JCTC Journal of Chemical Theory and Computation

LETTER

pubs.acs.org/JCTC

Nonmonotonic Recursive Polynomial Expansions for Linear Scaling Calculation of the Density Matrix

Emanuel H. Rubensson*

Division of Scientific Computing, Department of Information Technology, Uppsala University, Box 337, SE-751 05 Uppsala, Sweden

ABSTRACT: As it stands, density matrix purification is a powerful tool for linear scaling electronic structure calculations. The convergence is rapid and depends only weakly on the band gap. However, as will be shown in this letter, there is room for improvements. The key is to allow for nonmonotonicity in the recursive polynomial expansion. On the basis of this idea, new purification schemes are proposed that require only half the number of matrix−matrix multiplications compared to previous schemes. The speedup is essentially independent of the location of the chemical potential and increases with decreasing band gap.

During the last two decades, methods have been developed that make it possible to apply electronic structure calculations, using Hartree−Fock, Kohn−Sham density functional theory, or tight-binding models, to systems with many thousands of atoms.[1−5] Although the computational cost of these methods increases only linearly with system size, such calculations are extremely demanding. Therefore, there is a need to improve existing linear scaling methods in order to reduce the computational cost and make best use of modern computer resources.

In linear scaling electronic structure calculations, efficient computation of the one-particle density matrix D for a given effective Hamiltonian F is an important ingredient. Many methods for linear scaling computation of the density matrix have been proposed. A common approach is to employ a polynomial expansion of the function $D = \theta(\mu I - F)$, where θ is the Heaviside step function and μ is the chemical potential. The expansion may be built up serially by a Chebyshev series[6−9] or recursively by density matrix purification[10−14] or sign matrix methods.[15,16] Another approach is to minimize an energy functional with respect to the density matrix.[17−20]

For the isolated problem of computing the density matrix for a fixed Hamiltonian, the recursive density matrix purification schemes are highly efficient. The convergence is rapid, and the computational cost scales as $\mathcal{O}(\ln(\Delta\varepsilon/\xi))$, where $\Delta\varepsilon$ is the spectral width of the effective Hamiltonian matrix and ξ is the band gap.[11,21] This should be compared to an $\mathcal{O}((\Delta\varepsilon/\xi)^{1/2})$ cost for the serial polynomial expansion[9] and minimization[1,21] methods. However, despite the excellent performance of previously proposed density matrix purification schemes, substantial improvements are still possible, as will be shown in this letter.

In density matrix purification, the effective Hamiltonian matrix is first shifted and scaled so that the eigenvalues end up in the $[0, 1]$ interval in reverse order. After that, low order polynomials with fixed points at 0 and 1 are recursively applied to build up the desired step function. The general iterative procedure can be formulated as

$$\begin{aligned} X_0 &= f_0(F) \\ X_i &= f_i(X_{i-1}), i = 1, 2, \ldots \end{aligned} \qquad (1)$$

where f_0 is the initial linear transformation and f_i, $i = 1, 2, \ldots$ is a sequence of low order polynomials. The iterative procedure is stopped as soon as all eigenvalues of X_i are sufficiently close to their desired values of 0 and 1. The stopping criterion can for example be set in terms of the last eigenvalues to converge,[22] but other criteria are possible as well.[10,14]

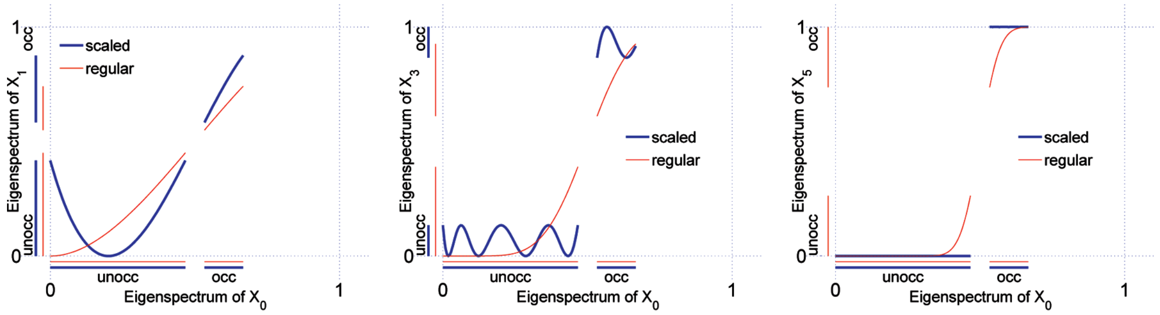
Purification can either be carried out with fixed or varying chemical potential μ. In the case of fixed-μ purification, a single polynomial with an unstable fixed point in $]0, 1[$ is typically used for all f_i, $i > 0$. The initial transformation f_0 maps the chemical potential to the unstable fixed point. The purification process then brings the eigenvalues to their desired values of 0 and 1. In the case of varying-μ purification, the chemical potential is allowed to move during the iterations. This flexibility can be used to automatically adjust the expansion so that the correct number of electrons is obtained, as in canonical[10] and trace-correcting[11] purification.

In any case, the idea has been to use polynomials that increase monotonically in $[0, 1]$ and have fixed points and vanishing derivatives at 0 and 1. As discussed by Niklasson,[11] it can be understood that a recursive expansion using such polynomials will converge toward a step function. In the following, we shall use the notation $P_{i,j}(x)$ for the polynomial of degree $1 + i + j$ with fixed points at 0 and 1 and with i and j vanishing derivatives at 0 and 1, respectively. Many previously proposed purification polynomials can be written in this form.[23]

In this letter, we withdraw from the idea of using monotonically increasing purification polynomials. A scale and fold technique giving nonmonotonic purification transformations is proposed that results in improved performance of both fixed- and varying-μ purification schemes. The new idea is the following: Before each iteration, the eigenspectrum is stretched out outside the $[0, 1]$ interval. Some of the polynomials of the form $P_{i,j}$ can then be used to fold the eigenspectrum over itself. For example, the polynomial $P_{1,0}(x) = x^2$ can be used to fold the unoccupied part of the eigenspectrum if the eigenspectrum is stretched out below 0 before its application. Similarly, the polynomial $P_{0,1}(x) = 2x - x^2$ can be used to fold the occupied part. In general, the scale and fold technique can for a polynomial $P_{i,j}$ be used for the unoccupied part if i is odd and for the occupied part if j is odd.

Received: March 11, 2011
Published: April 20, 2011

dx.doi.org/10.1021/ct2001705



(a) First iteration (b) Third iteration (c) Fifth iteration

Figure 1. Mapping of the eigenspectrum after 1, 3, and 5 iterations of McWeeny based fixed-μ purification with and without the use of scaling. In this illustrative example, $\Delta\varepsilon/\xi = 10$, and the chemical potential μ is located at $\lambda_{\min} + 0.25(\lambda_{\max} - \lambda_{\min})$.

Similar scaling techniques have previously been employed to improve the convergence of Newton iterations for sign matrix evaluations.[24,25] However, in this case, the regular unscaled iteration keeps the eigenvalues outside the interval, and the scaling is used to shrink rather than stretch out the eigenspectrum.

We will first apply the scale and fold technique to fixed-μ purification using a polynomial $P_{m,m}$ with m being odd. For such polynomials, the technique can be used to fold both the unoccupied and occupied parts of the eigenspectrum in each iteration. In this case, the nonmonotonic purification transformation

$$f_i(X_{i-1}) = P_{m,m}(\alpha(X_{i-1} - 0.5I) + 0.5I) \qquad (2)$$

where $\alpha \geq 1$, determines the amount of scaling around the unstable fixed point at 0.5. The complete algorithm for the special case $m = 1$ is given in Algorithm 1, where $\lambda_{\min}$ and $\lambda_{\max}$ are the extremal eigenvalues of F or bounds thereof. For simplicity, it is assumed here that the band gap is located symmetrically around μ. The expression for α can be derived by solving

$$P_{m,m}(\alpha(\beta - 0.5) + 0.5) = P_{m,m}(0.5(1 - \alpha)) \qquad (3)$$

for $\alpha \geq 1$. Here, β is a parameter depending on the eigenvalue closest to 0.5, see Algorithm 1. The behavior of Algorithm 1 is illustrated in Figure 1 for a case with $\Delta\varepsilon/\xi = 10$ and $\mu = \lambda_{\min} + 0.25(\lambda_{\max} - \lambda_{\min})$. The behavior of the regular grand-canonical purification algorithm,[10] corresponding to Algorithm 1 with $\alpha = 1$, is shown for reference. Note how the scaled variant is able to take advantage of the additional flexibility given by allowing for nonmonotonicity, resulting in much faster convergence. Fixed-μ purification schemes with scaling can also be derived for other polynomials of the form $P_{i,j}$, where i and j are both odd and larger than 0. Note that the scaling should be performed around the unstable fixed point of the polynomial, which will differ from 0.5 if $i \neq j$.

Algorithm 1. McWeeny-Based Fixed-μ Purification

```
Input: $F$, $\lambda_{\min}$, $\lambda_{\max}$, $\mu$, $\xi$
  $\gamma = 2\max(\lambda_{\max} - \mu, \mu - \lambda_{\min})$
  $X_0 = (\mu I - F)/\gamma + 0.5I$
  $\beta = 0.5(1 - \xi/\gamma)$
  for $i = 1, 2, ..., n$, do
     $\alpha = 3/(12\beta^2 - 18\beta + 9)^{1/2}$
     $X_s = \alpha(X_{i-1} - 0.5I) + 0.5I$
     $X_i = 3X_s^2 - 2X_s^3$
     $\beta_s = \alpha(\beta - 0.5) + 0.5$
     $\beta = 3\beta_s^2 - 2\beta_s^3$
  end for
  return $D = X_n$
```

The scale and fold technique can also be used together with varying-μ purification. We shall here focus on purification based on the polynomials $P_{0,1}$ and $P_{1,0}$.[26] These polynomials can be used to adjust the occupation count:[11] if the occupation is too high, the $P_{1,0}$ polynomial is applied; otherwise, $P_{0,1}$ is applied. The scaling should in this case be chosen to stretch out the eigenspectrum below 0 before application of x^2 and above 1 before application of $2x - x^2$. The purification transformations are

$$f_i(X_{i-1}) = P_{1,0}(\alpha X_{i-1} + (1 - \alpha)I) \qquad (4)$$

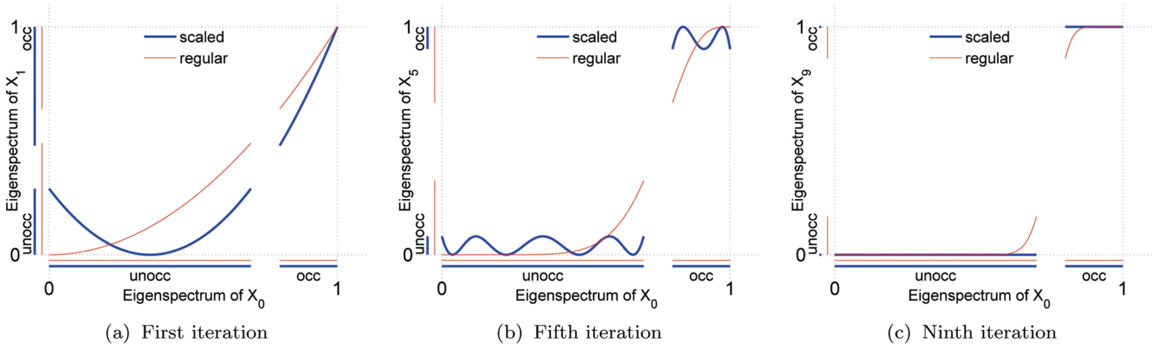
and

$$f_i(X_{i-1}) = P_{0,1}(\alpha X_{i-1}) \qquad (5)$$

where $\alpha \geq 1$ determines the amount of scaling. A complete algorithm is given in Algorithm 2, where λ_{lumo} and λ_{homo} are the eigenvalues closest above and below the band gap, respectively, and n_{occ} is the number of occupied orbitals. Without scaling, i.e., $\alpha = 1$, this algorithm is equivalent to the second order trace correcting purification scheme by Niklasson.[11] The choice of polynomial on line 5 of the algorithm is based on the trace of the current density matrix approximation, just as in the original trace correcting scheme. However, other ways to choose the polynomial can be used as well.[22,27] The behavior of Algorithm 2 is illustrated in Figure 2. The regular scheme with $\alpha = 1$ is shown for reference.

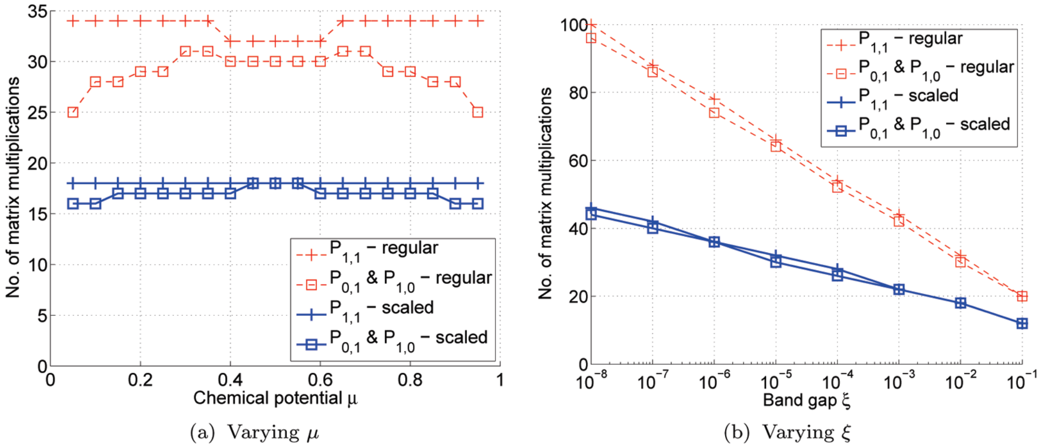
Algorithm 2. $P_{0,1}$ and $P_{1,0}$-Based Varying-μ Purification

```
Input: $F$, $n_{\text{occ}}$, $\lambda_{\min}$, $\lambda_{\max}$, $\lambda_{\text{lumo}}$, $\lambda_{\text{homo}}$
   1: $X_0 = f_0(F) = (\lambda_{\max}I - F)/(\lambda_{\max} - \lambda_{\min})$
   2: $\underline{\beta} = f_0(\lambda_{\text{lumo}})$
   3: $\bar{\beta} = f_0(\lambda_{\text{homo}})$
   4: for $i = 1, 2, ..., n$, do
   5:    if $\text{Tr}[X_{i-1}] > n_{\text{occ}}$, then
   6:    $\alpha = 2/(2 - \beta)$
   7:    $X_i = (\alpha X_{i-1} + (1 - \alpha)I)^2$
   8:    $\underline{\beta} = (\alpha\underline{\beta} + 1 - \alpha)^2$
   9:    $\bar{\beta} = (\alpha\bar{\beta} + 1 - \alpha)^2$
  10:  else
  11:    $\alpha = 2/(1 + \bar{\beta})$
  12:    $X_i = 2\alpha X_{i-1} - \alpha^2 X_{i-1}^2$
  13:    $\underline{\beta} = 2\alpha\underline{\beta} - \alpha^2\underline{\beta}^2$
  14:    $\bar{\beta} = 2\alpha\bar{\beta} - \alpha^2\bar{\beta}^2$
  15:  end if
  16: end for
  17: return $D = X_n$
```



(a) First iteration (b) Fifth iteration (c) Ninth iteration

Figure 2. Mapping of the eigenspectrum after 1, 5, and 9 iterations respectively of $P_{0,1}$- and $P_{1,0}$-based varying-μ purification with and without the use of scaling. In this illustrative example, $\Delta\varepsilon/\xi = 10$, and the chemical potential μ is located at $\lambda_{\min} + 0.25(\lambda_{\max} - \lambda_{\min})$.



(a) Varying μ (b) Varying ξ

Figure 3. Number of matrix−matrix multiplications needed to reach an accuracy of $||\tilde{D} - D||_2 \leq 10^{-9}$, where $\tilde{D}$ is the computed approximation of the exact density matrix D. The test calculations presented in panel a were performed on test Hamiltonians with band gaps $\xi = 0.01$ and varying chemical potential μ. The test calculations presented in panel b were performed on test Hamiltonians with chemical potentials $\mu = 0.5$ and varying band gap ξ. In all cases, the spectral widths of the test Hamiltonians were $\Delta\varepsilon = 1$. The test cases in panel a are essentially equivalent to the test cases presented in Figure 2 of ref 11.

Figures 1 and 2 show that the use of scaling results in more rapid convergence. In order to closer study the performance enhancement given by the scaling technique, we shall consider diagonal test Hamiltonians with varying chemical potentials and band gaps. As previously discussed by Mazziotti,[14] the results for a given chemical potential and a given band gap are valid for any Hamiltonian with that band gap and chemical potential.

Figure 3a shows that the proposed scaling techniques give significant speedup independently of the location of the chemical potential. As can be seen in Figure 3b, the costs of the scaled purification schemes scale as $\mathcal{O}(\ln(1/\xi))$ with the band gap ξ, just as for the regular schemes. However, the convergence for the scaled schemes is around twice as fast as for the regular schemes.

The scaling technique requires some information about the band gap. More precisely, a lower bound of the lower edge and an upper bound of the upper edge of the band gap are needed. These bounds can be used in place of λ_{homo} and λ_{lumo} in Algorithm 2. It should be noted that incorrect bounds can lead to a mix-up between occupied and unoccupied states. However, even if the bounds are not tight, the scaling technique can be used, although the effect will not be as good as it could have been. Tight bounds can be obtained by some technique for calculation of interior eigenvalues.[22,28,29]

The performance was here measured by the number of matrix−matrix multiplications needed to reach a certain accuracy. In practical linear scaling calculations, an efficient way to bring about sparsity is critical for the performance. Since the proposed schemes are on the standard form given by eq 1, it is possible to combine them with previously suggested schemes to control the forward error.[22] To achieve forward error control, information about the band gap is needed. Fortunately, this is the same information as needed for the proposed scaling techniques.

In this letter, nonmonotonic recursive polynomial expansions for calculation of the density matrix were proposed. We have withdrawn from the idea that the approximation of the step function should be monotonically increasing and show that this makes it possible to find new, more efficient nonmonotonic purification transformations. The scaled purification variants of this work represent a substantial improvement compared to previous purification schemes. The reduction in computational cost is essentially independent of the location of the chemical potential, and the proposed schemes are particularly efficient in the case of small band gaps.

AUTHOR INFORMATION

Corresponding Author
*E-mail: emanuel.rubensson@it.uu.se.

ACKNOWLEDGMENT

Discussions with Sara Zahedi and support from the Swedish Research Council under Grant No. 623-2009-803 are gratefully acknowledged.

REFERENCES

(1) Goedecker, S. *Rev. Mod. Phys.* **1999**, *71*, 1085–1123.

(2) Bowler, D. R.; Miyazaki, T.; Gillan, M. J. *J. Phys.: Condens. Matter* **2002**, *14*, 2781–2781.

(3) Saad, Y.; Chelikowsky, J. R.; Shontz, S. M. *SIAM Rev.* **2010**, *52*, 3–54.

(4) Hine, N. D. M.; Haynes, P. D.; Mostofi, A. A.; Skylaris, C.-K.; Payne, M. C. *Comput. Phys. Commun.* **2009**, *180*, 1041.

(5) Rudberg, E.; Rubensson, E. H.; Sałek, P. *J. Chem. Theory Comput.* **2011**, *7*, 340.

(6) Goedecker, S.; Colombo, L. *Phys. Rev. Lett.* **1994**, *73*, 122–125.

(7) Goedecker, S.; Teter, M. *Phys. Rev. B* **1995**, *51*, 9455–9464.

(8) Baer, R.; Head-Gordon, M. *J. Chem. Phys.* **1997**, *107*, 10003–10013.

(9) Liang, W.; Saravanan, C.; Shao, Y.; Baer, R.; Bell, A. T.; Head-Gordon, M. *J. Chem. Phys.* **2003**, *119*, 4117–4125.

(10) Palser, A. H. R.; Manolopoulos, D. E. *Phys. Rev. B* **1998**, *58*, 12704–12711.

(11) Niklasson, A. M. N. *Phys. Rev. B* **2002**, *66*, 155115.

(12) Niklasson, A. M. N.; Tymczak, C. J.; Challacombe, M. *J. Chem. Phys.* **2003**, *118*, 8611–8620.

(13) Holas, A. *Chem. Phys. Lett.* **2001**, *340*, 552–558.

(14) Mazziotti, D. A. *Phys. Rev. E* **2003**, *68*, 066701.

(15) Beylkin, G.; Coult, N.; Mohlenkamp, M. J. *J. Comput. Phys.* **1999**, *152*, 32–54.

(16) Németh, K.; Scuseria, G. E. *J. Chem. Phys.* **2000**, *113*, 6035–6041.

(17) Li, X.-P.; Nunes, R. W.; Vanderbilt, D. *Phys. Rev. B* **1993**, *47*, 10891–10894.

(18) Haynes, P. D.; Payne, M. C. *Phys. Rev. B* **1999**, *59*, 12173–12176.

(19) Helgaker, T.; Larsen, H.; Olsen, J.; Jørgensen, P. *Chem. Phys. Lett.* **2000**, *327*, 397–403.

(20) Shao, Y.; Saravanan, C.; Head-Gordon, M.; White, C. A. *J. Chem. Phys.* **2003**, *118*, 6144–6151.

(21) Rudberg, E.; Rubensson, E. H. *J. Phys.: Condens. Matter* **2011**, *23*, 075502.

(22) Rubensson, E. H.; Rudberg, E.; Sałek, P. *J. Chem. Phys.* **2008**, *128*, 074106.

(23) The McWeeny polynomial[30] is $P_{1,1}(x) = 3x^2 - 2x^3$. This polynomial is equivalent to the Newton−Schulz iteration polynomial $^1/_2x(3 - x^2)$ for sign matrix evaluation.[25] The polynomials suggested by Holas[13] can be written in the form $P_{m,m}(x)$. Niklasson[11] proposed purification schemes based on polynomials $P_{1,m}(x)$ and $P_{m,1}(x)$. Mazziotti[14] suggested use of asymmetric polynomials $P_{m,m+1}(x)$ and $P_{m+1,m}(x)$.

(24) Kenney, C.; Laub, A. J. *SIAM J. Matrix Anal. Appl.* **1992**, *13*, 688–706.

(25) Higham, N. J. *Functions of matrices: theory and computation*; Society for Industrial and Applied Mathematics: Philadelphia, PA, 2008.

(26) Mazziotti, D. A. *J. Chem. Phys.* **2001**, *115*, 8305–8311.

(27) Niklasson, A. M. N.; Weber, V. *J. Chem. Phys.* **2007**, *127*, 064105.

(28) Vömel, C.; Tomov, S. Z.; Marques, O. A.; Canning, A.; Wang, L.-W.; Dongarra, J. J. *J. Comput. Phys.* **2008**, *227*, 7113–7124.

(29) Rubensson, E. H.; Zahedi, S. *J. Chem. Phys.* **2008**, *128*, 176101.

(30) McWeeny, R. *Proc. R. Soc. London, Ser. A* **1956**, *235*, 496–509.

JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Validation of the GROMOS 54A7 Force Field with Respect to *β*-Peptide Folding

Wei Huang, Zhixiong Lin, and Wilfred F. van Gunsteren*

Laboratory of Physical Chemistry, Swiss Federal Institute of Technology, ETH, 8093 Zürich, Switzerland

ABSTRACT: The recently developed GROMOS 54A7 force field, a modification of the 53A6 force field, is validated by simulating the folding equilibrium of two *β*-peptides which show different dominant folds, i.e., a 3_{14}-helix and a hairpin, using three different force fields, i.e., GROMOS 45A3, 53A6, and 54A7. The 54A7 force field stabilizes both folds, and the agreement of the simulated NOE atom−atom distances with the experimental NMR data is slightly improved when using the 54A7 force field, while the agreement of the ^3J couplings with experimental results remains essentially unchanged when varying the force field. The 54A7 force field developed to improve the stability of α-helical structures in proteins can thus be safely used in simulations of *β*-peptides.

1. INTRODUCTION

Molecular dynamics simulation is an efficient method used to understand and predict biological or chemical processes at the atomic level. The simulation results, however, depend on the quality of the force field used, which describes the interactions between particles in the system. Several force fields have been developed for biomolecular simulation, such as AMBER,[1−3] CHARMM,[4−6] GROMOS,[7−11] and OPLS.[12,13] Over the years, successive GROMOS force-field parameter sets have been introduced.[9−11] The most widely used versions of this force field are the 43A1 force field[8,9] of 1996, the 45A3 force field[10] of 2001, and the 53A6 force field[11] of 2004.

The force field 43A1 contains 43 individual atom types to describe van der Waals interactions.[9] The force field 45A3 introduced two additional atom types for branched and cyclic alkanes and reparameterized the aliphatic CH_n groups based on thermodynamic data for long alkane chains.[10] The force field 53A6 reparameterized a number of polar groups also against thermodynamic data, including several (co)solvents,[14−17] renumbered all atom, bond, bond−angle, and torsional dihedral-angel types, and added eight new van der Waals atom types. The recently developed force field 54A7[18] is a modification of the 53A6 force field. It contains four modifications: (1) The torsional angle energy term for the polypeptide *φ*- and *ψ*-dihedral angles is modified. Four different torsional dihedral angle types are added, and the repulsive van de Waals $C_{12}^{1/2}(\mathrm{I}, \mathrm{I})$ parameter for the O−N pair is changed to be smaller than that in the 53A6 parameter set. (2) A new van der Waals nonbonded atom type for a charged $-CH_3$ group is introduced in order to generate a larger repulsion between the partly charged $-CH_3$ groups of the choline moiety and the negatively charged OM oxygen atoms of the phosphate moiety in DPPC type lipids.[19] (3) The van der Waals nonbonded interaction parameters for the Na^+ and Cl^- ions are changed. (4) Two additional improper dihedral angle types are defined in order to facilitate free energy difference calculations involving chirality changes. The first modification was introduced[18] in order to redress the tendency of the thermodynamically calibrated 53A6 force field to slightly destabilize α-helical structures in proteins. Application to four different proteins shows that the new 54A7 force field has the intended effect.[18] However, a more stringent test than simulating folded proteins would be the simulation of a folding equilibrium, which is only possible for short polypeptides.

β-Peptides are non-natural polypeptides which exhibit a strong tendency to form stable, well-defined secondary structures.[20−22] Their resistance to degradation by proteases makes them attractive as potential pharmaceuticals.[23,24] *β*-Peptides can form stable secondary structure motifs even at much shorter sequence lengths than those needed in α-peptides.[25,26] This feature makes them ideal cases to study the folding process and test the quality of the force field in molecular dynamics simulations[27−29] of the folding equilibrium. Since the modification of the peptidic *φ*- and *ψ*-angle torsional angle energy terms and the change in $C_{12}^{1/2}(\mathrm{I}, \mathrm{I})$ repulsive van der Waals parameters for backbone N−O atom pairs would influence the folding equilibrium of *β*-peptides, a test of the 54A7 force field with respect to its reproduction of *β*-peptide folding is necessary.

Here, we test the new GROMOS force field parameter set 54A7 with respect to *β*-peptide folding using two *β*-peptides which in methanol fold into different secondary structures: peptide **1**,[30−32] whose dominant fold is a 3_{14}-helix, and peptide **2**,[33] whose dominant fold is a hairpin (Figure 1). Both *β*-peptides were previously used to study the effect of the use of a polarizable methanol solvent model on their folding equilibrium.[34] The results of the simulations using the 54A7 parameter set are compared to the results obtained with two earlier parameter sets, 53A6 and 45A3, as well as to the NMR experimental data of these two *β*-peptides.

2. METHODS

2.1. 54A7 Parameter Set.

The torsional angle energy term for the polypeptide *φ*- and *ψ*-dihedral angles is modified in conjunction with a change of the combination prescription of the C_{12}

Received: December 24, 2010
Published: April 20, 2011

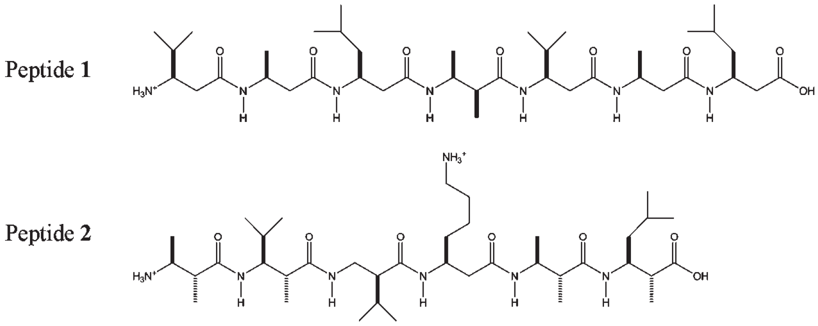


Figure 1. Chemical formulas of the two β-peptides: peptide **1**, H_2^+-β^3-HVal-β^3-HAla-β^3-HLeu-(S,S)-β^3-HAla(αMe)-β^3-HVal-β^3-HAla-β^3-HLeu-OH; Peptide **2**, H_2^+-(S,R)-β^3-HAla(αMe)-(S,R)-β^3-HVal(αMe)-β^2-HVal-β^3-HLys-(S,R)-β^3-HAla(αMe)-(S,R)-β^3-HLeu(αMe)-OH.

Table 1. New Torsional Dihedral Angle Parameters[a] in the GROMOS 54A7 Force Field

type code	K_{ϕ_n} kJ mol^{-1}	cos(δ_n)	m_n	example
42	3.50	−1	2	$-CH_n-C-$
43	2.80	+1	3	$-CH_n-N-$
44	0.70	−1	6	$-CH_n-N-$
45	0.40	+1	6	$-CH_n-C-$

[a] The definition of the parameters can be found in ref 11, Table 5.

van der Waals parameters for the atom type pair O(IAC=1)−N(IAC=6):

(a) In the selection table (Table 8 of ref 11), for the repulsive van der Waals $C_{12}^{1/2}$(I, I) parameters, the type for the O(IAC=1)−N(IAC=6) pair is changed from "2" to "1". This means that the smaller $C_{12}^{1/2}$(O, O) value of 1.000 × 10^{-3} [kJ mol^{-1} nm^{12}]$^{1/2}$ for the O atom (IAC=1) is selected for the interaction with an N atom (IAC=6) compared to the $C_{12}^{1/2}$(O, O) value of 1.130 × 10^{-3} [kJ mol^{-1} nm^{12}]$^{1/2}$ in 53A6.

(b) Four different torsional dihedral angle types are added to Table 5 of ref 11, see Table 1. In the molecular topology building blocks for α-peptides and β-peptides, the dihedral angle type 39 (53A6) in the backbone C−N−CA−C dihedral (α-residue) or the backbone C−N−CB−CA dihedral (β-residue) is to be changed to type 44 (54A7), and the same dihedral angle with type 43 (54A7) is added. In addition, the dihedral angle type 40 (53A6) for the backbone N−CA−C−N dihedral (α-residue) or the backbone CB−CA−C−N dihedral (β-residue) is to be changed into type 45 (54A7), and the same dihedral angle with type 42 (54A7) is added.

These changes increase the hydrogen bonding between the N−H and the C═O groups in the polypeptide backbone and bring the φ- and ψ-angle distributions for a number of proteins more in line[18] with the preferences observed in PDB protein structures.

2.2. Simulations. Six molecular dynamics simulations, of the two β-peptides and based on the 45A3, 53A6, and 54A7 parameter sets, were carried out using the GROMOS05 software,[35] see Table 2. The backbone termini of peptides **1** and **2** and the Lys side chain of peptide **2** were protonated. No counterions were used. The methanol model included in the 45A3 force field is slightly different from the one used in the 53A6 and 54A7 force fields.[11,14] For peptide **1** and parameter set 45A3, a trajectory of a previous simulation was used.[36]

Table 2. Overview of the MD Simulations

			charge	
	simulation		state	no. solvent
peptide	name	force field	[e]	molecules
Peptide **1**	$\mathbf{1}_{45A3}$[a]	45A3	+1	1090
	$\mathbf{1}_{53A6}$	53A6	+1	1090
	$\mathbf{1}_{54A7}$	54A7	+1	1096
Peptide **2**	$\mathbf{2}_{45A3}$	45A3	+2	1409
	$\mathbf{2}_{53A6}$	53A6	+2	1366
	$\mathbf{2}_{54A7}$	54A7	+2	1409

[a] The trajectory of this simulation is described in ref 36.

The folded conformations of the two peptides were used as initial structures. Each peptide was solvated in a periodic, rectangular box with methanol as the solvent. The minimum distance from any peptide atom to the box wall was set to 1.4 nm in both cases. The resulting numbers of solvents are listed in Table 2.

Both simulations were carried out for 200 ns at a constant temperature of 340 K and a constant pressure of 1 atm using the weak coupling algorithm.[37] The temperature coupling time was set to 0.1 ps and the pressure coupling time to 0.5 ps, and an isothermal compressibility of 4.575 × 10^{-4} (kJ mol^{-1} nm^{-3})$^{-1}$ was used.[8] All bond lengths were kept rigid at ideal bond lengths using the SHAKE algorithm,[38] as was the H−CH_3 distance in methanol, allowing a time step of 2 fs in the leapfrog algorithm to integrate the equations of motion. Nonbonded interactions were calculated using a twin-range cutoff scheme with cutoff radii of 0.8/1.4 nm. Interactions within 0.8 nm were evaluated every time step. The intermediate range interactions were updated every fifth time step, and the long-range electrostatic interactions beyond 1.4 nm were approximated by a reaction field force[39] representing a dielectric continuum with a dielectric permittivity of either 17.7 for the methanol model of the 45A3 force field or 19.8 for that of the 53A6 and 54A7 force fields.[14]

2.3. Analysis. Trajectory coordinates and energies stored at 0.5 ps intervals were used for analysis. Backbone atom-positional root-mean-square deviations (RMSD) were calculated after translational superposition of the solute centers of mass and least-squares rotational fitting of atomic positions, using all backbone atoms (N, CB, CA, C) of residues 2−6 for peptide **1** and residues 2−5 for peptide **2**. The backbone atom-positional RMSD criteria required to separate the folded conformation (C_F) from the unfolded ones (C_U) are 0.1 nm for peptide **1** and

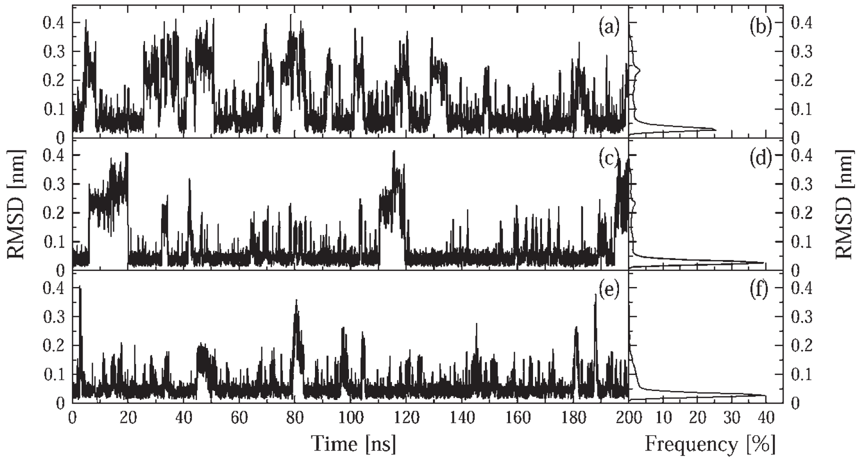


Figure 2. Time evolution (left panels) and distribution (right panels) of atom-positional RMSD from the 3_{14}-helical model structure for the backbone atoms of residues 2−6 in simulations of peptide **1**. Upper panels, $\mathbf{1}_{45A3}$; middle panels, $\mathbf{1}_{53A6}$; lower panels, $\mathbf{1}_{54A7}$.

