rounding. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) MacKerell, A. D., Jr. Atomistic Models and Force Fields. In *Computational Biochemistry and Biophysics*; Becker, O. M., MacKerell, A. D., Jr., Roux, B., Watanabe, M., Eds.; Marcel Dekker: New York, 2001; p 7.

(2) MacKerell, A. D., Jr. *J. Comput. Chem.* **2004**, *25*, 1584.

(3) Saiz, L.; Klein, M. L. *Acc. Chem. Res.* **2002**, *35*, 482.

(4) Halgren, T. A.; Damm, W. *Curr. Opin. Struct. Biol.* **2001**, *11*, 236.

(5) Rick, S. W.; Stuart, S. J. *Rev. Comput. Chem.* **2002**, *18*, 89.

(6) Sprik, M.; Klein, M. L. *J. Chem. Phys.* **1988**, *89*, 7556.

(7) Caldwell, J.; Dang, L. X.; Kollman, P. A. *J. Am. Chem. Soc.* **1990**, *112*, 9144.

(8) Wallqvist, A.; Berne, B. J. *J. Phys. Chem.* **1993**, *97*, 13841.

(9) Bernardo, D. N.; Ding, Y.; Krogh-Jespersen, K.; Levy, R. M. *J. Phys. Chem.* **1994**, *98*, 4180.

(10) Dang, L. X. *J. Phys. Chem. B* **1998**, *102*, 620.

(11) Rick, S. W.; Berne, B. J. *J. Am. Chem. Soc.* **1996**, *118*, 672.

(12) Bryce, R. A.; Vincent, M. A.; Malcolm, N. O. J.; Hillier, I. H.; Burton, N. A. *J. Chem. Phys.* **1998**, *109*, 3077.

(13) Patel, S.; Brooks, C. L., III. *J. Comput. Chem*. **2004**, *25*, 1.

(14) Patel, S.; MacKerell, A. D., Jr.; Brooks, C. L., III. *J. Comput. Chem.* **2004**, *25*, 1504.

(15) Chen, B.; Xing, J.; Siepmann, I. J. *J. Phys. Chem. B* **2000**, *104*, 2391.

(16) Stern, H. A.; Rittner, F.; Berne, B. J.; Friesner, R. A. *J. Chem. Phys.* **2001**, *115*, 2237.

(17) Stern, H. A.; Kaminski, G. A.; Banks, J. L.; Zhou, R.; Berne, B. J.; Friesner, R. A. *J. Phys. Chem. B* **1999**, *103*, 4730.

(18) Ren, P.; Ponder, J. W. *J. Phys. Chem. B* **2003**, *107*, 5933.

(19) Grossfield, A.; Ren, P.; Ponder, J. W. *J. Am. Chem. Soc.* **2003**, *125*, 15671.

(20) Shelley, J. C.; Sprik, M.; Klein, M. L. *Langmuir* **1993**, *9*, 916.

(21) Gao, J.; Habibollazadeh, D.; Shao, L. *J. Phys. Chem.* **1995**, *99*, 16460.

(22) Caldwell, J. W.; Kollman, P. A. *J. Phys. Chem.* **1995**, *99*, 6208.

(23) Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 4177.

(24) Freindorf, M.; Gao, J. *J. Comput. Chem.* **1996**, *17*, 386.

(25) Cieplak, P.; Caldwell, J. W.; Kollman, P. A. *J. Comput. Chem.* **2001**, *22*, 1048.

(26) Dang, L. X. *J. Phys. Chem. B* **1999**, *103*, 8195.

(27) Kaminski, G. A.; Stern, H. A.; Berne, B. J.; Friesner, R. A.; Cao, Y. X.; Murphy, R. B.; Zhou, R.; Halgren, T. A. *J. Comput. Chem.* **2002**, *23*, 1515.

(28) Stuart, S. J.; Berne, B. J. *J. Phys. Chem.* **1996**, *100*, 11934.

(29) van Maaren, P. J.; van der Spoel, D. *J. Phys. Chem. B* **2001**, *105*, 2618.

(30) Lamoureux, G.; MacKerell, A. D., Jr.; Roux, B. *J. Chem. Phys.* **2003**, *119*, 5185.

(31) Drude, P.; Mann, C. R.; Millikan, R. A. *The theory of optics*; Longmans, Green, and Co.: New York [etc.], 1902.

(32) Lamoureux, G.; Roux, B. *J. Chem. Phys.* **2003**, *119*, 3025.

(33) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.

(34) MacKerell, A. D., Jr.; Brooks, B.; Brooks, C. L., III.; Nilsson, L.; Roux, B.; Won, Y.; Karplus, M. CHARMM: The Energy Function and Its Paramerization with an Overview of the Program. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., III, Schreiner, P. R., Eds.; John Wiley & Sons: Chichester, 1998; Vol. 1, p 271.

(35) *CRC Handbook Chemistry and Physics*, 84th ed.; Lide, D. R., Ed.; CRC Press: Boca Raton, 2003.

(36) Giese, T. J.; York, D. M. *J. Chem. Phys.* **2004**, *120*, 9903.

(37) Kaminski, G. A.; Stern, H. A.; Berne, B. J.; Friesner, R. A. *J. Phys. Chem. A* **2004**, *108*, 621.

(38) Morita, A. *J. Comput. Chem.* **2002**, *23*, 1466.

(39) Morita, A.; Kato, S. *J. Chem. Phys.* **1999**, *110*, 11987.

(40) in het Panhuis, M.; Popelier, P. L. A.; Munn, R. W.; Angyan, J. G. *J. Chem. Phys.* **2001**, *114*, 7951.

(41) Tu, Y. Q.; Laaksonen, A. *Chem. Phys. Lett.* **2000**, *329*, 283.

(42) Rick, S. W.; Stuart, S. J.; Berne, B. J. *J. Chem. Phys.* **1994**, *101*, 6141.

(43) van Belle, D.; Couplet, I.; Prevost, M.; Wodak, S. J. *J. Mol. Biol.* **1987**, *198*, 721.

(44) Banks, J. L.; Kaminski, G. A.; Zhou, R. H.; Mainz, D. T.; Berne, B. J.; Friesner, R. A. *J. Chem. Phys.* **1999**, *110*, 741.

(45) Press: W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. *Numerical Recipes in C*; Cambridge University Press: Cambridge, 1988.

(46) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269.

(47) Connolly, M. L. *Science* **1983**, *221*, 709.

(48) Bonin, K. D.; Kresin, V. V. *Electric-dipole polarizabilities of atoms, molecules, and clusters*; World Scientific: Singapore River Edge, NJ, 1997.

(49) 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.; 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.

(50) Moller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.

(51) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. *Chem. Phys. Lett.* **1988**, *153*, 503.

(52) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta (Berlin)* **1973**, *28*, 213.

(53) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. v. R. *J. Comput. Chem.* **1983**, *4*, 294.

(54) Foloppe, N.; MacKerell, A. D., Jr. *J. Comput. Chem.* **2000**, *21*, 86.

(55) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.

(56) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(57) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(58) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007.

(59) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.

(60) Darden, T. A.; York, D.; Pedersen, L. G. *J. Chem. Phys.* **1993**, *98*, 10089.

(61) Steinbach, P. J.; Brooks, B. R. *J. Comput. Chem.* **1994**, *15*, 667.

(62) Lague, P.; Pastor, R. W.; Brooks, B. R. *J. Phys. Chem. B* **2004**, *108*, 363.

(63) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Clarendon Press: Oxford, 1987.

(64) Pan, Y.; MacKerell, A. D., Jr. *Nucl. Acid Res.* **2003**, *31*, 7131.

(65) Miller, K. J. *J. Am. Chem. Soc.* **1990**, *112*, 8533.

(66) Applequist, J.; Carl, J. R.; Fung, K.-K. *J. Am. Chem. Soc.* **1972**, *94*, 2952.

(67) Thole, B. T. *Chem. Phys.* **1981**, *59*, 341.

(68) van Duijnen, P. T.; Swart, M. *J. Phys. Chem. A* **1998**, *102*, 2399.

(69) Ewig, C. S.; Waldman, M.; Maple, J. R. *J. Phys. Chem. A* **2002**, *106*, 326.

(70) Stout, J. M.; Dykstra, C. E. *J. Phys. Chem. A* **1998**, *102*, 1576.

(71) Zhou, T.; Dykstra, C. E. *J. Phys. Chem. A* **2000**, *104*, 2204.

(72) Ding, Y.; Bernardo, D. N.; Krogh-Jespersen, K.; Levy, R. M. *J. Phys. Chem.* **1995**, *99*, 11575.

(73) Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 735.

(74) Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 1833.

(75) Leach, A. R. *Molecular Modelling: Principles and Applications*; Longman: Harlow, 1996.

(76) Mahan, G. D. *Phys. Rev. A* **1980**, *22*, 1780.

(77) Huang, N.; MacKerell, A. D., Jr. *J. Phys. Chem. B* **2002**, *106*, 7820.

(78) Berman, H. M.; Olson, W. K.; Beveridge, D. L.; Westbrook, J.; Gelbin, A.; Demeny, T.; Hsieh, S.-H.; Srinivasan, A. R.; Schneider, B. *Biophys. J.* **1992**, *63*, 751.

(79) Foresman, J. B.; Frisch, A. *Exploring Chemistry with Electronic Structure Methods*; Gaussian, Inc: Pittsburgh, PA, 1996.

(80) Banavali, N. K.; MacKerell, A. D., Jr. *J. Am. Chem. Soc.* **2001**, *128*, 6747.

CT049930P

JCTC Journal of Chemical Theory and Computation

Efficient Simulation Method for Polarizable Protein Force Fields: Application to the Simulation of BPTI in Liquid Water

Edward Harder, Byungchan Kim, Richard A. Friesner, and B. J. Berne*

Department of Chemistry and Center for Bimolecular Simulation, Columbia University, 3000 Broadway, New York, New York 10027

Received October 18, 2004

Abstract: A methodology for large scale molecular dynamics simulation of a solvated polarizable protein, using a combination of permanent and inducible point dipoles with fluctuating and fixed charges, is discussed and applied to the simulation of water solvated bovine pancreatic trypsin inhibitor (BPTI). The electrostatic forces are evaluated using a generalized form of the P3M Ewald method which includes point dipoles in addition to point charge sites. The electrostatic configuration is propagated along with the nuclei during the course of the simulation using an extended Lagrangian formalism. For the system size studied, 20000 atoms, this method gives only a marginal computational overhead relative to nonpolarizable potential models (1.23−1.45) per time step of simulation. The models employ a newly developed polarizable dipole force field for the protein[1] with two commonly used water models TIP4P-FQ and RPOL. Performed at constant energy and constant volume (NVE) using the velocity Verlet algorithm, the simulations show excellent energy conservation and run stably for their 2 ns duration. To characterize the accuracy of the solvation models the protein structure is analyzed. The simulated structures remain within 1 Å of the experimental crystal structure for the duration of the simulation in line with the nonpolarizable OPLS-AA model.

I. Introduction

The goal in force field development for biomolecular systems is to retain chemical accuracy while taking advantage of computational expediency by employing the simplest potential function. Expressing the electrostatic potential energy using a system of fixed point charges interacting via Coulomb's law is certainly simple and is the approach taken for the most popular models used in biomolecular force field simulation.[2−4] However, such nonpolarizable force fields do not reflect the dependence of a molecules electronic structure on its environment. This dependency is clearly manifest in water where the magnitude of the average dipole moment is approximately 40% larger in the liquid compared to the gas phase. For homogeneous systems, such as neat fluids, the exclusion of polarization to model the electrostatic energy may be sufficient for some purposes. However, the electrostatic environments found in solvated biomolecules range from nonpolar near hydrophobic residues to highly polar in the vicinity of hydrophilic and charged residues to a nearly bulk water like environment far from the protein. A rescaling of the partial charges to reflect the mean field response is one way to deal with the average effects of condensed phase environments; however, inhomogeneous systems with spatially varying fields necessitate the explicit inclusion of polarization to properly treat the electrostatic potential.

How best to incorporate polarization in a simple manner is an ongoing quest. The distributed polarizability analysis of Stone[5] is an approach that incorporates highly distributed inducible sites occupied by high order point multipoles. The main drawback to such an approach is the additional complexity of the potential function and the corresponding increase in computational cost. Recently the particle mesh Ewald method has been extended to include multipole

* Corresponding author phone: (212)854-2186; fax: (212)854-7454; e-mail: berne@chem.columbia.edu.

10.1021/ct049914s CCC: $30.25
Published on Web 12/15/2004

interactions up to hexadecapole-hexadecapole.[6] Although providing a considerable improvement in efficiency relative to regular multipole Ewald,[7] the algorithm is still a factor of 8.5 slower than simply using point charges alone.

The most common approach to including polarization in a simple force field describes the electrostatic configuration using a system of fixed point charges/dipoles and inducible dipoles.[8−18] Another approach, referred to as fluctuating charges (FQ), uses variable charge magnitudes to model polarization and has been growing in popularity due to the ease of its implementation and related computational speed.[19−24] Recent successful efforts have led to parametrized models for small molecules that use a combination of fixed charges/dipoles, fluctuating charges, and inducible dipoles using techniques for deriving electrostatic parameters from ab initio electronic structure calculations.[25,26] Using this QM technique, Kaminski et al.[1] have now developed a complete polarizable dipole model for proteins that has shown good accuracy in gas-phase experiments.

Before the Kaminski model can be used as a predictive tool, it will be necessary to validate and refine the model for simulation in liquid water. To do so it is critical to develop an efficient formalism for simulating polarizable condensed phase biomolecular systems. It is the long range electrostatic interactions in biomolecular systems which make these simulations computationally intensive. To minimize surface effects we use periodic boundary conditions. Spherical or minimum image truncation of the long-range electrostatic forces is a method that reduces the computational cost of the simulation but gives rise to unphysical effects.[27,28] The Ewald sum[29] provides a tractable solution to the accurate evaluation of the electrostatic forces but has a computational complexity of $O(N^{3/2})$. For system sizes between $10^4 < N < 10^5$, that are necessary to simulate solvated proteins, mesh based Ewald methods such as SPME[30] and P3M Ewald[31−33] have been shown to be particularly effective. These algorithms have favorable $O(NlogN)$ scaling and lead to approximately an order of magnitude improvement in the computational cost of evaluating the electrostatic forces for systems the size of 20 000 atoms (the relative merits and similarities of the SPME vs P3M Ewald methodology have been discussed at length elsewhere[32,34]). Further computational gains may be realized by using multiple time scale integration algorithms that allow for the evaluation of the expensive long range electrostatic interactions less frequently than in standard Verlet integration schemes.[33,35] The result of these advances is a formalism for atom-detail nonpolarizable potential functions that allows for the MD simulation of solvated proteins on nanosecond time scales using reasonable computational resources.

Explicit inclusion of polarizability in a simple potential function adds additional computational complexity which must be solved in order to realize size scales that are currently accessible to nonpolarizable molecular models. The additional computational burden is 3-fold: (1) Use of inducible point dipole sites in addition to fixed or variable charges requires additional charge-dipole and dipole−dipole interactions. (2) Resolving the electrostatic configuration and therefore the field at each molecular dynamics time step necessitates the self-consistent solution to a system of coupled linear equations. (3) A transparent application of a multiple time scale algorithm in a fashion similar to that applied to nonpolarizable potential models is not straightforward. Recent developments have laid the groundwork for efficient large scale simulations of polarizable systems. In this article we focus on addressing the first two points in constructing an efficient simulation methodology. Work on the problem of combining multiple time scale integration schemes with polarizable potential functions is ongoing in our lab.

In a following study[66] we will be interested in studying the dynamic properties of water solvent in the vicinity of the protein. To ensure an accurate evaluation of these properties our simulations are conducted at constant energy and constant volume free of artificial perturbations necessary to simulate in the isothermal/isobaric ensemble.[57,58]

The article is organized as follows. In section 2.1 we introduce the polarizable models based on the inducible dipole model of Kaminski et al.[1] for the protein. We choose two solvent models for comparison, a fluctuating charge water model, TIP4P-FQ,[19] and an inducible dipole water model, RPOL.[8] To efficiently evaluate the electrostatic potential for a system of charges and dipoles we have generalized the P3M Ewald method in a fashion similar to the methodology developed by Toukmaji et al.[36] To efficiently resolve the electrostatic configuration and therefore the nuclear forces at each molecular dynamics time step, an extended Lagrangian method[19,37] combined with the generalized P3M Ewald method is used to dynamically propagate the electrostatic variables during the course of the simulation. This is discussed in sections 2.2.1 and 2.2.2. In section 3 we apply this technology to the simulation of bovine pancreatic trypsin inhibitor in water. The computational complexity and simulation accuracy is discussed in sections 3.2.1 and 3.2.2. As an initial study of the quality of the polarizable protein model and the proposed solvation models, the structure of the protein is compared with the experimental structures in section 3.2.3.

II. Methodology

A. Model. The approach taken to include polarization in the force fields applied in this study replaces the usual fixed point charge representation for the electrostatic energy with a combination of fixed and variable point charges and dipoles that respond to perturbations in the electric field according to a parametrized potential energy. The charges and dipoles are located relative to the atomic positions of the molecules, either coincident with the atom position or on off-atom virtual sites.

Adding inducible point dipoles to a system of fixed charges is the most common method for introducing explicit polarizability into a molecular force field. The energy for an induced dipole moment on site i is

$$E_{\mu_i} = \gamma_i \cdot \mu_i + \frac{1}{2}\mu_i \cdot \alpha_i^{-1} \cdot \mu_i \qquad (1)$$

where γ_i and the component of the polarizability tensor of the dipole site, α_i, are treated as fitting parameters. The

parameter γ_i is a means of introducing a permanent dipole moment on the isolated site *i*. Through a simple transformation eq 1 can be expressed as the familiar self-energy of the induced dipole relative to the isolated site.