0.08 nm for peptide **2**. Conformational clustering was performed using the approach of Daura et al.[40] on the set of peptide structures taken at 10 ps intervals from the complete 200 ns trajectories of the simulations. The RMSD values described above were also used as the cutoffs for the conformational clustering. Only the clusters that make up 95% of a trajectory were selected and counted as a function of time.[41] Hydrogen bonds were defined by a maximum hydrogen−acceptor distance of 0.25 nm and a minimum donor−hydrogen−acceptor angle of 135°. Only hydrogen bonds with a population larger than 5% are reported. Distributions of the size of the solute dipole moment calculated using all atoms or the backbone atoms of the peptides are reported. Since the dipole moment of a set of atoms carrying a nonzero total charge depends on the position of the origin, the center of geometry of the solute was used as such.

Folding kinetics were studied by calculating the total residence time and mean residence time in the folded conformation C_F and the number of time periods for which the solute remains folded. The folding free enthalpy was calculated as

$$\Delta G_{\text{folding}} = -k_B T \ln(P_{C_F}/P_{C_U}) \qquad (1)$$

where k_B is the Boltzmann constant, T is the temperature, and P_{C_F} and P_{C_U} are the relative probabilities of the system in the folded and unfolded conformational states, respectively. P_{C_F} and P_{C_U} are obtained by counting the relative number of folded and unfolded structures respectively in a trajectory. The total residence time is the product of P_{C_F} and the total simulation time, and the mean residence time is the total residence time divided by the number of folded periods. The number of folded periods was calculated using structures taken at 10 ps intervals from the simulations. If the peptide changed from one conformation to the other and stayed there for at least 20 ps, it was considered a transition between a folded period and an unfolded period.

Interproton distances extracted from the NOE intensities measured in the NMR experiments were compared with the average interproton distances in the simulations calculated using $\langle r^{-6}\rangle^{-1/6}$, where r is the instantaneous interproton distance. The hydrogen−hydrogen distances involving aliphatic hydrogen atoms were calculated by defining virtual (CH_1), prochiral (stereospecific CH_2), and pseudo- (CH_3 and nonstereospecific CH_2) atomic positions, and pseudoatom corrections were added to the distance bounds for the latter, 0.1 nm for nonstereospecific CH_2, 0.15 nm for CH_3, and 0.29 nm for nonstereospecific rotating methyls.[42] ^3J-coupling constants were calculated using the Karplus relation,[43]

Table 3. Folding Dynamics and Thermodynamics of Peptides 1 and 2

simulation	$\mathbf{1}_{45A3}$	$\mathbf{1}_{53A6}$	$\mathbf{1}_{54A7}$	$\mathbf{2}_{45A3}$	$\mathbf{2}_{53A6}$	$\mathbf{2}_{54A7}$
number of folded periods	184	129	205	104	200	363
total residence time [ns]	125	157	174	10	23	37
mean residence time [ps]	680	1218	849	96	115	102
fraction folded [%]	63	79	87	5	12	19
free enthalpy of folding [kJ mol^{-1}]	−1.5	−3.7	−5.4	8.3	5.8	4.2

Table 4. Intramolecular Hydrogen Bond Populations of Peptide 1 (in %)

donor···acceptor	$\mathbf{1}_{45A3}$	$\mathbf{1}_{53A6}$	$\mathbf{1}_{54A7}$
NH(1)···O(3)	20	22	16
NH(2)···O(4)	57	77	87
NH(3)···O(5)	60	80	90
NH(4)···O(6)	57	73	74
NH(5)···O(7)	15	24	27

$$^3J(H,H) = a\cos^2\theta + b\cos\theta + c \qquad (2)$$

where a = 6.4 Hz, b = −1.4 Hz, and c = 1.9 Hz for the calculation of $^3J_{H_N,H_C}$,[44] and a = 9.5 Hz, b = −1.6 Hz, and c = 1.8 Hz for the calculation of $^3J_{H_C,H_C}$.[45]

3. RESULTS

3.1. Peptide 1. The atom-positional RMSD of the backbone atoms of residues 2 to 6 with respect to the 3_{14}-helical structure are shown in Figure 2 for MD simulations of peptide **1** as a function of the simulation time together with their distributions. The results show that the folding equilibrium of peptide **1** varies between the different force fields. Although the distributions of RMSD have the same pattern and the location of the major peak

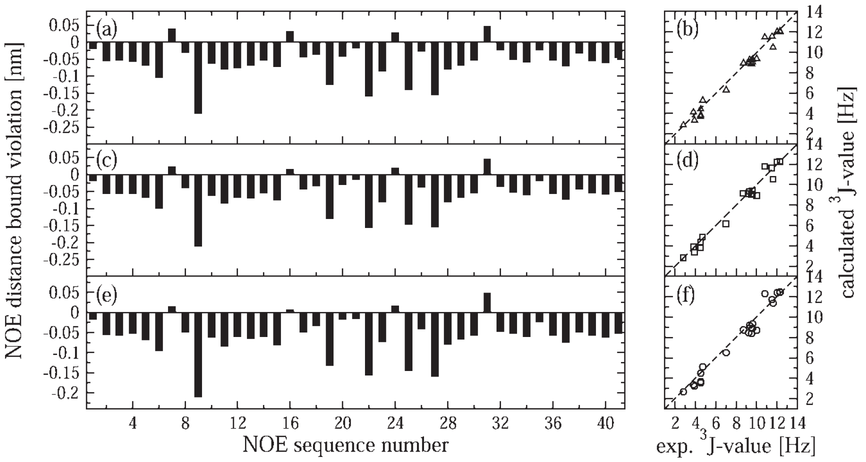


Figure 3. Comparison of $\langle r^{-6}\rangle^{-1/6}$ averaged NOE distance bound violations (left panels) and average ^3J-coupling constants (right panels) as obtained from simulations and experimental data[30] of peptide **1**. Upper panels, $\mathbf{1}_{45A3}$; middle panels, $\mathbf{1}_{53A6}$; lower panels, $\mathbf{1}_{54A7}$. For the specification of the NOE atom pairs and the ^3J-coupling constants, we refer to Tables S1 and S2 in ref 34.

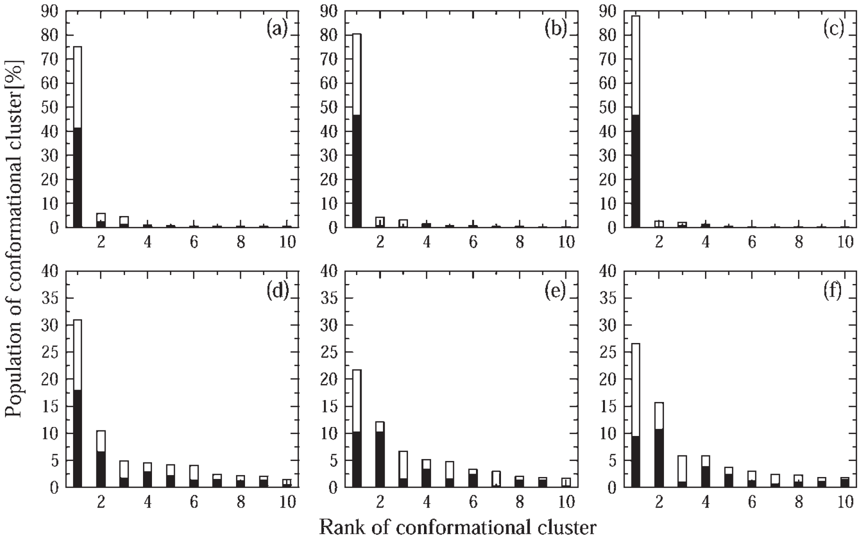


Figure 4. Conformational clustering analysis of the combined trajectories of three force fields for peptide **1** (panels a–c) and peptide **2** (panels d–f). The population in percentage per cluster and the portion of structures per cluster that belong to the trajectories generated using each of the three force fields is shown in decreasing order. (a) $\mathbf{1}_{45A3}$ (white)–$\mathbf{1}_{53A6}$ (black); (b) $\mathbf{1}_{45A3}$ (white)–$\mathbf{1}_{54A7}$ (black); (c) $\mathbf{1}_{53A6}$ (white)–$\mathbf{1}_{54A7}$ (black); (d) $\mathbf{2}_{45A3}$ (white)–$\mathbf{2}_{53A6}$ (black); (e) $\mathbf{2}_{45A3}$ (white)–$\mathbf{2}_{54A7}$ (black); (f) $\mathbf{2}_{53A6}$ (white)–$\mathbf{2}_{54A7}$ (black).

in the distributions is below 0.1 nm in all three simulations, the simulation using the force field 54A7 samples less unfolded conformations than the simulations using the force fields 53A6 and 45A3, see Table 3.

Populations of intramolecular hydrogen bonds in the simulations of peptide **1** are listed in Table 4. The populations of the 3_{14}-helical hydrogen bonds, i.e., those between NH(i) and O(i+2), increase from $\mathbf{1}_{45A3}$ to $\mathbf{1}_{53A6}$ to $\mathbf{1}_{54A7}$, except for the one between NH(1) and O(3), including the terminal residue. The free enthalpy of folding also shows an increased stability of the 3_{14}-helical fold for the successive force fields (Table 3).

Kinetic properties are listed in Table 3. The mean residence time for $\mathbf{1}_{54A7}$ is shorter than that for $\mathbf{1}_{53A6}$ due to the higher frequency of folding events during the former simulation.

The proton–proton NOE distance bound violations and ^3J-coupling constants calculated from the simulations of peptide **1** with the three different force fields are shown in Figure 3. There are four slightly positive violations for the same NOEs in all three simulations. Three of the four positive violations of $\mathbf{1}_{54A7}$ are smaller than those of $\mathbf{1}_{53A6}$ and $\mathbf{1}_{45A3}$. The average ^3J-coupling constants also agree well with the experimental data, with average absolute deviations of 0.4, 0.4, and 0.5 Hz for $\mathbf{1}_{45A3}$, $\mathbf{1}_{53A6}$, and $\mathbf{1}_{54A7}$, respectively, well within the accuracy of the Karplus relation.

To investigate whether the simulations using the different force fields sample the same conformational space, we performed a conformational clustering analysis on combined trajectories of simulations of the same peptide using different force fields. The results are shown in Figure 4. The populations of the clusters in panels a–c show that the conformational spaces sampled in simulations using the three different force fields are not very different.

The cumulative number of conformational clusters that make up 95% of the trajectory in the simulations as a function of time is shown in Figure 5. All of the numbers of clusters converge within 200 ns, indicating that the simulations converged to a particular set of conformations. Only two clusters comprise 95% of the trajectory in the simulation $\mathbf{1}_{54A7}$. A total of 93% of the trajectory of $\mathbf{1}_{54A7}$ is concentrated in the first cluster. These data also indicate that peptide **1** is more stable with the 54A7 force field than with the 53A6 or 45A3 force fields.

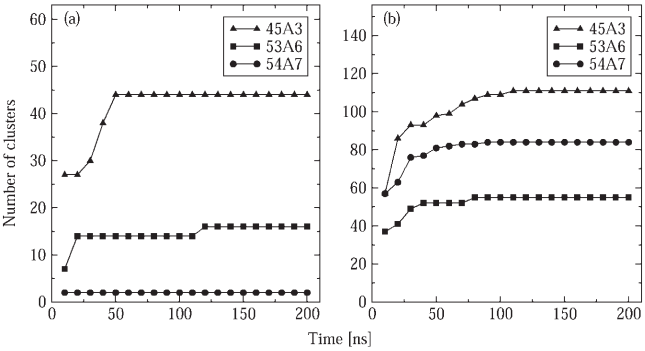


Figure 5. Cumulative number of conformational clusters that make up 95% of the trajectory in the simulations as a function of time. (a) Peptide **1**. (b) Peptide **2**. Triangles, force field 45A3; squares, force field 53A6; circles, force field 54A7.

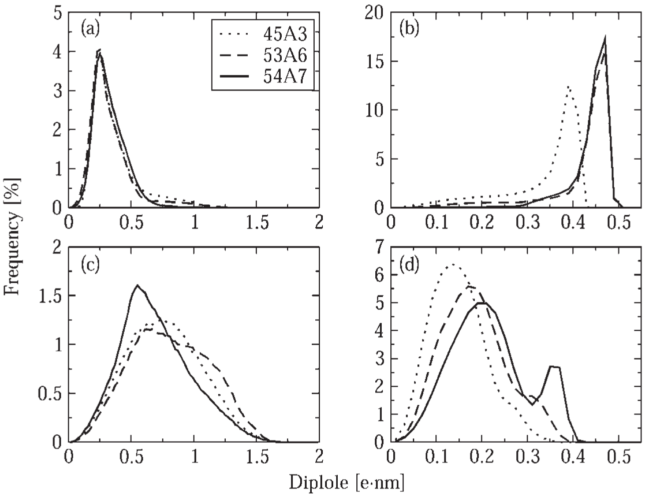


Figure 6. Distributions of solute dipole moment of peptide **1** (upper panels) and peptide **2** (lower panels) calculated using all atoms (left panels) or only the backbone atoms (right panels). Dotted lines, force field 45A3; dashed lines, force field 53A6; solid lines, force field 54A7.

Distributions of the solute dipole moment of peptide **1** are shown in Figure 6. The dipole moment distributions of backbone atoms are different between simulations $\mathbf{1}_{45A3}$ and $\mathbf{1}_{53A6}$ or $\mathbf{1}_{54A7}$, which is due to the larger backbone atomic partial charges in the latter two force fields.[11] Interestingly, the distributions of the dipole moment of the whole peptide, i.e., including the terminal NH_3^+ and COOH groups,[46] are similar.

3.2. Peptide 2. The atom-positional RMSD of the backbone atoms of residues 2 to 5 with respect to the model hairpin structure (X-PLOR structure number 1[33]) are shown in Figure 7 for the simulations of peptide **2** as a function of simulation time together with their distributions. The location of the peak of the distribution is around 0.08 nm in simulation $\mathbf{2}_{54A7}$, while the ones for the other two simulations are both around 0.15 nm. This indicates that the new force field 54A7 samples more ideal hairpin conformations for peptide **2** than the force fields 45A3 and 53A6. Both the total residence time and the fraction of folded conformation gradually increase from $\mathbf{2}_{45A3}$ to $\mathbf{2}_{53A6}$ to $\mathbf{2}_{54A7,}$ see Table 3. Thus, the free enthalpy of folding gradually decreases in this order. It indicates that the new 54A7 force field stabilizes the folded hairpin structure over the other two force fields.

Populations of intramolecular hydrogen bonds in the simulations of peptide **2** are listed in Table 5. The populations of the hairpin hydrogen bonds NH(2)···O(5) and NH(3)···O(4) are larger for $\mathbf{2}_{54A7}$ than for $\mathbf{2}_{45A3}$, while being similar to those in $\mathbf{2}_{53A6}$. The sum of the two populations is 50.2% in $\mathbf{2}_{54A7}$ and is 52.8% in $\mathbf{2}_{53A6}$.

Table 5. Intramolecular Hydrogen Bond Populations of Peptide 2 (in %)

donor···acceptor	$\mathbf{2}_{45A3}$	$\mathbf{2}_{53A6}$	$\mathbf{2}_{54A7}$
NH(2)···O(5)	1.8	5.1	11.6
NH(3)···O(4)	30.0	47.6	38.6
NH(4)···O(1)	7.6	16.1	5.5
NH(5)···O(3)	3.5	2.7	7.0
NH(5)···O(6)	5.0	0.1	1.8

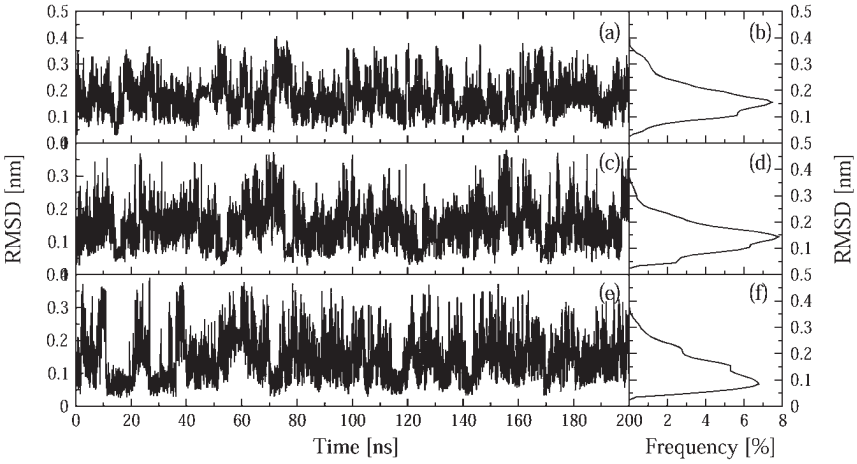


Figure 7. Time evolution (left panels) and distribution (right panels) of the atom-positional RMSD from the model hairpin structure for the backbone atoms of residues 2−5 in simulations of peptide **2**. Upper panels, $\mathbf{2}_{45A3}$; middle panels, $\mathbf{2}_{53A6}$; lower panels, $\mathbf{2}_{54A7}$.

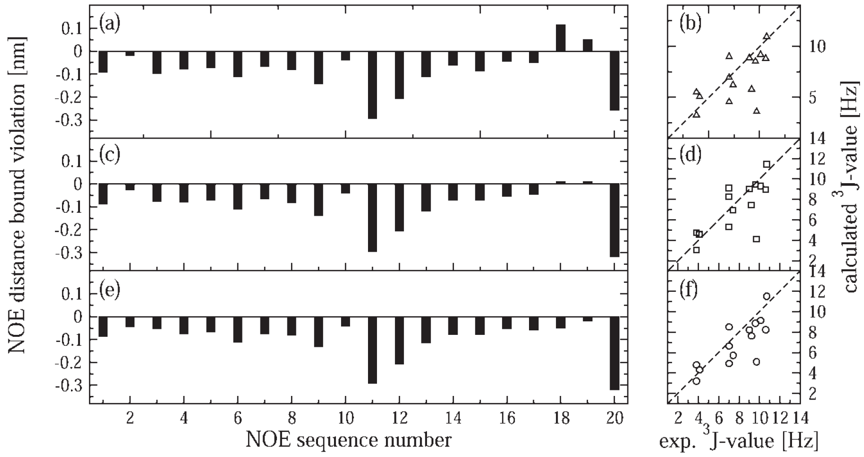


Figure 8. Comparison of $\langle r^{-6}\rangle^{-1/6}$ averaged NOE distance bound violations (left panels) and average ^3J-coupling constants (right panels) as obtained from simulations and experimental data[33] of peptide **2**. Upper panels, $\mathbf{2}_{45A3}$; middle panels, $\mathbf{2}_{53A6}$; lower panels, $\mathbf{2}_{54A7}$. For the specification of the NOE atom pairs and the ^3J-coupling constants, we refer to Tables S4 and S5 in ref 34.

Kinetic properties are listed in Table 3. The mean residence times do not differ much between the three force fields.

The proton−proton NOE distance bound violations and ^3J-coupling constants calculated from the simulations of peptide **2** are shown in Figure 8. There are two positive violations in $\mathbf{2}_{45A3}$, 0.11 and 0.05 nm. Both of these violations are reduced to 0.01 nm in $\mathbf{2}_{53A6}$ and disappear in $\mathbf{2}_{54A7}$. The average absolute deviations of ^3J-coupling constants are 1.6, 1.3, and 1.4 Hz for $\mathbf{2}_{45A3}$, $\mathbf{2}_{53A6}$, and $\mathbf{2}_{54A7}$, respectively. These deviations from the experimental values are larger than those observed for peptide **1**.

The results of the conformational clustering analyses of combined trajectories are shown in Figure 4. It seems that the conformational spaces sampled are most different between $\mathbf{2}_{45A3}$ and $\mathbf{2}_{53A6}$ on the one hand and $\mathbf{2}_{54A7}$ on the other. This is in line with the deviations from the ideal hairpin structure shown in Figure 7. The RMSDs of the central member structures of the first two clusters from the ideal hairpin structure are 0.10 and 0.07 nm for the combined trajectories of $\mathbf{2}_{45A3}$ and $\mathbf{2}_{54A7}$ and are 0.11 and 0.05 nm for the combined trajectories of $\mathbf{2}_{53A6}$ and $\mathbf{2}_{54A7}$. The central member structures of the first clusters are partly folded, while those of the second clusters are fully folded. This indicates that the new 54A7 force field samples more hairpin conformations than the other two force fields.

The cumulative number of conformational clusters that make up 95% of the trajectory in the simulations as a function of time is shown in Figure 5. The numbers of clusters converge in 200 ns for all simulations. Peptide **2** samples a larger conformational space than peptide **1**.

Distributions of the solute dipole moment of peptide **2** are shown in Figure 6. The dipole moment distributions of the whole peptide are broader than those of peptide **1**. The dipole moment of backbone atoms of peptide **1** is larger than that of peptide **2**, because of the stronger alignment of the NH and CO dipoles in the 3_{14}-helix compared to the hairpin.

4. CONCLUSION

To test the performance of the new GROMOS 54A7 force field, we compared the folding behavior of two β-peptides that show a helical and a hairpin as the dominant fold using three different force fields, GROMOS 45A3, 53A6, and 54A7, respectively. The 54A7 force field samples more 3_{14}-helical or hairpin conformations than the 53A6 or 45A3 force fields, which is due to the slightly enhanced capacity of the backbone NH and CO groups to form hydrogen bonds with each other in the 54A7 parameter set. The agreement with the experimental NOE data was slightly improved by using the 54A7 force field, while the experimental ^3J couplings were reproduced equally well in the simulations of the three different force fields.

Overall, the new 54A7 force field reproduces the folding equilibria equally well or slightly better than the 53A6 and 45A3 force fields. Yet, the distributions of conformations are slightly different for the different force fields as are the folding kinetics. The GROMOS 54A7 force field developed to enhance the stability of α-helical structures in proteins can thus be safely used in simulations of folding equilibria of β-peptides.

AUTHOR INFORMATION

Corresponding Author

*E-mail: wfvgn@igc.phys.chem.ethz.ch.

ACKNOWLEDGMENT

This work was financially supported by the National Center of Competence in Research (NCCR) in Structural Biology and by grant number 200020-121913 of the Swiss National Science Foundation, by grant number 228076 of the European Research Council (ERC), and by grants number IZLCZ2-123884 of the Sino-Swiss Science and Technology Cooperation Program, which is gratefully acknowledged.

REFERENCES

(1) Weiner, P. K.; Kollman, P. A. *J. Comput. Chem.* **1981**, *2*, 287–303.

(2) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.

(3) 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*, 1–41.

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

(5) MacKerell, A. D.; Wiorkiewiczkuczera, J.; Karplus, M. *J. Am. Chem. Soc.* **1995**, *117*, 11946–11975.

(6) 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*, 3586–3616.

(7) van Gunsteren, W. F.; Berendsen, H. J. C. *Groningen Molecular Simulation (GROMOS) Library Manual*; Biomos: Groningen, 1987.

(8) van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*; Vdf Hochschulverlag AG an der ETH Zürich: Zürich, Switzerland, 1996.

(9) Daura, X.; Mark, A. E.; van Gunsteren, W. F. *J. Comput. Chem.* **1998**, *19*, 535–547.

(10) Schuler, L. D.; Daura, X.; van Gunsteren, W. F. *J. Comput. Chem.* **2001**, *22*, 1205–1218.

(11) Oostenbrink, C.; Villa, A.; Mark, A. E.; van Gunsteren, W. F. *J. Comput. Chem.* **2004**, *25*, 1656–1676.

(12) Jorgensen, W. L.; TiradoRives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657–1666.

(13) Jorgensen, W. L.; Maxwell, D. S.; TiradoRives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.

(14) Walser, R.; Mark, A. E.; van Gunsteren, W. F.; Lauterbach, M.; Wipff, G. *J. Chem. Phys.* **2000**, *112*, 10450–10459.

(15) Geerke, D. P.; Oostenbrink, C.; van der Vegt, N. F. A.; van Gunsteren, W. F. *J. Phys. Chem. B* **2004**, *108*, 1436–1445.

(16) Smith, L. J.; Berendsen, H. J. C.; van Gunsteren, W. F. *J. Phys. Chem. B* **2004**, *108*, 1065–1071.

(17) Fioroni, M.; Burger, K.; Mark, A. E.; Roccatano, D. *J. Phys. Chem. B* **2000**, *104*, 12347–12354.

(18) Schmid, N.; Eichenberger, A.; Choutko, A.; Riniker, S.; Winger, M.; Mark, A. E.; van Gunsteren, W. F. *Eur. Biophys. J.* **2010**in press.

(19) Poger, D.; van Gunsteren, W. F.; Mark, A. E. *J. Comput. Chem.* **2010**, *31*, 1117–1125.

(20) Cubberley, M. S.; Iverson, B. L. *Curr. Opin. Chem. Biol.* **2001**, *5*, 650–653.

(21) Cheng, R. P. *Curr. Opin. Struct. Biol.* **2004**, *14*, 512–520.

(22) Hecht, E.; Huc, I. *Foldamers: structure, properties, and applications*; Wiley-VCH, New York, 2007.

(23) Hintermann, T.; Seebach, D. *Chimia* **1997**, *51*, 244–247.

(24) Seebach, D.; Abele, S.; Schreiber, J. V.; Martinoni, B.; Nussbaum, A. K.; Schild, H.; Schulz, H.; Hennecke, H.; Woessner, R.; Bitsch, F. *Chimia* **1998**, *52*, 734–739.

(25) Seebach, D.; Matthews, J. L. *Chem. Commun.* **1997**, *21*, 2015–2022.

(26) Cheng, R. P.; Gellman, S. H.; DeGrado, W. F. *Chem. Rev.* **2001**, *101*, 3219–3232.

(27) van Gunsteren, W. F.; Bürgi, R.; Peter, C.; Daura, X. *Angew. Chem., Int. Ed.* **2001**, *40*, 351–355.

(28) Daura, X.; Glattli, A.; Gee, P.; Peter, C.; van Gunsteren, W. F. *Adv. Protein Chem.* **2002**, *62*, 341–360.

(29) van Gunsteren, W. F.; Gattin, Z. In *Simulation of Folding Equilibria in Foldamers: Structure, Properties, and Applications*; Hecht, S., Huc, I., Eds.; Wiley, Weinheim, Germany: 2007, p 173—192.

(30) Seebach, D.; Ciceri, P. E.; Overhand, M.; Jaun, B.; Rigo, D.; Oberer, L.; Hommel, U.; Amstutz, R.; Widmer, H. *Helv. Chim. Acta* **1996**, *79*, 2043–2066.

(31) Daura, X.; van Gunsteren, W. F.; Rigo, D.; Jaun, B.; Seebach, D. *Chem.—Eur. J.* **1997**, *3*, 1410–1417.

(32) Daura, X.; Jaun, B.; Seebach, D.; van Gunsteren, W. F.; Mark, A. E. *J. Mol. Biol.* **1998**, *280*, 925–932.

(33) Daura, X.; Gademann, K.; Schafer, H.; Jaun, B.; Seebach, D.; van Gunsteren, W. F. *J. Am. Chem. Soc.* **2001**, *123*, 2393–2404.

(34) Lin, Z. X.; Schmid, N.; van Gunsteren, W. F. *Mol. Phys.* **2011**, *109*, 493–506.

(35) Christen, M.; Hünenberger, P. H.; Bakowies, D.; Baron, R.; Bürgi, R.; Geerke, D. P.; Heinz, T. N.; Kästenholz, M. A.; Kräutler, V.; Oostenbrink, C.; Peter, C.; Trzesniak, D.; van Gunsteren, W. F. *J. Comput. Chem.* **2005**, *26*, 1719–1751.

(36) Lin, Z. X.; van Gunsteren, W. F. *Phys. Chem. Chem. Phys.* **2010**, *12*, 15442–15447.

(37) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Dinola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

(38) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. phys.* **1977**, *23*, 327–341.

(39) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1995**, *102*, 5451–5459.

(40) Daura, X.; van Gunsteren, W. F.; Mark, A. E. *Proteins: Struct. Funct. Genet.* **1999**, *34*, 269–280.

(41) Wang, D.; Jaun, B.; van Gunsteren, W. F. *ChemBioChem* **2009**, *10*, 2032–41.

(42) Wüthrich, K.; Billeter, M.; Braun, W. *J. Mol. Biol.* **1983**, *169*, 949–961.

(43) Karplus, M. *J. Chem. Phys.* **1959**, *30*, 11–15.

(44) Pardi, A.; Billeter, M.; Wüthrich, K. *J. Mol. Biol.* **1984**, *180*, 741–751.

(45) de Marco, A.; Llinas, M.; Wüthrich, K. *Biopolymers* **1978**, *17*, 617–636.

(46) Gee, P. J.; van Gunsteren, W. F. *Proteins Struct. Funct. Bioinf.* **2006**, *63*, 136–143.

JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Boxed Molecular Dynamics: Decorrelation Time Scales and the Kinetic Master Equation

David R. Glowacki,*[,†] Emanuele Paci,[‡] and Dmitrii V. Shalashilin[§]

[†]Centre for Computational Chemistry, University of Bristol, Bristol BS8 1TS, United Kingdom
[‡]Institute of Molecular and Cellular Biology, University of Leeds, Leeds LS2 9JT, United Kingdom
[§]School of Chemistry, University of Leeds, Leeds LS2 9JT, United Kingdom

ABSTRACT: A number of methods proposed in the past few years have been aimed at accelerating the sampling of rare events in molecular dynamics simulations. We recently introduced a method called Boxed Molecular Dynamics (BXD) for accelerating the calculation of thermodynamics and kinetics (*J. Phys. Chem. B* **2009**, *113*, 16603−16611). BXD relies upon confining the system in a series of adjacent "boxes" by inverting the projection of the system velocities along the reaction coordinate. The potential of mean force along the reaction coordinate is obtained from the mean first passage times (MFPTs) for exchange between neighboring boxes, simultaneously providing both kinetics and thermodynamics. In this paper, we investigate BXD in the context of its natural relation to a kinetic master equation and show that the BXD first passage times (FPTs) include different time scales—a fast short time decay due to correlated dynamical motion and slower long time decay arising from phase space diffusion. Correcting the FPTs to remove the fast correlated motion yields accurate thermodynamics and master equation kinetics. We also discuss interrelations between BXD and a recently described Markovian milestoning technique and use a simple application to show that, despite each method producing distinct nonstatistical effects on time scales on the order of dynamical decorrelation, both yield similar long-time kinetics.

1. INTRODUCTION

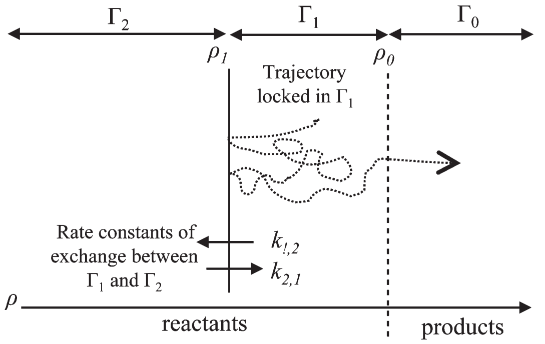
Accurate sampling of rare events remains a significant challenge to molecular dynamics (MD) simulations, and a number of methods have been proposed to address it.[1−27] In a recent publication, we introduced a simple and exact technique for accelerating molecular dynamics (MD) simulations.[28] By slicing a reaction coordinate into "boxes" along some reaction coordinate, short-time dynamics within each box may be analyzed in order to obtain mean first passage times (MFPTs) between neighboring boxes. This method, which we named "boxed" molecular dynamics (BXD), is an extension of the method of classical dynamics accelerated by phase space constraints[29,30] and has its origin in Intramolecular Dynamics Diffusion Theory (IDDT).[31−34] IDDT shows that Newtonian classical molecular dynamics written in the form of the Liouville equation can effectively be replaced by an equation for diffusion along the reaction coordinate.[33,34] In IDDT, the reaction coordinate is divided into boxes, and coefficients for the diffusion equation or the equivalent Langevin equation[31] are determined by short-time MD with initial conditions sampled box by box. Reactions on longer time scales may then be reconstructed from the set of short time simulations by integration. A number of modern techniques, including BXD, are similar in spirit.

Free energies and rate coefficients are among the most important calculable quantities that may be obtained from MD; however, unlike BXD, few methods simultaneously provide both kinetic and thermodynamic information. In this paper, we consider in detail the relationship between the kinetic master equation and the BXD technique. This connection naturally arises because kinetics and thermodynamics are linked in BXD—i.e., thermodynamics are obtained from BXD using box-to-box mean first passage times (MFPTs). Through consideration of the loop formation dynamics of a simple 10-ALA peptide, we show that the distribution of FPTs obtained from BXD includes distinct time scales: fast dynamically correlated motion at short times, with slower diffusional decay at longer times. This separation in FPT time scales is akin to the sorts of dynamical recrossing corrections to transition state theory (TST) rate coefficients which arise from the fluctuation−dissipation theorem.[35−38] In this paper, we show that free energy surfaces obtained with BXD are largely independent of fast correlated motion; however, an accurate calculation of phenomenological kinetics using a kinetic master equation requires that the box-to-box MFPTs are corrected for the fast dynamical motion.

We also consider the relationships between BXD and a recent modification of milestoning[39−41] called Markovian milestoning (MM).[7,8] The distribution of FPTs used to solve the kinetic master equation in Markovian milestoning is distinct from that obtained using BXD, and we show that the difference primarily arises in how each treats the fast correlated motion. Whereas the BXD FPTs include very fast recrossing events, the Markovian milestoning FPTs have an initial short time lag. When the BXD FPTs are corrected for recrossing, they give kinetic master equation solutions in close agreement with those obtained from Markovian milestoning, in addition to providing accurate thermodynamics.

Received: January 5, 2011
Published: April 19, 2011

Scheme 1. Illustration of the AXD Approach (Boxed Molecular Dynamics with Two Boxes) for Calculating Accelerated Reaction Rates[a]


Γ_2
Γ_1
Γ_0
ρ_1
ρ_0
Trajectory locked in Γ_1
Rate constants of exchange between Γ_1 and Γ_2
$k_{1,2}$
$k_{2,1}$
ρ
reactants
products

[a] The trajectory simulation locked in Γ_1 provides the accelerated rate coefficient. The correction factor is expressed through phase volumes Γ_1 and Γ_2 (or rate constants k_{12} and k_{21}) as described in the text.

2. BRIEF OVERVIEW OF BOXED DYNAMICS FORMALISM

The technique we introduced previously was composed of two parts (AXD and BXD), which we named differently according to the circumstances of their usage—AXD is an abbreviation of "accelerated dynamics" and is generally intended for accelerating the calculation of rate coefficients in MD, while BXD is intended for accelerating potential of mean force calculations. However, both techniques amount to the same procedure, despite their different intents. The only significant difference is that AXD has two boxes, while BXD has an arbitrary number of boxes. Thus, it is appropriate to refer to both methods as "boxed molecular dynamics". To keep this paper reasonably self-contained, a brief summary of AXD and BXD is included below.

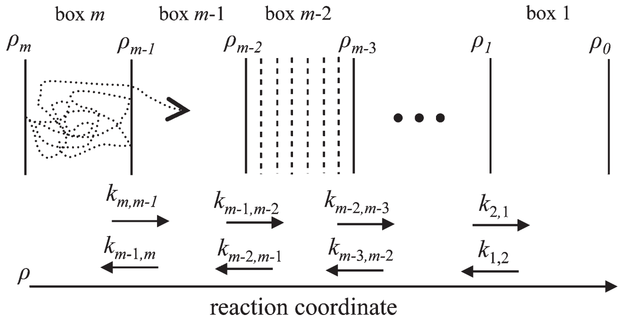
Boxed molecular dynamics relies on the assumption that the classical mechanics of anharmonic systems on time scales above the dynamical decorrelation time, τ_{corr}, are essentially ergodic. In such systems, equilibrium is quickly established between neighboring regions of the system phase space. So long as this is the case, dynamics can be accurately approximated using Markov models (obtained via Monte Carlo methods or short MD trajectories) for describing the exchange between different regions of the phase space.[35−38] Scheme 1 illustrates the main ideas of AXD.

According to classical TST, the phase space of the system is separated into reactant and product regions by a dividing surface at ρ_0 along some reaction coordinate, and the reaction rate coefficient is then calculated as a flux through the dividing surface. AXD accelerates passage through ρ_0 by splitting the reactant phase space into two boxes: Γ_1, which spans ρ_0 to ρ_1, and Γ_2, which is bounded by ρ_1. By locking the dynamics within Γ_1, the trajectory crosses the transition state more often, yielding an accelerated rate coefficient, k^{AXD}. The actual rate coefficient, $k(T)$, of going from the reactant region, $\Gamma_1 + \Gamma_2$, to the product region, Γ_0, may then be recovered as

$$k(T) = k^{\mathrm{AXD}} \times P^{\mathrm{CORR}} \tag{1}$$

where the P^{CORR} correction factor is the probability of finding the system in Γ_1. Trajectories are confined within Γ_1 by utilizing a velocity inversion algorithm which conserves the total energy, linear momentum, and angular momentum. At each integration time step t, we calculate the trajectory's position along the reaction coordinate. If it moves outside ρ_1 at time step $t + dt$, then we return to the previous step t and invert the projections of the velocities along the reaction coordinate.

Scheme 2. Illustration of the BXD Procedure for a System Partitioned into m "Boxes"[a]



[a] In this simple picture, the trajectory penetrates from box m into box $m - 1$ after two inversions at the ρ_{m-1} boundary.

Provided the assumption of equilibrium between boxes Γ_1 and Γ_2 is valid, P^{CORR} is calculated simply as the fraction of the phase volume Γ_1 to the total reactant phase volume.

$$P^{\mathrm{CORR}} = \frac{\Gamma_1}{\Gamma_1 + \Gamma_2} \tag{2}$$

The phase volume ratio in eq 2 may be estimated from a Monte Carlo random walk or by running a trajectory in $\Gamma_1 + \Gamma_2$. Another way to calculate the correction factor is to recognize that the ratio of the two phase volumes is simply the equilibrium constant k_{21} of exchange between Γ_1 and Γ_2, which can be estimated from classical molecular dynamics as a ratio of the box-to-box rate constants k_{12} and k_{21}:

$$P^{\mathrm{CORR}} = \frac{1}{1 + \dfrac{\Gamma_2}{\Gamma_1}} = \frac{1}{1 + K_{21}} = \frac{1}{1 + \dfrac{k_{12}}{k_{21}}} \tag{3}$$

The fundamental efficiency gain of AXD derives from the fact that it is less expensive to converge k^{AXD} and P^{CORR} separately than their small product $k(T)$, and the bulk of this paper is concerned with how to accurately calculate k_{12} and k_{21}.

If motion along the reaction coordinate involves rare events, then boxed dynamics (BXD), which is a simple extension of AXD, offers a tractable means for obtaining P^{CORR}. The BXD approach is illustrated in Scheme 2, which shows the reaction coordinate ρ split into m intervals by boundaries at ρ_0, ρ_1, ..., ρ_{m-1}, ρ_m. Velocity inversion is carried out at each of the boundaries. By counting the total time t the trajectory spends in a particular box as well as the number of inversions at a particular boundary, the kinetics for "exchange" between the neighboring boxes may be obtained, as described further below.

If the box boundaries are positioned so that there is no irreversible flux through border ρ_0 or ρ_m (i.e., $k_{1,0} = k_{m,m+1} = 0$), one may obtain thermodynamic information along the entire reaction coordinate of a system like that described in Scheme 2 using the box-to-box forward and reverse rate constants. So long as a temperature may be defined, equilibrium constants between the neighboring boxes n and $n - 1$ may be obtained as follows:

$$K_{n-1,n} = \frac{k_{n-1,n}}{k_{n,n-1}} = \exp\left(\frac{-\Delta G_{n-1,n}}{kT}\right) \qquad (4)$$

The free energy difference between each neighboring box, $\Delta G_{n-1,n}$, may then be found by rearranging eq 4. With respect to an arbitrary zero, each box averaged free energy, ΔG_n, may then be determined together with p_n, the equilibrium probability of residing in box n:

$$p_n = \frac{1}{\sum_n \exp(-\Delta G_n/kT)} \exp(-\Delta G_n/kT) \qquad (5)$$

The time that the trajectory spends in each box is determined by how long it takes for the rate coefficients in each box to converge. As described in our previous article,[28] the statistics obtained within each box may be placed into smaller histogram bins (represented by the dashed lines in Scheme 2) and then exactly renormalized to provide the probability distribution along the entire reaction coordinate, $p(\rho)$, to a higher resolution.

3. DYNAMICAL DECORRELATION IN AXD

Accurate calculation of the rate constants of "exchange" between the boxes is the key to the AXD and BXD techniques. In this section, we consider in detail how dynamical decorrelation time scales affect the two-box AXD procedure. AXD is based upon eq 1, whose origin may be understood through consideration of the classical canonical TST expression for calculating the reactive flux from the reactant phase space, R, across a dividing surface in phase space. Dividing surfaces in coordinate space are more common than coordinates in phase space because of their relative simplicity; however, we note that the TST hypersurface may be defined by both coordinates, $\mathbf{q}$, and momenta, $\mathbf{p}$. For the purposes of this article, we simply and generally write the TST rate coefficient as

$$k_{R\rightarrow}^{\mathrm{TST}}(T) = \frac{\langle |v|\delta(\mathbf{q},\mathbf{p})\ \Theta(\mathbf{q},\mathbf{p})\rangle}{\Gamma_R} \qquad (6)$$

where $|v|$ is the magnitude of the velocity vector normal to the transition state dividing surface in phase space, $\Theta(\mathbf{q},\mathbf{p})$ is an indicator function which is unity when the system is in state R and zero otherwise, and $\delta(\mathbf{q},\mathbf{p})$ is a Dirac δ function which is unity at the dividing surface. The numerator in eq 6 uses angled brackets $\langle ... \rangle$ to indicate that it is the Boltzmann weighted average of velocities perpendicular to the TS dividing surface which are leaving R. The integral in the denominator is the reactant phase space volume Γ_R. The logic of Scheme 1 and the corresponding basis for eq 1 derives from the fact that the reactant phase space $\Gamma_R = \Gamma_1 + \Gamma_2$, which allows eq 6 to be written as

$$\begin{aligned} k_{R\rightarrow}^{\mathrm{TST}}(T) &= \frac{\langle |v|\delta(\mathbf{q},\mathbf{p})\ \Theta(\mathbf{q},\mathbf{p})\rangle}{\Gamma_1 + \Gamma_2} \\ &= \frac{\langle |v|\delta(\mathbf{q},\mathbf{p})\ \Theta(\mathbf{q},\mathbf{p})\rangle}{\Gamma_1} \times \frac{\Gamma_1}{\Gamma_1 + \Gamma_2} \\ &= k^{\mathrm{AXD}} \times P^{\mathrm{CORR}} \end{aligned} \qquad (7)$$

Neither eq 6 nor eq 7 include corrections for motion on time scales smaller than that of τ_{corr},[35−38] which enters in the form of a scaling factor, κ, reducing the magnitude of the flux through the TST dividing surface:

$$\begin{aligned} \kappa k_{R\rightarrow}^{\mathrm{TST}}(T) &= \kappa \frac{\langle |v|\delta(\mathbf{q},\mathbf{p})\ \Theta(\mathbf{q},\mathbf{p})\rangle}{\Gamma_1} \times \frac{\Gamma_1}{\Gamma_1 + \Gamma_2} \\ &= \kappa k^{\mathrm{AXD}} \times P^{\mathrm{CORR}} \end{aligned} \qquad (8)$$

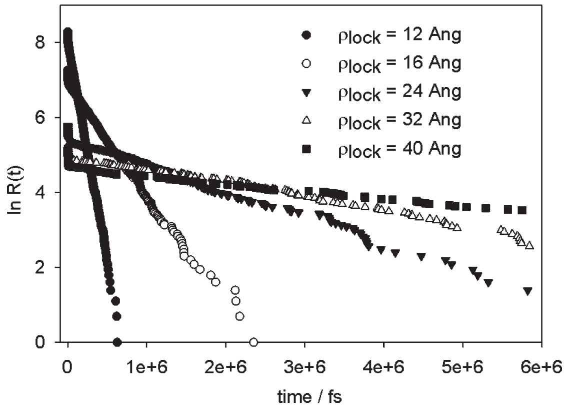


Figure 1. AXD decay traces at varying values of ρ_1 where ρ_1 is equivalent to ρ_{lock}. Each trace shows a fast decay at short times deriving from recrossing and a slow decay at longer times.

Equation 8 shows that κ is an effective scaling factor to the accelerated rate coefficient, k^{AXD}, in the same way that it is for k^{TST}; this is because the rapid dynamical motion for which κ accounts is localized in the phase space neighborhood of the dividing surface. The correction factor P^{CORR}, being an equilibrium property of the system, is applied in exactly the same fashion whether or not κ is included, and this holds so long as Γ_1 is larger than the dynamical decorrelation length scale near the dividing surface ρ_0 (see Scheme 1).

Consideration of fast dynamical motion on passage through the dividing surface ρ_0 is critical to obtaining an accurate estimate of the unbiased rate coefficient. Below, we illustrate this point for the loop formation dynamics of a 10-alanine peptide where the reaction coordinate, ρ, is the distance between the main-chain nitrogen on the N-terminal residue and the carboxyl carbon of the C terminal residue—i.e., the peptide extension.[42,43] We ran dynamics using a CHARMM19 force field and a simple implicit solvation model.[44] With this force field, 10-alanine turns out to have a strong helical propensity. The MD was run for 1 μs using Langevin dynamics with a friction coefficient of 1 ps^{-1} and a time step of 1 fs. The transition state ρ_0 between reactants (the extended peptide) and the product (the peptide which formed a loop) was chosen at a separation of 4.7 Å, near the free energy barrier to contact formation, which was determined in our previous article.[28]

In what follows, we consider the survival probability (or decay trace), $R(t)$, obtained from the lifetime distribution $N(t)$ for passage across a particular boundary within a particular box. In general, $R(t) = \int_0^{t_{\max}} N(t')\,dt' - \int_0^t N(t')\,dt'$ where $t_{\max}$ is the maximum lifetime in the distribution. Figure 1 nicely illustrates the principles discussed above in eqs 6−8. It shows $R(t)$ obtained from FPTs through ρ_0, where ρ_0 is the loop formation TS dividing surface separating Γ_1 from Γ_0 in Scheme 1. $R(t)$ traces were obtained by running a single long trajectory constrained to remain within Γ_1 and Γ_0 and are shown at different values of ρ_1. Each of the traces in Figure 1 clearly shows two different decay time scales—typical of those often observed in studies of

Table 1. AXD Results of the Loop Formation Rate Obtained with Different Values of ρ_1[a]

ρ_1/Å	k^{AXD}/s^{-1}	P^{CORR}	$(k^{\mathrm{AXD}}P^{\mathrm{CORR}}) \pm \sigma$/s^{-1}
10	2.87×10^{10}	0.021	$(1.29 \pm 0.65) \times 10^{9}$
12	8.77×10^{9}	0.045	$(1.05 \pm 0.74) \times 10^{9}$
16	1.50×10^{9}	0.099	$(1.05 \pm 0.50) \times 10^{9}$
20	7.93×10^{8}	0.649	$(7.90 \pm 3.60) \times 10^{8}$
24	1.12×10^{9}	0.985	$(1.12 \pm 0.36) \times 10^{9}$

[a] The corrected rate coefficients, $k^{\mathrm{AXD}}P^{\mathrm{CORR}}$, are given in the final column and plotted in Figure 2.

nonstatistical effects in trajectory simulations of chemical reactions.[29,45] The very fast decay of $R(t)$ at short times in Figure 1 arises from fast recrossing through the ρ_0 surface within the dynamical decorrelation time scale, τ_{corr}, and does not vary significantly with different locations of ρ_1 (in Figure 1, ρ_1 is equivalent to ρ_{lock} to maintain consistency with notation used in our previous paper).[28] The slower decay at long times arises from trajectories that have lost dynamical memory within Γ_1 before subsequent ρ_0 crossings. The long time $R(t)$ decay profiles in Figure 1 are inversely proportional to Γ_1. Only by removing the very fast decay events in Figure 1 is it possible to obtain MFPTs and rate coefficients that are statistically identical when scaled by the appropriate values of P^{CORR}, and this is the procedure that we adopted for the results reported in our previous paper.[46] Table 1 and Figure 2 give the corrected rate coefficients, $k^{\mathrm{AXD}}P^{\mathrm{CORR}}$, where k^{AXD} has been determined from the long time $R(t)$ decay.

4. DYNAMICAL DECORRELATION IN BXD

Having described how to account for nonergodic dynamical motion in two-box AXD simulations, we now turn to multiple-box BXD simulations. At the outset, we note that the methodology described above is similarly applicable. A convenient starting point for the discussion in this section begins by considering BXD's straightforward relation to the kinetic master equation. Assuming for the moment that BXD has provided us with a set of average box-to-box rate coefficients, then the global time dependence of any box population may be described using a set of coupled first-order differential equations:

$$\begin{aligned}
\frac{dn_1(t)}{dt} &= -(k_{21}+k_{10})n_1(t) + k_{21}n_2(t) \\
\frac{dn_2(t)}{dt} &= k_{12}n_1(t) + k_{32}n_2(t) - (k_{21}+k_{23})n_2(t) \\
&\ldots \\
\frac{dn_m(t)}{dt} &= (k_{m-1,m}n_{m-1}(t) - k_{m,m-1}n_m(t)
\end{aligned} \qquad (9)$$

Equation 9 is a discretized kinetic master equation (ME),[47−54] where k_{ij} is an average rate coefficient for transfer from box i to its neighboring box j. The equation for n_1 is written assuming that passage across boundary ρ_0 is irreversible.

The whole set of coupled differential equations may be expressed as a matrix eigenvalue problem

$$\frac{d\mathbf{n}(t)}{dt} = \mathbf{M}\mathbf{n}(t) \qquad (10)$$

where $\mathbf{n}(t)$ is a vector containing the time dependent populations of each box and $\mathbf{M}$ is the matrix of rate coefficients in eq 9.

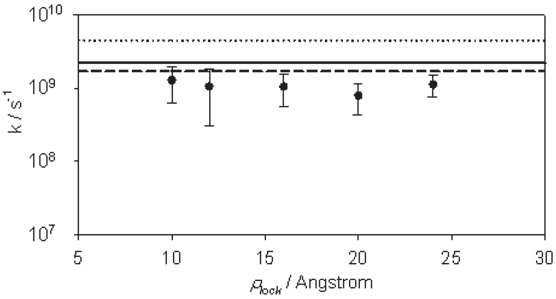


Figure 2. Corrected rate coefficients for contact formation at different values of ρ_{lock}. Also shown are the eigenvalues of smallest magnitude obtained from solving the kinetic master equation using the following methods to obtain transition probabilities: from MFPTs in column 3 of Table 2 (·····), from corrected MFPTs in column 4 of Table 2 (—), and from milestoning MFPTs in column 3 of Table 3 (----).

Solution of eq 10 provides the time dependence of $\mathbf{n}(t)$ and has the form

$$\mathbf{n}(t) = \mathbf{U}e^{\lambda t}\mathbf{U}^{-1}\mathbf{n}(0) \qquad (11)$$

where $\mathbf{n}(0)$ contains the initial conditions for each box, $\mathbf{U}$ is the eigenvector matrix obtained from diagonalization of $\mathbf{M}$, and λ is a vector of the corresponding eigenvalues, where the total number of eigenvalues is equal to m, the number of boxes. An important property of the matrix $\mathbf{M}$ is that the lowest few eigenvalues are often isolated from the other eigenvalues. In such cases, the smallest eigenvalues determine the time scale of kinetic evolution—an example of which is given below.

In boxed MD simulations, practical determination of the rate coefficients for transfer from a particular box to neighboring boxes utilizes the inverse of the mean first passage time (MFPT), $\langle\tau\rangle$. The simplest way to calculate $\langle\tau\rangle$ is by keeping track of (1) how many times a trajectory is inverted at a particular boundary, h, and (2) the lifetime of the trajectory in a particular box. For example, if we run constrained dynamics in box i, which is bounded by ρ_i and ρ_{i-1}, the respective rate coefficients for transfer from box $i \rightarrow i-1$ and box $i \rightarrow i+1$ are

$$\begin{aligned}
k_{i,i-1} &= \langle\tau_{i,i-1}\rangle^{-1} = \frac{h_{i,i+1}}{t_i} \\
k_{i,i+1} &= \langle\tau_{i,i+1}\rangle^{-1} = \frac{h_{i,i-1}}{t_i}
\end{aligned} \qquad (12)$$

where t_i is the lifetime of the trajectory in box i and $h_{i,i-1}$ and $h_{i,i+1}$ are the respective numbers of hits (i.e., velocity inversions) at the walls ρ_{i-1} and ρ_{i+1}. When eq 12 is corrected for the effects of fast correlated motion in the same way as AXD, eq 12 may be rewritten as

$$\begin{aligned}
\kappa k_{i,i-1} &= \kappa\langle\tau_{i,i-1}\rangle^{-1} \\
\kappa k_{i,i+1} &= \kappa\langle\tau_{i,i+1}\rangle^{-1}
\end{aligned} \qquad (13)$$

When BXD is used in order to obtain free energy surfaces by rearranging eq 4, the box-to-box MFPTs *per se* are never required—only their ratio. Thus, the effects of dynamical decorrelation in the forward and reverse directions cancel if they are approximately equal—i.e., $\kappa \approx \kappa_{i,j} \approx \kappa_{j,i}$, an assumption which is discussed further below.

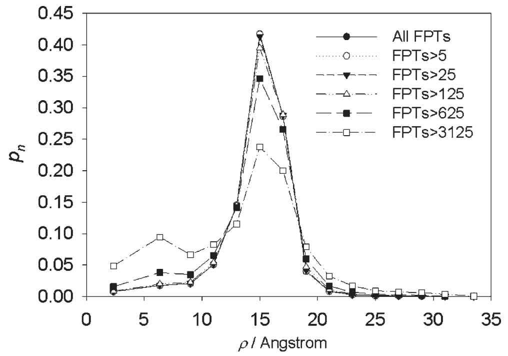


Figure 3. Box averaged probabilities along the reaction coordinate ρ obtained using eq 15 and eq 5. The plot compares box averaged probabilities obtained using values for τ_{corr} of 0, 5, 25, 125, 625, and 3125 fs. The latter two free energy profiles deviate significantly from the others.

The role of κ is to remove fast events on the dynamical decorrelation time scale, τ_{corr}, from the inverse MFPTs, $\langle\tau\rangle^{-1}$, in eq 12. Another way to accomplish this is to replace $\langle\tau\rangle$ with $\langle\tau'\rangle$ using the following equation:

$$\kappa k_{i,i-1} = \langle\tau'_{i,i-1}\rangle^{-1} = \left(\frac{\sum_k \tau_{i,i-1}\theta(\tau^k_{i,i-1} - \tau_{\text{corr}})}{\sum_k \theta(\tau^k_{i,i-1} - \tau_{\text{corr}})}\right)^{-1}$$
$$\kappa k_{i,i+1} = \langle\tau'_{i,i+1}\rangle^{-1} = \left(\frac{\sum_k \tau_{i,i+1}\theta(\tau^k_{i,i+1} - \tau_{\text{corr}})}{\sum_k \theta(\tau^k_{i,i+1} - \tau_{\text{corr}})}\right)^{-1} \tag{14}$$

where k is an index that runs over all of the individual first passage times obtained in going from box i toward respective boxes $i - 1$ and $i + 1$, and $\theta(t)$ is an indicator function which is zero when $t < 0$ and unity otherwise. Equation 14 calculates $\langle\tau'\rangle$ as a MFPT which excludes events shorter than τ_{corr}. In the limit that $\tau_{\text{corr}} = 0$, then $\langle\tau\rangle$ in eq 12 is identical to $\langle\tau'\rangle$ in eq 14.

Replacing the values of k in eq 4 with the values of $\langle\tau'\rangle$ in eq 14 yields

$$K_{n-1,n} = \frac{\langle\tau'_{n-1,n}\rangle^{-1}}{\langle\tau'_{n,n-1}\rangle^{-1}} = \exp\left(\frac{-\Delta G_{n-1,n}}{kT}\right) \tag{15}$$

enabling one to subsequently calculate box averaged probabilities, p_n, with eq 5. The difference between eq 4 and eq 15 is that the latter includes corrections to the MFPTs which account for motion on the time scale of τ_{corr}. Figure 3 shows the box averaged probability distributions calculated using eq 15 and eq 5 at values for τ_{corr} ranging from 0 to 3125 fs with τ_{corr} arbitrarily chosen as 5^0, 5^1, 5^2, etc. In these simulations, the reaction coordinate was split into 16 boxes, with indices 0−15, using the box boundaries given in Table 2. The box-to-box values for $\langle\tau'\rangle^{-1}$ (with $\tau_{\text{corr}} = 0$ and 125 fs) are given in Table 2. The probability distributions in Figure 3 are statistically indistinguishable for values of τ_{corr} from 0 to 125 fs. Beyond 125 fs, the results start to change significantly. The decorrelation time scale, τ_{corr}, will vary from system to system; however, analysis of the sort shown in Figure 3 helps to place a quantitative upper limit on its value for any particular system and suggests that BXD is a robust method for obtaining PMFs in molecular systems.

Table 2. Inverse MFPTs Obtained Using eq 14 for the 15 Box System Described in the Text[a]

box index			
i	j	$\langle\tau'_{ij}\rangle^{-1}$/fs^{-1}	$\langle\tau'_{ij}\rangle^{-1}$/ fs^{-1}
1 (4.7−8 Å)	0 (0−4.7 Å)	4.33×10^{-7}	2.17×10^{-7}
1 (4.7−8 Å)	2 (8−10 Å)	9.73×10^{-7}	4.94×10^{-7}
2 (8−10 Å)	1 (4.7−8 Å)	8.36×10^{-7}	4.37×10^{-7}
2 (8−10 Å)	3 (10−12 Å)	1.86×10^{-6}	9.33×10^{-7}
3 (10−12 Å)	2 (8−10 Å)	7.53×10^{-7}	3.91×10^{-7}
3 (10−12 Å)	4 (12−14 Å)	1.88×10^{-6}	9.20×10^{-7}
4 (12−14 Å)	3 (10−12 Å)	6.59×10^{-7}	3.44×10^{-7}
4 (12−14 Å)	5 (14−16 Å)	1.77×10^{-6}	8.55×10^{-7}
5 (14−16 Å)	4 (12−14 Å)	6.14×10^{-7}	3.12×10^{-7}
5 (14−16 Å)	6 (16−18 Å)	1.63×10^{-6}	8.43×10^{-7}
6 (16−18 Å)	5 (14−16 Å)	2.38×10^{-6}	1.15×10^{-6}
6 (16−18 Å)	7 (18−20 Å)	3.87×10^{-7}	1.87×10^{-7}
7 (18−20 Å)	6 (16−18 Å)	2.76×10^{-6}	1.14×10^{-6}
7 (18−20 Å)	8 (20−22 Å)	4.39×10^{-7}	2.04×10^{-7}
8 (20−22 Å)	7 (18−20 Å)	2.13×10^{-6}	8.92×10^{-7}
8 (20−22 Å)	9 (22−24 Å)	6.27×10^{-7}	2.93×10^{-7}
9 (22−24 Å)	8 (20−22 Å)	2.00×10^{-6}	8.54×10^{-7}
9 (22−24 Å)	10 (24−26 Å)	6.91×10^{-7}	3.30×10^{-7}
10 (24−26 Å)	9 (22−24 Å)	1.46×10^{-6}	6.58×10^{-7}
10 (24−26 Å)	11 (26−28 Å)	1.05×10^{-6}	4.93×10^{-7}
11 (26−28 Å)	10 (24−26 Å)	1.34×10^{-6}	6.19×10^{-7}
11 (26−28 Å)	12 (28−30 Å)	9.92×10^{-7}	4.52×10^{-7}
12 (28−30 Å)	11 (26−28 Å)	1.54×10^{-6}	6.77×10^{-7}
12 (28−30 Å)	13 (30−32 Å)	7.40×10^{-7}	3.24×10^{-7}
13 (30−32 Å)	12 (28−30 Å)	2.72×10^{-6}	1.08×10^{-6}
13 (30−32 Å)	14 (32−35 Å)	1.87×10^{-7}	8.70×10^{-8}
14 (32−35 Å)	13 (30−32 Å)	5.94×10^{-6}	2.05×10^{-6}
smallest eigenvalue		4.50×10^{-9}	2.20×10^{-9}

[a] The inverse MFPTs in column 3 were obtained with $\tau_{\text{corr}} = 0$ fs. Those in the fourth column were obtained with $\tau_{\text{corr}} = 125$ fs. The eigenvalues of smallest absolute magnitude, obtained from solution of the kinetic master equation with each set of corresponding inverse MFPTs are given in the last row.

Figure 3 is consistent with the fact that local dynamical motion on the order of the τ_{corr} is approximately identical in both the forward and backward directions. A simple rationalization of the agreement between the box averaged probability distributions in Figure 3 begins from writing the MFPT for transfer from box i to j as

$$\langle t_{ij}\rangle = \alpha^{\text{corr}}_{ij}\langle t^{\text{corr}}_{ij}\rangle + \alpha^{\text{diff}}_{ij}\langle t^{\text{diff}}_{ij}\rangle \tag{16}$$

where $\langle t^{\text{corr}}_{ij}\rangle$ is the average decorrelation time, $\langle t^{\text{diff}}_{ij}\rangle$ is the average decay at times longer than τ_{corr}, which corresponds to diffusive box-to-box motion, and $\langle\alpha^{\text{corr}}_{ij}\rangle$ and $\langle\alpha^{\text{diff}}_{ij}\rangle$ are the respective fractions of these decay times. So long as the box is large enough, then $\langle t_{ij}\rangle$ is dominated by the longer time motion—i.e.,

$$\langle t_{ij}\rangle \approx \alpha^{\text{diff}}_{ij}\langle t^{\text{diff}}_{ij}\rangle \tag{17}$$

In the ergodic approximation, the same incoming trajectories are balanced by outgoing trajectories on the shared boundary between boxes i and j. In this limit, it is reasonable to assume

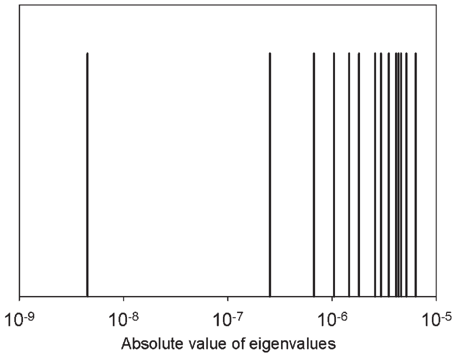


Figure 4. A typical spectrum of the absolute values of the eigenvalues obtained from diagonalization of **M** for the 10-ALA example considered in this work. Loop formation kinetics are dominated by the eigenvalue of smallest absolute magnitude, which is well separated from the others.

that the fraction of diffusive trajectories (those that decorrelate following inversion at the boundary) is approximately the same in both the forward and reverse directions across a particular boundary—i.e., $\alpha_{ij}^{\text{diff}} \approx \alpha_{ji}^{\text{diff}}$ which implies that $\alpha_{ij}^{\text{corr}} \approx \alpha_{ji}^{\text{corr}}$. Hence, the ratio of the rate constants required in eq 4 depends only on the long time decay

$$\frac{k_{ij}}{k_{ji}} \approx \frac{\langle \tau_{ij}^{\text{diff}} \rangle^{-1}}{\langle \tau_{ji}^{\text{diff}} \rangle^{-1}} \qquad (18)$$

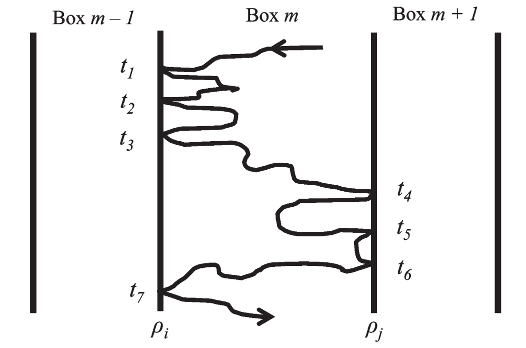
The same is not, however, true for the kinetics: solutions to the kinetic master equation (eqs 9—11) are rather more sensitive to dynamical corrections than the box averaged probabilities. If **M** is constructed using MFPTs corrected for dynamical decorrelation, then the eigenvalue of smallest absolute magnitude (i.e., the rate coefficient for loop formation kinetics) is in significantly better agreement with the unbiased rate coefficients shown in Table 1 and Figure 2. To demonstrate this fact, we used the data in Table 2 to solve the kinetic master equation. Passage over ρ_0 (the loop formation boundary at 4.7 Å for transfer from box 0→1) was treated as an irreversible channel. Other than the boxes at the extrema of the extension reaction coordinate, which have only one set of outgoing transition probabilities, each box in Table 2 has both an ingoing and an outgoing $\langle \tau' \rangle^{-1}$. The final row of Table 2 gives the eigenvalue of smallest magnitude obtained from diagonalization of **M** using MFPTs obtained with eq 14 and τ_{corr} equal to 0 and 125 fs.

The inverse MFPTs in Table 2 result in **M** having one zero eigenvalue, with all of the others negative. Inspection of the eigenvalue spectrum of **M** in Figure 4 shows a single eigenvalue that is well separated from all the others by more than an order of magnitude. Given the separation in eigenvalues, it is a good approximation to represent the system loop formation kinetics using a single exponential term containing the unique eigenvalue. The recrossing corrected inverse MFPTs in the fourth column of table two (τ_{corr} = 125 fs.) are roughly half as large as those which are uncorrected (i.e., τ_{corr} = 0 fs), and the effect on the corresponding eigenvalues is approximately the same, with the corrected eigenvalues in better agreement with results obtained from unbiased simulations (see Figure 1).

5. SHORT TIME BEHAVIOR IN BXD AND MILESTONING

Another method that shares similarities with the BXD technique is milestoning,[7,8,14] which uses short time dynamics to recover local kinetic information about the motion along the reaction coordinate in order to solve a kinetic state-to-state master equation. In particular, a recent variant of milestoning called Markovian milestoning[7] has been proposed wherein the dynamics initiated between a set of milestones are locked so that they cannot escape; however, the kinetic master equation is different for each technique: BXD describes box-to-box transfer, whereas milestoning describes boundary-to-boundary transfer. BXD's formulation arises from a sort of transition state theory (TST) intuition typical in chemistry where we tend to think of chemical reactions in terms of a reactant phase space bounded by a TS dividing surface. This allows us to easily think in terms of a “box averaged free energy”.

Scheme 3. Illustration of the Time Propagation, *t*, of a Constrained Trajectory, along with a Table That Illustrates the Manner in Which FPTs Are Calculated for Both BXD and Milestoning



BXD FPTs		Milestoning FPTs	
$m \rightarrow m+1$	$m \rightarrow m-1$	$m \rightarrow m+1$	$m \rightarrow m-1$
$t_5 - t_4$	$t_2 - t_1$	$t_4 - t_1$	$t_7 - t_4$
$t_6 - t_5$	$t_3 - t_2$		
	$t_7 - t_3$		

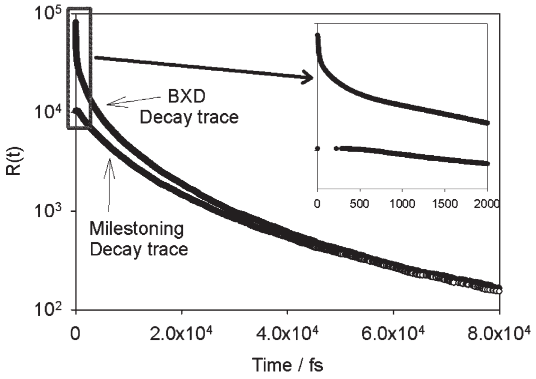


Figure 5. Typical comparison of the decay traces, $R(t)$, generated from the definition of BXD FPTs and milestoning FPTs for the 10-ALA model system. The BXD decay is for transfer from box 2 (spanning ρ = 4.7 Å to ρ = 8.0 Å) to box 1 (spanning ρ = 4.7 Å to ρ = 0.0 Å), whereas the milestoning decay corresponds to transit from milestone 2 (ρ = 8.0 Å) to milestone 1 (ρ = 4.7 Å). The inset shows the different short time behavior for each decay trace.