The fluctuating charge model[19] introduces variable charges that respond to fluctuations in the electrostatic potential according to the principle of electronegativity equalization. By this principle the charges will distribute so that the electronegativity on each variable charge site is the same subject to appropriate charge constraints.[19] The energy of creating a fluctuating charge is

$$E_{q_i} = \chi_i q_i + \frac{1}{2} q_i^2 J_i^0 \tag{2}$$

where the Mulliken electronegativity, χ_i and J_i^0 are treated as fitting parameters. The parameter J_i^0 is twice the hardness of the electronegativity of the isolated site.[19]

The electrostatic potential energy in an interacting system relative to a system of isolated molecules can be expressed as

$$V^{el} = E_{elec} + E_{\mu} + E_q - E_{gp} \tag{3}$$

where E_{gp} is the system energy in the gas phase and E_{elec} is the electrostatic energy resulting from the interaction of different sites:

$$E_{elec} = \sum_i \sum_{j \neq i} \left[\frac{1}{2} q_i J(r_{ij}) q_j + q_i \mathbf{S}_{ij} \cdot \mu_j + \frac{1}{2} \mu_i \cdot \mathbf{T}_{ij} \cdot \mu_j \right]$$

$$E_{\mu} = \sum_i E_{\mu i}$$

$$E_q = \sum_i E_{qi} \tag{4}$$

If the electrostatic sites are well separated, the coupling terms can be expressed as Coulomb interactions, $J(r_{ij}) = 1/(r_{ij})$, $\boldsymbol{S}_{ij} = \mathbf{r_{ij}}/(r_{ij}^3)$, and the tensor $\mathbf{T}_{ij} = \mathbf{1}/(r_{ij}^3) - 3\mathbf{r_{ij}}\mathbf{r_{ij}}/(r_{ij}^5)$. At short distances the point multipole approximation for the electron charge distribution breaks down, and the above Coulombic potential diverges. Where deemed necessary screening functions are used at close intermolecular site distances.[26] Intramolecular electrostatic interaction between neighbors (1,2 interactions) and one neighbor removed (1,3) are omitted in the model for the polarizable protein. For water, the (1,2) and (1,3) distances are fixed, and the Coulombic interactions are treated as either additional fitted electrostatic parameters[19] or omitted.[8]

The equilibrium charge/dipole configuration is determined at each set of nuclear coordinates by minimizing the potential energy with respect to the electrostatic variables subject to a charge conservation constraint:

$$\frac{\partial V^{el}}{\partial q_i} = 0 \tag{5}$$

$$\nabla_{\mu_i} V^{el} = 0 \tag{6}$$

In the models presented in this report the charges are constrained to give charge neutral molecules

$$\sum_{i=1}^{N_\beta} q_{i\beta} = 0 \tag{7}$$

where N_β is the number of charges in molecule β. One can equivalently express the charge constraint implicitly by transforming to a set of generalized charge coordinates.[21]

The polarizable model used for the polypeptide (PFF) comes from the work of Kaminski et al.[1] The model places fixed partial charges on all atomic positions and on massless virtual sites representing the lone pairs of the oxygen and sulfur atoms. The electrostatic parameters are fit from gas-phase electronic structure calculations[25] using density-functional theory (DFT) with the B3LYP method[38,39] and the cc-pVTZ(-f) basis set. The choice of basis set, which does not include diffuse functions, is based on evidence that including contributions from such functions in gas-phase DFT calculations results in an overpolarization of the parametrized model in the condensed phase.[26,40,41] The effect is likely a result of an energetic cost, hindering polarization, that results from Pauli repulsion between neighboring molecules in the condensed phase.[42] The polarizabilities of the atomic sites are parametrized by a series of electrostatic perturbations, using dipolar probes applied to the target molecule. The resulting change in the electrostatic potential is measured at a set of grid points outside the van der Waals surface of the molecule. Polarizabilities (α_i) are chosen to minimize deviations from the DFT calculation. The fixed charges and the parameters γ_i are chosen to best approximate the electrostatic potential from the unperturbed DFT calculation.

Stretching and bending energies for PFF are retained from the OPLS-AA force field,[3] while the torsional energy is reparametrized.[1] Further details can be found in the respective references. The electrostatic energy consists of a system of fixed point charges and point polarizable dipoles described by eq 3. The (1,2) and (1,3) interactions are omitted owing to the breakdown of the bare Coulomb potential at such short intersite distances. No intermolecular screening of the Coulomb potential is included in the original model formulation. Short-range repulsion and dispersion is represented by a Lennard-Jones function

$$U_{ij} = \sum_{i \neq j} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] f_{ij} \tag{8}$$

where we apply the geometric sum rule ($\sigma_{ij} = (\sigma_i \sigma_j)^{1/2}$ and $\epsilon_{ij} = (\epsilon_i \epsilon_j)^{1/2}$) for the interaction between particle *i* and *j*. The function f_{ij} is a scaling factor equal to zero for particles connected by a valence bond or angle, set to 0.5 for intramolecular 1,4 interactions and is 1.0 for all other pairs. The Lennard-Jones parameters are derived from ab initio dimer energies of organic compound analogues of the residues and from the OPLS-AA force field.

We employ three commonly used water models to solvate the polypeptide. A fixed charge TIP4P[43] model (for the fixed charge OPLS-AA protein[3,44]) and two polarizable water models, TIP4P-FQ[19] fluctuating charge model and an inducible point dipole model, RPOL.[8] All three models employ an interacting Lennard-Jones site placed on the oxygen atom.

Table 1: Screening Radius Applied to Selected Dipole Sites on the Polarizable Peptide Molecule in Solvated TIP4P-FQ Simulations (Å)

residue	(*) site of screened dipole	screening radius
glutamic acid	-C*O−O	2.5
aspartic acid	-C*O−O	2.5
aspartic acid	-CO*−O*	1.8
methionine	-S−C*H_3	2.0
tyrosine	-C*OH−	2.5

Intermolecular interactions between electrostatic sites is described by the bare Coulomb potential (see eq 4). The TIP4P-FQ model includes an intramolecular interaction between the charges within the molecule that is parametrized along with the other electrostatic parameters empirically. The RPOL model places point polarizable dipoles on the oxygen and hydrogen atoms and omits intramolecular electrostatic coupling.

In practice using this energy function for the simulation of the TIP4P-FQ model with the polarizable protein results in a polarization catastrophe where the electrostatic variables between an interacting molecular pair mutually enhance to infinite polarization. The polarization catastrophe arose from interactions between TIP4P-FQ water and specific residues on the protein (see Table 1). This problem is a direct consequence of the point charge/dipole approximation to the electron charge distribution. For point polarizable models the pair interaction energy diverges at intersite distances proportional to the molecules' polarizability[1/3]. A similar phenomena is found in fluctuating charge and combined fluctuating charge/polarizable dipole models. A simple illustration is a pair of isotropic interacting point inducible dipoles where the singularity occurs at[45]

$$r_{critical} = (4\alpha_i\alpha_j)^{1/6} \tag{9}$$

Models that incorporate a large molecular polarizability by variable point charges/dipoles may have values for $r_{critical}$ that approach physically relevant interaction distances. When this is the case, it is necessary to replace the Coulomb potential with a more accurate representation of the true potential at small interaction distances. A scaling factor may be applied to the Coulomb function, or more rigorously a screening function[46,26] may be used to effectively smear the point multipole and more accurately represent the potential of an electron charge distribution.

One should note the difficulty in accurately modeling polarization with a simple potential function may lead to an exaggerated polarizability and therefore a large $r_{critical}$.[42,40] It is interesting to point out that the polarizable RPOL water model did not result in a polarization catastrophe when used to solvate the polarizable protein. A more detailed study is necessary to definitively resolve whether the catastrophe in the TIP4P-FQ solvation model is a product of an unrealistically large polarizability of the water and protein or the use of the Coulombic coupling between the electrostatic sites. Providing evidence for the former explanation, a recent study[47] has shown a significant overpolarization response for hydrogen bonding configurations in the neat fluid using the TIP4P-FQ model, which may result from exaggerated components of the molecular polarizability along the molecular plane. Settling potential problems with the TIP4P-FQ water model is a prerequisite to dealing with the source of the polarization catastrophe in our solvated protein simulation. This work is ongoing in our lab. In lieu of a satisfactory resolution on the TIP4P-FQ water model, we adopt measures to dampen the polarization response between specific residues and TIP4P-FQ water by applying a screening function for close range intermolecular interactions. The cubic spline $f(x)$ is chosen such that $f(0) = 1$, $f'(0) = 0$, $f(1) = 1$ and $f'(1) = -1, -2, -3$, correspond to the value of the functions $1/x$, $1/x^2$, $1/x^3$, respectively, at $x = 1$.[26] A "screening radius" is applied to specific dipole sites on the protein which affects the charge-dipole interactions. A summary of the sites and the respective screening radii is given in Table 1. The Coulomb potential $u(r) = 1/r^2$ is replaced with

$$u(r) = \frac{1}{s^2} f(r/s) \tag{10}$$

when $r < s$ and s is the sum of the screening radii on the pair of interacting sites.

B. Polarization and MD. 1. *P3M Ewald with Dipoles.* The models studied in this report include point dipoles in addition to point charges to describe the polarizable system; therefore, in addition to interactions between charges the electrostatic potential needs to describe the interactions between charge−dipole and dipole−dipole sites.

The Ewald sum for evaluating the Coulombic energy for a system of point charges has been extended to a system of multipoles by Smith.[7] The electrostatic energy for a periodic system of point charges and point dipoles is

$$E_{elec} = \frac{1}{2}\sum_{\mathbf{n}}\sum_{i}\sum_{j}{}' \frac{(q_i + \mu_i\cdot\nabla_i)(q_j + \mu_j\cdot\nabla_j)}{|\mathbf{r_{ij}} + \mathbf{n}\cdot\mathbf{L}|} \tag{11}$$

The Ewald sum with the metallic boundary condition follows from ref 7

$$E^r = \frac{1}{2}\sum_{i}\sum_{j\neq 1}(q_i + \mu_i\cdot\nabla_i)(q_j + \mu_j\cdot\nabla_j)\frac{\mathrm{erfc}(\eta r_{ij})}{r_{ij}} \tag{12}$$

$$E^k = \frac{1}{2}\sum_{i}\sum_{j}\sum_{\mathbf{k}\neq 0}\frac{1}{V}\frac{4\pi}{k^2}(q_i + \mu_i\cdot\nabla_i)(q_j + \mu_j\cdot\nabla_j)e^{-k^2/4\eta^2}e^{-i\mathbf{k}\cdot\mathbf{r_{ij}}} \tag{13}$$

$$E^s = -\frac{\eta}{\sqrt{\pi}}\sum_{i}\left(q_i^2 + \frac{2\eta^2}{3}\left|\mu_i\right|^2\right) \tag{14}$$

where $\mathbf{L}$ is the unit cell dimensions and the Ewald splitting parameter η modulates the relative weight of E^r and E^k to the total potential energy. For excluded i,j pairs we subtract $1/r_{ij}$ which is equivalent to replacing $\mathrm{erfc}(\eta r_{ij})/r_{ij}$ in E^r with $-\mathrm{erf}(\eta r_{ij})/r_{ij}$. Mesh based approximations to the regular Ewald sum discretize space on a regular grid reducing the Fourier series transforms in E^k to finite Fourier transforms which can be evaluated by fast Fourier transform (FFT)

algorithms.[48,49] The Fourier space portion of the electrostatic energy is evaluated using a suitable extension of the P3M Ewald (P3ME) method,[33,32] as discussed below. As is the case for a system of point charges the real space part and self-energy remain unchanged when using this generalized P3ME method.

Recently Toukmaji et al.[36] have extended the SPME method to include dipole−charge and dipole−dipole interactions. We adopt a similar extension to the P3ME method for the applications presented in this study. Following a similar formalism to that used in ref 36 we define a charge dipole array and follow the same four step procedure for generating the forces outlined in refs 32 and 33 for a fixed charge system:

$$Q_i = q_i + \mu_i \cdot \nabla_i \tag{15}$$

We assign this array to a grid using the same spline function (P is the order of the spline) applied to the fixed charge system

$$\rho_M(\mathbf{r}_p) = \frac{1}{V_h}\sum_{i=1}^{N}\sum_{\mathbf{p}=0}^{\mathbf{P}} Q_i W^{\mathbf{P}}(\mathbf{r}_p - \mathbf{r}_i)$$

$$= \frac{1}{V_h}\sum_{i=1}^{N}\sum_{\mathbf{p}=0}^{\mathbf{P}} (q_i W^{\mathbf{P}}(\mathbf{r}_p - \mathbf{r}_i) + \mu_i \cdot \nabla_i W^{\mathbf{P}}(\mathbf{r}_p - \mathbf{r}_i)) \tag{16}$$

where $\mathbf{r_p}$ are the positions of the grid sites and M denotes the spatial grid. In a similar spirit to the "analytic" scheme for evaluating the spatial gradients in the field calculation discussed in ref 32 the dipole gradients are evaluated by analytic differentiation of the spline function, W^P, which can be factorized into a product of its Cartesian components.[31] The spline functions up to order $P = 7$ are tabulated in ref 32. After assignment we apply a forward FFT to get the Fourier space charge/dipole density

$$\hat{\rho}_M(\mathbf{k}) = V_c \sum_{r_p \epsilon M} \rho_M(\mathbf{r}_p) e^{-i\mathbf{k}\cdot\mathbf{r}_p} \tag{17}$$

and the reciprocal space potential is

$$\hat{\phi}_M^k(\mathbf{k}) = \hat{G}(\mathbf{k})\hat{\rho}_M(\mathbf{k}) \tag{18}$$

where the wave vectors are periodic with values $\mathbf{k} = 2\pi\mathbf{m}/\mathbf{L}$ and $\mathbf{m}$ is an integer vector with values between $-\mathbf{N}_p/2 \le \mathbf{m} < \mathbf{N}_p/2$ and $\mathbf{N}_p$ is the number of grid points along each Cartesian axis. We use the optimized function $\hat{G}$ corresponding to that derived for a system of point charges:[31,33]

$$\hat{G}_{opt}(\mathbf{k}) = \frac{\sum_{\mathbf{n}} \hat{\mathbf{D}}(\mathbf{k_n}) \cdot \hat{\mathbf{R}}(\mathbf{k_n}) [\hat{W}(\mathbf{k_n})/V_c]^2}{\sum_{\mathbf{n}} [\hat{W}(\mathbf{k_n})/V_c]^2 \sum_{\mathbf{n}} |\hat{\mathbf{D}}(\mathbf{k_n})|^2 [\hat{W}(\mathbf{k_n})/V_c]^2} \tag{19}$$

The function $\mathbf{k_n} = \mathbf{k} + 2\pi\mathbf{n}/\mathbf{h}$. The parameter V_c is the volume of the grid cell and $\mathbf{h}$ is the cell width. The function $\hat{D}(\mathbf{k_n})$ is the Fourier transform of the differential operator and is $i\mathbf{k}$ in this study. The function $\hat{R}(\mathbf{k_n})$ is the Fourier transform of the true reference force

$$\hat{\mathbf{R}}(\mathbf{k}) = -i\mathbf{k}\frac{4\pi}{k^2} e^{-k^2/4\eta^2} \tag{20}$$

In principle an optimized function, $\hat{G}$, can be found that corresponds specifically to charge−charge, charge−dipole, and dipole−dipole interactions. However this is not a practical solution for an efficient algorithm, and the possible gains in accuracy are negligible. This is discussed further in Appendix I.

Using an inverse FFT we get the potential on the real space grid.

$$\phi_M^k(\mathbf{r}_p) = \frac{1}{V}\sum_{\mathbf{m}=-\mathbf{N}_p/2}^{\mathbf{N}_p/2-1} \hat{\phi}_M^k(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}_p} \tag{21}$$

The polarizable models studied require the potential, field, and force to propagate the fluctuating charges, polarizable dipoles, and nuclei, respectively (see eq 26). To evaluate the forces and the field it is necessary to interpolate the potential back to the particles. This is done in a similar way to the analytic differentiation method of ref 32. The resultant equations are

$$\phi^k(\mathbf{r}_i) = \sum_{\mathbf{p}=0}^{\mathbf{P}} \phi_M^k(\mathbf{r}_p) W^P(\mathbf{r}_i - \mathbf{r}_p) \tag{22}$$

$$\mathbf{E}(\mathbf{r}_i) = -\nabla_i \phi^k(\mathbf{r}_i)$$

$$= -\sum_{\mathbf{p}=0}^{\mathbf{P}} \phi_M^k(\mathbf{r}_p) \nabla_i W^P(\mathbf{r}_i - \mathbf{r}_p) \tag{23}$$

$$\mathbf{F}(\mathbf{r}_i) = (q_i + \mu_i \cdot \nabla_i)\mathbf{E}(\mathbf{r}_i)$$

$$= -\sum_{\mathbf{p}=0}^{\mathbf{P}} \{q_i \phi_M^k(\mathbf{r}_p) \nabla_i W^P(\mathbf{r}_i - \mathbf{r}_p) + \phi_M^k(\mathbf{r}_p)\, \mu_i \cdot \nabla_i (\nabla_i W^P (\mathbf{r}_i - \mathbf{r}_p))\} \tag{24}$$

This method requires only 2 FFT's in order to evaluate the forces.

2. Extended Lagrangian Formalism for Polarization. Eqs 5 and 6 lead to a set of coupled linear equations in the total electric field and the total potential at site i which can be solved iteratively until self-consistency is achieved. Upward of six iterative calculations of the electric field and potential may be necessary at each step in the molecular dynamics simulation in order to conserve energy.[36] Considering that the field calculation is the most expensive portion of a molecular dynamics simulation, the calculation of polarization in this manner leads to at least a 6-fold increase in the computational complexity. For polarizable force fields an alternative to iteratively solving for the electric degrees of freedom at each time step in the simulation is an approximate method similar in spirit to the Car-Parrinello ab initio MD method.[10,19,37,50] This method treats the electric degrees of freedom as dynamical variables by defining an extended Lagrangian for the equations of motion

$$L = \frac{1}{2}\sum_{i}^{N} [m_i \dot{r}_i^2 + m_q \dot{q}_i^2 + m_\mu \dot{\mu}_i^2] - V - \sum_{\beta} \lambda_\beta \sum_{i=1}^{N_\beta} q_{i\beta} \tag{25}$$

where V is the total potential energy, λ_β is a Lagrange multiplier necessary to satisfy charge neutrality on each

molecule, and N_β is the number of charge sites on molecule β. The charge/dipole dynamics are fictitious and serve solely to keep these variables near the minimum energy, in a computationally efficient manner. The corresponding Lagrangian equations of motion are[19,10]

$$m_q \ddot{q}_j = -\frac{1}{N_\beta}\sum_{i=1}^{N_\beta}(\phi_j - \phi_i)$$

$$m_\mu \ddot{\mu}_j = -\alpha_j^{-1}\cdot\mu_j + \mathbf{E}_j \tag{26}$$

where ϕ_j and $\mathbf{E}_j$ are the electrostatic potential and total field on site j respectively and $\sum_i$ is over all fluctuating charge sites on the molecule containing site j.