A milestone, on the other hand, is an infinitesimal slice of phase space along some reaction coordinate. Markovian milestoning utilizes different sets of transition probabilities, depending

Table 3. Rate Coefficients Used to Solve the Kinetic Master Equation Obtained by Obtaining MFPTs from the Milestoning Decay Traces[a]

milestone position		
i	j	k_{ij}/fs^{-1}
1 (8 Å)	0 (4.7 Å)	9.00×10^{-8}
0 (4.7 Å)	1 (8 Å)	1.44×10^{-7}
2 (10 Å)	1 (8 Å)	3.31×10^{-7}
1 (8 Å)	2 (10 Å)	6.07×10^{-7}
3 (12 Å)	2 (10 Å)	3.02×10^{-7}
2 (10 Å)	3 (12 Å)	6.08×10^{-7}
4 (14 Å)	3 (12 Å)	2.54×10^{-7}
3 (12 Å)	4 (14 Å)	5.96×10^{-7}
5 (16 Å)	4 (14 Å)	1.87×10^{-7}
4 (14 Å)	5 (16 Å)	4.62×10^{-7}
6 (18 Å)	5 (16 Å)	7.10×10^{-7}
5 (16 Å)	6 (18 Å)	1.42×10^{-7}
7 (20 Å)	6 (18 Å)	6.62×10^{-7}
6 (18 Å)	7 (20 Å)	1.47×10^{-7}
8 (22 Å)	7 (20 Å)	5.75×10^{-7}
7 (20 Å)	8 (22 Å)	2.17×10^{-7}
9 (24 Å)	8 (22 Å)	6.16×10^{-7}
8 (22 Å)	9 (24 Å)	2.47×10^{-7}
10 (26 Å)	9 (24 Å)	5.54×10^{-7}
9 (24 Å)	10 (26 Å)	4.23×10^{-7}
11 (28 Å)	10 (26 Å)	5.41×10^{-7}
10 (26 Å)	11 (28 Å)	3.80×10^{-7}
12 (30 Å)	11 (28 Å)	5.44×10^{-7}
11 (28 Å)	12 (30 Å)	2.74×10^{-7}
13 (32 Å)	12 (30 Å)	9.48×10^{-7}
12 (30 Å)	13 (32 Å)	8.00×10^{-8}
smallest eigenvalue		1.70×10^{-9}

[a] Eigenvalues were obtained by diagonalization of a 14 × 14 matrix.

on what is being calculated: (1) box-to-box MFPTs of the sort written in eq 12 are used to construct free energies, and (2) milestone-to-milestone FPTs are used to construct the kinetic master equation. These different approaches result in different counting algorithms for computing FPTs in the kinetic master equation, which are illustrated in Scheme 3. In the BXD approach, the passage times from box to box (i.e., $m \rightarrow m + 1$, and $m \rightarrow m - 1$ in Scheme 3) are defined as the times between subsequent hits at a particular box boundary; in Markovian milestoning, they are defined as the time it takes to go from milestone to milestone (i.e., $\rho_i \rightarrow \rho_j$, and $\rho_j \rightarrow \rho_i$ in Scheme 3).

Figure 5 shows a typical comparison of the decay traces generated from the definition of BXD FPTs and milestoning FPTs for the 10-ALA model system. The BXD decay shown in Figure 5 corresponds to that from box 2 (spanning ρ = 4.7 Å to ρ = 8.0 Å) to box 1 (spanning ρ = 4.7 Å to ρ = 0.0 Å), whereas the milestoning decay corresponds to transit from milestone 2 (ρ = 8.0 Å) to milestone 1 (ρ = 4.7 Å). The difference in the short time decays between milestoning and BXD arises from the fact that the BXD decays include very fast short time events on the time scale of dynamical decorrelation, which are schematically illustrated in Scheme 3. By contrast, milestone-to-milestone FPTs show a lag in the decay at short times as the trajectory makes a transit from one box edge to another, effectively eliminating the short time dynamical motion. At time scales longer than the short-time milestoning plateau, Figure 5 clearly illustrates that the BXD and milestoning decay traces converge. Table 3 gives the milestone to milestone rate coefficients for the simple 10 ALA model system, as well as the eigenvalue of smallest magnitude obtained from diagonalization of the corresponding rate coefficient matrix.

6. CHARACTERISTIC TIMES

Several characteristic times have emerged in the analysis we have presented in this paper, and their consideration sheds further light on the general applicability of BXD, Markovian milestoning, and other related techniques. The shortest time scale considered in this work is that of dynamical decorrelation, τ_{corr}, during which a trajectory has not lost memory of its initial conditions. τ_{corr} is short compared to the characteristic diffusion time, τ_{diff}, from the interior of a box to its boundary. In the analysis presented in this paper, τ_{diff} varies from box to box, and we have related τ_{diff} to the slower time scales for box-to-box diffusion. So long as

$$\tau_{\mathrm{diff}} \gg \tau_{\mathrm{corr}} \tag{19}$$

then the "diffusional" picture of box-to-box dynamics is accurate. Inequality 19 may always be met if the boxes are large enough and is the condition of Markovian kinetics when only the transitions between neighboring boxes need to be considered. It is reasonable to assume that eq 19 will hold for large systems with significant anharmonic coupling. The time τ_{kin} for phenomenological kinetics is determined by the eigenvalues of smallest absolute magnitude determined from diagonalization of the matrix **M** and

$$\tau_{\mathrm{kin}} \gg \tau_{\mathrm{diff}} \tag{20}$$

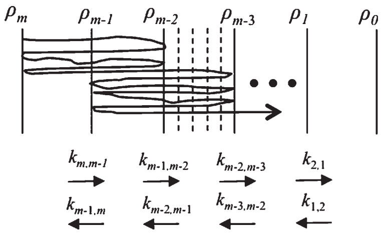
Along with Figure 4, the data in Tables 2 and 3 show that inequality 20 holds. If both conditions, eqs 19 and 20, are met, then the phenomenological kinetics obtained from **M** are largely insensitive to the number of boxes and their size.

7. DISCUSSION AND CONCLUSIONS

In this paper, we have considered in further detail the recently developed boxed molecular dynamics method and shown that a proper accounting for dynamical effects on time scales shorter than that of dynamical decorrelation is important for obtaining accurate kinetics. However, in calculating free energy surfaces, the dynamical effects largely cancel, making BXD a robust method for calculating free energies in the ergodic limit.

We have also discussed the interrelations between BXD and Markovian milestoning (MMS). The manner in which the free energy surfaces are calculated in both BXD and MMS is not sensitive to dynamical decorrelation, which is interesting *per se*, and both methods give similar long time kinetics decay traces. The primary difference between the two methods is in their uncorrected short time behavior, where each is subject to different types of nonstatistical effects: BXD overestimates the decay rate for going from one box to another, while MMS underestimates it. For the system considered in this work, overestimation introduces more error than underestimation, although it is in principle possible to imagine a situation where the MMS lag may lead to overestimation of the box-to-box transition time.

Scheme 4. Extended BXD Scheme That We Will Investigate in Further Work[a]



[a] The method is similar to that illustrated in Scheme 2, with the primary difference being the overlap in consecutive box dynamics.

Removing the very fast short-time dynamical effects from BXD provides a set of box-to-box MFPTs that allows an accurate calculation of both the free energy and the kinetics along a particular reaction coordinate. This insight should prove useful to users of BXD following its recent implementation in CHARMM[55] and is initutively appealing insofar as it is provides a set of mean first passage times which are compatible with the manner in which molecular modellers typically tend to think of a reaction: i.e., transition from a reactant configuration space volume to a product configuration space volume. This enables one to avoid using separate sets of rate coefficients for the free energies and kinetics, distinct from the current MMS protocol. Of further interest is the fact that BXD includes a specification for an exact renormalization procedure,[28] which we have found suffers from few of the numerical problems often associated with the WHAM procedure used to reweight umbrella sampling simulations. The BXD renormalization procedure should also be applicable to MMS.

Another feature of AXD and BXD which we believe to be of practical convenience is that the velocity inversion procedure preserves angular momentum, linear momentum, and energy, making it useful for accelerating both equilibrium and nonequilibrium dynamics. This method for doing velocity inversion has recently enabled application of AXD to nonequilibrium molecular dynamics simulations (NEMD)[56] to accelerate solution phase reaction dynamics. Along the same lines, we note that there may be cases where the initial nonstatistical dynamical behavior, which is preserved by AXD and BXD, is of fundamental interest.

Figure 5 illustrates another point regarding the restricted dynamics used by both BXD and MMS. At times longer than the MMS lag or the BXD τ_{corr}, the logarithmic decay plot is not a perfectly straight line, slowing noticeably at times longer than $\sim 2 \times 10^4$ fs. More generally, this points to the fact that associating the BXD and MMS decays with a single rate constant may be an oversimplification. For this reason, we chose herein to emphasize mean first passage times in the description of BXD. As shown in our previous article, this procedure results in BXD giving free energies that agree well with those obtained in unbiased simulations.[28]

In further work, we plan to extend BXD to multidimensional reaction coordinates; the recent use of Voronoi tessellations in MMS offers a possible way forward.[7,57] Additionally, we plan to investigate an extended version of BXD, which is shown in Scheme 4. In this extension of BXD, $\Delta G_{n-1,n}$ (the free energy difference between two boxes $n-1$ and n) may be determined by constrained dynamics in overlapping boxes. This will allow free passage between consecutive boxes which is unperturbed by inversion. Exploiting ergodicity, the free energy difference between regions in any box may then be written as a ratio of the times spent by the trajectory in each region of the box:

$$\exp\left(\frac{-\Delta G_{n-1,n}}{kT}\right) = \frac{P_n}{P_{n-1}} = \frac{\tau_n}{\tau_{n-1}} \tag{21}$$

By matching up the probabilities for each of the overlapping boxes, it will be possible to construct a free energy profile to higher resolution along the entire reaction coordinate.

AUTHOR INFORMATION

Corresponding Author

*E-mail: david.r.glowacki@bristol.ac.uk.

ACKNOWLEDGMENT

We thank Eric Vanden-Eijnden, Ron Elber, Maddalena Venturoli, and Benoit Roux for a useful exchange that motivated this study.

REFERENCES

(1) Maragliano, L.; Vanden-Eijnden, E.; Roux, B. Free Energy and Kinetics of Conformational Transitions from Voronoi Tessellated Milestoning with Restraining Potentials. *J. Chem. Theory Comput.* **2009**, *5*, 2589.

(2) Zuckerman, D. M.; Woolf, T. B. Efficient dynamic importance sampling of rare events in one dimension. *Phys. Rev. E* **2001**, *63*, 10.

(3) Henin, J.; Chipot, C. Overcoming free energy barriers using unconstrained molecular dynamics simulations. *J. Chem. Phys.* **2004**, *121*, 2904.

(4) Kinnear, B. S.; Jarrold, M. F.; Hansmann, U. H. E. All-atom generalized-ensemble simulations of small proteins. *J. Mol. Graphics Modell.* **2004**, *22*, 397.

(5) Frenkel, D.; Smit, B. *Understanding Molecular Simulation*, 2nd ed.; Academic Press: London, 2002.

(6) Darve, E.; Rodriguez-Gomez, D.; Pohorille, A. Adaptive biasing force method for scalar and vector free energy calculations. *J. Chem. Phys.* **2008**, *128*, 144120/1.

(7) Vanden-Eijnden, E.; Venturoli, M. Markovian milestoning with Voronoi tessellations. *J. Chem. Phys.* **2009**, *130*, 194101/1.

(8) Vanden-Eijnden, E.; Venturoli, M.; Ciccotti, G.; Elber, R. On the assumptions underlying milestoning. *J. Chem. Phys.* **2008**, *129*, 174102/1.

(9) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.* **1992**, *13*, 1011.

(10) Voter, A. F. A method for accelerating the molecular dynamics simulation of infrequent events. *J. Chem. Phys.* **1997**, *106*, 4665.

(11) Voter, A. F. Hyperdynamics: accelerated molecular dynamics of infrequent events. *Phys. Rev. Lett.* . **1997**, *78*, 3908.

(12) Voter, A. F.; Montalenti, F.; Germann, T. C. Extending the time scale in atomistic simulation of materials. *Annu. Rev. Mater. Res.* **2002**, *32*, 321.

(13) Torrie, G. M.; Valleau, J. P. Non-physical sampling distributions in monte-carlo free-energy etsimation - umbrella sampling. *J. Comput. Phys.* **1977**, *23*, 187.

(14) Faradjian, A. K.; Elber, R. Computing time scales from reaction coordinates by milestoning. *J. Chem. Phys.* **2004**, *120*, 10880.

(15) Chipot, C.; Henin, J. Exploring the free-energy landscape of a short peptide using an average force. *J. Chem. Phys.* **2005**, *123*, 244906/1.

(16) Laio, A.; Parrinello, M. Escaping free-energy minima. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 12562.

(17) Carter, E. A.; Ciccotti, G.; Hynes, J. T.; Kapral, R. Constrained reaction coordinate dynamics for the simulation of rare events. *Chem. Phys. Lett.* **1989**, *156*, 472.

(18) Huber, T.; Torda, A. E.; van Gunsteren, W. F. Local elevation: a method for improving the searching properties of molecular dynamics simulation. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 695.

(19) Sprik, M.; Ciccotti, G. Free energy from constrained molecular dynamics. *J. Chem. Phys.* **1998**, *109*, 7737.

(20) Paci, E.; Ciccotti, G.; Ferrario, M.; Kapral, R. Activation Energies by Molecular Dynamics with Constraints. *Chem. Phys. Lett.* **1991**, *176*, 581.

(21) Ciccotti, G.; Ferrario, M. Blue moon approach to rare events. *Mol. Simul.* **2004**, *30*, 787.

(22) Grubmuller, H. Predicting Slow Structural Transitions in Macromolecular Systems - Conformational Flooding. *Phys. Rev. E* **1995**, *52*, 2893.

(23) Hummer, G. Fast-growth thermodynamic integration: Error and efficiency analysis. *J. Chem. Phys.* **2001**, *114*, 7330.

(24) Woods, C. J.; Essex, J. W.; King, M. A. The development of replica-exchange-based free-energy methods. *J. Phys. Chem. B* **2003**, *107*, 13703.

(25) Jarzynski, C. Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach. *Phys. Rev. E* **1997**, *56*, 5018.

(26) Warmflash, A.; Bhimalapuram, P.; Dinner, A. R. Umbrella sampling for nonequilibrium processes. *J. Chem. Phys.* **2007**, *127*, 8.

(27) Vanden-Eijnden, E. Some Recent Techniques for Free Energy Calculations. *J. Comput. Chem.* **2009**, *30*, 1737.

(28) Glowacki, D. R.; Paci, E.; Shalashilin, D. V. Boxed molecular dynamics: a simple and general technique for accelerating rare event kinetics and mapping free energy in large molecular systems. *J. Phys. Chem. B* **2009**, *113*, 16603.

(29) Martinez-Nunez, E.; Shalashilin, D. V. Acceleration of Classical Mechanics by Phase Space Constraints. *J. Chem. Theory Comput.* **2006**, *2*, 912.

(30) Shalashilin, D. V.; Thompson, D. L. Monte Carlo Variational Transition-State Theory Study of the Unimolecular Dissociation of RDX. *J. Phys. Chem. A* **1997**, *101*, 961.

(31) Guo, Y.; Shalashilin, D. V.; Krouse, J. A.; Thompson, D. L. Intramolecular dynamics diffusion theory approach to complex unimolecular reactions. *J. Chem. Phys.* **1999**, *110*, 5521.

(32) Guo, Y.; Shalashilin, D. V.; Krouse, J. A.; Thompson, D. L. Predicting nonstatistical unimolecular reaction rates using Kramers' theory. *J. Chem. Phys.* **1999**, *110*, 5514.

(33) Shalashilin, D. V.; Thompson, D. L. Method for predicting IVR-limited unimolecular reaction rate coefficients. *J. Chem. Phys.* **1997**, *107*, 6204.

(34) Shalashilin, D. V.; Thompson, D. L. Intramolecular dynamics diffusion theory: nonstatistical unimolecular reaction rates. *ACS Symp. Ser.* **1997**, *678*, 81.

(35) Skinner, J. L.; Wolynes, P. G. Relaxation Processes and Chemical Kinetics. *J. Chem. Phys.* **1978**, *69*, 2143.

(36) Chandler, D. Statistical Mechanics of isomerization Dynamics in Liquids and Transition State Approximation. *J. Chem. Phys.* **1978**, *68*, 2959.

(37) Montgomery, J. A.; Chandler, D.; Berne, B. J. Trajectory analysis of a kinetic theory for isomerization dynamics in condensed phases. *J. Chem. Phys.* **1979**, *70*, 4056.

(38) Voter, A. F.; Doll, J. D. Dynamical Corrections to Transition State Theory for Multistate Systems - Surface Self-Diffusion in the Rare-Event Regime. *J. Chem. Phys.* **1985**, *82*, 80.

(39) Elber, R. A milestoning study of the kinetics of an allosteric transition: Atomically detailed simulations of deoxy Scapharca hemoglobin. *Biophys. J.* **2007**, *92*, L85.

(40) Kuczera, K.; Jas, G. S.; Elber, R. Kinetics of Helix Unfolding: Molecular Dynamics Simulations with Milestoning. *J. Phys. Chem. A* **2009**, *113*, 7461.

(41) West, A. M. A.; Elber, R.; Shalloway, D. Extending molecular dynamics time scales with milestoning: example of complex kinetics in a solvated peptide. *J. Chem. Phys.* **2007**, *126*, 145104.

(42) Feige, M. J.; Paci, E. Rate of Loop Formation in Peptides: A Simulation Study. *J. Mol. Biol.* **2008**, *382*, 556.

(43) Paci, E.; Lindorff-Larsen, K.; Dobson, C. M.; Karplus, M.; Vendruscolo, M. Transition State Contact Orders Correlate with Protein Folding Rates. *J. Mol. Biol.* **2005**, *352*, 495.

(44) Ferrara, P.; Apostolakis, J.; Caflisch, A. Evaluation of a fast implicit solvent model for molecular dynamics simulations. *Proteins* **2002**, *46*, 24.

(45) Lourderaj, U.; Hase, W. L. Theoretical and Computational Studies of Non-RRKM Unimolecular Dynamics. *J. Phys. Chem. A* **2009**, *113*, 2236.

(46) There are a number of schemes for removing recrossing. In this particular case, we found that it was easy to remove recrossing for the decays in Figure 1 using fits to a biexponential of the form $R(t) = A \exp(-k_f t) + B \exp(-k_s t)$ where the sum $A + B$ is constrained to give $R(0)$, and k_f and k_s are the respective rate coefficients for the fast and slow components of the $R(t)$ decay.

(47) Buchete, N. V.; Hummer, G. Coarse master equations for peptide folding dynamics. *J. Phys. Chem. B* **2008**, *112*, 6057.

(48) Chodera, J. D.; Singhal, N.; Pande, V. S.; Dill, K. A.; Swope, W. C. Automatic discovery of metastable states for the construction of Markov models of macromolecular conformational dynamics. *J. Chem. Phys.* **2007**, *126*, 155101/1.

(49) Bowman, G. R.; Beauchamp, K. A.; Boxer, G.; Pande, V. S. Progress and challenges in the automated construction of Markov state models for full protein systems. *J. Chem. Phys.* **2009**, *131*, 124101/1.

(50) Glowacki, D. R.; Pilling, M. J. Unimolecular Reactions of Peroxy Radicals in Atmospheric Chemistry and Combustion. *ChemPhysChem* **2010**, *11*, 3836.

(51) Gillespie, D. T. Stochastic simulation of chemical kinetics. *Annu. Rev. Phys. Chem.* **2007**, *58*, 35.

(52) Miller James, A.; Klippenstein Stephen, J. Master equation methods in gas phase chemical kinetics. *J. Phys. Chem. A* **2006**, *110*, 10528.

(53) Wales, D. J. Energy landscapes: calculating pathways and rates. *Int. Rev. Phys. Chem.* **2006**, *25*, 237.

(54) Bartis, J. T.; Widom, B. Stochastic Models of Interconversion of 3 or more Chemical Species. *J. Chem. Phys.* **1974**, *60*, 3474.

(55) Brooks, B. R.; Brooks, C. L., III; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: The biomolecular simulation program. *J. Comput. Chem.* **2009**, *30*, 1545.

(56) Greaves, S. J.; Rose, R. A.; Oliver, T. A. A.; Glowacki, D. R.; Ashfold, M. N. R.; Harvey, J. N.; Clark, I. P.; Greetham, G. M.; Parker, A. W.; Towrie, M.; Orr-Ewing, A. J. Vibrationally Quantum-State—Specific Reaction Dynamics of H Atom Abstraction by CN Radical in Solution. *Science* **2011**, *331*, 1423.

(57) Vanden-Eijnden, E.; Venturoli, M. Revisiting the finite temperature string method for the calculation of reaction tubes and free energies. *J. Chem. Phys.* **2009**, *130*, 17.

JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

SHARC: *ab Initio* Molecular Dynamics with Surface Hopping in the Adiabatic Representation Including Arbitrary Couplings

Martin Richter,[†] Philipp Marquetand,[*,†] Jesús González-Vázquez,[*,†,§] Ignacio Sola,[‡] and Leticia González[†]

[†]Institut für Physikalische Chemie, Friedrich-Schiller-Universität Jena, Helmholtzweg 4, 07743 Jena, Germany

[‡]Departamento de Química Física I, Universidad Complutense, 28040 Madrid, Spain

ABSTRACT: We present a semiclassical surface-hopping method which is able to treat arbitrary couplings in molecular systems including all degrees of freedom. A reformulation of the standard surface-hopping scheme in terms of a unitary transformation matrix allows for the description of interactions like spin−orbit coupling or transitions induced by laser fields. The accuracy of our method is demonstrated in two systems. The first one, consisting of two model electronic states, validates the semiclassical approach in the presence of an electric field. In the second one, the dynamics in the IBr molecule in the presence of spin−orbit coupling after laser excitation is investigated. Due to an avoided crossing that originates from spin−orbit coupling, IBr dissociates into two channels: I + Br(^{2}P$_{3/2}$) and I + Br*(^{2}P$_{1/2}$). In both systems, the obtained results are in very good agreement with those calculated from exact quantum dynamical simulations.

1. INTRODUCTION

The ultimate goal of chemistry is to precisely steer chemical reactions. However, to improve the state-of-the-art control, an understanding of molecular and atomic processes including all kinds of couplings and interactions is crucial. These two points, understanding and control of chemical processes, become increasingly challenging when looking at complex systems of growing size. Experiment and theory have to work hand in hand as, e.g., spectra become very complicated and the use of exact equations is not feasible anymore. Focusing on the theoretical part, an exact description of the coupled motion of nuclei and electrons is offered by the time-dependent Schrödinger equation, but it can be solved only for the simplest systems.[1] To study bigger systems, different approximations have been developed and implemented, leading to methods like, e.g., the multiconfigurational time-dependent Hartree method (MCTDH),[2−4] multiple spawning,[5,6] or similar techniques.[7−15] An alternative is to use *ab initio* molecular dynamics (MD), where the electronic structure is treated quantum mechanically and the nuclear motion is subject to classical mechanics.[16,17] Such a classical nuclear trajectory can only be subject to one electronic potential at a time. However, several potentials are often necessary to provide a correct description of the system's dynamics. The Tully's fewest switches algorithm of surface hopping (SH) method[18,19] is one of the most prominent solutions to this predicament.

SH was initially developed to account for nonadiabatic couplings between different states. Yet, couplings like those induced by electric fields or spin−orbit couplings (SOC) are also relevant to the treatment of light-induced processes. Especially the description of the interaction between light and matter is essential for simulating many spectroscopic experiments. Moreover, the whole field of quantum control is built on these foundations.[20−29] SOC also plays a major role in modern photochemistry. Processes like intersystem crossing or phosphorescence, i.e., transitions between triplet and singlet states, determine the outcome of photochemical reactions.[30,31] Major effects are expected in molecules including heavy atoms. More unexpectedly, they also strongly affect the photochemisty of organic or biomolecules such as DNA, see, e.g., ref 32.

Despite the importance of these couplings, only few methods exist which incorporate one or the other of those effects in MD.[33−37] To our knowledge, no MD method is able so far to handle all of the couplings simultaneously in a straightforward way. In this paper, we present a method which, derived from the original SH scheme, allows one to treat arbitrary couplings without introducing any further approximations. In this way, a semiclassical description of the coupled electronic and nuclear motion with laser interaction in complex molecular systems including all degrees of freedom is feasible. In principle, not only SOC or laser interactions can be included but all imaginable couplings can be straightforwardly considered by our surface-hopping-in-adiabatic-representation-including-arbitrary-couplings (SHARC) method. The SH probabilities are calculated in terms of a unitary transformation matrix which diagonalizes the matrix containing the considered electronic potentials and all possible couplings at once.

In order to demonstrate the effectiveness of SHARC, two test systems are chosen in this paper: two coupled harmonic oscillators first and then the excited-state dissociation in the IBr molecule. IBr exhibits strong SOC, leading to an avoided crossing between two excited states, the $1^3\Pi_{0+}$ and the $1^3\Sigma^-_{0+}$ states, see ref 38 and references therein. These two states are responsible for the product channels resulting in I + Br and I + Br*, respectively.[39] The asterisk denotes the $^2P_{1/2}$ excited spin-state of the dissociating Br atom, while the ground state has the configuration $^2P_{3/2}$. The outcome from SHARC simulations is compared with results from exact quantum-dynamical calculations.

Received: January 28, 2011
Published: March 29, 2011

The methodology and the theoretical description are presented in section 2. The numerical results are contained in section 3, and a summary is given in section 4.

2. METHODOLOGY

The time evolution of the system is followed using a mix of quantum and classical dynamics, where the electrons are treated quantum mechanically and the nuclei classically. The interaction between the quantum and the classical part is described in two ways. On the one hand, the nuclei follow the quantum potential created by the electrons using classical trajectories. These trajectories are defined by the position $\vec{R}(t)$ and velocity $\vec{v}(t)$ of the nuclei at every time. On the other hand, the electronic wave function depends on the electronic coordinates $(\vec{r})$ and parametrically on the nuclei coordinates. The electronic wave function is defined by $|\Psi[\vec{R}(t);\vec{r},t]\rangle$, and the potential that governs the evolution of the classical trajectory is given by the expectation value of an effective Hamiltonian $V(t) = \langle\Psi(\vec{R}(t);\vec{r},t)|\hat{H}_{\text{eff}}[\vec{R}(t),\vec{r}]|\Psi(\vec{R}(t);\vec{r},t)\rangle$. $\hat{H}_{\text{eff}}$ depends parametrically on the nuclear coordinates, and it includes the nucleus−nucleus repulsion, the electron−nucleus attraction, the electron−electron, and the electronic kinetic energy.

In quantum mechanics, the position and velocity of the nuclei cannot be simultaneously known. This uncertainty is described in our classical dynamics by considering a set of initial conditions for the trajectories. This set mimics the initial nuclear quantum probability creating a swarm of trajectories. Every single trajectory is propagated using Newton's equations with the Velocity Verlet algorithm.[40,41] In this algorithm, the time evolution of the nuclear coordinates $\vec{R}(t)$ is driven by the gradient of the potential at time t:

$$\vec{R}(t+\Delta t) = \vec{R}(t) + \vec{v}(t)\Delta t + \frac{1}{2M}\nabla_{\vec{R}}\, V(t)\, \Delta t^2 \tag{1}$$

where M represents the mass of the nuclei. Finally, the velocity is propagated using the gradient of the potential at times t and $t + \Delta t$:

$$\vec{v}(t+\Delta t) = \vec{v}(t) + \frac{1}{2M}\nabla_{\vec{R}}\, V(t)\, \Delta t + \frac{1}{2M}\nabla_{\vec{R}} V(t+\Delta t)\Delta t \tag{2}$$

The definition of the potential is given by the time evolution of the electronic wavepacket following the time-dependent Schrödinger equation:

$$i\hbar\frac{\partial|\Psi[\vec{R}(t);\ \vec{r},t]\rangle}{\partial t} = \hat{H}_{\text{eff}}[\vec{R}(t);\ \vec{r}]|\Psi[\vec{R}(t);\ \vec{r},t]\rangle \tag{3}$$

In order to solve this equation, we expand the wavepacket as a linear combination of basis functions at different $\vec{R}(t)$:

$$|\Psi[\vec{R}(t);\ \vec{r},t]\rangle = \sum_{\alpha} c_\alpha(t)|\phi_\alpha[\vec{R}(t);\ \vec{r}]\rangle \tag{4}$$

Using this formalism, the time evolution of the coefficients is described by

$$\frac{\partial c_\beta(t)}{\partial t} = -\sum_{\alpha}\left\{\frac{i}{\hbar}H_{\beta\alpha}[\vec{R}(t)] + K_{\beta\alpha}[\vec{R}(t)]\right\}c_\alpha(t) \tag{5}$$

where $H_{\beta\alpha}[\vec{R}(t)] = \langle\phi_\beta[\vec{R}(t);\vec{r}]|\hat{H}_{\text{eff}}[\vec{R}(t);\vec{r}]|\phi_\alpha[\vec{R}(t);\vec{r}]\rangle$ represents the diabatic Hamiltonian whose diagonal elements are the different potentials and the off-diagonal elements are the diabatic couplings. The second term,

$$K_{\beta\alpha}[\vec{R}(t)] = \langle\phi_\beta[\vec{R}(t);\ \vec{r}]|\partial/\partial t|\phi_\alpha[\vec{R}(t);\ \vec{r}]\rangle \tag{6}$$

evaluates the change of the electronic basis functions with time, which is equivalent to the variation of the basis with the nuclear coordinates times the velocity:

$$\begin{aligned} K_{\beta\alpha}[\vec{R}(t)] &= \langle\phi_\beta[\vec{R}(t);\ \vec{r}]|\partial/\partial t|\phi_\alpha[\vec{R}(t);\ \vec{r}]\rangle \\ &= \langle\phi_\beta[\vec{R}(t);\ \vec{r}]|d/d\vec{R}(t)|\phi_\alpha[\vec{R}(t);\ \vec{r}]\rangle\,\vec{v}(t) \end{aligned} \tag{7}$$

This equation is solved using a simple Runge−Kutta algorithm of fourth order.

The definition of the basis functions is very important in this methodology. The most common way to define the basis functions of the electronic wavepacket is using the eigenfunctions of the time-independent Schrödinger equation for every $\vec{R}(t)$. In this way, the effective Hamiltonian is the adiabatic energy of the different electronic states $H_{\beta\alpha}[\vec{R}(t)] = V_\alpha[\vec{R}(t)]\delta_{\beta\alpha}$, while $K_{\beta\alpha}[\vec{R}(t)]$ is related to the nonadiabatic coupling elements that break the Born−Oppenheimer approximation. However, the trajectories cannot be spread over several electronic states, and hence, it is necessary to assign the electronic state that governs the trajectory dynamics at each time. In this work, we employ the SH method proposed by Tully,[18] where the classical trajectory is propagated in a single potential β. In order to take into account nonadiabatic effects, the trajectory can jump from one to another state. The probability for such a hop is calculated using the time-dependent coefficients of the electronic wave function:

$$P_{\beta\alpha} = \frac{2\mathscr{R}\left\{c_\beta^*(t)\, c_\alpha(t)\left[\frac{i}{\hbar}H_{\beta\alpha}[\vec{R}(t)] + K_{\beta\alpha}[\vec{R}(t)]\right]\right\}}{c_\beta^*(t)\, c_\beta(t)}\Delta t \tag{8}$$

This methodology is widely used to simulate relaxation dynamics via conical intersections where large kinetic couplings are localized around the degeneration points.[42,43] In this work, we extend SH to the situation where SOCs and/or the interaction with an electric field must be included in the dynamical simulation. These two terms are typically evaluated in the diabatic representation, and thus they are included in the potential part of the Hamiltonian, introducing a new $H^d[\vec{R}(t)]$ matrix with elements (where the index d indicates that additional nondiagonal terms are included):

$$H^d_{\beta\alpha}[\vec{R}(t),t] = H_{\beta\alpha}[\vec{R}(t)] - \vec{\mu}_{\beta\alpha}[\vec{R}(t)]\ \vec{\mathscr{E}}(t) + \hat{H}^{\text{SO}}_{\beta\alpha}[\vec{R}(t)] \tag{9}$$

In this equation, $\vec{\mu}_{\beta\alpha}[\vec{R}(t)]$ and $\hat{H}^{\text{SO}}_{\beta\alpha}[\vec{R}(t)]$ are the dipole moment and the relativistic spin−orbit coupling between the states β and α, respectively.

In contrast to the relaxation dynamics via conical intersections (nonadiabatic couplings), this new Hamiltonian contains off-diagonal elements which may be extremely extended in space, making the solution of the corresponding equations not an easy task. As such off-diagonal elements are responsible for the jumps, spatially delocalized couplings might induce jumps at all geometries contradicting the idea of the fewest switches criterion. Especially SOCs are usually very much extended in space, see, e.g., ref 38. In SHARC, this problem is solved by translating the coupling elements to the $K[\vec{R}(t)]$ matrix. In this adiabatic (index a) approach, the $H^d[\vec{R}(t),t]$ matrix is diagonalized, and afterwards, the $K[\vec{R}(t)]$ matrix is recalculated, leading to localized couplings in geometries where the electronic states are (nearly) degenerated.

The idea exploited here is the substitution of the basis set of electronic wave functions $|\phi^d[\vec{R}(t);r]\rangle$ for a linear combination:

$$|\phi^a_\beta[\vec{R}(t);\ \vec{r},t]\rangle = \sum_\alpha U_{\beta\alpha}[\vec{R}(t),t]|\phi^d_\alpha[\vec{R}(t);\ \vec{r}]\rangle \tag{10}$$

where $U[\vec{R}(t),t]$ is the unitary matrix that diagonalizes the Hamiltonian $H^d[\vec{R}(t),t]$ matrix at every time t. In this new basis, the elements of the $H^a[\vec{R}(t),t]$ matrix are defined as

$$H^a_{\beta\alpha}[\vec{R}(t),t] = V^a_\alpha[\vec{R}(t),t]\delta_{\beta\alpha} \tag{11}$$

where $V^a_\alpha[\vec{R}(t),t]$ are the diagonal elements of $H^a[\vec{R}(t),t]$. The nonadiabatic coupling comes from the derivative of the $|\phi^a[\vec{R}(t);\vec{r},t]\rangle$:

$$\begin{aligned} K^a_{\beta\alpha}[\vec{R}(t),t] &= \left\langle \phi^{a*}_\beta[\vec{R}(t);\ \vec{r},t]\left|\frac{\partial}{\partial t}\right|\phi^a_\alpha[\vec{R}(t);\ \vec{r},t]\right\rangle \\ &= K^\phi_{\beta\alpha}[\vec{R}(t),t] + K^U_{\beta\alpha}[\vec{R}(t),t] \end{aligned} \tag{12}$$

where $K^\phi_{\beta\alpha}[\vec{R}(t),t]$ and $K^U_{\beta\alpha}[\vec{R}(t),t]$ are the nonadiabatic terms in the original basis $|\phi^d[\vec{R}(t);\vec{r}]\rangle$ and those induced via the rotation matrix $U[\vec{R}(t),t]$.

The first term $K^\phi[\vec{R}(t),t]$ is just the rotation of the original nonadiabatic term to the new basis:

$$\begin{aligned} &K^\phi_{\beta\alpha}[\vec{R}(t),t] \\ &= \sum_{\lambda\gamma} U^*_{\lambda\beta}[\vec{R}(t),t]K_{\lambda\gamma}[\vec{R}(t)]U_{\gamma\alpha}[\vec{R}(t),t] \\ &= \vec{v}(t)\sum_{\lambda\gamma} U^*_{\lambda\beta}[\vec{R}(t),t]\langle\phi^d_\gamma[\vec{R}(t);\ \vec{r}]|\nabla_{\vec{R}}|\phi^d_\gamma[\vec{R}(t);\ \vec{r}]\rangle U_{\gamma\alpha}[\vec{R}(t),t] \end{aligned} \tag{13}$$

and the other component comes from the variation of the rotation matrix:

$$\begin{aligned} K^U_{\beta\alpha}[\vec{R}(t),t] &= \sum_\lambda U^*_{\lambda\beta}[\vec{R}(t),t]\frac{\partial}{\partial t}U_{\lambda\alpha}[\vec{R}(t),t] \\ &= \vec{v}(t)\sum_\lambda U^*_{\lambda\beta}[\vec{R}(t),t]\nabla_{\vec{R}}U_{\lambda\alpha}[\vec{R}(t),t] \end{aligned} \tag{14}$$

To obtain the new potentials $V^a[\vec{R}(t),t]$ and the nonadiabatic coupling elements $K^a[\vec{R}(t),t]$, the matrix $H^d[\vec{R}(t),t]$ is diagonalized at distances $\vec{R}(t)$,$\vec{R}(t)+\Delta\vec{R}$, and $\vec{R}(t)-\Delta\vec{R}$. In this fashion, the gradient of the potential and the gradient of the $U[\vec{R}(t),t]$ matrix are evaluated. These new matrices are used in eqs 5 and 8 to calculate the nonadiabatic dynamics.