The extended Lagrangian approximation requires that the fictitious dynamics run approximately adiabatically for the duration of the simulation. This is satisfied with an appropriately small choice of $m_{q/\mu}$. If the frequency ω_q and ω_μ of the electric variables is sufficiently larger than the fastest nuclear frequency, the degree of thermal coupling between the fictitious and real dynamics will be small. The weaker this coupling, the longer the simulation will progress with the variable charges and dipoles remaining near the minimum energy surface. However, values for the fictitious frequency that are very large require a time step in the simulation that is proportionally small resulting in a computationally costly simulation. It is necessary to strike a balance in the selection of the fictitious frequency parameters between the degree of thermal coupling to the nuclear dynamic bath and the computational cost of the simulation. A typical time step for integrating a biomolecular system is on the order of a femtosecond. Simulations that require a time step significantly smaller than this will be prohibitively slow. The frequency for the electrostatic variables is approximately $\omega_q = (J^0/m_q)^{1/2}$ for the fluctuating charges and $\omega_\mu = 1/(\alpha m_\mu)^{1/2}$ for the polarizable dipoles with an isotropic polarizability. We choose a charge and dipole mass (m_q, m_μ) such that ω_q and $\omega_\mu = 53,333\ \mathrm{cm}^{-1}$. By applying a stability analysis to the leapfrog algorithm, Hockney and Eastwood[31] have derived that the relationship between the time step used in the finite difference integration algorithm and the largest frequency in the system is $\omega_{max}\Delta t_{max} = 2$. The charge/dipole frequency used in this study ($\omega_{max} = 53,333\ \mathrm{cm}^{-1}$) gives a $\Delta t_{max} = 1.25$ fs. However, the motion of the charge/dipole variables are coupled by the electric potential/field leading to frequency modes in the system larger than $53,333\ \mathrm{cm}^{-1}$. In practice we found a system dependent time step ranging between 0.75 fs and 1 fs was necessary for stable integration of the equations of motion. The extended Lagrangian temperature remained below 0.5 K for all the polarizable model simulations indicating these degrees of freedom remain near the minimum energy surface. Thus on the time scale of the extended Lagrangian simulations (approximately 300 ps between iterative minimization of the electrostatic configuration) there is no appreciable transfer of energy to the fictitious degrees of freedom. For very long simulations (much greater than 300 ps), not tested in this study, it may be necessary to systematically minimize the charge/dipole configuration in order to ensure a more accurate representation of the model system. Considering the infrequency that such a minimization would be needed, the added computational cost is essentially zero. Previous studies have found it necessary to apply thermostats and in some cases a restraining potential, to keep the charge/dipole variables near the minimum energy surface.[20,24,40] It is not clear how such an approach affects the dynamics of the system. A faithful representation of the nuclear dynamics requires that the energy flow to the fictitious degrees of freedom is negligible. This condition can be transparently satisfied when using a constant NVE simulation with the extended Lagrangian protocol applied in this report, by monitoring the extended Lagrangian temperature as a function of simulation time.

III. Application: Water Solvated BPTI

Bovine pancreatic trypsin inhibitor (BPTI) has been used as a benchmark for force field simulations for some time,[51,52] and for this reason is our choice of protein for this study. We compare equilibrium simulations of BPTI using a newly developed polarizable force field for polypeptides[1] (PFF), solvated in TIP4P-FQ and RPOL water, to a fixed charge representation of BPTI using the OPLS-AA force field with fixed charge TIP4P water. Regarding the accuracy of their model, Kaminski et al.[1] make note that their model is a first generation attempt at developing a quantitatively accurate force field for biomolecular simulation which includes polarization. Further development and testing in the condensed phase are prerequisites to refining the model. Incorporating an efficient methodology for large-scale polarizable condensed phase simulation, as is applied in this study, is a step toward that goal. As such we are interested in a stable efficient simulation for long time scales and a reasonable representation of native state stability.

A. Simulation Procedure. The simulation procedure is as follows. The starting structure, obtained from the 4PTI structure in the Brookhaven Protein Data Bank,[53] included a protein monomer and 60 water molecules. Hydrogen atoms were added using the MAESTRO software package.[54] Six counter chloride ions were added using GENION,[55] to neutralize the system. BPTI was then solvated in a 60 Å cubic unit cell of water generated from a preequilibrated box of neat TIP4P. Removing water molecules that overlap with the protein left 6377 water molecules in the system. The equilibration and production simulations were performed using the program SIM developed in our lab.[56] The initial equilibration procedure used nonpolarizable force fields and proceeded as follows. Using the OPLS-AA force field and keeping the protein structure fixed, the water solvent was equilibrated at constant temperature (298 K) and pressure (1 atm) for 20 ps using Nose-Hoover chain thermostats[57] and Andersen-Hoover type barostats,[58] giving a cubic unit cell of length $L = 58.8$ Å. To generate the starting configurations for the polarizable model production simulations, an additional 10 ps of simulation at constant volume and constant temperature (NVT) with a fixed protein preceded an additional 50 ps of simulation in NVT allowing the protein and water to relax. The production simulations were run in the microcanonical ensemble using the velocity Verlet algorithm for 2 ns with a 1 fs time step (0.75 fs for

PFF/TIP4P-FQ). A further 500 ps of equilibration was found to be necessary leaving the final 1.5 ns for analysis. The RATTLE[59] constraint algorithm was used to keep the water molecular geometry rigid, and the bonds between the protein heavy atoms and hydrogens were held fixed. The Ewald parameters for the simulations were $\eta = 0.37$ Å^{-1}, a spherical truncation of the real space potential at $R_{cut} = 10$ Å applied to the electrostatic site positions, a grid spacing of 0.75 Å, and an assignment order $P = 6$. The minimum energy electrostatic configuration is solved iteratively at the outset of the simulation. The extended Lagrangian method was used to propagate these variables during the simulation. The availability of computing time dictated the duration of each simulation segment (65−600 ps). To generate the 2 ns trajectories the simulation segments were run in sequential order using the final nuclear configuration of the previous segment as the initial configuration for the following segment. For computational convenience the variable charges/dipoles are iteratively minimized at the outset of each segment. It should be noted that the minimizations are not motivated by a drift of these variables from the minimum energy surface. For example the increase in the extended Lagrangian temperature is only ≈0.01 K for the 315 ps segment using PFF/TIP4P-FQ.

B. Results and Discussion. *1. Efficiency.* To analyze the speedup from adopting the extended Lagrangian formalism and the P3ME approximation we compare to a regular Ewald calculation of the electrostatic energy with an iterative solution to the electrostatic variables. For a consistent level of accuracy with the P3ME simulation (relative rms Force $\approx 10^{-5}$) an efficient parameter set[60] for the regular Ewald sum corresponds to $\eta = 0.25$ Å^{-1}, a spherical truncation of the real space potential at $R_{cut} = 15$ Å and of the wave vectors $k_{max} = 14$ Å. The P3ME method is approximately 8 times faster than Ewald in evaluating the forces. An iterative simulation requires greater than six iterations of the field calculation per time step.[36] In contrast, the overhead in using the extended Lagrangian formalism is less than a factor of 1.1. A conservative estimate of the computational gain is therefore on the order of 40 for an extended Lagrangian/P3ME polarizable simulation compared to an iterative/Ewald scheme even for this relatively small protein. For a similarly sized system of neat RPOL water, Toukmaji et al.[36] have reported a speedup of 100 using an extended Lagrangian/SPME simulation. The difference lies not in the relative efficiency of SPME vs P3M Ewald, which is very similar, but in the choice of suboptimal parameters for the regular Ewald calculation used in the comparison between SPME and Ewald. To compare the computational cost of a polarizable simulation to fixed charge models we present timing data for the execution of one molecular dynamics time step. Remarkably the fully polarizable model is only a factor of 1.23−1.45 more expensive than the fully fixed charge model (the ratios are relative to the fixed charge simulation). The bulk of the computational effort comes from the evaluation of the electrostatic interactions. The TIP4P-FQ model requires no new interactions and thus requires the smaller computational effort (1.23). Using truncation methods and the regular Ewald sum, models that incorporate inducible dipoles have been shown to be a factor of 2 more expensive than analogous fixed charge models.[10,26] However in the P3ME formalism there is no additional expense in the evaluation of the FFT's when using a model with shared charge/dipole sites compared to charges alone. For this reason the RPOL polarizable dipole solvation model gives the improved scaling (1.45).

Table 2: Time Averaged RMSD between the Average NMR Structure and the Average Simulation Structure[a] (Å)

model	C_α	heavy atoms
OPLS-AA/TIP4P	0.7	1.17
PFF/TIP4P-FQ	0.85	1.47
PFF/RPOL	0.91	1.39

[a] Terminal residues are excluded from the comparison.

The molecular dynamics program used in this study is benchmarked against a popular MD program in order to put the preceding timing experiments into a familiar context. SIM shows comparable computational speed relative to AMBER 7.[4] Our SIM molecular dynamics program is approximately 1.3 times slower than AMBER 7 on a 2.4 Ghz Pentium IV processor for a 23558 atom sized system of fixed charges.