In the original SH method, the velocity of a trajectory is adjusted after a jump in order to conserve energy. This, however, is not reasonable for laser transitions. Therefore, we do not adjust the kinetic energy as long as the resonance condition is fulfilled, i.e., as long as the potential energy difference of the involved states lies within the laser bandwidth.

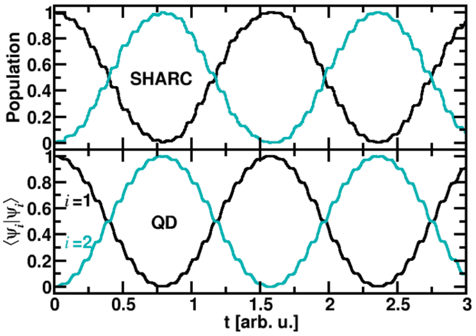


Figure 1. Comparison of the population dynamics calculated via SHARC (upper panel) or QD (lower panel) in a system consisting of two harmonic oscillators. Rabi oscillations of the populations in the ground state (i = 1; black) and the excited state (i = 2; turquoise) are clearly visible in both cases.

3. NUMERICAL RESULTS

In what follows, two model systems are investigated. First, we consider the modeling of Rabi oscillations between two harmonic oscillators using SHARC. Second, the branching ratio of excited-state dissociation products of IBr is examined. The motivation behind these rather simple models is to be able to compare the results of SHARC with those from exact quantum dynamics (QD). For the quantum part, the time-dependent Schrödinger equation is solved employing the split-operator method.[44] Solutions of the stationary Schrödinger equation are obtained by imaginary time propagation.[45] From these solutions, Wigner distributions are obtained and used to establish the initial conditions in the MD simulations.

As a first model system, we consider two vertically displaced one-dimensional harmonic oscillators defined by

$$V_1(R) = \frac{1}{2}kR^2 \tag{15}$$

$$V_2(R) = \frac{1}{2}kR^2 + D_{12} \tag{16}$$

where k and D_{12} are 1 and 40 in dimensionless units. The reduced mass is taken to be 1, and the transition dipole moment between the states is μ_{12} = 1. The two potentials are coupled by a *cw* field:

$$E(t) = A\sin(\omega t) \tag{17}$$

where A = 4 and ω = D_{12} to induce a resonant transition.

For both MD and QD calculations, a time step of 0.002 was employed. A set of 500 trajectories was used in the MD simulations, although the results are converged already after 100 trajectories.

The time-dependent populations calculated from SHARC and QD are shown in Figure 1. Rabi oscillations are clearly visible in both cases. More precisely, the curves from SHARC and QD show exactly the same behavior. Therefore, the SHARC approach is capable of describing laser-induced processes. However, this approach is not limited to dipole-type couplings but can treat simultaneously any other kind of coupling, as will be shown in the next example.

As a second model system, we look at the IBr molecule and its excited-state dissociation. In order to compare our MD results directly with those of QD simulations, we restrict ourselves to the

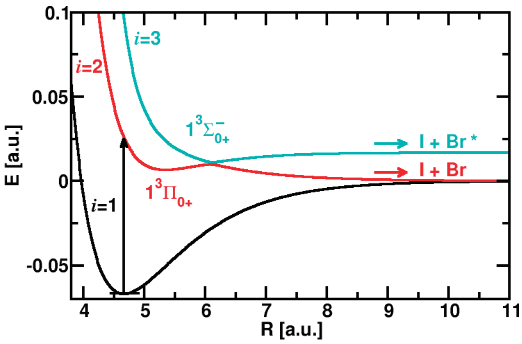


Figure 2. Potential energy curves of the IBr molecule and excitation scheme. IBr molecules initially in the electronic ground state (black) are excited to the $1^3\Pi_{0+}$ excited electronic state (red) and can undergo dissociation into two different channels due to an avoided crossing introduced by SOC with the $1^3\Sigma^-_{0+}$ excited state (turquoise).

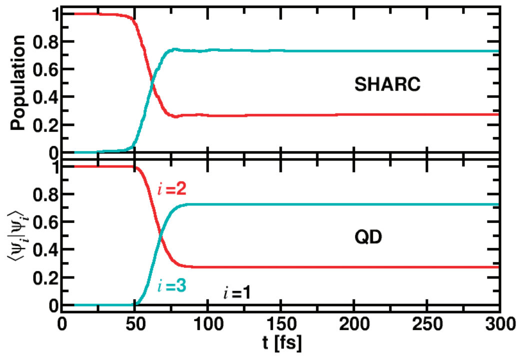


Figure 3. Population dynamics in the excited states of IBr after excitation with a δ pulse computed with the SHARC algorithm (upper panel) and QD (lower panel). After around 65 fs, an avoided crossing is passed, which gives rise to a branching ratio $Q = 72\%$ of the products in the different dissociation channels (I + Br in state $i = 2$ and I + Br* in state $i = 3$) with both simulation types. The ground state ($i = 1$; black) is not populated.

model potentials from ref 46. Note that even if here we limit ourselves to analytical potentials in order to benchmark SHARC in the presence of SOC and laser interactions, more complex systems can be tackled easily using "on the fly" electronic structure calculations.

IBr is a good candidate since it is a diatomic, which can be treated conveniently with QD, and it exhibits large SOC. Due to the latter coupling, the $1^3\Pi_{0+}$ and the $1^3\Sigma^-_{0+}$ state potentials show an avoided crossing at around $R = 6.1$ au = 3.2 Å, see Figure 2, which otherwise would not exist since the two states are of different symmetry. Moreover, it is only because of SOC that the excited-state dissociation channels differ in energy, see ref 38. In the present model, the dissociation products are I + Br or I + Br*, respectively.

To test whether the SOC is treated correctly within SHARC, we excite the ground state population first with a δ pulse to the $1^3\Pi_{0+}$ state and look at the population dynamics. In both the MD and QD simulations, the propagation time step is 0.02 fs. Again, 500 trajectories are considered in the MD case. It is gratifying to see (Figure 3) that the results of both calculations are in perfect agreement. The branching ratio of the dissociation products, defined as $Q = ([\mathrm{I + Br^*}])/([\mathrm{I + Br}] + [\mathrm{I + Br^*}])$, is equal to 72% in both cases.

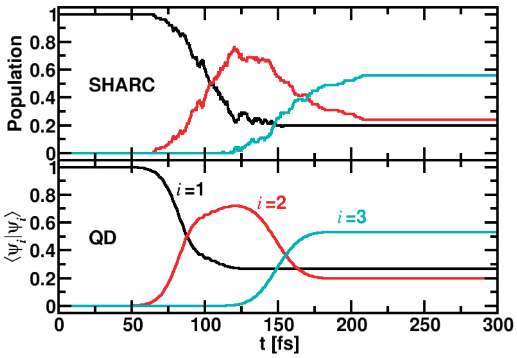


Figure 4. Population dynamics in IBr after excitation with a 50 fs fwhm pulse computed with the SHARC algorithm (upper panel) and QD (lower panel). The population in the ground state ($i = 1$) is depicted in black, the first excited state ($i = 2$) in red, and the second excited state ($i = 3$) in turquoise. The excitation yield from the SHARC simulation ($Y = 80\%$) is in good agreement with the one from the QD calculation ($Y = 73\%$). Also, the branching ratio of excited state products is in very good agreement ($Q = 70\%$ from SHARC vs $Q = 73\%$ from QD).

As a second scenario, we treat the spin—orbit and strong laser-field induced couplings at the same time in the MD simulation. The laser pulse has a finite duration, and it is of Gaussian shape with the electric field's fwhm (full width at half-maximum) of 50 fs centered at t = 100 fs and a field strength of 0.00534 au (corresponding to an intensity of 1 TW/cm^2). The wavelength for the laser excitation from the electronic ground state is set to 493 nm to satisfy the resonance condition with the $1^3\Pi_{0+}$ electronic state. As in the previous simulation, we use a time step of 0.02 fs and 500 trajectories in SHARC.

Figure 4 shows the time-dependent populations calculated via SHARC (upper panel) and via QD (lower panel). Very good agreement is also found between the two simulations. First, we look at the excitation yield Y. After the laser pulse is over, $Y = 80\%$ of the population has been excited according to SHARC and $Y = 73\%$ according to QD. Second, we compare the resulting branching ratios; this is calculated as $Q = 70\%$ from SHARC and $Q = 73\%$ from QD. Although we used 500 trajectories to get this number in the MD case, a branching ratio of $Q = 69\%$ is already obtained after only 100 trajectories. It is very encouraging that the deviations are very small, especially taking into account that there is no coherent interaction possible between the individual trajectories. Even the momentum distributions in the respective states are almost identical for the different algorithms, see Figure 5. This latter property is of great importance in numerous applications, e.g., the simulation of velocity map imaging.[47] In Figure 5, the momentum distribution of the ground state is the fraction centered around p = 0 au. After the laser excitation, when the $1^3\Pi_{0+}$ state is populated, there is a quick gain of momentum in this state. The passing through the avoided crossing is visible in the form of a splitting of the momentum distribution at later times (ca. 160 fs). The momentum distribution clearly indicates that the largest amount of trajectories (SHARC) or the wavepacket (QD) changes the adiabatic state to yield I + Br*; see that due to the climbing of the potential slope, the momentum diminishes. The rest of the excited state momentum distribution stems from the state resulting in I + Br, where additional momentum is collected during the dissociation.

Albeit rather simple at first sight, IBr turns out to be quite an extreme test case. It possesses an intermediate SOC strength

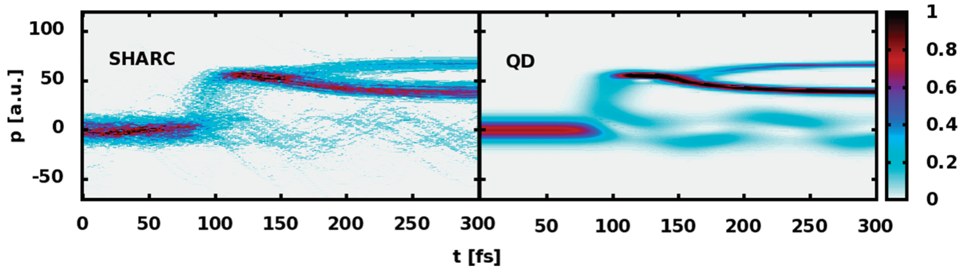


Figure 5. Time-dependent momentum probability distribution as calculated from SHARC (left panel) and QD (right panel).

such that neither the pure adiabatic nor the pure diabatic picture completely describe its dynamics.[48] In a two-level system, this is indeed the worst-case scenario. Moreover, the excited state potentials are extremely steep in the Franck–Condon region. As a consequence, even a narrow distribution of initial conditions in the ground state will result in the most different momenta in the excited state at the time the avoided crossing is reached. Finally, the ground state potential is very anharmonic due to the SOC. Despite these hurdles, SHARC is able to correctly describe the complete dynamics influenced by laser interactions and SOC.

4. CONCLUSION

To summarize, here, we present a new surface-hopping-in-adiabatic-representation-including-arbitrary-couplings (SHARC) algorithm, where the surface hopping probabilities are calculated in terms of a unitary transformation matrix. Within this semiclassical scheme, a matrix containing the considered electronic potentials and all possible couplings is diagonalized at once. In this way, we are able to treat all kinds of couplings in molecular systems including all degrees of freedom on the same footing. While the motion of the nuclei is treated classically, the potentials entering into the propagation can stem from simple analytical functions (as it was exemplarily done here to compare explicitly with exact quantum dynamics), complex parametrized force fields, or semiempirical or state-of-the-art *ab initio* methods.

We have therefore shown that, besides nonadiabatic couplings, field-induced transitions to triplet states can henceforth be treated within molecular dynamics. In this way, the often neglected influence of triplet states in the dynamics of most different molecules can be investigated using semiclassical simulations. The treatment of large systems, which may even include molecules in solution, is straightforward by computing the potential energies "on the fly".

AUTHOR INFORMATION

Corresponding Author

*E-mail: p.marquetand@uni-jena.de (P.M.); jgv@tchiko.quim.ucm.es (J.G.-V.).

Present Addresses

[§]Departamento de Química Física I, Universidad Complutense, 28040 Madrid, Spain.

ACKNOWLEDGMENT

This work has been supported by the Deutsche Forschungsgemeinschaft (DFG) within the project GO 1059/6-1, the German Federal Ministry of Education and Research within the research initiative PhoNa, the Dirección General de Investigación of Spain under Project No. CTQ2008-06760, the Friedrich-Schiller-Universität Jena, and a Juan de la Cierva contract. Generous allocation of computer time at the Computer center of the Friedrich-Schiller-Universität is gratefully acknowledged.

REFERENCES

(1) Tannor, D. *Introduction to Quantum Mechanics: A Time-Dependent Perspective*; University Science Books: Sausalito, CA, 2006.
(2) Beck, M. H.; Jäckle, A.; Worth, G. A.; Meyer, H. D. *Phys. Rep* **2000**, *324*, 1–105.
(3) Bowman, J. M.; Carrington, T.; Meyer, H. *Mol. Phys.* **2008**, *106*, 2145–2182.
(4) Worth, G. A.; Meyer, H. D.; Köppel, H.; Cederbaum, L. S.; Burghardt, I. *Int. Rev. Phys. Chem.* **2008**, *27*, 569–606.
(5) Virshup, A. M.; Punwong, C.; Pogorelov, T. V.; Lindquist, B. A.; Ko, C.; Martínez, T. J. *J.Phys. Chem. B* **2009**, *113*, 3280–3291.
(6) Levine, B. G.; Coe, J. D.; Virshup, A. M.; Martínez, T. *J. Chem. Phys.* **2008**, *347*, 3–16.
(7) Worth, G. A.; Robb, M. A.; Burghardt, I. *Faraday Discuss* **2004**, *127*, 307–323.
(8) Rassolov, V. A.; Garashchuk, S. *Phys. Rev. A* **2005**, *71*, 032511.
(9) Li, J.; Woywod, C.; Vallet, V.; Meier, C. *J. Chem. Phys.* **2006**, *124*, 184105.
(10) Spezia, R.; Burghardt, I.; Hynes, J. T. *Mol. Phys.* **2006**, *104*, 903–914.
(11) Lasorne, B.; Robb, M. A.; Worth, G. A. *Phys. Chem. Chem. Phys.* **2007**, *9*, 3210–3227.
(12) Shalashilin, D. V.; Child, M. S.; Kirrander, A. *Chem. Phys.* **2008**, *347*, 257–262.
(13) Yonehara, T.; Takahashi, S.; Takatsuka, K. *J. Chem. Phys.* **2009**, *130*, 214113.
(14) Yonehara, T.; Takatsuka, K. *J. Chem. Phys.* **2010**, *132*, 244102.
(15) Granucci, G.; Persico, M.; Zoccante, A. *J. Chem. Phys.* **2010**, *133*, 134111.
(16) Marx, D., Hutter, J. *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*; Cambridge University Press: Cambridge, U. K., 2009.
(17) Doltsinis, N.; Marx, D. *J. Theory Comput. Chem.* **2002**, *1*, 319–349.
(18) Tully, J. C. *J. Chem. Phys.* **1990**, *93*, 1061–1071.
(19) Doltsinis, N. In *Computational Nanoscience: Do It Yourself!*; Grotendorst, J., Ed.; John von Neumann-Institut für Computing: Jülich, Germany, 2006; pp 389—409.
(20) Brumer, P.; Shapiro, M. *Annu. Rev. Phys. Chem.* **1992**, *43*, 257–282.
(21) Gordon, R. J.; Rice, S. A. *Annu. Rev. Phys. Chem.* **1997**, *48*, 601–641.
(22) Rice, S. A. *Adv. Chem. Phys.* **1997**, *101*, 213–283.
(23) Tannor, D. J.; Kosloff, R.; Bartana, A. *Faraday Discuss.* **1999**, *113*, 365–383.
(24) Shapiro, M.; Brumer, P. *Rep. Prog. Phys.* **2003**, *66*, 859–942.
(25) Rice, S. A.; Zhao, M. *Optical Control of Molecular Dynamics*; Wiley: New York, 2000.

(26) Shapiro, M.; Brumer, P. *Principles of Quantum Control of Molecular Processes*; Wiley: New York, 2003.

(27) Brixner, T.; Damrauer, N. H.; Gerber, G. *Adv. At. Mol. Opt. Phys.* **2001**, *46*, 1–54.

(28) Nuernberger, P.; Vogt, G.; Brixner, T.; Gerber, G. *Phys. Chem. Chem. Phys.* **2007**, *9*, 2470–2497.

(29) Engel, V.; Meier, C.; Tannor, D. J. *Adv. Chem. Phys.* **2009**, *141*, 29–101.

(30) Reiher, M.; Wolf, A. *Relativistic Quantum Chemistry*; Wiley-VCH: Weinheim, Germany, 2009.

(31) Fedorov, D. G.; Koseki, S.; Schmidt, M. W.; Gordon, M. S. *Int. Rev. Phys. Chem.* **2003**, *22*, 551–592.

(32) González-Luque, R.; Climent, T.; González-Ramírez, I.; Merchán, M.; Serrano-Andrés, L. *J. Chem. Theory Comput.* **2010**, *6*, 2103–2114.

(33) Maiti, B.; Schatz, G. C.; Lendvay, G. *J. Phys. Chem. A* **2004**, *108*, 8772–8781.

(34) Yagi, K.; Takatsuka, K. *J. Chem. Phys.* **2005**, *123*, 224103.

(35) Jones, G. A.; Acocella, A.; Zerbetto, F. J. *Phys. Chem. A* **2008**, *112*, 9650–9656.

(36) Mitrić, R.; Petersen, J.; Bonačić-Koutecký, V. *Phys. Rev. A* **2009**, *79*, 053416.

(37) Tavernelli, I.; Curchod, B. F. E.; Rothlisberger, U. *Phys. Rev. A* **2010**, *81*, 052508.

(38) Patchkovskii, S. *Phys. Chem. Chem. Phys.* **2006**, *8*, 926–940.

(39) Sussman, B. J.; Townsend, D.; Ivanov, M. Y.; Stolow, A. *Science* **2006**, *314*, 278–281.

(40) Verlet, L. *Phys. Rev.* **1967**, *159*, 98–103.

(41) Verlet, L. *Phys. Rev.* **1968**, *165*, 201–214.

(42) Barbatti, M.; Aquino, A. J. A.; Szymczak, J. J.; Nachtigallová, D.; Hobza, P.; Lischka, H. P. *Natl. Acad. Sci. USA* **2010**, *107*, 21453 −21458.

(43) González-Vázquez, J.; González, L. *Chem. Phys. Chem.* **2010**, *11*, 3617–3624.

(44) Feit, M. D.; Fleck, J. A., Jr.; Steiger, A. *J. Comput. Phys.* **1982**, *47*, 412–433.

(45) Kosloff, R.; Tal-Ezer, H. *Chem. Phys. Lett.* **1986**, *127*, 223–230.

(46) Guo, H. *J. Chem. Phys.* **1993**, *99*, 1685–1692.

(47) Whitaker, B. *Femtosecond Chemistry*; Cambridge University Press: Cambridge, U. K., 2003; Vols. I, II.

(48) Shapiro, M.; Vrakking, M. J. J.; Stolow, A. *J. Chem. Phys.* **1999**, *110*, 2465–2473.

JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Predicting Fixation Tendencies of the H3N2 Influenza Virus by Free Energy Calculation

Keyao Pan[†] and Michael W. Deem[*,†,‡]

[†]Department of Bioengineering and [‡]Department Physics & Astronomy, Rice University, Houston, Texas 77005, United States

ABSTRACT: The influenza virus evolves to escape from immune system antibodies that bind to it. We used free energy calculations with Einstein crystals as reference states to calculate the difference of antibody binding free energy ($\Delta\Delta G$) induced by amino acid substitution at each position in epitope B of the H3N2 influenza hemagglutinin, the key target for antibodies. A substitution with a positive $\Delta\Delta G$ value decreases the antibody binding constant and increases viral fitness. On average, an uncharged to charged amino acid substitution generates the highest $\Delta\Delta G$ values. Also, on average, substitutions between small amino acids generate $\Delta\Delta G$ values near zero. The 21 sites in epitope B have varying expected free energy differences for a random substitution. Historical amino acid substitutions in epitope B for the A/Aichi/2/1968 strain of influenza A show that most fixed and temporarily circulating substitutions generate positive $\Delta\Delta G$ values. We propose that the observed pattern of H3N2 virus evolution is affected by the free energy landscape, the mapping from the free energy landscape to the virus fitness landscape, and random genetic drift of the virus. Monte Carlo simulations of virus evolution are presented to support this view.

1. INTRODUCTION

The influenza A virus causes annual global epidemics resulting in 5−15% of the population being infected, 3−5 million severe cases, and 250 000−500 000 fatalities.[1] The subtype of influenza A is determined by two surface glycoproteins—hemagglutinin (H) and neuraminidase (N). The H3N2 virus has been one of the dominant circulating subtypes since its emergence in 1968. The antibodies IgG and IgA are the major components of the immune system that control influenza infection, binding to the influenza hemagglutinin.[2] There are five epitopes at the antibody binding sites on the top of H3 hemagglutinin, namely, epitopes A−E. The epitope bound most prolifically by antibodies is defined as the dominant epitope, and it is central to the process of virus neutralization by antibody and virus escape substitution.[3] The cellular immune system, on the other hand, plays a relatively less recognized role in handling the invasive influenza virus.[2] The cellular system along with the innate immune system exerts a somewhat more homogeneous immune reaction against genetically distinct influenza strains.[2,4]

Vaccination is currently the primary method to prevent and control an influenza epidemic in the human population.[1] Influenza vaccination raises the level of antibodies specific for hemagglutinin and significantly enhances the binding affinity between antibodies and hemagglutinin. Vaccine effectiveness depends on the antigenic distance between the hemagglutinin of the administered vaccine strain and that of the dominant circulating strain in the same season.[3,5] Memory immune response from the virus in previous seasons as well as vaccination in the current and previous seasons impose selective pressure on the current circulating virus to force it to evolve away from the virus strains recognized by memory antibodies that selectively bind to hemagglutinin.

As a result of the immune pressure and the escape evolution of the influenza virus, which is largely substitution in the dominant epitope of hemagglutinin, the influenza vaccine must be redesigned and administered each year, and vaccine effectiveness has been suboptimal in some flu seasons.[3,6] The escape evolution in the dominant epitope is at a higher rate than that in the amino acid sites outside the dominant epitope.[7] Sites in the dominant epitope also show higher Shannon entropy of the 20 amino acids than do those outside the dominant epitope.[8] A high substitution rate and Shannon entropy in the dominant epitope of hemagglutinin suggest that the dominant epitope is under the strongest positive selection by human antibodies. The immune pressure against each genotype of the dominant epitope can be at least partially quantified by the binding constant between the antibody and hemagglutinin.

The H3N2 virus and human immune system in this work are simplified to be a system consisting of the H3 hemagglutinin and the corresponding human antibody. Exposure by infection or vaccination produces an affinity-matured antibody with the binding constant to the corresponding hemagglutinin equal to 10^6-10^7 M^{-1}, while the binding constant of an antibody uncorrelated to the hemagglutinin is below 10^2 M^{-1}.[2] Escape substitutions may decrease the binding constant by changing the antibody binding free energy ΔG. Some substitutions decrease the antibody binding constant more than others and have higher probabilities to be fixed, because a decrease in the antibody binding constant is favorable to the virus. Here, we define the difference of antibody binding free energy as $\Delta\Delta G = \Delta G_{42} - \Delta G_{31}$ in which ΔG_{31} and ΔG_{42} are antibody-wildtype hemagglutinin binding free energy and antibody-evolved hemagglutinin binding free energy, respectively, as shown in Figure 1. The fixation tendency of each substitution is a function of the difference of the antibody binding free energy[9] of the escape substitution.

Received: September 21, 2010
Published: March 30, 2011

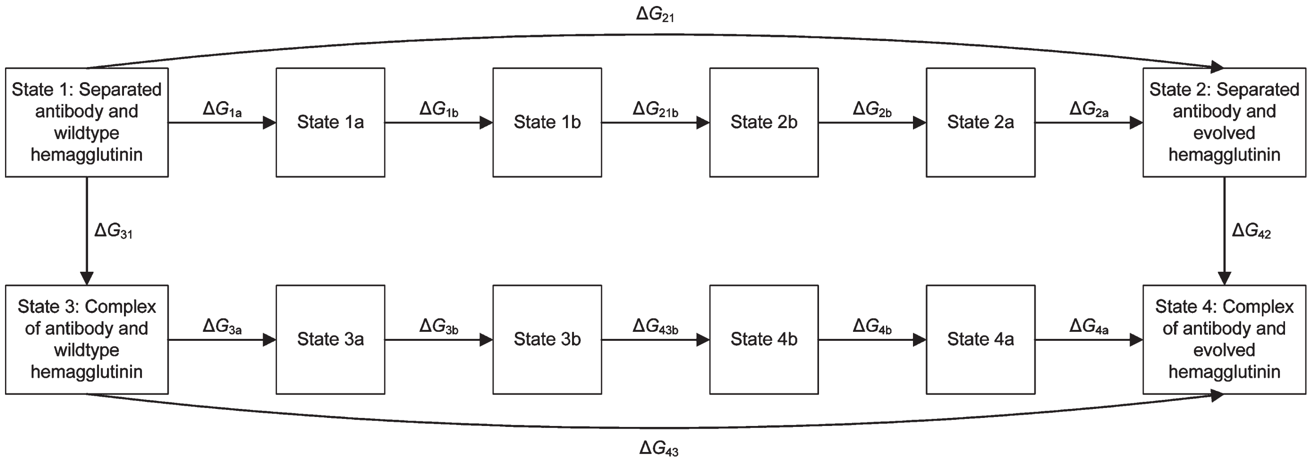

ΔG21
State 1: Separated antibody and wildtype hemagglutinin
ΔG1a
State 1a
ΔG1b
State 1b
ΔG21b
State 2b
ΔG2b
State 2a
ΔG2a
State 2: Separated antibody and evolved hemagglutinin
ΔG31
ΔG42
State 3: Complex of antibody and wildtype hemagglutinin
ΔG3a
State 3a
ΔG3b
State 3b
ΔG43b
State 4b
ΔG4b
State 4a
ΔG4a
State 4: Complex of antibody and evolved hemagglutinin
ΔG43

Figure 1. The scheme of the free energy calculation. The free energy difference of one substitution is calculated by $\Delta\Delta G = \Delta G_{43} - \Delta G_{21}$. State n, $n = 1-4$, is the real system. State na has the same configuration of atoms as state n except that all the hydrogen atoms have a mass of 16.000 amu. Compared to state na, state nb contains one additional Einstein crystal of product atoms ($n = 1, 3$) or reactant atoms ($n = 2, 4$). The mass of hydrogen atoms in state nb is also 16.000 amu. Free energy $\Delta G_{21\text{b}}$ and $\Delta G_{43\text{b}}$ are obtained by thermodynamic integration.

Epitope A or B of the H3N2 virus was dominant in most influenza seasons.[3] Epitope B of the H3N2 virus was the dominant epitope presenting more substitutions than any other epitope in the recent years. Epitope B was also dominant in 1968 when the H3N2 virus emerged. Thus, during these periods of time, the substitutions in epitope B directly affect the antibody binding constant and reflect the direction of the virus escape substitution. To attain a global view of the effects of substitutions in epitope B, it is necessary to compute a matrix containing the differences of antibody binding free energy caused by each possible single substitution in epitope B. There are 21 amino acid sites in epitope B, and each residue in the wild type strain may substitute to any of the 19 different types to amino acid residues. Hence, we need to calculate a 19 × 21 matrix with 399 elements. Such a matrix is a free energy landscape quantifying the immune selection over each evolved influenza strain. In this free energy landscape, the virus tends to evolve to a position with a low binding affinity of the antibody to evade antibodies and reduce the immune pressure. Calculation of this landscape will enable us to study the mechanism of immune escape from a quantitative viewpoint, providing a criterion to describe and foresee the evolution of influenza virus.

This paper is organized as follows: In the Materials and Methods section, we describe the protocol for the free energy calculation and the system of hemagglutinin and antibodies. In the Results section, we present and analyze the calculated free energy landscape. The substitutions observed in history are also compared with the results of the calculation. In the Discussion section, a general picture of H3N2 virus evolution under the selection pressure of the immune system is discussed, and simulation results are discussed. Finally, our work is summarized in the Conclusion section.

2. MATERIALS AND METHODS

2.1. Scheme of the Free Energy Calculation. The expression of the binding constant K depends on the antibody binding free energy ΔG, $K = \exp(-\Delta G/RT)$. The Boltzmann constant $R = 1.987 \times 10^{-3}$ kcal/mol/K. The temperature is fixed to the normal human body temperature $T = 310$ K. Shown in Figure 1, one substitution in hemagglutinin changes the antibody binding free energy from ΔG_{31} to ΔG_{42}. The first and second subscripts define the end state and the starting state of the binding process, respectively. The ratio of the antibody binding constant after and before substitution is written as

$$\frac{K_1}{K_0} = \exp(-\Delta\Delta G/RT) \qquad (1)$$

where K_1 and K_0 are the antibody binding constant to substituted and wildtype hemagglutinin, respectively.

The difference in antibody binding free energy $\Delta\Delta G = \Delta G_{42} - \Delta G_{31} = \Delta G_{43} - \Delta G_{21}$ is calculated by applying Hess' Law to the thermodynamic cycle defined by states 1−4 in Figure 1. The processes corresponding to ΔG_{43} and ΔG_{21} are unphysical but more convenient to simulate. We calculated ΔG_{21} and ΔG_{43} for each amino acid substitution in the unbound hemagglutinin and hemagglutinin bound by antibodies, respectively. On the surface of the virus particle, hemagglutinin exists in the form of a trimer in which three monomers are encoded by the same virus gene. Thus, we simultaneously substituted the amino acids in three hemagglutinin monomers in the trimer. The antibody has a Y-shaped structure with two heavy chains and two light chains. In the resolved structure (PDB code: 1KEN), the hemagglutinin trimer is bound by two Fab fragments. Thus, we incorporated the Fab dimer into the system for MD simulation.

Using the software CHARMM,[10] we calculated ΔG_{21} and ΔG_{43} using thermodynamic integration.[11] We used molecular dynamics (MD) simulation to obtain the ensemble averages of the integrand from which ΔG_{21} and ΔG_{43} are calculated. The potential energy for the MD algorithm to sample the conformation space of the system is

$$U(r, \lambda) = (1 - \lambda)\, U_{\text{reac}}(r) + \lambda U_{\text{prod}}(r) \qquad (2)$$

in which r represents the coordinates of all of the atoms, λ is the variable of integration, U_{reac} is the potential energy of the system corresponding to wildtype hemagglutinin, and U_{prod} is the potential energy of the system corresponding to substituted hemagglutinin. The value of ΔG_{21} or ΔG_{43} is

$$\Delta G = \int_0^1 \left\langle \frac{\partial U(r,\lambda)}{\partial \lambda} \right\rangle_\lambda \mathrm{d}\lambda = \int_0^1 \langle U_{\mathrm{prod}}(r) - U_{\mathrm{reac}}(r) \rangle_\lambda \, \mathrm{d}\lambda \tag{3}$$

The integrand $\langle U_{\mathrm{prod}}(r) - U_{\mathrm{reac}}(r)\rangle_\lambda$ is the ensemble average with fixed λ of potential energy difference between the system after and before substitution. The interval of integration $\lambda \in (0,1)$ was equally divided into four subintervals in each of which a 16-point Gauss-Legendre quadrature was applied to numerically integrate the ensemble averages. The ensemble averages with 64 distinct $\lambda \in (0,1)$ were calculated by MD simulation with the potential energy defined in eq 2.

2.2. Einstein Crystal.

We introduce the Einstein crystals to calculate the free energy of the reference state in the dual topology at both end points of the thermodynamic integration. To illustrate the function of the Einstein crystals, we analyze the free energy of the dual topology without Einstein crystals when $\lambda = 0$ as an example. We denote with n_1, n_2, and n_0 the numbers of the reactant atoms, product atoms, and all of the remaining atoms in the system, respectively. We denote with r, r_{product}, and x the coordinates of the reactant atoms, product atoms, and all of the remaining atoms in the system. The momenta of reactant atoms, product atoms, and the remaining atoms are denoted by $p_{r,i}$, $p_{p,i}$, and $p_{x,i}$. The masses are similarly denoted by $m_{r,i}$, $m_{p,i}$, and $m_{x,i}$. The Hamiltonian of the system with $\lambda = 0$ is

$$H = \sum_{i=1}^{n_0} \frac{p_{x,i}^2}{2m_{x,i}} + \sum_{i=1}^{n_1} \frac{p_{r,i}^2}{2m_{r,i}} + \sum_{i=1}^{n_2} \frac{p_{p,i}^2}{2m_{p,i}} + U_{n_0}(x) + (1-\lambda)U_{n_0+n_1}(x,r) + \lambda U_{n_0+n_2}(x, r_{\mathrm{product}}) \tag{4}$$

The partition function is

$$Q = \prod_{i=1}^{n_0} \left(\frac{2\pi m_{x,i}}{h^2\beta}\right)^{3/2} \prod_{i=1}^{n_1} \left(\frac{2\pi m_{r,i}}{h^2\beta}\right)^{3/2} \prod_{i=1}^{n_2} \left(\frac{2\pi m_{p,i}}{h^2\beta}\right)^{3/2} \int \mathrm{d}x\,\mathrm{d}r\, \exp[-\beta U_{n_0}(x) - \beta U_{n_0+n_1}(x,r)] \times \int \mathrm{d}r_{\mathrm{product}}\, 1 = Q_{\mathrm{real}} \times \prod_{i=1}^{n_2} \left[\left(\frac{2\pi m_{p,i}}{h^2\beta}\right)^{3/2} V\right] = Q_{\mathrm{real}} \times Q_{\mathrm{product}} \tag{5}$$

When $\lambda = 0$, this partition function is the product of Q_{real}, the partition function of the real system without product atoms, and Q_{product}, the partition function of the product atoms when $\lambda = 0$.

The free energy is given by $-1/\beta$ times the logarithm of the above partition function. The free energy is

$$G = G_{\mathrm{real}} - \frac{1}{\beta}\sum_{i=1}^{n_2} \frac{3}{2}\log\left(\frac{2\pi m_{p,i}}{h^2\beta}\right) - \frac{1}{\beta} n_2 \log V \tag{6}$$

As shown in the above equation, the effect on the translational entropy from the product atoms is proportional to the logarithm of system size V. It diverges in the thermodynamic limit. This divergence exists, no matter what λ scaling is performed. Note that we do not use the Einstein crystals to handle the translational entropy a ligand loses or gains when binding a flexible biomolecular receptor, which is taken into account by the thermodynamic cycle in Figure 1. The translational entropy, proportional to log V in eq 6, is that of the dummy product atoms, not that of the bound or unbound complex.