2. Simulation Accuracy. Energy conservation in the NVE ensemble is one measure of how faithfully our simulation represents the model Hamiltonian given the P3ME and extended Lagrangian approximations employed. The total energy fluctuations, ΔV, provide a measure of the energy conservation and the total energy drift

$$\Delta V = \frac{1}{N_T} \sum_{i=1}^{N_T} \left| \frac{V_0 - V_i}{V_0} \right| \tag{27}$$

where V_i is the total energy at step i, V_0 is the initial energy under the extended Lagrangian dynamics, and N_T is the total number of time-steps. This quantity has been interpreted as a reasonable measure of accuracy,[61−63] and a value of $\Delta V \leq 0.003$, i.e., $\log(\Delta V) < -2.5$, gives an acceptable numerical accuracy. Another parameter that measures the simulation accuracy is the ratio of root-mean-square fluctuations between the total energy (ΔV_{rms}) and the kinetic energy (ΔKE_{rms})

$$R = \frac{\Delta V_{rms}}{\Delta KE_{rms}} \tag{28}$$

A value of $R < 0.05$ has been correlated with good energy conservation.[64] In Table 3 we present these parameters for the various simulations. The largest values for $\log(\Delta V)$ and R are −4.66 and 0.004, respectively, far less than the acceptable minimum level of accuracy, indicating very good energy conservation. One should note the excellent performance of our simulations indicates a conservative array of P3ME and integration parameters. A less strict tolerance may be sufficient and result in a faster simulation.

The accuracy of the configurational trajectory for the fluctuating charge and polarizable dipoles is related to the extended Lagrangian temperature. If the electrostatic variables begin to drift from the potential energy minimum, the corresponding extended Lagrangian temperature will begin to increase. The extended Lagrangian temperature remains

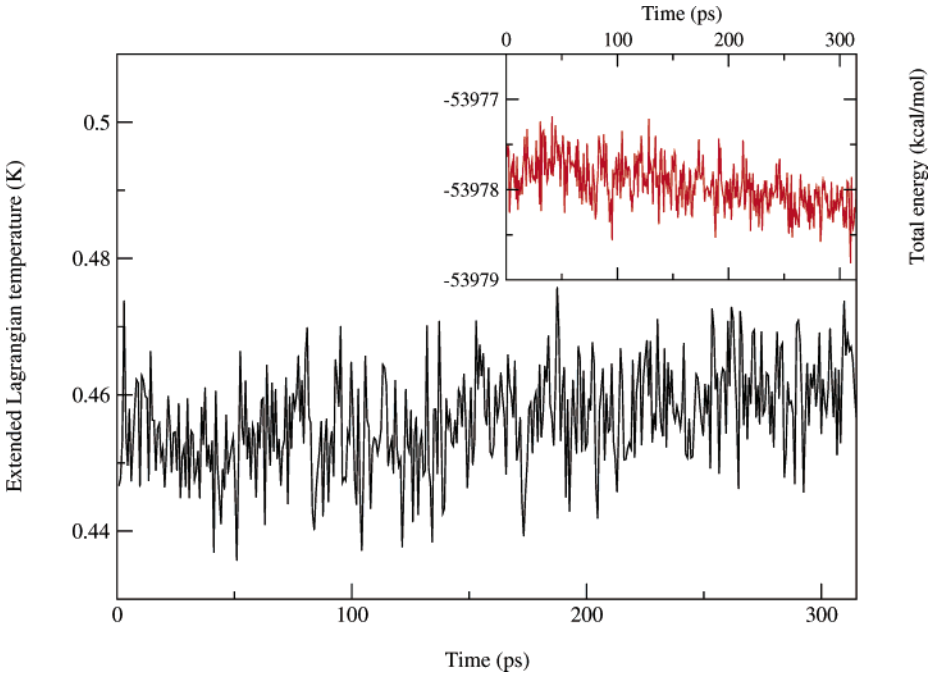


Figure 1. Trajectory of the total energy and extended Lagrangian temperature for PFF/TIP4P-FQ simulation. The 315 ps duration is the longest simulation segment used between minimizations for this model combination. The temperature stays approximately constant for the duration indicating little thermal coupling with the nuclei on this time scale. The total energy for the same trajectory is also plotted showing fluctuations about a consistent value for the simulation length.

Table 3: Summary of the Simulation Accuracy

model	$log(\Delta V)$[a]	R[b]	extended Lagrangian T (K)[c]	q_{rmsd} (e)	μ_{rmsd} (Debye)
OPLS-AA/TIP4P	−5.21	0.004	N/A	N/A	N/A
PFF/TIP4P-FQ	−4.96	0.003	0.5	0.002	0.012
PFF/RPOL	−5.30	0.004	0.1	N/A	0.003

[a] $log(\Delta V)$ measures fluctuations about a reference energy near the beginning of the simulation. [b] R measures the ratio of root-mean-square fluctuations between the total and kinetic energy. Accurate fixed charge simulations have been correlated with a value of $log(\Delta V) < -2.5$ and a value of $R < 0.05$. [c] The extended Lagrangian temperature measures the degree of thermal coupling between the electrostatic variables and the nuclear variables. Typically temperatures less than 6 K have been correlated with a good representation of the potential ground state.

near the value at the beginning of the polarizable simulations never rising above 0.5 K (see Table 3), which indicates these variables remain near the potential energy minimum for the duration of the simulation (for the TIP4P-FQ model a temperature < 6 K has been correlated with a good representation of the minimum energy electrostatic configuration[19]). Figure 1 shows the extended Lagrangian temperature for the longest simulation segment between iterative minimizations (315 ps) using the TIP4P-FQ model. Table 3 also shows the root-mean-square deviations of the electrostatic variables from the minimum energy configuration for the last nuclear configuration corresponding to the largest simulation segment between minimizations (400 ps for RPOL). The deviations are small providing further evidence the generated trajectories are representative of the minimum energy surface.

3. Force Field Accuracy. How well simulations represent the native structure of the protein provides a coarse measurement of the quality of the polarizable force field models. Assuming the model simulation begins in the real native state, a poor representation will lead to large deviations as it relaxes to the erroneous native state of the model. Measuring small deviations from experiment over long simulation periods (nanoseconds) is a positive indication for the model. In this study the experimental crystal structure is used to represent the native state. NMR experiments of BPTI in liquid water[65] have shown small root-mean-square deviations (RMSD) between the average NMR structure and the crystal structure (RMSD=0.85 Å) indicating the native protein structure does not change significantly between the liquid and crystal. We present the RMS deviations of the simulated peptide backbones relative to the experimental crystal structure for 2 ns simulations in Figure 2. The terminal residues show large fluctuations in the liquid water simulations and are not included in the RMSD analysis in line with the results from the NMR experiments.[65] The polarizable force field simulations are similar to the fixed charge model with the average RMS deviation being ≈0.8 Å for all the models. For a direct comparison to an experimental solvated structure, time averaged RMSD values between simulation and the average NMR structure are evaluated and summarized in Table 2. The differences between the polarizable and nonpolarizable models remain small and within the experimental error.

We also investigate the RMS deviations of the heavy atoms of the protein (see Figure 3). The results show close similarity between the polarizable and nonpolarizable model simulations. It is interesting to note that equilibrium of the protein including the side chains requires an additional 500 ps of simulation as evidenced by the slope at the start of the

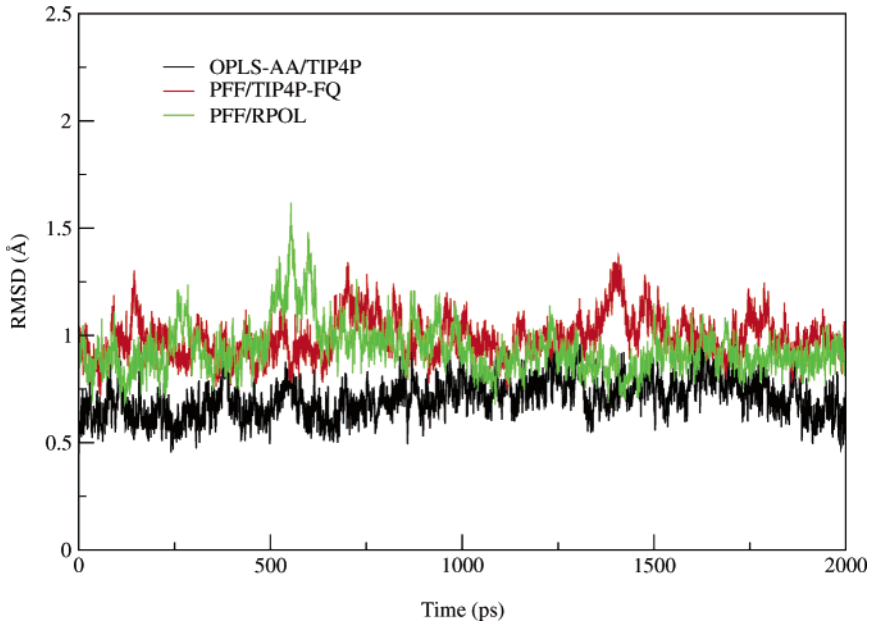


Figure 2. Root-mean-square deviation of C_α atoms for the simulation structures from the experimental crystal structure as a function of simulation time. Terminal residues, which show large fluctuations from NMR experiments as well as simulation, are not included in this analysis. All models do a reasonable job representing the protein native state for the 2 ns duration of the simulation.

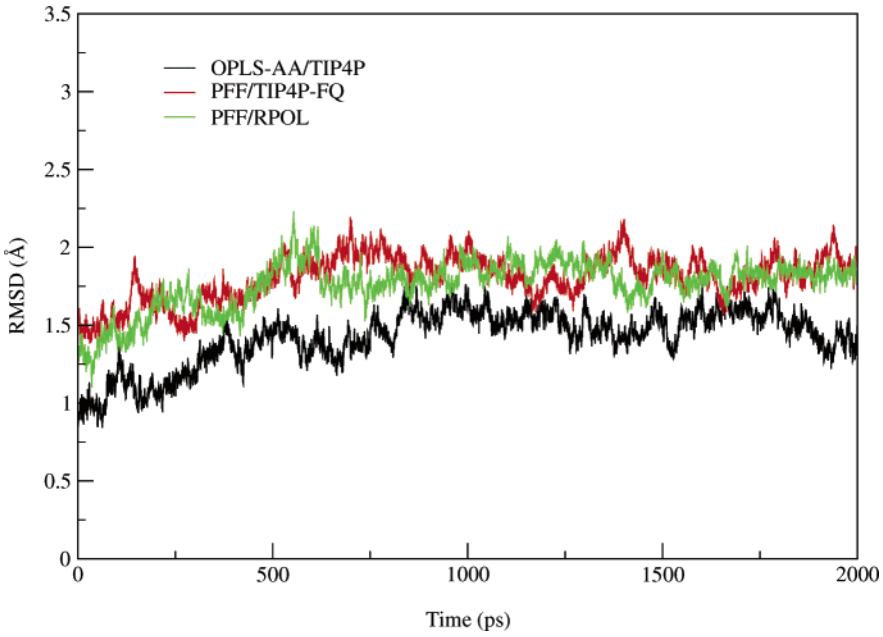


Figure 3. Root-mean-square deviation, including all heavy atoms, for the simulation structures from the experimental crystal structure as a function of simulation time.

simulations in Figure 3. Figure 4 shows the time averaged RMS deviation corresponding to each residue in the protein from simulation and NMR experiment relative to their average structures. The results are again similar for all three simulations.

Considering the small structural deviations between simulation and the protein crystal structure are similar to that from the nonpolarizable force field one can only conclude that like the fixed charge models the polarizable simulations are a reasonable representation of this water/BPTI system. The results are similar for averaged RMSD values between simulation and the average NMR structure (see Table 2). A more sensitive experimental probe is necessary to resolve the relative accuracy of the nonpolarizable OPLS-AA simulation and the proposed polarizable solvation models studied in this report.

IV. Conclusions

We have presented a computationally efficient and accurate methodology for the simulation of large polarizable systems using a combination of fluctuating charges and polarizable dipoles. The method requires only a modest overhead relative to nonpolarizable force fields and the simulations run stably for 2 ns. The method does not need thermostats or the

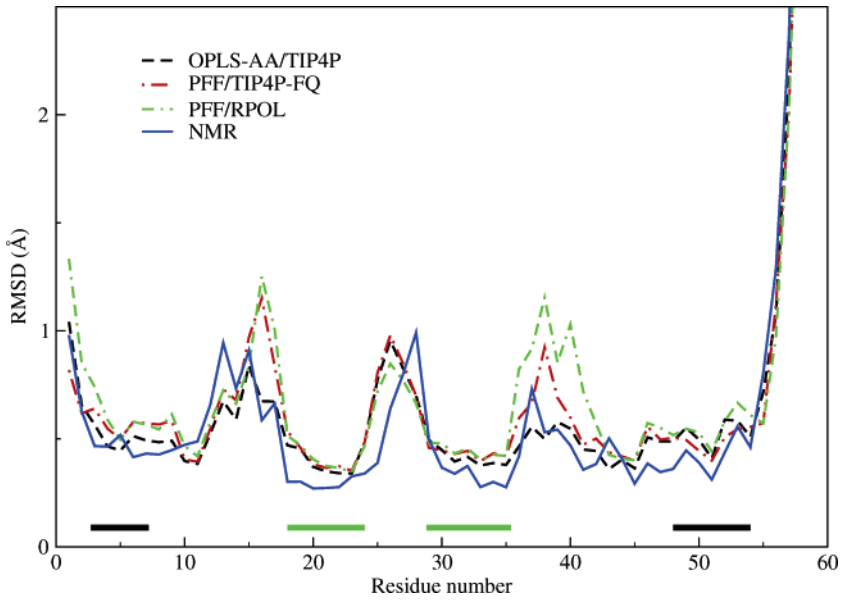


Figure 4. Time averaged RMSD over main chain atoms of individual residues from the respective average simulation and NMR structure. The β-strands are marked in green, and the α-helices are marked in dark gray. The residue dependence of the protein fluctuations in the liquid show correspondence between the simulation models and the NMR experiment.

imposition of restraint potentials, which may obscure the dynamics of the system, to keep the electrostatic degrees of freedom near the minimum energy surface. Thus, the method allows a reliable representation of the system dynamics at constant energy and constant volume. Our tests of a newly developed polarizable protein force field combined with the TIP4P-FQ and RPOL models for water gave promising levels of accuracy compared to the experimental structures. However, a more detailed study is necessary to resolve the relative accuracy of this protein polarizable model to that of fixed charge polypeptide models. Experiments that probe the hydrogen bonding environment of these solvated peptides, such as time-resolved infrared spectroscopy, are a promising tool for evaluating the quality of biomolecular force fields.[47] Although the structural quantities investigated in this report show little to differentiate the polarizable and nonplarizable simulations, one should not conclude that the simulations are similar. Significant differences between the nonpolarizable and polarizable simulations are found in the hydrogen bonding patterns of the protein structure and in the structural and dynamic properties of the solvent surrounding the protein. This will be presented in a following publication.[66] One should also keep in mind that the polarizable peptide model is in the first stage of development. Further refinement of the model, by including condensed phase data into the parametrization, is now possible and will certainly lead to more accurate polarizable models.

Acknowledgment. The authors thank Dr. Tom Young and Dr. Shenglong Wang for helpful discussions. This work was supported by a grant to B.J. Berne from NSF (CHE-03-16896), to Richard A. Friesner from NIH (GM52018) and by a grant of computer time from the EMSL at Pacific Northwest National Laboratories.

V. Appendix: Optimized Coulomb Propagator for a System of Point Charges and Point Dipoles

Following the procedure of Hockney and Eastwood[31] one can derive optimized Coulomb propagators on a discretized space for charge−dipole and dipole−dipole interactions. Defining the mean square difference between the force calculated on the grid, **F**, and the true reference force for the continuous space problem, **R**, to be

$$Q = \frac{1}{V_h}\int_{V_h} d^3\mathbf{r}_1 \int_V d^3\mathbf{r} |\mathbf{F}(\mathbf{r};\mathbf{r}_1) - \mathbf{R}(\mathbf{r})|^2 \qquad (29)$$

The Fourier space representations of the reference forces for the interacting charge/dipole sites are

$$\hat{\mathbf{R}}_{qq} = q_i q_j i\mathbf{k}\frac{4\pi}{k^2}e^{-k^2/4\eta^2}$$

$$\hat{\mathbf{R}}_{qp} = q_i(i\mathbf{k}\cdot\mu_j)i\mathbf{k}\frac{4\pi}{k^2}e^{-k^2/4\eta^2}$$

$$\hat{\mathbf{R}}_{pq} = -(i\mathbf{k}\cdot\mu_i)q_j i\mathbf{k}\frac{4\pi}{k^2}e^{-k^2/4\eta^2}$$

$$\hat{\mathbf{R}}_{pp} = -(i\mathbf{k}\cdot\mu_i)(i\mathbf{k}\cdot\mu_j)i\mathbf{k}\frac{4\pi}{k^2}e^{-k^2/4\eta^2} \qquad (30)$$

where *qq* denotes the force between two charges, *qp* denotes the force on charge *q* from dipole *p*, *pq* denotes the force on *p* from *q*, and *pp* denotes the force between two dipoles. The calculated force from the discretized space is

$$\hat{\mathbf{F}}_{qq} = q_i q_j \hat{U}\hat{\mathbf{D}}\hat{G}\sum_n \hat{U}e^{i(\mathbf{k}-\mathbf{k}_n)\cdot\mathbf{r}_1}$$

$$\hat{\mathbf{F}}_{qp} = q_i \hat{U}\hat{\mathbf{D}}\hat{G}\sum_n \hat{U}(i\mathbf{k}_n\cdot\mu_j)e^{i(\mathbf{k}-\mathbf{k}_n)\cdot\mathbf{r}_1}$$

$$\hat{\mathbf{F}}_{pq} = -(i\mathbf{k}\cdot\mu_i)q_j\hat{U}\hat{\mathbf{D}}\hat{G}\sum_n \hat{U}e^{i(\mathbf{k}-\mathbf{k}_n)\cdot\mathbf{r}_1}$$

$$\hat{\mathbf{F}}_{pp} = (\mathbf{k}\cdot\mu_i)\hat{U}\hat{\mathbf{D}}\hat{G}\sum_n \hat{U}(\mathbf{k}_n\cdot\mu_j)e^{i(\mathbf{k}-\mathbf{k}_n)\cdot\mathbf{r}_1} \quad (31)$$

where $\hat{U} = \hat{W}/V_h$ and $\hat{\mathbf{D}}$ depends on the method of potential differentiation for the field gradient. The wave vector $\mathbf{k} = 2\pi\mathbf{m}/\mathbf{L}$ and $\mathbf{k}_n = \mathbf{k} + 2\pi\mathbf{n}/\mathbf{h}$. Minimizing the functional derivative of $\hat{Q}$ with respect to $\hat{G}$ we can get the optimized functional parameter $\hat{G}$ corresponding to this set of interactions. However since $i\mathbf{k}$ is not periodic in the alias sum over $\mathbf{n}$, the Fourier transform of the dipole gradient remains within the alias sum (see page 274 of ref 31 for details of the derivation), and the parameters $\hat{G}_{pq}$, $\hat{G}_{qp}$, and $\hat{G}_{pp}$ remain explicit functions of the particular dipole site, μ. Obviously this is not a tenable solution. A finite difference approximation for differentiating the dipole gradients is periodic, and the result reduces to eq 19 for all interacting pairs. Of course as was shown in ref 33 the difference in $\hat{G}$ between keeping $i\mathbf{k}$ within the alias sum and factoring it out all but disappears for the assignment orders and grid densities used in this study, which means to an excellent approximation, $\hat{G}_{qq} = \hat{G}_{pq} = \hat{G}_{qp} = \hat{G}_{pp}$.