The value of G depends on the identity of the product atoms. Thus, the contribution to the thermodynamic integration is different at the two end points, i.e., $-kT \log Q_{\mathrm{reactant}} \neq -kT \log Q_{\mathrm{product}}$, in which Q_{reactant} is the partition function of the reactant atoms when $\lambda = 1$. Note also that the expression of the partition function contains the factor Q_{product} for the product atoms. Relating the conventional expression for thermodynamic integration, eq 3, to $\Delta\Delta G$ of eq 1 requires one to account for this term. This term arises from the use of a dual topology in CHARMM, and this term is typically ignored. While the contribution from the decoupled atoms is not constant, it can be exactly calculated if the restricted partition function over the decoupled atoms can be calculated. This calculation is what the Einstein crystal performs, using an Einstein crystal for the reference state rather the ideal gas in eq 4.

In four 16-window thermodynamic integrations, the smallest variable of integration is $\lambda = 1.32 \times 10^{-3}$. Since λ is close to zero, product atoms in the system have potential energy near zero and behave as ideal gas atoms, with translational entropy proportional to the logarithm of system size, see eq 6. Exact calculation of the translational entropy terms of product atoms at $\lambda = 0$ by explicit dynamics seems difficult, because the translational entropy of the product atoms grows as the logarithm of the system size. These relatively free product atoms destabilize the system. This entropy divergence is a fundamental feature of the statistical mechanics, not a numerical artifact. Unrestrained product atoms induce large fluctuation of the Hamiltonian in the MD algorithm. These fluctuations increase the standard error of the quantity $U_{\mathrm{prod}}(r) - U_{\mathrm{reac}}(r)$, which is defined in eq 3 and is computed from the trajectory of the MD simulation. These fluctuations often cause the numerical integration algorithm in the MD simulation to be unstable.[12] In this case, the energy of the simulated system increases rapidly. This phenomenon causes CHARMM to terminate abnormally. The translational entropy introduced by the free atoms at $\lambda = 0$ and 1 affects the result. Reactant atoms cause the same problem near $\lambda = 1$.

We noticed that the nonlinear scaling, i.e., using a high power of λ such as the fourth power of λ, in eq 2[13,14] did not work. The high power of the smallest λ is extremely close to zero, and the product atoms are almost free, which cause the MD simulation to terminate abnormally at several windows with small λ. Additionally, the issue of translational entropy of reactant and product atoms needs to be addressed. Even when the MD algorithm with the nonlinear scaling of λ[13,14] terminates and appears to have generated a converged simulation trajectory, this does not necessarily imply that the translational entropy of reactant or product atoms has been properly controlled. In fact, the λ scaling approach may hide the entropy divergence at $\lambda = 0$ or $\lambda = 1$ by letting the algorithm terminate due to numerical roundoff error, rather than building statistical mechanical reference states for each of $\lambda = 0$ and $\lambda = 1$ to account for or control the effect of translational entropy.

An alternative to λ scaling introduces the soft-core potential as a way to turn off the potential.[15,16] The soft-core approach, like the λ-scaling approach, does not address the translation entropy of the atoms at $\lambda = 0$ or $\lambda = 1$. Previous studies with non-constrained atoms at both end points have been performed.[17−23] Besides the classical molecular dynamics with a nonideal-gas reference state introduced into the dual topology, quantum molecular dynamics via metadynamics has been used to analyze a deamidation process.[24] Other applications of quantum-molecular-dynamics-based free energy calculation include chorismate conversion to prephenate,[25] isomerization of glycine,[26] and histone lysine methylation.[27] As illustrated in eq 6, the translational entropy of the uncoupled atoms causes error in the final free energy results if it is not accounted for.

One way to calculate the free energy change exactly is to use a nonideal-gas reference state. This is quite natural, since the protein is not composed of ideal gas atoms. Deng and Roux introduced restraint potentials to confine the translational and rotational motion of a bound ligand to accelerate convergence of the simulation.[28] We use this idea to exactly include the contribution from the restrained states and built two Einstein crystals as the reference states for reactant and product atoms, respectively. Our calculation allows a theoretically exact determination of the free energy due to amino acid substitution.

To handle these two difficulties at both end points of the integration in a theoretically exact way, we use two Einstein crystals as the reference states for reactant and product atoms, respectively. The Einstein crystal has been used as a reference state for free energy calculations. Frenkel and Ladd computed free energy of solids by building a path connecting the real solid and the reference Einstein crystal.[29] Noya et al. showed that a restrained Einstein crystal is a suitable reference in the free energy calculation of biomolecules.[30] The Einstein crystal, a solid state model, is consistent with the nature of antibody binding processes in the liquid phase. First, although the importance of biomolecular flexibility in protein–protein binding processes is well-accepted, and is fully and exactly included in our calculation, we simply need to localize the product atoms when $\lambda = 0$ and the reactant atoms when $\lambda = 1$. Moreover, we need to calculate the contribution to the free energy of these localized atoms.

The choice of Einstein crystals as the reference states removes the singularity in thermodynamic integration in eq 3. As an example, an Einstein crystal was used as the reference state for the free energy calculation of hard-sphere fluid in order to remove the singularity in eq 3 at the end point $\lambda = 0$.[31] In this example, the reference Einstein crystal was achieved by harmonically coupling the particles to their equilibrium positions and removing all interactions between particles.[32]

We here use Einstein crystals as the reference states to calculate the binding free energy change due to amino acid substitution. The Einstein crystal is a model for localized atoms. The free energy of the Einstein crystal can be exactly calculated. One Einstein crystal contains distinguishable and noninteracting atoms under harmonic constraints around reference positions fixed in space. In the Einstein crystal, the atom i with coordinates r_i has potential energy

$$U_i(r_i) = \frac{K_i}{2} || r_i - r_{i0} ||^2 \tag{7}$$

in which r_i and r_{i0} are the actual and reference positions of the atom, respectively, and K_i is the force constant of the harmonic constraint. We denote by m_i the mass of atom i. The canonical partition function of an Einstein crystal is

$$\begin{aligned} Q_E(N, V, T) &= \frac{1}{h^{3N}} \int \exp\left(\sum_{i=1}^{N} \frac{-\beta p_i^2}{2m_i}\right) \exp\left(\sum_{i=1}^{N} \frac{-\beta K_i \, || \, r_i - r_{i0} \, ||^2}{2}\right) \mathrm{d}p \, \mathrm{d}r \\ &= \left(\frac{2\pi}{h\beta}\right)^{3N} \prod_{i=1}^{N} \left(\frac{m_i}{K_i}\right)^{3/2} \end{aligned} \tag{8}$$

The spatial fluctuation of atom i in the Einstein crystal is

$$\langle (\delta r_i)^2 \rangle = \frac{3}{\beta K_i} \tag{9}$$

In our system, we let the potential energy for MD simulation defined by eq 2 become

$$U(r, \lambda) = (1 - \lambda) U_{\mathrm{reac}}(r) + \lambda U_{\mathrm{prod}}(r) + \lambda U_{\mathrm{ein,\,reac}}(r) + (1 - \lambda) U_{\mathrm{ein,\,prod}}(r) \tag{10}$$

Therefore, reactant and product atoms are localized at both $\lambda = 0$ and $\lambda = 1$. The reference positions of atoms in Einstein crystals are the equilibrium positions of corresponding reactant and product atoms. To minimize the numerical error during the thermodynamic integration calculation, we minimized the fluctuation of the integrand of thermodynamic integration $\langle \partial U(r,\lambda)/\partial\lambda \rangle_\lambda = \langle U_{\mathrm{ein,reac}}(r) - U_{\mathrm{reac}}(r) \rangle_\lambda + \langle U_{\mathrm{prod}}(r) - U_{\mathrm{ein,prod}}(r) \rangle_\lambda$. Minimization of the terms on the right-hand side is approximately achieved by letting the average spatial fluctuation of each atom in Einstein crystals equal that of the corresponding reactant or product atom, i.e.

$$\langle (\delta r_i)^2 \rangle_{\mathrm{reac}} = \langle (\delta r_i)^2 \rangle_{\mathrm{ein,\,reac}} = \frac{3}{\beta K_i^{\mathrm{reac}}} \tag{11}$$

$$\langle (\delta r_i)^2 \rangle_{\mathrm{prod}} = \langle (\delta r_i)^2 \rangle_{\mathrm{ein,\,prod}} = \frac{3}{\beta K_i^{\mathrm{prod}}} \tag{12}$$

For each atom in the Einstein crystal, the force constant of harmonic constraint, K_i^{reac} or K_i^{prod}, was calculated from the monitored fluctuations of the corresponding reactant or product atom with eq 11 or 12. In the scheme in Figure 1, the states with Einstein crystals are states 1b, 2b, 3b, and 4b.

2.3. Modified Hydrogen Atoms. The frequency of atom vibration depends on its mass. Hydrogen atoms generally have the highest vibration frequencies in the system. Such high frequencies require a short time step in MD simulation and increase the computational load. To limit vibration frequencies and allow a longer time step, one can apply the SHAKE algorithm to fix the length of any bond involving hydrogen atoms.[33] The SHAKE algorithm decreases the degrees of freedom in the system by introducing additional constraints between atoms. Instead, we artificially changed the mass of hydrogen atoms from 1.008 to 16.000 amu in order to preserve the degree of freedom in the system following the suggestion by Bennett.[34] A larger mass of hydrogen atoms allows a longer time step in the MD algorithm. Pomes and McCammon showed that changing the hydrogen mass to 10 amu allows the use of a 0.01 ps time step to simulate a system which consists of 215 TIP3P water molecules, smaller than our system.[35] Feenstra et al. change the mass of hydrogen atoms to 4 amu to increase the simulation stability of a system which contains protein and water molecules and resembles our system.[36] We set the time step as 0.001 ps, a value widely used in simulations with physical masses for all atoms, to gain higher stability in the simulation of our large system with a hemagglutinin trimer, a Fab dimer, and water molecules. As with the Einstein crystals, we exactly calculated and subtracted off the contribution of the change to the hydrogen mass to $\Delta\Delta G$. Note that the modification of hydrogen mass is independent of the reference states in the simulation, which is selected to be Einstein crystals in this project. In fact, most of the hydrogen atoms in the system are neither reactant nor product atoms. In Figure 1, the states with Einstein crystals and modified hydrogen atoms are states 1a, 2a, 3a, 4a, 1b, 2b, 3b, and 4b.

2.4. Expressions of Free Energies. Introducing two Einstein crystals and heavier hydrogen atoms changes the potential energy in the system, as well as the canonical partition functions. After the modification of hydrogen atoms, the mass of atoms changed from $m_{r,i}$ to $m'_{r,i}$, from $m_{p,i}$ to $m'_{p,i}$, or from $m_{x,i}$ to $m'_{x,i}$. Canonical partition functions of the states in Figure 1 are

$$Q_3(n_0+n_1,V,T) = \frac{1}{h^{3(n_0+n_1)}}\prod_{i=1}^{n_0}\left(\frac{2\pi m_{x,i}}{\beta}\right)^{3/2}\prod_{i=1}^{n_1}\left(\frac{2\pi m_{r,i}}{\beta}\right)^{3/2} \times Z_3(n_0+n_1,V,T) \qquad (13)$$

$$Q_{3a}(n_0+n_1,V,T) = \frac{1}{h^{3(n_0+n_1)}}\prod_{i=1}^{n_0}\left(\frac{2\pi m'_{x,i}}{\beta}\right)^{3/2}\prod_{i=1}^{n_1}\left(\frac{2\pi m'_{r,i}}{\beta}\right)^{3/2} \times Z_3(n_0+n_1,V,T) \qquad (14)$$

$$Q_{3b}(n_0+n_1+n_2,V,T) = \frac{1}{h^{3(n_0+n_1)}}\prod_{i=1}^{n_0}\left(\frac{2\pi m'_{x,i}}{\beta}\right)^{3/2}\prod_{i=1}^{n_1}\left(\frac{2\pi m'_{r,i}}{\beta}\right)^{3/2} \times Z_3(n_0+n_1,V,T)\left(\frac{2\pi}{h\beta}\right)^{3n_2}\prod_{i=1}^{n_2}\left(\frac{m'_{p,i}}{K_i^{\text{prod}}}\right)^{3/2} \qquad (15)$$

$$Q_4(n_0+n_2,V,T) = \frac{1}{h^{3(n_0+n_2)}}\prod_{i=1}^{n_0}\left(\frac{2\pi m_{x,i}}{\beta}\right)^{3/2}\prod_{i=1}^{n_2}\left(\frac{2\pi m_{p,i}}{\beta}\right)^{3/2} \times Z_4(n_0+n_2,V,T) \qquad (16)$$

$$Q_{4a}(n_0+n_2,V,T) = \frac{1}{h^{3(n_0+n_2)}}\prod_{i=1}^{n_0}\left(\frac{2\pi m'_{x,i}}{\beta}\right)^{3/2}\prod_{i=1}^{n_2}\left(\frac{2\pi m'_{p,i}}{\beta}\right)^{3/2} \times Z_4(n_0+n_2,V,T) \qquad (17)$$

$$Q_{4b}(n_0+n_1+n_2,V,T) = \frac{1}{h^{3(n_0+n_2)}}\prod_{i=1}^{n_0}\left(\frac{2\pi m'_{x,i}}{\beta}\right)^{3/2}\prod_{i=1}^{n_2}\left(\frac{2\pi m'_{p,i}}{\beta}\right)^{3/2} \times Z_4(n_0+n_2,V,T)\left(\frac{2\pi}{h\beta}\right)^{3n_1}\prod_{i=1}^{n_1}\left(\frac{m'_{r,i}}{K_i^{\text{reac}}}\right)^{3/2} \qquad (18)$$

in which the states are denoted by the subscripts. Contribution of the potential energy part of the Hamiltonian to the partition function is

$$Z_3(n_0+n_1,V,T) = \int \exp(-\beta U_{n_0+n_1}(r))\,\mathrm{d}r \qquad (19)$$

$$Z_4(n_0+n_2,V,T) = \int \exp(-\beta U_{n_0+n_2}(r))\,\mathrm{d}r \qquad (20)$$

From the partition functions, free energies defined in Figure 1 are calculated:

$$\Delta G_{3a} = -\frac{3}{2\beta}\sum_{i=1}^{n_0}\ln\left(\frac{m'_{x,i}}{m_{x,i}}\right) - \frac{3}{2\beta}\sum_{i=1}^{n_1}\ln\left(\frac{m'_{r,i}}{m_{r,i}}\right) \qquad (21)$$

$$\Delta G_{4a} = \frac{3}{2\beta}\sum_{i=1}^{n_0}\ln\left(\frac{m'_{x,i}}{m_{x,i}}\right) + \frac{3}{2\beta}\sum_{i=1}^{n_2}\ln\left(\frac{m'_{p,i}}{m_{p,i}}\right) \qquad (22)$$

$$\Delta G_{3b} = -\frac{3n_2}{\beta}\ln\left(\frac{2\pi}{h\beta}\right) - \frac{3}{2\beta}\sum_{i=1}^{n_2}\ln\left(\frac{m'_{p,i}}{K_i^{\text{prod}}}\right) \qquad (23)$$

$$\Delta G_{4b} = \frac{3n_1}{\beta}\ln\left(\frac{2\pi}{h\beta}\right) + \frac{3}{2\beta}\sum_{i=1}^{n_1}\ln\left(\frac{m'_{r,i}}{K_i^{\text{reac}}}\right) \qquad (24)$$

$$\Delta G_{43b} = -\frac{1}{\beta}\ln\left[\frac{\prod_{i=1}^{n_1}(m'_{r,i}/K_i^{\text{reac}})^{3/2}Z_4(n_0+n_2,V,T)}{\prod_{i=1}^{n_2}(m'_{p,i}/K_i^{\text{prod}})^{3/2}Z_3(n_0+n_1,V,T)}\right] \qquad (25)$$

The free energy between states 3 and 4 is

$$\Delta G_{43} = \Delta G_{43b} - \frac{1}{\beta}\ln\frac{(2\pi/h\beta)^{3n_2}\sum_{i=1}^{n_2}(m_{p,i}/K_i^{\text{prod}})^{3/2}}{(2\pi/h\beta)^{3n_1}\sum_{i=1}^{n_1}(m_{r,i}/K_i^{\text{reac}})^{3/2}}$$

$$= \Delta G_{43b} - \frac{1}{\beta}\ln\frac{Q_{E2}(n_2,V,T)}{Q_{E1}(n_1,V,T)} \qquad (26)$$

in which Q_{E1} and Q_{E2} are the partition functions of the Einstein crystals for product atoms and reactant atoms, respectively. The free energy ΔG_{43b} was calculated by thermodynamic integration, while ΔG_{43} was used to calculate the free energy difference of one substitution. Note that the correction term between ΔG_{43b} and ΔG_{43} is independent of the masses of atoms. Canonical partition functions as well as free energies of states 1, 1a, 1b, 2, 2a, and 2b are calculated in a similar way.

2.5. Implementation of Free Energy Calculation Algorithm. The above discussion is the theoretical basis for the implementation of our free energy calculation algorithm. The free energy calculation protocol consists of four steps. First, we built the dual topology with reactant and product atoms in the amino acid substitution site in separated antibodies and hemagglutinin or an antibody–hemagglutinin complex. We then solvated the protein system and modified the mass of hydrogen atoms. Second, two Einstein crystals were introduced as the reference states for the reactant and product atoms, respectively. Third, the MD simulation was run at 64 windows. The thermodynamic integration algorithm obtained the free energy values ΔG_{21} for separated antibodies and hemagglutinin or ΔG_{43} for the antibody–hemagglutinin complex, as in Figure 1. This step gave the $\Delta\Delta G$ value. Fourth, we calculated the error bar of the $\Delta\Delta G$ value obtained in the last step. The technical details of these four steps are illustrated in the text below. Also described are the verification of the free energy calculation protocol, the software and hardware information, and the CPU hours consumed by the protocol.

The hemagglutinin trimer of H3N2 virus strain A/Aichi/2/1968 with bound dimer antibody HC63 (PDB code: 1KEN) was used in our calculation. For each amino acid substitution, we built the dual topology with side chains of both amino acids prior to the simulation. Reactant and product atoms were defined as the side chains in the original and substituting amino acid, respectively. All of the covalent and nonbonded interactions between reactant and product atoms were removed. The protein was in an explicit water box with periodic boundary conditions. The mass of hydrogen atoms was changed from 1.008 to 16.000 amu.

All of the simulations were performed by CHARMM c33b2 with a CHARMM22 force field.[10] We first fixed the positions of the hemagglutinin trimer, except for reactant atoms, and minimized the system with a 200 steps of steepest descent (SD) algorithm and a 5000 steps of adopted basis Newton−Raphson (ABNR) algorithm. We ran a 5 ps MD simulation of the system, the trajectory of which gave the spatial fluctuation $\langle(\delta r_i)^2\rangle$ of each reactant atom. Then, we fixed reactant atoms, released product atoms, and ran a 5 ps MD simulation to obtain the spatial fluctuation of each product atom. Final positions of both reactant and product were adopted as the reference positions of the corresponding Einstein crystal. The force constant K_i of each atom in Einstein crystals was obtained from $\langle(\delta r_i)^2\rangle$ using eqs 11 and 12. With modified hydrogen atoms and two Einstein crystals as the reference states of reactant and product atoms, states 1b, 2b, 3b, and 4b in Figure 1 were generated for thermodynamic integration.

In thermodynamic integration, MD simulations were run at 64 windows with distinct λ's. In each window, the pressure of the system was first calibrated with a 10 ps MD simulation in an isothermal−isobaric (NPT) ensemble. The duration of 10 ps is appropriate because it is long enough to equilibrate the pressure and short enough to prevent the protein from drifting away from the original location. We fixed coordinates of the residues and water molecules except for those within 15 Å from the three α carbons. Then, we removed amino acid residues and water molecules other than those within 27.5 Å from the three α carbons of substituted residues in the hemagglutinin trimer to reduce the system size, because the fixed atoms are not included in the topology of movable atoms and the cutoff of the nonbonded forces is 12 Å. The Ewald sum was used to calculate charge interactions. Note that this substantial reduction of the system relies on the assumption that the free energy change due to the amino acid substitution is mostly affected by atoms near the binding site after the system reaches equilibrium. This assumption is based on two facts: the conformations of hemagglutinin and antibodies are stable once the system reaches equilibrium, and all of the removed or fixed atoms have invariant interactions with the substituting amino acid residues. The stable protein conformation means amino acid residues far away from the substituting residue do not move during the amino acid substitution process. In the CHARMM22 force field used in this project, the cutoff of nonbonded force is 12 Å and less than the 15 Å threshold for system reduction. The system reduction does not directly affect the force on the substituted residue because of an absence of the long-range nonbonded force between the substituted residue and atoms removed from the system. This system reduction method was also applied to compute the binding free energies of subtilisin,[37] tripsin,[21] and the Src SH2 domain.[22] Robust results were obtained in all of these applications. Generally, this system reduction strategy can produce reliable result if the reduced system contains the residues and molecules critical to the binding process.[21] We note that the system reduction method could be a limitation of the free energy calculation model. The fixing of amino acid residues and water molecules described in section 2.5 substantially reduced the CPU time needed but is an approximation to the real system containing the whole proteins. This limitation reflects the trade-off between model accuracy and required computational resource. In the canonical ensemble, the new system was equilibrated for 200 ps and simulated for another 900 ps as the data production phase. The integrand of thermodynamic integration is the ensemble average of the sampled trajectory $\langle\partial U(r,\lambda)/\partial\lambda\rangle_\lambda = \langle U_{\mathrm{ein,reac}}(r) - U_{\mathrm{reac}}(r) - U_{\mathrm{ein,prod}}(r) + U_{\mathrm{prod}}(r)\rangle_\lambda$. The free energies ΔG_{21} and ΔG_{43} between the real states were calculated by adding a correction term of the Einstein crystals in eq 26. Finally, the difference of antibody binding free energy is $\Delta\Delta G = \Delta G_{43} - \Delta G_{21}$.

Error bars of $\Delta\Delta G$ are also given. The convergence behavior of the simulation was analyzed using the block average method developed by Flyvbjerg and Petersen.[38] As mentioned above, the MD simulation for either the unbound hemagglutinin or the hemagglutinin−antibody complex contains 64 windows with distinct λ. The 900 ps data production phase contains 9×10^5 simulation steps. The values $A = U_{\mathrm{prod}}(r) - U_{\mathrm{reac}}(r)$, as in equation eq 3, computed in consecutive simulation steps were grouped into bins, and consecutive bins were merged progressively. The quantity $\sigma^2(A)/(n-1)$, in which $\sigma^2(A)$ is the variance of the average of each bin $A_1, A_2, ..., A_n$ and n is the number of bins, increases with the bin size and reaches a plateau when the bin size is 1×10^4 steps. We fixed the bin size to 1×10^4 steps and estimate the variance of ensemble average $\langle A\rangle$ as $\sigma^2(A)/(n-1)$, following Flyvbjerg and Petersen's method.[38]

This protocol, without the Einstein crystal contribution, was verified by recalculating published free energy differences of amino acid substitution T131I.[9] Without the Einstein crystal contribution, our protocol gave $\Delta\Delta G = 5.69 \pm 0.07$ kcal/mol, compared to the $\Delta\Delta G = 5.20 \pm 0.94$ kcal/mol in the published work.[9] Theoretically exact results presented here include the Einstein crystal contribution. We note that the theoretically exact $\Delta\Delta G$ for T131I, including the Einstein crystal contribution, is 3.71 ± 0.07 kcal/mol.

The simulation was performed using a CHARMM22 force field at three clusters: tg-steele.purdue.teragrid.org (Intel Xeon E5410, 2.33 GHz), sugar.rice.edu (Intel Xeon E5440, 2.83 GHz), and biou.rice.edu (IBM POWER7, 3.55 GHz), as well as at the condor pool tg-condor.rcac.purdue.edu at Purdue University. Simulation of each substitution took approximately 7.5 thousand CPU hours on average, and so this work consumed about 3 million CPU hours.

3. RESULTS

3.1. Free Energy Landscape. For each of the 21 amino acid sites in epitope B, we substituted from alanine to each one of the 19 other amino acids, in which we used the neutral histidine (CHARMM code: Hse) as the model of histidine. The free energy difference and standard error of each substitution were calculated using the MD simulation (see Materials and Methods). The wildtype amino acid in each site of epitope B was extracted from the hemagglutinin sequence of the H3N2 strain A/Aichi/2/1968. The free energy difference and standard error of the substitution from the wildtype amino acid in each site were then calculated from the values for the change from the wildtype amino acid to alanine and from alanine to the new amino acid. The values are listed in Table 1.

As described in eq 26, each $\Delta\Delta G$ value listed in Table 1 contains the contribution of two Einstein crystals. The contribution of Einstein crystals to the final $\Delta\Delta G$ values was calculated for each of the 399 amino acid substitutions in epitope B. The average fraction of the contribution of Einstein crystals in the calculated $\Delta\Delta G$ values is 44%. The contribution of Einstein crystals is far greater than that of the statistical error of our free energy calculation in Table 1, which is 4.5% on average. Thus, the Einstein crystal contribution is both theoretically exact and practically important. In 371 of the 399 substitutions, the absolute values of the contribution of Einstein crystals is greater than 1.96 standard errors of the final $\Delta\Delta G$ values. That is, the

Table 1. Summary of the Calculated Free Energy Differences ΔΔG in Each Amino Acid Site in Epitope B from the Wildtype Amino Acid to All 20 Amino Acids[a]

positions	128	129	155	156	157	158	159
Ala	−13.12 ± 0.27	3.33 ± 0.29	2.78 ± 0.20	1.19 ± 0.33	2.48 ± 0.21	4.27 ± 0.31	5.18 ± 0.21
Arg	22.57 ± 0.46	2.31 ± 0.45	16.98 ± 0.37	0.08 ± 0.50	−4.19 ± 0.44	−1.61 ± 0.48	7.07 ± 0.42
Asn	−4.80 ± 0.36	5.83 ± 0.42	−7.83 ± 0.30	10.72 ± 0.40	5.64 ± 0.34	3.41 ± 0.42	10.97 ± 0.35
Asp	4.52 ± 0.38	19.12 ± 0.42	16.28 ± 0.32	11.06 ± 0.42	9.95 ± 0.37	18.37 ± 0.40	15.34 ± 0.36
Cys	−11.83 ± 0.34	12.64 ± 0.37	−2.37 ± 0.30	5.32 ± 0.38	−2.72 ± 0.29	−7.88 ± 0.40	7.92 ± 0.32
Gln	−12.37 ± 0.40	7.34 ± 0.42	−4.29 ± 0.36	13.14 ± 0.41	−0.45 ± 0.36	11.47 ± 0.43	6.54 ± 0.40
Glu	11.15 ± 0.38	10.50 ± 0.42	17.77 ± 0.34	26.54 ± 0.43	4.68 ± 0.36	8.58 ± 0.48	5.19 ± 0.39
Gly	−9.93 ± 0.39	0.00 ± 0.00	17.00 ± 0.34	0.11 ± 0.44	0.21 ± 0.36	0.00 ± 0.00	−4.19 ± 0.41
His	4.43 ± 0.42	0.15 ± 0.43	2.47 ± 0.36	−6.89 ± 0.43	12.18 ± 0.38	5.54 ± 0.46	1.06 ± 0.39
Ile	−16.03 ± 0.41	0.54 ± 0.40	1.55 ± 0.33	8.33 ± 0.42	11.22 ± 0.37	8.09 ± 0.43	18.96 ± 0.39
Leu	−23.58 ± 0.41	−4.27 ± 0.43	−8.92 ± 0.33	2.64 ± 0.45	−6.26 ± 0.39	1.61 ± 0.45	4.08 ± 0.38
Lys	3.57 ± 0.45	11.18 ± 0.46	14.58 ± 0.37	0.00 ± 0.00	6.24 ± 0.40	−1.60 ± 0.48	5.39 ± 0.46
Met	−13.38 ± 0.39	−2.59 ± 0.39	1.23 ± 0.35	10.11 ± 0.43	16.15 ± 0.36	14.49 ± 0.44	−6.38 ± 0.37
Phe	−10.21 ± 0.45	6.12 ± 0.43	9.39 ± 0.35	0.30 ± 0.45	10.28 ± 0.40	5.17 ± 0.48	12.33 ± 0.42
Pro	−9.36 ± 0.36	−2.43 ± 0.42	−1.86 ± 0.31	2.32 ± 0.43	5.69 ± 0.30	17.09 ± 0.40	6.08 ± 0.36
Ser	−14.55 ± 0.34	3.36 ± 0.37	−1.09 ± 0.29	−1.45 ± 0.38	0.00 ± 0.00	2.76 ± 0.39	0.00 ± 0.00
Thr	0.00 ± 0.00	7.35 ± 0.36	0.00 ± 0.00	−1.08 ± 0.41	6.34 ± 0.32	8.36 ± 0.41	15.32 ± 0.32
Trp	9.82 ± 0.47	4.81 ± 0.47	19.84 ± 0.43	23.26 ± 0.48	16.14 ± 0.45	3.52 ± 0.62	−1.35 ± 0.45
Tyr	−14.83 ± 0.43	2.72 ± 0.42	7.25 ± 0.36	−2.18 ± 0.46	−8.37 ± 0.44	18.42 ± 0.51	5.95 ± 0.43
Val	−19.13 ± 0.37	3.56 ± 0.38	8.57 ± 0.31	−3.01 ± 0.39	7.63 ± 0.32	3.77 ± 0.42	6.45 ± 0.32
positions	160	163	165	186	187	188	189
Ala	4.16 ± 0.22	−0.24 ± 0.22	4.15 ± 0.24	−3.19 ± 0.19	−4.03 ± 0.23	3.45 ± 0.25	−9.01 ± 0.28
Arg	9.70 ± 0.44	5.97 ± 0.39	14.58 ± 0.41	21.01 ± 0.38	8.12 ± 0.42	−0.06 ± 0.45	−0.39 ± 0.48
Asn	2.07 ± 0.34	−2.32 ± 0.32	0.00 ± 0.00	4.67 ± 0.30	−10.07 ± 0.34	0.00 ± 0.00	−3.18 ± 0.37
Asp	13.50 ± 0.32	12.64 ± 0.32	25.01 ± 0.31	24.54 ± 0.28	7.78 ± 0.35	19.77 ± 0.37	6.77 ± 0.35
Cys	15.82 ± 0.31	1.84 ± 0.30	1.93 ± 0.29	−2.30 ± 0.25	−11.09 ± 0.32	4.07 ± 0.34	6.23 ± 0.33
Gln	3.04 ± 0.39	−8.29 ± 0.35	4.27 ± 0.36	5.16 ± 0.33	−2.87 ± 0.37	12.36 ± 0.39	0.00 ± 0.00
Glu	15.48 ± 0.36	2.17 ± 0.35	15.74 ± 0.34	33.29 ± 0.31	14.41 ± 0.35	10.10 ± 0.37	12.16 ± 0.39
Gly	1.22 ± 0.39	−5.83 ± 0.38	9.11 ± 0.37	0.13 ± 0.27	−0.60 ± 0.30	−5.06 ± 0.32	−5.69 ± 0.32
His	0.52 ± 0.38	6.31 ± 0.33	7.44 ± 0.33	18.15 ± 0.30	3.69 ± 0.39	−1.95 ± 0.40	−8.53 ± 0.40
Ile	1.51 ± 0.34	10.62 ± 0.36	3.85 ± 0.33	−1.85 ± 0.30	−2.51 ± 0.33	−4.77 ± 0.37	3.65 ± 0.37
Leu	−1.39 ± 0.40	3.85 ± 0.35	−9.20 ± 0.37	1.07 ± 0.30	−0.40 ± 0.38	−1.30 ± 0.37	−6.91 ± 0.39
Lys	5.91 ± 0.44	10.37 ± 0.38	1.93 ± 0.41	−1.15 ± 0.39	24.91 ± 0.41	8.42 ± 0.44	9.48 ± 0.64
Met	10.78 ± 0.38	7.22 ± 0.35	1.63 ± 0.36	13.06 ± 0.33	−5.11 ± 0.36	6.97 ± 0.38	6.86 ± 0.40
Phe	7.90 ± 0.41	−0.86 ± 0.36	13.87 ± 0.38	6.94 ± 0.33	−7.23 ± 0.39	2.05 ± 0.39	4.37 ± 0.43
Pro	4.51 ± 0.32	12.50 ± 0.34	18.96 ± 0.33	11.82 ± 0.29	10.69 ± 0.32	−10.24 ± 0.35	−8.98 ± 0.36
Ser	7.13 ± 0.29	9.07 ± 0.30	−0.92 ± 0.28	0.00 ± 0.00	−4.88 ± 0.31	8.09 ± 0.33	−5.09 ± 0.34
Thr	0.00 ± 0.00	9.18 ± 0.30	10.35 ± 0.31	−14.79 ± 0.27	0.00 ± 0.00	3.53 ± 0.38	9.30 ± 0.35
Trp	0.86 ± 0.44	12.34 ± 0.35	19.02 ± 0.43	−7.69 ± 0.38	−11.04 ± 0.48	7.20 ± 0.40	−9.19 ± 0.45
Tyr	−5.43 ± 0.39	1.06 ± 0.34	14.76 ± 0.37	11.90 ± 0.33	5.29 ± 0.42	1.57 ± 0.40	4.81 ± 0.41
Val	7.99 ± 0.34	0.00 ± 0.00	9.79 ± 0.32	2.97 ± 0.29	3.08 ± 0.33	3.73 ± 0.34	−7.89 ± 0.36
positions	190	192	193	194	196	197	198
Ala	−18.12 ± 0.24	−0.86 ± 0.23	−5.20 ± 0.20	2.37 ± 0.23	5.95 ± 0.23	−2.40 ± 0.29	0.00 ± 0.00
Arg	4.97 ± 0.41	23.07 ± 0.44	32.33 ± 0.41	−13.66 ± 0.37	−25.38 ± 0.44	−17.94 ± 0.47	3.99 ± 0.37
Asn	−16.44 ± 0.30	−2.56 ± 0.32	8.24 ± 0.30	−3.81 ± 0.31	13.27 ± 0.36	−6.58 ± 0.38	0.05 ± 0.28
Asp	18.75 ± 0.32	2.92 ± 0.32	15.29 ± 0.29	26.72 ± 0.35	9.25 ± 0.34	5.58 ± 0.39	5.17 ± 0.24
Cys	−20.36 ± 0.32	−1.45 ± 0.32	−9.79 ± 0.26	1.91 ± 0.30	1.30 ± 0.31	6.70 ± 0.36	5.91 ± 0.22
Gln	−17.37 ± 0.37	−6.00 ± 0.37	4.87 ± 0.34	−0.83 ± 0.32	7.68 ± 0.36	0.00 ± 0.00	1.41 ± 0.31
Glu	0.00 ± 0.00	1.18 ± 0.35	45.40 ± 0.34	38.35 ± 0.33	3.60 ± 0.36	11.34 ± 0.41	2.37 ± 0.30
Gly	−17.09 ± 0.29	−13.46 ± 0.30	−13.89 ± 0.27	−18.59 ± 0.30	8.08 ± 0.31	4.11 ± 0.36	3.65 ± 0.28
His	−26.26 ± 0.35	−0.96 ± 0.38	−0.96 ± 0.35	9.95 ± 0.34	18.42 ± 0.37	−2.62 ± 0.40	−3.27 ± 0.35

Table 1. Continued

positions	190	192	193	194	196	197	198
Ile	−16.45 ± 0.37	−5.57 ± 0.37	−3.80 ± 0.31	−6.91 ± 0.32	0.77 ± 0.34	1.23 ± 0.41	0.01 ± 0.32
Leu	−17.27 ± 0.36	−7.97 ± 0.37	10.76 ± 0.34	0.00 ± 0.00	10.07 ± 0.39	−0.03 ± 0.40	−11.18 ± 0.29
Lys	−9.33 ± 0.38	5.67 ± 0.42	39.36 ± 0.39	−16.67 ± 0.38	0.49 ± 0.40	−16.50 ± 0.47	1.98 ± 0.37
Met	−26.63 ± 0.34	6.82 ± 0.36	−2.91 ± 0.32	7.75 ± 0.35	4.08 ± 0.37	−7.79 ± 0.40	15.57 ± 0.32
Phe	−31.89 ± 0.39	1.56 ± 0.40	16.46 ± 0.59	2.78 ± 0.34	−1.99 ± 0.37	1.05 ± 0.44	8.73 ± 0.34
Pro	−17.85 ± 0.33	−2.28 ± 0.33	9.84 ± 0.31	8.01 ± 0.31	15.42 ± 0.35	−5.34 ± 0.40	0.70 ± 0.29
Ser	−14.75 ± 0.31	−7.79 ± 0.30	0.00 ± 0.00	6.62 ± 0.29	6.91 ± 0.29	1.97 ± 0.36	−2.40 ± 0.22
Thr	−4.17 ± 0.32	0.00 ± 0.00	−2.04 ± 0.27	12.40 ± 0.31	7.81 ± 0.33	−7.91 ± 0.36	6.79 ± 0.24
Trp	−22.93 ± 0.39	2.31 ± 0.44	17.92 ± 0.42	−1.30 ± 0.40	8.17 ± 0.43	−7.73 ± 0.44	−7.23 ± 0.38
Tyr	−13.82 ± 0.38	7.63 ± 0.42	16.16 ± 0.38	9.73 ± 0.36	2.92 ± 0.40	6.10 ± 0.44	−4.82 ± 0.32
Val	−9.12 ± 0.31	−6.80 ± 0.32	−6.92 ± 0.30	2.59 ± 0.29	0.00 ± 0.00	4.16 ± 0.39	−4.22 ± 0.24

[a] The standard errors are also listed. The free energy difference and its standard error of the substitution from the wildtype amino acid to itself are both zero. The units of free energy differences and their standard errors are in kcal/mol.