References

(1) Kaminski, G. A.; Stern, H. A.; Berne, B. J.; Friesner, R. A.; Cao, Y. X.; Murphy, R. B.; Zhou, R.; Halgren, T. A. *J. Comput. Chem.* **2002**, *23*, 1515.

(2) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fisher, 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, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewic-Kuczera, D.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586.

(3) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225.

(4) 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. *J. Am. Chem. Soc.* **1995**, *117*, 5179.

(5) Stone, A. *Chem. Phys. Lett.* **1981**, *83*, 233.

(6) Sagui, C.; Pedersen, L. G.; Darden, T. A. *J. Chem. Phys.* **2004**, *120*, 73.

(7) Smith, W. *CCP5 Newsletter* **1998**, *46*, 18.

(8) Smith, D. E.; Dang, L. X. *J. Chem. Phys.* **1994**, *100*, 3757.

(9) Dang, L. X.; Chang, T. *J. Chem. Phys.* **1997**, *106*, 8149.

(10) van Belle, D.; Froeyen, M.; Lippens, G.; Wodak, S. J. *Mol. Phys.* **1992**, *77*, 239.

(11) Mountain, R. D. *J. Chem. Phys.* **1995**, *103*, 3084.

(12) Wallqvist, A.; Berne, B. J. *J. Phys. Chem.* **1993**, *97*, 13841.

(13) Bernardo, D. N.; Ding, Y.; Krogh-Jespersen, K.; Levy, R. M. *J. Phys. Chem.* **1994**, *98*, 4180.

(14) Ding, Y.; Bernardo, D. N.; Krogh-Jespersen, K.; Levy, R. M. *J. Phys. Chem.* **1995**, *99*, 11575.

(15) Meng, E. C.; Caldwell, J. W.; Kollman, P. A. *J. Phys. Chem.* **1996**, *100*, 2367.

(16) Gao, J.; Habibollazadeh, D.; Shao, L. *J. Phys. Chem.* **1995**, *99*, 16460.

(17) Cieplak, P.; Caldwell, J.; Kollman, P. *J. Comput. Chem.* **2001**, *22*, 1048.

(18) Kaminski, G. A.; Jorgensen, W. L. *J. Chem. Soc., Perkin. Trans. 2* **1999**, *11*, 2365.

(19) Rick, S. W.; Stuart, S. J.; Berne, B. J. *J. Chem. Phys.* **1994**, *101*, 6141.

(20) Rick, S. W.; Berne, B. J. *J. Am. Chem. Soc.* **1996**, *118*, 672.

(21) Liu, Y.-P.; Kim, K.; Berne, B. J.; Friesner, R. A.; Rick, S. W. *J. Chem. Phys.* **1998**, *108*, 4739.

(22) Stuart, S. J.; Berne, B. J. *J. Phys. Chem.* **1996**, *100*, 11934.

(23) Banks, J. L.; Kaminski, G. A.; Zhou, R.; Mainz, D. T.; Berne, B. J.; Friesner, R. A. *J. Chem. Phys.* **1999**, *110*, 741.

(24) Patel, S.; Brooks, C. *J. Comput. Chem.* **2004**, *25*, 1.

(25) Stern, H. A.; Kaminski, G. A.; Banks, J. L.; Zhou, R.; Berne, B. J.; Friesner, R. A. *J. Phys. Chem. B* **1999**, *103*, 4730.

(26) Stern, H. A.; Berne, B. J.; Friesner, R. A. *J. Chem. Phys.* **2001**, *115*, 2237.

(27) York, D.; Darden, T.; Pedersen, L. *J. Chem. Phys.* **1993**, *99*, 8345.

(28) Belhadj, M.; Alper, H. E.; Levy, R. M. *Chem. Phys. Lett.* **1991**, *179*, 13.

(29) Ewald, P. *Ann. Phys* **1921**, *64*, 253.

(30) Essmann, U.; Perera, L.; Berkowitz, M.; Darden, T.; Lee, H.; Pedersen, L. *J. Chem. Phys.* **1995**, *103*, 8577.

(31) Hockney, R. W.; Eastwood, J. W. *Computer Simulations Using Particles;* IOP: Bristol, 1988.

(32) Deserno, M.; Holm, C. *J. Chem. Phys.* **1998**, *109*, 7678.

(33) Zhou, R.; Harder, E.; Xu, H.; Berne, B. J. *J. Chem. Phys.* **2001**, *115*, 2348.

(34) Sagui, C.; Darden, T. *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 155.

(35) Procacci, P.; Darden, T.; Marchi, M. *J. Chem. Phys.* **1996**, *100*, 10464.

(36) Toukmaji, A.; Sagui, C.; Board, J.; Darden, T. *J. Chem. Phys.* **2000**, *113* (24), 10913.

(37) Sprik, M.; Klein, M. L. *J. Chem. Phys.* **1988**, *89*, 7556.

(38) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.

(39) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(40) Lamoureux, G.; MacKerell, A. D.; Roux, B. *J. Chem. Phys.* **2003**, *119*, 5185.

(41) Patel, S.; MacKerell, A. D., Jr.; Brooks, C. L. *J. Comput. Chem* **2004**, *25*, 1504.

(42) Morita, A.; Kato, S. *J. Chem. Phys.* **1999**, *110*, 11987.

(43) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. *J. Chem. Phys.* **1983**, *79*, 926.

(44) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **2001**, *105*, 6474.

(45) Bottcher, C. F. J. *Theory of Electric Polarization, Vol I*, 2nd ed.; Elsevier: 1973.

(46) Thole, B. T. *Chem. Phys.* **1981**, *59*, 341.

(47) Harder, E.; Eaves, J.; Tokmakoff, A.; Berne, B. J. manuscript in preparation.

(48) Brigham, E. O.; Morrow, R. E. *IEEE Spectrum* **1967**, *4*, 63.

(49) Frigo, M.; Johnson, S. G. *ICASS conference proceedings* **1998**, *3*, 1381.

(50) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471.

(51) York, D.; Wlodawer, A.; Pedersen, L. G.; Darden, T. *Proc. Natl. Acad. Sci.* **1994**, *91*, 8715.

(52) Daggett, V.; Levitt, M. *Annu. Rev. Biophys. Biomol. Struct.* **1993**, *22*, 353.

(53) Deisenhofer, J.; Steigemann, W. *Acta Crystallogr. B* **1975**, *31*, 238.

(54) v. 6.0, M. *Schrödinger, Inc.;* Portland OR, 2003.

(55) van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Huenenberger, P. H.; Krueger, P.; Mark, A.; Scott, W.; Tironi, I. *Biomolecular Simulation: The GROMOS96 Manual and User Guide;* Hochshulverlag AG/ETH Zurich, 1996.

(56) Stern, H.; Rittner, F.; Pavese, M.; Harder, E.; Xu, H.; Kim, B. *SIM: Molecular Dynamics Simulation Program*.

(57) Tuckerman, M. E.; Martyna, G. J.; Klein, M. L. *J. Chem. Phys.* **1992**, *97*, 2635.

(58) Martyna, G. J.; Tobias, D. J.; Klein, M. L. *J. Chem. Phys.* **1994**, *101*, 4177.

(59) Andersen, H. C. *J. Comput. Phys* **1983**, *52*, 24.

(60) Petersen, H. G. *J. Chem. Phys.* **1995**, *103*.

(61) Watanabe, M.; Karplus, M. *J. Chem. Phys.* **1993**, *99*, 8063.

(62) Zhou, R.; Berne, B. J. *J. Chem. Phys.* **1995**, *103*.

(63) Figueirido, F.; Zhou, R.; Levy, R.; Berne, B. J. *J. Chem. Phys.* **1997**, *106*, 9835.

(64) van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Phys.* **1977**, *34* (5), 1311.

(65) Berndt, K. D.; Guntert, P.; Orbons, L. P. M.; Wuthrich, K. *J. Mol. Biol.* **1992**, *227*, 757.

(66) Kim, B.; Young, T.; Harder, E.; Berne, B. J.; Friesner, R. A. manuscript in preparation.

CT049914S