Table 2. Rank of the Average Binding Free Energy Difference of the Single Substitution from Alanine to Another Amino Acid over All 21 Amino Acid Sites in Epitope B of Hemagglutinin Trimer[a]

rank	amino acid	$\Delta\Delta G$(kcal/mol)	charged	hydrophobic	large	medium	small	relative frequency
1	Glu	14.612 ± 0.061	×		×			0.029
2	Asp	14.533 ± 0.055	×			×		0.051
3	Arg	6.018 ± 0.078	×		×			0.052
4	Lys	5.766 ± 0.078	×		×			0.057
5	Trp	4.458 ± 0.081		×	×			0.016
6	Tyr	3.984 ± 0.071		×	×			0.035
7	Thr	3.981 ± 0.050				×		0.078
8	Pro	3.912 ± 0.054		×		×		0.060
9	Met	3.562 ± 0.062		×	×			0.009
10	Phe	3.522 ± 0.073		×	×			0.030
11	His	2.654 ± 0.064			×			0.020
12	Gln	1.985 ± 0.063			×			0.042
13	Ile	1.396 ± 0.060		×	×			0.070
14	Asn	1.150 ± 0.054				×		0.085
15	Val	1.147 ± 0.051		×		×		0.055
16	Cys	0.888 ± 0.046				×		0.028
17	Ser	0.469 ± 0.044					×	0.096
(18)	(Ala)	(0.000 ± 0.000)		×			×	0.046
19	Gly	−1.612 ± 0.055		×			×	0.070
20	Leu	−2.273 ± 0.064		×	×			0.071

[a] The rank correlates with the charge and the size of the amino acid, and it is relatively uncorrelated to the hydrophobicity. Here, we applied classifications of RasMol for the biochemical properties of the 20 amino acids.[43] The relative frequencies of 20 amino acids were counted from the H3 sequences in the NCBI database from 1968 to 2009.

contribution of Einstein crystals is significant with $p < 0.05$ in 93.0% of all of the amino acid substitutions. Consequently, it is essential to incorporate Einstein crystals in the free energy calculation to eliminate the error caused by the methods that neglect the unknown effect of the translational entropy of the free atoms in thermodynamic integration. The contribution of the translational entropy of ideal gas-like atoms ($\lambda = 0$ or $\lambda = 1$) needs to be either calculated or removed by a theoretically exact method to perform an exact free energy calculation.

The obtained $\Delta\Delta G$ values allow us to analyze the character of each of the 20 amino acids. We first averaged over all of the 21 amino acid sites in epitope B the $\Delta\Delta G$ value caused by the single substitutions from alanine to the other amino acids. The averaged $\Delta\Delta G$ values are listed in Table 2. The largest $\Delta\Delta G$'s are caused by the negatively charged amino acids (Glu, Asp) and the positively charged amino acids (Arg, Lys), indicating that introduction of charged amino acids in the dominant epitope decreases the binding affinity between antibody and hemagglutinin. Note that amino acid substitutions that change the charge of hemagglutinin significantly affect the calculated free energy values.[39−41] The issue of how to best calculate free energy differences when charge changes has been debated over the years. In the present paper, we are using the standard Ewald approach with explicit solvent. We note that the evolutionary history of H3 hemagglutinin since 1968 shows an increasing trend of the number of charged amino acids in epitope B,[42] which agrees with the results that the introduction of charged amino facilitates virus evasion from antibodies, as illustrated in Table 2.

Table 3. Rank of the Average Free Energy Difference $\langle\Delta\Delta G\rangle_i$ Generated by a Substitution in Each Amino Acid Site *i* of Epitope B

rank	site	$\langle\Delta\Delta G\rangle_i$ (kcal/mol)
1	193	8.074 ± 0.081
2	159	7.792 ± 0.094
3	165	7.741 ± 0.086
4	158	6.128 ± 0.108
5	196	5.444 ± 0.088
6	160	4.956 ± 0.090
7	186	4.754 ± 0.076
8	163	4.722 ± 0.085
9	129	4.690 ± 0.103
10	155	4.471 ± 0.081
11	156	4.029 ± 0.106
12	157	3.944 ± 0.090
13	188	2.945 ± 0.092
14	194	1.886 ± 0.080
15	187	1.182 ± 0.087
16	198	0.531 ± 0.072
17	189	-0.631 ± 0.098
18	192	-1.737 ± 0.087
19	197	-1.967 ± 0.099
20	128	-7.746 ± 0.098
21	190	-12.666 ± 0.084

The result that the introduction of charged amino acids on average increases $\Delta\Delta G$ is not an artifact, is supported by data from the influenza evolution, and is expected on the basis that charge is hydrophilic. In addition to the charge, the rank of free energy differences also largely correlated to the size of amino acid. By the definition used by RasMol,[43] the 16 uncharged amino acids are tagged as hydrophobic (Ala, Gly, Ile, Leu, Met, Phe, Pro, Trp, Tyr, Val), large (Gln, His, Ile, Leu, Met, Phe, Trp, Tyr), medium (Asn, Cys, Pro, Thr, Val), and small (Ala, Gly, Ser), as shown in Table 2. The ranks of small amino acids are lower than those of medium amino acids (p = 0.036, Wilcoxon rank-sum test) and those of large amino acids (p = 0.085, Wilcoxon rank-sum test). In contrast, the hydrophobicity of the uncharged amino acids is largely uncorrelated with their ranks by $\Delta\Delta G$. As a result, charged amino acids in the dominant epitope are essential to the immune evasion, while the virus escape substitution among small amino acids has minimal effect.

Epitope B comprises 21 amino acid sites in the top of the hemagglutinin trimer. Taking the probability for one substituting amino acid to exist at each site to be proportional to the relative frequency of this amino acid in H3 hemagglutinin, the weighted average free energy difference in each of the 21 sites was calculated. The relative frequencies of 20 amino acids were obtained from 6896 H3 hemagglutinin sequences deposited between 1968 and 2009 in the NCBI database[44] and are listed in Table 2. Also using the $\Delta\Delta G$ values in Table 1, we calculated and tabulated in Table 3 for each site i the value of $\langle\Delta\Delta G\rangle_i$, which is the average $\Delta\Delta G$ weighted by the probability for each different amino acid to be introduced, where the probability is proportional to the relative frequencies of 20 amino acids counted from the H3 sequences in the NCBI database from 1968 to 2009.

As shown in Table 3, there is obvious variation among the expected free energy differences $\langle\Delta\Delta G\rangle_i$ caused by single substitutions at amino acid site i of epitope B. This variation is partly due to the wildtype amino acids in the sites. For instance, the wildtype amino acid in site 190 is Glu, which has the highest rank in Table 2. As shown in Table 3, any amino acid substitution in site 190 tends to have a negative $\Delta\Delta G$. Another cause of variation in $\langle\Delta\Delta G\rangle_i$ is that distinct sites affect differently the antibody binding process. Epitope B of the wildtype A/Aichi/2/1968 hemagglutinin sequence contains five sites with threonine: 128, 155, 160, 187, and 192. The mathematical expectancies $\langle\Delta\Delta G\rangle_i$ in these five sites are −7.746, 4.471, 4.956, 1.182, and −1.737 kcal/mol, respectively. Therefore, each site in epitope B has a specific effect on the virus escape substitution. A random substitution in epitope B affects the antibody binding free energy differently depending on the site and the substituting amino acids.

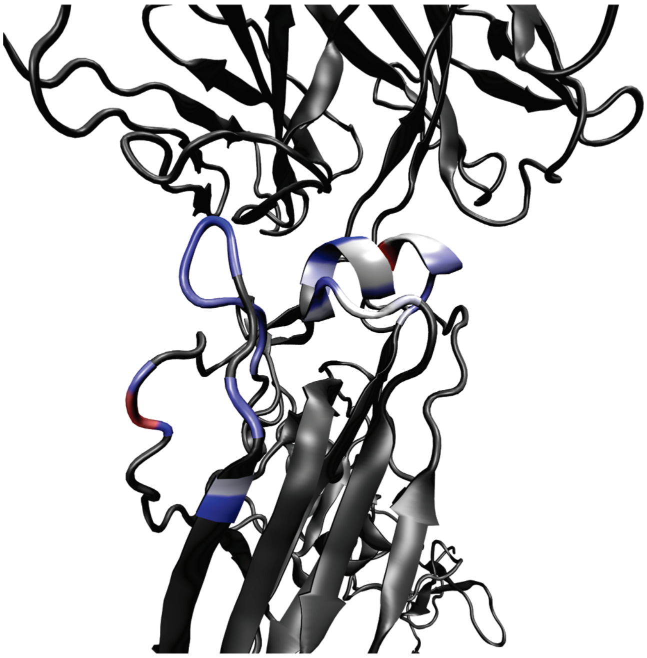


Figure 2. The tertiary structure of the interface between the HA1 domain of H3 hemagglutinin monomer A/Aichi/2/1968 (bottom) and the antibody HC63 (top) (PDB code: 1KEN). Water molecules are not shown. Epitope B of the HA1 domain is located in two loops and one α-helix, with the color scale modulated according to the expected free energy difference $\langle\Delta\Delta G\rangle_i$ of each site i in epitope B. The color scale ranges from red for the most negative $\langle\Delta\Delta G\rangle_i$ values to blue for the most positive $\langle\Delta\Delta G\rangle_i$ values. The sites i in epitope B with $\langle\Delta\Delta G\rangle_i$ near zero are colored white. The region outside epitope B is colored gray. The red site 128 is far from the antibody binding region, and the red site 190 possessed the original amino acid Glu, which is a charged amino acid. It may explain why these two sites show negative $\langle\Delta\Delta G\rangle_i$ with large absolute values.

The variation of $\langle\Delta\Delta G\rangle_i$ is also reflected by the tertiary structure of epitope B bound by the antibody. By looking into the structure of epitope B, shown in Figure 2, it is seen that epitope B resides in two protruding loops from amino acid site 128 to 129 and from site 155 to 165 and in an α-helix from site 186 to 198. Site 128 has a negative average free energy difference, $\langle\Delta\Delta G\rangle_{128} = -7.746 \pm 0.098$ kcal/mol. All of the other sites in these two loops show a positive $\langle\Delta\Delta G\rangle_i$ value of a random substitution, with the minimum $\langle\Delta\Delta G\rangle_{157} = 3.944 \pm 0.090$ kcal/mol in site 157. The α-helix is located between the hemagglutinin and antibody. In the α-helix,

Table 4. Substitutions Occurred in Epitope B of the Hemagglutinin A/Aichi/2/1968 (H3N2) as of 1975[a]

substitution	year	$\Delta\Delta G$ (kcal/mol)	rank (substituting)	rank (WT)
T128N	1971	-4.796 ± 0.361	8	7
T128I	1975	-16.026 ± 0.412	18	7
G129E	1970, 1972	10.500 ± 0.415	4	17
T155Y	1972−1973, fixed in 1973	7.254 ± 0.358	9	14
G158E	1971−1972	8.584 ± 0.479	6	17
S159N	1971, 1974−1975	10.969 ± 0.352	5	17
S159C	1972	7.923 ± 0.324	6	17
S159R	1972	7.065 ± 0.424	7	17
T160A	1973	4.160 ± 0.217	11	18
S186N	1975	4.673 ± 0.298	10	14
N188D	1971−1973, fixed in 1973	19.767 ± 0.367	1	14
Q189K	1975	9.484 ± 0.640	2	10
E190V	1972	-9.115 ± 0.310	5	3
E190D	1975	18.752 ± 0.324	1	3
S193N	1972−1975	8.239 ± 0.301	10	12
S193D	1975	15.285 ± 0.294	7	12
A198T	1972	6.793 ± 0.236	3	14

[a] Also listed are the years when the substitutions were observed, and the free energy differences with standard errors. In each site of epitope B, all 20 amino acids were sorted in descending order by the free energy differences introduced by a substitution from the wildtype amino acid to 20 amino acids. The ranks of the substituting amino acid and the wildtype amino acid in each substituted site are listed in the columns rank (substituting) and rank (WT), respectively.

the sites facing toward the antibody usually present large positive $\langle\Delta\Delta G\rangle_i$ values such as sites 193 and 196, while the sites facing toward the hemagglutinin show lower $\langle\Delta\Delta G\rangle_i$ values such as sites 189, 192, and 197. Thus, in the one-dimensional sequence from site 186 to 198, the $\langle\Delta\Delta G\rangle_i$ values oscillate with peaks and valleys corresponding to the sites in the α-helix alternatingly facing the antibody and hemagglutinin. Consequently, the variation of the expected free energy changes in distinct sites depends on the structure of the hemagglutinin−antibody complex.

3.2. Historical Substitutions in Epitope B. The simulation results are supported in two ways by amino acid sequence data of H3 hemagglutinin collected since 1968. These historical sequences are downloaded from the NCBI Influenza Virus Resource[45] and aligned. First, Pan et al. analyzed the number of charged amino acids in epitope B of H3 hemagglutinin in each year since 1968 and found an increasing trend of charged amino acids.[42] This finding supports the results that amino acid substitution introducing charged residues on average facilitates virus escape from antibodies, as illustrated in Table 2. Second, amino acid substitutions in epitope B between 1968 and 1975 also verified the free energy calculation, as shown below.

With the knowledge of the free energy landscape of the single substitutions, we are able to recognize favorable single substitutions in epitope B. Substitutions with large positive $\Delta\Delta G$ values enable the virus to evade the immune pressure and increase the virus fitness. Favorable substitutions grow in the virus population. Selection for substitutions with large $\Delta\Delta G$ is part of the evolutionary strategy of the virus. The results of free energy calculation can also explain the substituted virus strains collected in history.

We analyzed the hemagglutinin sequence information of H3N2 strains evolving from the A/Aichi/2/1968 strains. H3 hemagglutinin circulating from 1968 to 1971 was mainly in the HK68 antigenic cluster, while those circulating from 1972 to 1975 were mainly in the EN72 antigenic cluster.[46] Table 4 shows that in the dominant epitope B, 17 substitutions occurred in 12 sites collected between 1968 to 1975,[47] which contributed to immune evasion and corresponding virus evolution from the HK68 cluster to the EN72 cluster. Also listed in Table 4 are the free energy differences of these historical substitutions. The 17 substituting amino acids have significantly higher ranks compared to the corresponding wildtype amino acids (p = 0.0044, Wilcoxon signed-rank test). This significant difference is expected because 15 of 17 substituting amino acids have ranks between 1 and 10, while 10 of 12 wildtype amino acids in the substituted site have ranks between 11 and 20. In all of the 21 sites in epitope B, 15 of 21 wildtype amino acids have ranks between 11 and 20. Additionally, the $\Delta\Delta G$ values of these 17 substitutions listed in Table 4 are greater than the expected free energy differences $\langle\Delta\Delta G\rangle_i$ in Table 3 of random substitutions in the 12 substituted sites (p = 0.013, Wilcoxon signed-rank test).

We also looked into the historical escape substitutions in epitope B evading the immune pressure of the vaccine strains. For each influenza season, the amino acids in the administered vaccine strain were defined as the wildtype ones and those in the dominant circulating strain as the substituting amino acids. In each of the 19 seasons from 1971 to 2004 in which the H3N2 virus was the dominant subtype, the substitutions in epitope B were located,[3] and their $\Delta\Delta G$ values were obtained from Table 1. As shown in Table 5, escape substitutions in epitope B as of 1973 mostly had positive $\Delta\Delta G$ values and generated substituting amino acids with increased rank (p = 0.047, Wilcoxon signed-rank test). Such a tendency to introduce amino acids with higher ranks was not observed after 1973: the ranks of wildtype and substituting amino acids after 1973 present little significant difference (p = 0.28, Wilcoxon signed-rank test). The hemagglutinin of A/Aichi/2/1968 used in the free energy calculating is in the HK68 antigenic cluster. Perhaps after the virus evolved into the next EN72 cluster, a change in the virus antigenic character stimulated the immune system to produce new types of antibodies other than the HC63 antibody used in the calculation. A different binding antibody changes the free energy landscape of the substitutions in epitope B. Thus, the application of the

Table 5. Substitutions Occurring in Epitope B of H3 Hemagglutinin between the Vaccine Strain and the Dominant Circulating Strain in Each Season in Which the H3N2 Subtype Was Dominant[a]

year	substitution	$\Delta\Delta G$ (kcal/mol)	rank (vaccine)	rank (circulating)
1972	T155Y	7.254 ± 0.358	14	9
1972	G158E	8.584 ± 0.479	17	6
1972	S159C	7.923 ± 0.324	17	6
1972	E190V	−9.115 ± 0.310	3	5
1973	T160A	4.160 ± 0.217	18	11
1973	N188D	19.767 ± 0.367	14	1
1973	S193N	8.239 ± 0.301	12	10
1975	S157L	−6.256 ± 0.394	15	19
1975	A160T	−4.160 ± 0.217	11	18
1975	Q189K	9.484 ± 0.640	10	2
1975	N193D	7.046 ± 0.317	10	7
1984	E156K	−26.536 ± 0.429	1	15
1984	V163A	−0.243 ± 0.217	15	16
1984	D190E	−18.752 ± 0.324	1	3
1984	I196V	−0.768 ± 0.343	16	18
1987	Y155H	−4.782 ± 0.414	9	11
1987	E188D	9.669 ± 0.382	3	1
1987	K189R	−9.872 ± 0.697	2	11
1996	V190D	27.867 ± 0.299	5	1
1996	L194I	−6.914 ± 0.324	13	17
1997	K156Q	13.140 ± 0.413	15	3
1997	E158K	−10.187 ± 0.515	6	18
1997	V190D	27.867 ± 0.299	5	1
1997	L194I	−6.914 ± 0.324	13	17
1997	V196A	5.947 ± 0.229	18	11
2003	H155T	−2.472 ± 0.355	11	14
2003	Q156H	−20.028 ± 0.365	3	20
2003	S186G	0.132 ± 0.275	14	13

[a] The free energy difference with standard error of each substitution is obtained using the free energy landscape in Table 1. The ranks of free energy differences sorted in descending order are listed in column rank (vaccine) and in column rank (circulating) for the amino acids in the vaccine strain and the dominant circulating strain, respectively.

present free energy landscape should be limited within the HK68 and EN72 clusters. Free energy differences of substitutions in the EN72 cluster would need to be calculated using the updated antibody crystal structure.

4. DISCUSSION

4.1. Fitness of the Virus Strains. The free energy landscape shown in Table 1 gives the change of the antibody binding affinity, $K_1/K_0 = \exp(-\Delta\Delta G/RT)$, induced by each possible substitution in epitope B of the wildtype hemagglutinin. The majority of the substitutions lead to positive $\Delta\Delta G$ and yield a reduced binding affinity K_1 that is smaller than the binding affinity of the original mature antibody K_0. A decreased antibody binding constant grants the virus a higher chance of evading immune pressure and infecting host cells. We propose that virus fitness is positively correlated with the free energy difference $\Delta\Delta G$. The other factor affecting virus fitness is the capability of the hemagglutinin to maintain the normal biochemical functions, such as virus entry. Most sites in epitope B changed amino acid identities during 1968 to 2005 as the H3N2 virus kept circulating.[47] We therefore postulate that the substitutions in epitope B do not greatly interfere with the biochemical function of hemagglutinin, and virus fitness is dominantly determined by the free energy difference resulted from substitutions in epitope B.

The binding constant between the hemagglutinin and antibody after the first round of maturation is about 10^6 M^{-1}, and the binding constant of an uncorrelated antibody is below 10^2 M^{-1}.[2] On average, four substitutions in epitope B change the substituted hemagglutinin sufficiently so that the immune response of the original antibody binding to epitope B is abrogated.[3] Since this is a reduction of the binding constant from roughly 10^6 M^{-1} to 10^2 M^{-1}, one amino acid substitution that contributes to immune escape causes on average a 10-fold decrease in antibody binding constant, or equivalently $\Delta\Delta G_{\text{crit}}$ = 1.42 kcal/mol at 310 K. Assuming the effect of immune evasion can be broken into the sum of individual amino acid substitutions in the dominant epitope,[3] we define the virus fitness w as the sum of the contribution in each site of epitope B:

$$w = A_0 + \sum_{\text{epitope B}} \delta w_i \tag{27}$$

We denote by $\Delta\Delta G_i^{\alpha\gamma}$ the free energy difference when substituting amino acid α for amino acid γ at site i. We investigated two versions of the virus fitness landscape. The first defines δw_i as a linear function of the free energy difference of the substitution

$$\delta w_i = A_1 \frac{\Delta\Delta G_i^{\alpha\gamma}}{\Delta\Delta G_{\text{crit}}} \tag{28}$$

The second defines δw_i as a step function

$$\delta w_i = A_2 H(\Delta\Delta G_i^{\alpha\gamma} - \Delta\Delta G_{\text{crit}}) \tag{29}$$

in which H is the Heaviside step function. As illustrated in the simulation below, either definition of fitness is sufficient to explain the observed immune evasion of the H3N2 virus.

4.2. Selection in the Epitope. Evolution of the H3N2 virus is driven jointly by neutral evolution and selection.[48] Neutral evolution may be ongoing in sites outside the epitopes. The high substitution rate in epitope B suggests that selection is the major factor shaping the pattern of evolution in that epitope.[47] Shown in Tables 4 and 5 are the historical substitutions. The significantly increased ranks of free energy differences suggests the existence of selection by the immune pressure for substitutions that have increased the free energy difference $\Delta\Delta G$ and decreased the antibody binding constant. The immune selection is directional: certain types of amino acids such as charged ones were initially more likely to be added into epitope B[42] because they maximally decreased the antibody binding constant, as indicated in Table 2. The heterogeneity of the expected free energy difference of a random substitution in Table 3 shows that each site in epitope B has a specific weight with regard to immune escape.

Table 4 also illustrates that the immune selection did not necessarily pick the amino acid with the highest rank of $\Delta\Delta G$ as the substituting amino acid. Amino acids with moderate rank were introduced into epitope B even for the fixed substitution T155Y. Therefore, the historical evolution did not simply substitute amino acids by maximizing the free energy differences in Table 1. This phenomenon is possibly due to two causes. First, the virus fitness may be insensitive to the $\Delta\Delta G$ values; e.g., A_1 in eq 28 may be small, or amino acid substitutions with large

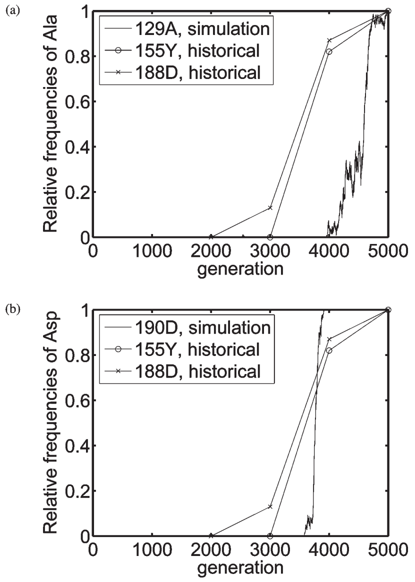


Figure 3. Two fixed substitutions G129A and E190D generated by Monte Carlo simulation of epitope B using eq 28. Also plotted are two historical fixed substitutions in epitope B: T155Y fixed between 1971 and 1973 and N188D fixed between 1970 and 1973. The frequency data of historical substitutions are from Shih et al.[47] The origin of time axis is 1968. 1000 generations of the H3N2 virus is approximately 1 year. (a) Substitution G129A causing the free energy difference $\Delta\Delta G$ = 3.33 ± 0.29 kcal/mol is fixed by the simulation. The rank of the free energy difference of G129A is 12 in 19 possible substitutions in site 129. (b) Substitution E190D with $\Delta\Delta G$ = 18.75 ± 0.32 kcal/mol. The rank is 1 in 19 possible substitutions in site 190.

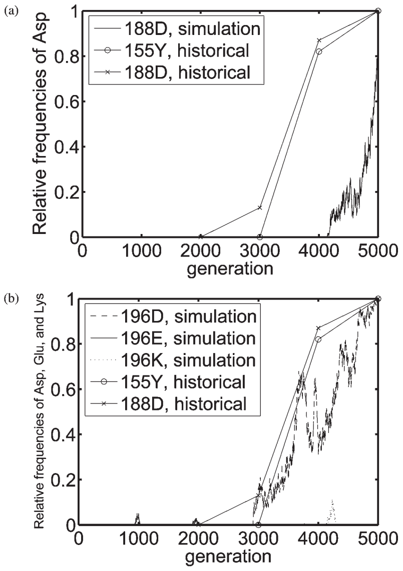


Figure 4. Two fixed substitutions N188D and V196D generated by Monte Carlo simulation of epitope B using eq 29. Two historical fixed substitutions T155Y and N188D are also plotted, and data are from Shih et al.[47] (a) Substitution N188D causing the free energy difference $\Delta\Delta G$ = 19.77 ± 0.37 kcal/mol is fixed by the simulation. The rank of the free energy difference of N188D is 1 in 19 possible substitutions in site 188. (b) Substitution V196D with $\Delta\Delta G$ = 9.25 ± 0.34 kcal/mol. The rank is 5 in 19 possible substitutions in site 196. The proportions of substituting amino acids are represented by different line types.

$\Delta\Delta G$ values may contribute equivalently to the fitness, as in eq 29. Second, only a small fraction of virus in one host is shed by the host and infects the next host, so the population size of the propagated virus from one host is smaller by several orders of magnitude than the total virus population size in the same host. Additionally, a seasonal bottleneck exists in the influenza virus circulation.[49] Both random mutation and small population sizes lead to dramatic randomness in the evolution. Consequently, the evolution of H3 hemagglutinin is not solely determined by maximizing the free energy differences in Table 1 and minimizing the antibody binding constant, even if the virus is under immune selection. Instead, randomness plays a key role in the H3N2 virus evolution.

4.3. A Picture of the H3N2 Virus Evolution. Selection depends on the fitness of each virus genotype that is quantified as a nondecreasing function of the free energy difference $\Delta\Delta G$. Moderate selection in epitope B requires that fitness improvement is limited when $\Delta\Delta G$ is large. One possibility is that the ratio A_1/A_0 in eq 28 is small. Another is that the fitness takes the form of eq 29 in which all substitutions with $\Delta\Delta G > \Delta\Delta G_{\mathrm{crit}}$ have equal fitness.

The virus evolution is also affected by the genetic drift. Genetic drift is a term which captures the random component of evolution due to the large size of the phase space of possible substitutions relative to the single set of substitutions that lead to the highest viral fitness. The effect of genetic drift is quantitatively reflected in the fixation process of a new strain, as shown in the simulation below. A narrow bottleneck of virus propagation allows only a small fraction of the progeny to survive, imposing a notable probability that a favorable substitution is lost in the next generation. The effect of genetic drift is to increase the randomness in the virus evolution so that observed substitutions are based on chance in addition to the fitness of these substitutions.

To model the H3N2 evolution discussed above, we ran two Monte Carlo simulations of the influenza evolution model. A population of N sequences of epitope B with 21 sites was created and initialized as the wildtype A/Aichi/2/1968 sequence. Here, $N = 10^3$ to account for a narrow genetic bottleneck of hemagglutinin and for tractability of the simulation. We iterated the simulation program for 5000 generations or about five years to recreate a pattern of evolution similar to that in history and shown in Table 4. The random substitution rate of H3 hemagglutinin is roughly 4.5×10^{-6} amino acid substitution/site/generation.[50] We let the number of substitutions follow a Poisson distribution with mean $\lambda = 21 \times 4.5 \times 10^{-6}N = 9.5 \times 10^{-5}N$ and randomly assigned the substitution sites. The substituting amino acid at each substitution site was randomly picked from the remaining 19 amino acids proportional to the historical frequencies observed in hemagglutinin. The fitness w in the first simulation was calculated for each sequence using eq 28 with $A_0 = 100$ and $A_1 = 3$, and that in the second simulation was calculated for each sequence using eq 29 with $A_0 = 100$, $A_2 = 9$, and $\Delta\Delta G_{\mathrm{crit}}$ = 1.42 kcal/mol. Note that by choosing $A_1 = 3$ for the first simulation, a random substitution causes the expected fitness to change from 100 to 104.9, and by choosing $A_2 = 9$ for the second simulation, a random substitution changes the expected fitness from 100 to 105.0. The size of the progeny of each sequence equals the fitness w of the sequence if $w > 0$ and equals 0 if $w \leq 0$. The next generation of sequences was initialized by randomly sampling N sequences from the progeny sequences.

The results of both simulations showed remarkable similarity to the observed substitutions in Table 4 with the bottleneck N equal to 10^3, see Figures 3 and 4. Amino acid substitutions generated in the simulation are usually distinct from those in Table 4, observed in history. The $\Delta\Delta G$ values of each substitution emerging in the simulation are nevertheless similar to those of the historical substitutions listed in Table 4. As was observed in history, in Table 4, most of the substituted strains in the simulations with a relative frequency greater than 1% have positive $\Delta\Delta G$ values with the ranks of the substituting amino acids ranging from 1 to 10. The fixation of a newly emerged substitution takes about 1000 generations or one year on average. Fixed substitutions mostly introduce amino acids with positive $\Delta\Delta G$ values in Table 1 and higher ranks in Table 2, and several of these fixed substitutions in simulation, such as E190D and N188D, have the highest $\Delta\Delta G$ values in the current site. However, fixed substitutions in the simulation are not always the substitutions with the highest $\Delta\Delta G$ values in Table 1. These observations suggest that the Monte Carlo simulation considering the effect of substitution, selection, and genetic drift is able to reproduce the pattern of evolution observed in history. This simulation also shows that besides the free energy difference of each substitution, the mapping from the free energy landscape to the fitness landscape as well as the random genetic drift are dominant factors of the evolution in virus epitopes.

Shown in Figures 3 and 4 for both simulations are the trajectories of relative frequencies of substituting amino acids. The trajectories are similar to historical observations of the human H3N2 virus data.[47] For influenza, 1000 generations roughly equal 1 year. The two substitutions T155Y and N188D were fixed in epitope B in 1968−1973. As indicated by Figures 3 and 4, substitution T155Y emerged between generations 3000 and 4000 or, equivalently, between 1971 and 1972 from the emergence of the H3N2 virus in 1968.[47] Substitution T155Y was fixed between generations 4000 and 5000. Similarly, substitution N188D emerged between generations 2000 and 3000 and was fixed between generations 4000 and 5000. The first simulation in which virus fitness is calculated using eq 28 generated two fixed substitutions, G129A, which emerged at generation 4000 and was fixed by generation 5000, and E190D, which emerged at generation 3600 and was fixed by generation 3900. The second simulation using eq 29 generated one fixed substitution, V196D emerging at generation 2900 and fixed by generation 5000, and one substitution that was nearly fixed, N188D, emerging at generation 4100 and acquiring the relative frequency 0.84 at generation 5000. The trajectories in both simulations resemble those of substitutions T155Y and N188D observed in history. From these results, the two Monte Carlo simulations appear to capture the main factors of immune selection and genetic drift in evolution of the H3N2 virus.

4.4. Multiple Substitutions. In this work, we calculated the free energy difference for each possible substitution in epitope B. The free energy calculation for multiple substitutions is intractable using the current technology due to the combinatorial increase in calculation load for multiple substitutions. The issue of multiple substitutions is here addressed by assuming that the effect of immune evasion is well represented by the sum of the contribution in each substituted site in epitope B. The data indicate the independence of the immune evasion effect of the sites in epitope B.[3] We may, thus, assume that the free energy difference of the multiple substitution is the sum of the individual $\Delta\Delta G$ values available in Table 1 plus a minor correction term.

4.5. Prediction of Future Virus Evolution. The result of this work quantifies the reduction of the binding constant of an antibody to a virus for substitutions in epitope B with larger $\Delta\Delta G$ values and higher ranks of substituting amino acids. A newly emerging virus strain with a larger antibody binding free energy difference has a greater probability to become the dominant strain in the next flu season. Note that due to random fluctuations in the large phase space of possible substitutions, actual trajectories deviate from the trajectory determined by choosing sites and substituting amino acids with the greatest free energy differences. With a three-dimensional structure of hemagglutinin of the current circulating virus and binding antibody, one is able to calculate the free energy landscape for all of the possible single substitutions in the dominant epitope and estimate the *a priori* escape probabilities in the next season. The dominant circulating influenza strain usually possesses amino acid substitutions from the vaccine strain against which memory antibodies are generated. Usually these substitutions disrupt the antibody binding process by decreasing the binding constant, as shown in Table 5. Thus, one can predict vaccine effectiveness by evaluating the antibody binding constant against the dominant circulating strain, which is acquired by calculating the free energy difference of the amino acid substitutions between the vaccine strain and the dominant circulating strain.[3] More accurate predictions of the evolutionary pattern of a virus as well as epidemiological data such as vaccine effectiveness may be obtained by optimally mapping the free energy landscape to the fitness landscape and taking into account random factors such as genetic drift in the evolution process.

5. CONCLUSION

We introduced the Einstein crystal as a technology to improve the results of free energy calculation. By calculating the free energy difference of each amino acid substitution, we obtained the free energy landscape for substitutions in epitope B of hemagglutinin. There is notable variation between the values of free energy differences of different substitutions at different sites, because the identities of original and substituting amino acids, as well as the locations of amino acid substitutions, affect to differing degrees the antibody binding process. In this free energy landscape, we suggest that a virus tends to evolve to higher $\Delta\Delta G$ values to escape binding of an antibody. Counterbalancing this selection is random drift. Historical amino acid substitutions in epitope B and Monte Carlo simulations of the virus evolution using the free energy based virus fitness, in which random genetic drift of the virus adds statistical noise into the virus evolution process, showed that selected substitutions are biased to those with positive $\Delta\Delta G$ values.

■ AUTHOR INFORMATION

Corresponding Author

*Fax: 713-348-5811. E-mail: mwdeem@rice.edu.

■ ACKNOWLEDGMENT

This research was supported in part by the National Science Foundation through TeraGrid resources provided by Purdue and Indiana Universities under grant number TG-MCA05S015. We specifically acknowledge the assistance of Phillip Cheeseman. Part of the calculation was performed at computing resources at Research Computing Support Group, Rice University. Keyao Pan's research was supported by a training fellowship from the Keck Center Nanobiology Training Program of the Gulf Coast

Consortia (NIH Grant No. R90 DK071504). This project was also partially supported by DARPA grant HR 00110910055.

■ REFERENCES

(1) World Health Organization Media Centre Influenza Fact Sheet 211. http://www.who.int/mediacentre/factsheets/fs211/en/index.html (accessed on August 10, 2010).

(2) Janeway, C.; Travers, P.; Walport, M.; Shlomchik, M. *Immunobiology: The immune system in health and disease*, 6th ed.; Garland Science: New York, 2005; p 430.

(3) Gupta, V.; Earl, D. J.; Deem, M. W. *Vaccine* **2006**, *24*, 3881–3888.

(4) Lee, L. Y. H.; Ha, D. L. A.; Simmons, C.; de Jong, M. D.; Chau, N. V. V.; Schumacher, R.; Peng, Y .C.; McMichael, A. J.; Farrar, J. J.; Smith, G. L.; Townsend, A. R. M.; Askonas, B. A.; Rowland-Jones, S.; Dong, T. *J. Clin. Invest.* **2008**, *118*, 3478–3490.

(5) Pan, K.; Subieta, K. C.; Deem, M. W. *Protein Eng., Des. Sel.* **2011**, *24*, 291–299.

(6) Pan, K.; Deem, M. W. *Vaccine* **2009**, *27*, 5033–5034.

(7) Ferguson, N. M.; Galvani, A. P.; Bush, R. M. *Nature* **2003**, *422*, 428–433.

(8) Deem, M. W.; Pan, K. *Protein Eng., Des. Sel.* **2009**, *22*, 543–546.

(9) Zhou, R. H.; Das, P.; Royyuru, A. K. *J. Phys. Chem. B* **2008**, *112*, 15813–15820.

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

(11) Frenkel, D.; Smit, B. *Understanding molecular simulation: from algorithms to applications*, 2nd ed.; Academic Press: San Diego, CA, 2002; p 168.

(12) Beveridge, D. L.; DiCapua, F. M. *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 431–492.

(13) Mezei, M.; Beveridge, D. L. *Ann. N.Y. Acad. Sci.* **1986**, *482*, 1–23.

(14) Cross, A. J. *Ann. N.Y. Acad. Sci.* **1986**, *482*, 89–90.

(15) Beutler, T. C.; Mark, A. E.; Vanschaik, R. C.; Gerber, P. R.; Vangunsteren, W. F. *Chem. Phys. Lett.* **1994**, *222*, 529–539.

(16) Zacharias, M.; Straatsma, T. P.; McCammon, J. A. *J. Chem. Phys.* **1994**, *100*, 9025–9031.

(17) Boresch, S.; Karplus, M. *J. Phys. Chem. A* **1999**, *103*, 103–118.

(18) Boresch, S.; Karplus, M. *J. Phys. Chem. A* **1999**, *103*, 119–136.

(19) Roux, B. *Biophys. J.* **1996**, *71*, 3177–3185.

(20) Nina, M.; Beglov, D.; Roux, B. *J. Phys. Chem. B* **1997**, *101*, 5239–5248.

(21) Essex, J. W.; Severance, D. L.; TiradoRives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **1997**, *101*, 9663–9669.

(22) Price, D. J.; Jorgensen, W. L. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 681–695.

(23) Zacharias, M.; Straatsma, T. P.; Mccammon, J. A.; Quiocho, F. A. *Biochemistry* **1993**, *32*, 7428–7434.

(24) Kaliman, I.; Nemukhin, A.; Varfolomeev, S. *J. Chem. Theory Comput.* **2010**, *6*, 184–189.

(25) Crespo, A.; Marti, M. A.; Estrin, D. A.; Roitberg, A. E. *J. Am. Chem. Soc.* **2005**, *127*, 6940–6941.

(26) Takahashi, H.; Kawashima, Y.; Nitta, T.; Matubayasi, N. *J. Chem. Phys.* **2005**, *123*, 124504.

(27) Wang, S. L.; Hu, P.; Zhang, Y. K. *J. Phys. Chem. B* **2007**, *111*, 3758–3764.

(28) Deng, Y Q.; Roux, B. *J. Chem. Theory Comput.* **2006**, *2*, 1255–1273.

(29) Frenkel, D.; Ladd, A. J. C. *J. Chem. Phys.* **1984**, *81*, 3188–3193.

(30) Noya, E. G.; Conde, M. M.; Vega, C. *J. Chem. Phys.* **2008**, *129*, 104704.

(31) Frenkel, D.; Smit, B. *Understanding molecular simulation: from algorithms to applications*, 2nd ed.; Academic Press: San Diego, CA, 2002; p 248.

(32) Meijer, E. J.; Frenkel, D.; Lesar, R. A.; Ladd, A. J. C. *J. Chem. Phys.* **1990**, *92*, 7570–7575.

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

(34) Bennett, C. H. *J. Comput. Phys.* **1975**, *19*, 267–279.

(35) Pomes, R.; McCammon, J. A. *Chem. Phys. Lett.* **1990**, *166*, 425–428.

(36) Feenstra, K. A.; Hess, B.; Berendsen, H. J. C. *J. Comput. Chem.* **1999**, *20*, 786–798.

(37) Rao, S. N.; Singh, U. C.; Bash, P. A.; Kollman, P. A. *Nature* **1987**, *328*, 551–554.

(38) Flyvbjerg, H.; Petersen, H. G. *J. Chem. Phys.* **1989**, *91*, 461–466.

(39) Morgan, B. R.; Massi, F. *J. Chem. Theory Comput.* **2010**, *6*, 1884–1893.

(40) Hünenberger, P. H.; McCammon, J. A. *J. Chem. Phys.* **1999**, *110*, 1856–1872.

(41) Figueirido, F.; Delbuono, G. S.; Levy, R. M. *J. Chem. Phys.* **1995**, *103*, 6133–6142.

(42) Pan, K.; Long, J.; Sun, H.; Tobin, G. J.; Nara, P. L.; Deem, M. W. *J Mol Evol* **2011**, 72, 90–103.

(43) Sayle, R. A.; Milnerwhite, E. J. *Trends Biochem. Sci.* **1995**, *20*, 374–376.

(44) Pan, K.; Deem, M. W. An Entropy Method to Quantifying Selection and Diversity in Viruses by Entropy Methods, with Application to the Hemagglutinin of H3N2 Influenza. *J. R. Soc. Interface*, in press.

(45) NCBI Influenza Virus Resource. http://www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html (accessed on August 10, 2010).

(46) Smith, D. J.; Lapedes, A. S.; de Jong, J. C.; Bestebroer, T. M.; Rimmelzwaan, G. F.; Osterhaus, A. D. M. E.; Fouchier, R. A. M. *Science* **2004**, *305*, 371–376.

(47) Shih, A. C.; Hsiao, T. C.; Ho, M. S.; Li, W. H. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 6283–6288.

(48) Koelle, K.; Cobey, S.; Grenfell, B.; Pascual, M. *Science* **2006**, *314*, 1898–1903.

(49) Rambaut, A.; Pybus, O. G.; Nelson, M. I.; Viboud, C.; Taubenberger, J. K.; Holmes, E. C. *Nature* **2008**, *453*, 615–U2.

(50) Nobusawa, E.; Sato, K. *J. Virol.* **2006**, *80*, 3675–3678. In the amino acid level, the average mutation rate of the influenza A virus is converted to 4.5×10^{-6} amino acid substitution/site/generation.

JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Path-Integral Calculations of Nuclear Quantum Effects in Model Systems, Small Molecules, and Enzymes via Gradient-Based Forward Corrector Algorithms

Asaf Azuri, Hamutal Engel, Dvir Doron, and Dan Thomas Major*

Department of Chemistry and the Lise Meitner-Minerva Center of Computational Quantum Chemistry, Bar-Ilan University, Ramat-Gan 52900, Israel

S *Supporting Information*

ABSTRACT: A practical approach to treat nuclear quantum mechanical (QM) effects in simulations of condensed phases, such as enzymes, is via Feynman path integral (PI) formulations. Typically, the standard primitive approximation (PA) is employed in enzymatic PI simulations. Nonetheless, these PI simulations are computationally demanding due to the large number of discretizations, or *beads*, required to obtain converged results. The efficiency of PI simulations may be greatly improved if higher order factorizations of the density matrix operator are employed. Herein, we compare the results of model calculations obtained employing the standard PA, the improved operator of Takahashi and Imada (TI), and several gradient-based forward corrector algorithms due to Chin (CH). The quantum partition function is computed for the harmonic oscillator, Morse, symmetric, and asymmetric double well potentials. These potentials are simple models for nuclear quantum effects, such as zero-point energy and tunneling. It is shown that a unique set of CH parameters may be employed for a variety of systems. Additionally, the nuclear QM effects of a water molecule, treated with density functional theory, are computed. Finally, we derive a practical perturbation expression for efficient computation of isotope effects in chemical systems using the staging algorithm. This new isotope effect approach is tested in conjunction with the PA, TI, and CH methods to compute the equilibrium isotope effect in the Schiff base-oxyanion keto−enol tautomerism in the cofactor pyridoxal-5′-phosphate in the enzyme alanine racemase. The study of the different factorization methods reveals that the higher-order actions converge substantially faster than the PA approach, at a moderate computational cost.

1. INTRODUCTION

Enzymes are remarkably efficient catalysts evolved to perform well-defined and highly specific chemical transformations.[1] Studying the nature of enzymatic rate enhancements is highly important from several aspects, including the rational design of synthetic catalysts and transition-state (TS) inhibitors. Isotope effects (IE) and particular equilibrium isotope effect (EIE) and kinetic isotope effect (KIE) are important tools in elucidating reaction mechanisms in enzymes.[2] The KIE is a fundamental phenomenon measuring the sensitivity of chemical reaction rates on isotopic substitutions and provides the most direct probe to the structure of the TS of the reaction. Moreover, KIE might provide insights into tunneling in enzymes.[2] EIE is an invaluable tool for insight into chemical reaction equilibrium, enzymatic binding, and hydrogen bonding.[3,4] The EIE is defined as

$$\mathrm{EIE} = \frac{K_{\mathrm{L}}}{K_{\mathrm{H}}} = \frac{Q_{\mathrm{L}}^{\mathrm{PS}}/Q_{\mathrm{L}}^{\mathrm{RS}}}{Q_{\mathrm{H}}^{\mathrm{PS}}/Q_{\mathrm{H}}^{\mathrm{RS}}} = e^{-\beta(\Delta G_{\mathrm{L}}^{\mathrm{r}} - \Delta G_{\mathrm{H}}^{\mathrm{r}})} \tag{1}$$

where Q is the partition function for the reactant state (RS) and product state (PS) for the light (L) and heavy (H) isotopes, and ΔG^{r} is the reaction free energy. $\beta = 1/k_{\mathrm{B}}T$ with k_{B} being Boltzmann's constant and T the temperature. Similarly, the KIE is defined as

$$\mathrm{KIE} = \frac{k^{\mathrm{L}}}{k^{\mathrm{H}}} \approx e^{-\beta(\Delta G_{\mathrm{L}}^{\neq} - \Delta G_{\mathrm{H}}^{\neq})} \tag{2}$$

where k is the rate constant, and $\Delta G^{\neq}$ is the free energy barrier. In quantum transition-state theory (QTST),[5] the exact rate constant is expressed by the QTST rate constant, k_{QTST}, multiplied by a transmission coefficient γ_q:

$$k = \gamma_q \cdot k_{\mathrm{QTST}} \tag{3}$$

where the QTST rate constant is given by

$$k_{\mathrm{QTST}} = \frac{1}{h\beta Q_{\mathrm{R}}} e^{-\beta G(z^{\neq})} \tag{4}$$

where h is Planck's constant. In the following, we assume that $\gamma_q = 1$. In eq 4, $G(z)$ is the free energy as a function of the centroid reaction coordinate $z[\bar{x}]$, $z^{\neq}$ is the value of $z[\bar{x}]$ at the free energy maximum. Specifically,

$$G(z^{\neq}) = -\frac{1}{\beta} \ln \left[\frac{Q^{\neq}}{(m/2\pi\hbar^2\beta)^{1/2}} \right] \tag{5}$$

where $Q^{\neq}$ is the reduced quantum phase space density at the dividing hyper-surface at $z^{\neq}$, $\hbar = h/2\pi$, and m is the mass.

The computational prediction of IE in enzymatic reactions presents a considerable challenge. First, classical statistic mechanics simulations cannot reproduce the observed KIE since

Received: December 13, 2010
Published: April 01, 2011

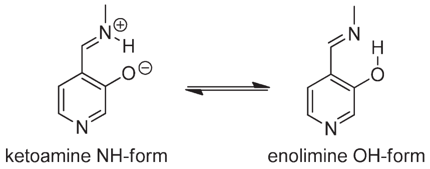
they ignore nuclear quantum mechanical (QM) effects (NQE), such as zero-point energy and tunneling. Thus, computationally expensive quantum dynamics simulations are required. Second, condensed phase simulations require extensive sampling of both solute and solvent degrees of freedom to obtain converged results, and simulations are prone to statistical noise. Finally, IE depends exponentially on the free energy differences between the light and heavy isotopes, making it a very difficult observable to predict. In particular, secondary and heavy atom KIE and EIE are small in magnitude and are extremely challenging to compute from condensed phase simulations.

Several simulation methods have been used to determine NQE in solution phase and enzymatic reactions. A practical approach to including these effects is via path-integral (PI) formulations which may be employed to calculate various properties of quantum or mixed quantum−classical systems.[6,7] Numerous examples of PI simulations of condensed phase reactions exist.[8−24] Additional approaches have been developed for condensed phase reactions, including the ensemble-averaged variational TS theory with multidimensional tunneling (EA-VTST/MT),[25,26] a wave function-based method,[27,28] and model reactions.[29] These methods have been applied to several enzymatic reactions with high-quality accord between the calculated and experimental KIEs. Recently, we developed a novel free energy mass-perturbation PI (PIFEP) method,[13,14] which has been successfully applied to numerous model and enzymatic reactions.[30−34] In particular, a combined PIFEP and EA-VTST/MT study has recently identified enhanced tunneling in the enzyme nitroalkane oxidase compared to the analogues uncatalyzed reaction.[34] Recently, additional approaches for computation of IE have been developed and applied to various chemical systems.[35,36] The modeling of IE, however, is computationally extremely demanding, and it is important to develop enhanced methods.

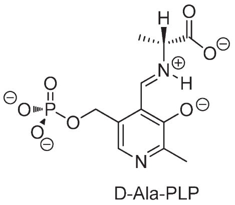
In PI simulations of enzymes the standard primitive approximation (PA) is typically employed. The PA is based on a primitive factorization of the canonical density operator and when combined with efficient sampling schemes, such as staging,[37] bisection,[38] or normal mode transformation,[39] yields satisfactory results. However, the application of these PA PI simulations to condensed phase reactions, employing fully QM or hybrid QM/molecular mechanical (MM) potential energy functions, is computationally demanding. The efficiency of PI simulations may be greatly improved if more accurate factorization schemes are employed. A higher-order PI approach was devised by Takahashi and Imada (TI) which may greatly enhance the efficiency of PI simulations.[40,41] Suzuki has formulated a higher-order composite factorization scheme which features an additional squared force term.[42,43] This strategy has been employed in studies targeting time-dependent classical dynamics,[44] real-time propagator,[45,46] solution of the Fokker−Planck equation,[47] and in PI simulations.[48−56]

Herein, we compare the results of model calculations obtained employing the standard PA, the TI,[40] and a novel higher-order factorization based on the symplectic algorithms developed by Chin.[46,57] These Chin-based factorizations have recently been employed in the study of quantum liquids.[58] In the current study, the quantum partition function is computed for several model one-dimensional systems, including the harmonic oscillator, Morse potential, and symmetric and asymmetric double wells. The computations employ two complementary methods free of sampling noise: the numerical matrix multiplication[59−61] and the direct matrix diagonalization methods.[62] Results emerging from this study on the different higher-order factorization methods reveal that the higher-order actions converge substantially faster than the PA approach, at a moderate computational cost. Moreover, we obtain a unique parametrization for the Chin factorization (CH) which is equally applicable to all the potentials employed herein. As a test case, we employ these higher-order Feynman PI formulations of the density matrix in conjunction with density functional theory (DFT) calculations to estimate the QM correction to vibrational free energy in a water molecule. Finally, we derive a perturbation expression for efficient computation of IEs in chemical systems using the staging algorithm. This new IE approach is tested in conjunction with the PA, TI, and CH methods to compute the EIE in the cofactor pyridoxal-5′-phosphate (PLP) Schiff base keto−enol tautomerism (Schemes 1 and 2) in the enzyme alanine racemase (AlaR).

Scheme 1. Ketoamine−Enolimine Tautomerism in a Model PLP System



Scheme 2. D-Ala PLP Employed in AlaR



2. THEORY

2.1. Basic Formal Expressions. The PI strategy is particularly suited for computing the quantum partition function Q in condensed phase systems since it may be obtained by applying classical simulation techniques.[63] The partition function is defined as the trace of the canonical density matrix:

$$Q = Tr(\rho) = \int \mathrm{d}x\rho(x, x; \beta) = \int \mathrm{d}x\langle x|e^{-\beta \mathrm{H}}|x\rangle \tag{6}$$

where x denotes the position of a particle in one dimension and extension to N dimensions is straightforward.

To express the partition function as a Feynman PI, we write the density matrix operator as a product of P exponents, each representing a time slice of length $\tau = \beta/P$:

$$Q = \int \mathrm{d}x\langle x|e^{-\tau \mathrm{H}}e^{-\tau \mathrm{H}}\cdots e^{-\tau \mathrm{H}}|x\rangle \tag{7}$$

and insert a complete set of $P - 1$ eigenstates, $\int dx_i|x_i\rangle\langle x_i| = 1$:

$$\begin{aligned}
Q &= \int dx_1\langle x_1|e^{-\tau H}\int dx_2|x_2\rangle\langle x_2|e^{-\tau H}\int dx_3|x_3\rangle\langle x_3|\cdots \\
&\quad \int dx_P|x_P\rangle\langle x_P|e^{-\tau H}|x_{P+1}\rangle \\
&= \int dx_1 dx_2\cdots dx_P\langle x_1|e^{-\tau H}|x_2\rangle\langle x_2|e^{-\tau H}|x_3\rangle\cdots\langle x_P|e^{-\tau H}|x_{P+1}\rangle \\
&= \int dx_1 dx_2\cdots dx_P\,\rho(x_1, x_2;\tau)\rho(x_2, x_3;\tau)\cdots\rho(x_P, x_{P+1};\tau) \\
&= \int dx_1\cdots dx_P\prod_{i=1}^{P}\rho(x_i, x_{i+1};\tau)
\end{aligned} \tag{8}$$

where $x_1 = x_{P+1}$. In the limit $P\rightarrow\infty$ and $\tau\rightarrow 0$, one can use the semiclassical PA:

$$\begin{aligned}
\rho(x_i, x_{i+1};\tau) &\cong \rho_{PA}(x_i, x_{i+1};\tau) \\
&= \rho_T(x_i, x_{i+1};\tau)\rho_V(x_i;\tau)
\end{aligned} \tag{9}$$

where ρ_T is the kinetic energy (T), i.e., free particle term:

$$\rho_T(x_i, x_{i+1};\tau) = \Omega\cdot\exp\left[-\tau\frac{m}{2\tau^2\hbar^2}(x_i - x_{i+1})^2\right] \tag{10}$$

where $\Omega = (m/2\pi\tau\hbar^2)^{1/2}$ and m is the mass, while ρ_V is the potential energy (V) term:

$$\rho_V(x_i;\tau) = \exp[-\tau V(x_i)] \tag{11}$$

where $V(x_i)$ is the potential at time slice i. The above expression (eq 9) is correct in the $P\rightarrow\infty$ limit due to the Trotter formula $e^{-\beta(T+V)} = \lim_{P\rightarrow\infty}(e^{-\tau T}e^{-\tau V})^P$.[64] The above quantum system is isomorphic to a classical system of ring polymers where each bead, i, in the polymer interacts with its neighbor, $i \pm 1$, via a harmonic potential (eq 10) and experiences only a fraction, $1/P$, of the full potential V (eq 11).

In order to obtain improved PI methods, various operator decomposition approaches may be employed. First, we note that the PA algorithm may be derived from the general operator splitting:

$$e^{-\tau(T+V)} = e^{-\tau V/2}\,e^{-\tau T}\,e^{-\tau V/2} + O(\tau^3) \tag{12}$$

where $O(\tau^3)$ is the big O notation describing the convergence as a function of τ. Summation of the exponential in computing a property (e.g., partition function) P times, yields an error order of $O(\tau^2)$ (i.e., the error in the partition function computed using PA decreases quadratically as P increases or β decreases). In the higher-order action due to Takahashi and Imada (TI),[40] the following operator decomposition is employed:

$$\exp\{-\tau(T + V)\} \cong \exp\{-\tau T\}\exp\{-\tau(V + \tau^2[V,[T,V]]/24)\} \tag{13}$$

This expression converges as $O(\tau^5)$ with respect to the diagonal terms (e.g., the partition function may be computed with fourth-order accuracy). This yields the density matrix elements:

$$\rho_{TI}(x_i, x_{i+1};\tau) = \Omega\cdot\exp\left\{-\tau\frac{m}{2\tau^2\hbar^2}(x_i - x_{i+1})^2 - \tau W_{TI}(x_i)\right\} \tag{14}$$

where $W_{TI}(x_i)$ is the effective one-dimensional TI potential:

$$\begin{aligned}
W_{TI}(x_i) &= V(x_i) + \frac{\hbar^2\tau^2}{24m}|F(x_i)|^2 \quad \text{where} \\
F(x_i) &= \frac{\partial V(x_i)}{\partial x_i}
\end{aligned} \tag{15}$$

The need for a correction term arises due to the fact that the kinetic and potential energy operators do not commute. Thus, in order to add the TI correction, all that is required is to compute the gradient of the potential.

Here we suggest employing a family of higher-order factorization methods based on the symplectic algorithms developed by Chin.[46,57] We start with the expression suggested by Chin (eq 29 in ref 46):

$$\begin{aligned}
e^{-\tau(T+V)} &\cong e^{-t_3\tau T}e^{-v_3\tau V(a_3\tau)}e^{-t_2\tau T}e^{-v_2\tau W(a_2\tau)} \\
&\quad e^{-t_1\tau T}e^{-v_1\tau V(a_1\tau)}e^{-t_0\tau T}
\end{aligned} \tag{16}$$

where W is an effective potential given by $W = V + (u_0/v_2)(\tau^2[V,[T,V]])$ and a_l, t_l, and v_l are positive coefficients which will be defined explicitly below. Setting $t_3 = t_0$ in eq 16 and redistributing the kinetic energy term at t_0:

$$e^{-\tau(T+V)} \cong e^{-v_3\tau V(a_3\tau)}e^{-t_2\tau T}e^{-v_2\tau W(a_2\tau)}e^{-t_1\tau T}e^{-v_1\tau V(a_1\tau)}e^{-2t_0\tau T} \tag{17}$$

Substituting for W yields:

$$\begin{aligned}
\exp\{-\tau(T + V)\} &\cong \exp\{-v_3\tau V(a_3\tau)\}\exp\{-t_2\tau T\} \\
&\quad \exp\{-v_2\tau(V(a_2\tau) + (u_0/v_2)\tau^2[V(a_2\tau),[T, V(a_2\tau)]])\} \\
&\quad \exp\{-t_1\tau T\}\exp\{-v_1\tau V(a_1\tau)\}\exp\{-2t_0\tau T\}
\end{aligned} \tag{18}$$

If the computation of the commuter in eq 18 is not the bottleneck of the calculation (e.g., QM/MM simulations), it is advantageous to distribute the commuter more evenly over the three V. Thus, we multiply the central commuter term by a factor of $1 - \lambda$ and add $\lambda/2$ times the commuter term to each potential operator on each side, as suggested by Chin,[44,46,65] obtaining:

$$\begin{aligned}
&\exp\{-\tau(T + V)\} \cong \\
&\quad \exp\{-v_3\tau(V(a_3\tau) + (\lambda u_0/2v_3)\tau^2[V(a_3\tau),[T, V(a_3\tau)]])\} \\
&\quad \exp\{-t_2\tau T\}\exp\{-v_2\tau(V(a_2\tau) + ((1-\lambda)u_0/v_2)\tau^2[V(a_2\tau), \\
&\qquad [T, V(a_2\tau)]])\}\exp\{-t_1\tau T\} \\
&\quad \exp\{-v_1\tau(V(a_1\tau) + (\lambda u_0/2v_1)\tau^2[V(a_1\tau),[T, V(a_1\tau)]])\} \\
&\quad \exp\{-2t_0\tau T\}
\end{aligned} \tag{19}$$

Setting $a_1 = t_0$, $a_2 = 1/2$, $a_3 = 1 - t_0$, and $v_3 = v_1$ we get

$$\begin{aligned}
&\exp\{-\tau(T + V)\} \cong \\
&\quad \exp\{-\tau(v_1 V((1 - t_0)\tau) + (\lambda u_0/2)\tau^2[V((1 - t_0)\tau), \\
&\qquad [T, V((1 - t_0)\tau)]])\}\exp\{-t_2\tau T\} \\
&\quad \exp\{-\tau(v_2 V(\tau/2) + (1 - \lambda)u_0\tau^2[V(\tau/2),[T, V(\tau/2)]])\} \\
&\qquad \exp\{-t_1\tau T\} \\
&\quad \exp\{-\tau(v_1 V(t_0\tau) + (\lambda u_0/2)\tau^2[V(t_0\tau),[T, V(t_0\tau)]])\} \\
&\qquad \exp\{-2t_0\tau T\}
\end{aligned} \tag{20}$$

Substituting for the commuter $[V,[T,V]] = (\hbar^2/m)|F|^2$, where F is the gradient as defined above in eq 15, yields

$$e^{-\tau(T+V)} = e^{-\nu_1 \tau W_i} e^{-t_2 \tau T} e^{-\nu_2 \tau W_j} e^{-t_1 \tau T} e^{-\nu_1 \tau W_k} e^{-2t_0 \tau T} \tag{21}$$

This expression represents a family of algorithms with fourth-order convergence, which may be modified by changing the parameters u_0, ν_1, ν_2, t_0, t_1, and t_2.[46] Interestingly, $O(\tau^6)$ convergence may be achieved by an optimal choice of factorization parameters, due to cancellation of higher-order error terms. Based on this CH, the corresponding density matrix then becomes

$$\rho_{\mathrm{CH}}(x_i, x_{i+1}; \tau) = \Omega^3 \cdot \left(\frac{1}{2t_1^2 t_0}\right)^{1/2} \cdot \int \mathrm{d}x_j \mathrm{d}x_k$$
$$\exp\left\{\begin{array}{l} -\tau \dfrac{m}{2\tau^2\hbar^2}\left(\dfrac{1}{t_1}(x_i - x_j)^2 + \dfrac{1}{t_1}(x_j - x_k)^2 + \dfrac{1}{2t_0}(x_k - x_{i+1})^2\right) \\ -\tau(W(x_i) + W(x_j) + W(x_k)) \end{array}\right\} \tag{22}$$

where i, j, and k correspond to time slices $(1 - t_0)\tau$, $\tau/2$, and $t_0\tau$, respectively, and $W(x_{i/j/k})$ are generalized effective one-dimensional TI-like potentials at time slices i, j, and k:

$$\begin{aligned} W(x_i) &= \nu_1 V(x_i) + \tau^2 \frac{\hbar^2 u_0 \lambda}{2m} |F(x_i)|^2 \\ W(x_j) &= \nu_2 V(x_j) + \tau^2 \frac{\hbar^2 u_0 (1-\lambda)}{m} |F(x_j)|^2 \\ W(x_k) &= \nu_1 V(x_k) + \tau^2 \frac{\hbar^2 u_0 \lambda}{2m} |F(x_k)|^2 \end{aligned} \tag{23}$$

Here u_0, ν_1, ν_2, t_1, and t_2 are parameters to be optimized via t_0:

$$\begin{aligned} &0 \le t_0 \le \frac{1}{2}\left(1 - \frac{1}{\sqrt{3}}\right); \quad t_1 = t_2 = \frac{1}{2} - t_0 \\ &\nu_1 = \frac{1}{6(1-2t_0)^2}; \quad \nu_2 = 1 - 2\nu_1; \\ &u_0 = \frac{1}{12}\left[1 - \frac{1}{1-2t_0} + \frac{1}{6(1-2t_0)^3}\right] \end{aligned} \tag{24}$$

and λ is a function of t_0 yielding an algorithm correctable to sixth order for the harmonic oscillator:[65]

$$\lambda = \frac{1 + 6t_0\{-3 + 4t_0[6 + t_0(-23 + 24t_0)]\}}{5[1 - 12t_0(1-2t_0)^2][1 - 6t_0(1 + 2t_0 - 4t_0^2)]} \tag{25}$$

This latter expression may be a useful starting point for reducing the fourth-order error. Herein λ will be limited to values between 0 and 1 to yield a forward algorithm (i.e., a negative exponent in eq 22 yields an expression with a bounded integral which may be evaluated directly or simulated using Monte Carlo, MC, methods). This may be seen by inspecting eq 23. This CH PI approach with a gradient-based forward correction converges as τ^6 in favorable cases, such as a harmonic potential, compared with the τ^4 convergence of TI and τ^2 for PA.

2.2. Condensed Phase Expressions. In condensed phase simulations it is useful to compute the QM effects as a correction to the classical mechanics (CM) results. Thus, we write the ratio between the classical and quantum partition functions:[9,10]

$$\frac{Q^{\mathrm{QM}}}{Q^{\mathrm{CM}}} = \frac{\int \mathrm{d}x \rho^{\mathrm{QM}}(x, x; \beta)}{\int \mathrm{d}x \rho^{\mathrm{CM}}(x, x; \beta)} \tag{26}$$

Here the QM density matrix may be described by PA, TI, or CH, as described above, while the CM density matrix may be written as an analogue of the PA, TI, and CH approaches, respectively. In general, we may write the high-temperature density matrices:

$$\rho^{\mathrm{QM}}(x_i, x_{i+1}; \tau) = \rho_T(x_i, x_{i+1}; \tau) \rho_V^{\mathrm{M}}(x_i; \tau) \tag{27}$$

$$\rho^{\mathrm{CM}}(x_i, x_{i+1}; \tau) = \rho_T(x_i, x_{i+1}; \tau) \rho_V^{\mathrm{PA}}(x_{\mathrm{c}}; \tau) \tag{28}$$

where M represents the PA, TI, or CH methods and x_{c} is the classical coordinate which coincides with the centroid, $\bar{x}$, which in discrete representation is defined as $\bar{x} = \frac{1}{P}\Sigma_{i=1}^{P} x_i$. Further we may write

$$\begin{aligned} \frac{Q^{\mathrm{QM}}}{Q^{\mathrm{CM}}} &= \frac{\int \mathrm{d}x \rho^{\mathrm{QM}}(x, x; \beta)}{\int \mathrm{d}x \rho^{\mathrm{CM}}(x, x; \beta)} \\ &= \frac{\int \mathrm{d}x_{\mathrm{c}} \int \mathrm{d}x_1 \cdots \mathrm{d}x_P \delta(x_{\mathrm{c}} - \bar{x}) \prod_{i=1}^{P} \rho^{\mathrm{QM}}(x_i, x_{i+1}; \tau)}{\int \mathrm{d}x_{\mathrm{c}} \int \mathrm{d}x_1 \cdots \mathrm{d}x_P \delta(x_{\mathrm{c}} - \bar{x}) \prod_{i=1}^{P} \rho^{\mathrm{CM}}(x_i, x_{i+1}; \tau)} \\ &= \frac{\int \mathrm{d}x_{\mathrm{c}} \int \mathrm{d}x_1 \cdots \mathrm{d}x_P \delta(x_{\mathrm{c}} - \bar{x}) \prod_{i=1}^{P} \rho_T(x_i, x_{i+1}; \tau) \rho_V^{\mathrm{M}}(x_i; \tau)}{\int \mathrm{d}x_{\mathrm{c}} \int \mathrm{d}x_1 \cdots \mathrm{d}x_P \delta(x_{\mathrm{c}} - \bar{x}) \prod_{i=1}^{P} \rho_T(x_i, x_{i+1}; \tau) \rho_V^{\mathrm{PA}}(x_{\mathrm{c}}; \tau)} \end{aligned} \tag{29}$$

where $\rho_T(x_i,x_{i+1};\tau)$ and $\rho_V^{\mathrm{M}}(x_i;\tau)$ have been defined above. The delta function, $\delta(x_{\mathrm{c}} - \bar{x})$, imposes the centroid constraint on the beads, assuring that the centroid coincides with the classical position. The classical analogue of the quantum potential energy density matrix is obtained in the limit $P = 1$ and is defined as $\rho_V^{\mathrm{PA}}(x_c;\tau) = \exp[-\tau V(x_{\mathrm{c}})]$.

Employing either the PA, TI, or CH potentials, the following useful expression may be derived

$$\begin{aligned} \frac{Q^{\mathrm{QM}}}{Q^{\mathrm{CM}}} &= \frac{\int \mathrm{d}x_{\mathrm{c}} \rho_V^{PA}(x_{\mathrm{c}}; \tau) \int \mathrm{d}x_1 \cdots dx_P \delta(x_{\mathrm{c}} - \bar{x}) \prod_{i=1}^{P} \rho_T(x_i, x_{i+1}; \tau)(\exp(-\tau(\sum_{i=1}^{P}(W^{\mathrm{M}}(x_i) - V(x_{\mathrm{c}})))))}{\int \mathrm{d}x_{\mathrm{c}} \rho_V^{\mathrm{PA}}(x_{\mathrm{c}}; \tau) \int \mathrm{d}x_1 \cdots \mathrm{d}x_P \delta(x_{\mathrm{c}} - \bar{x}) \prod_{i=1}^{P} \rho_T(x_i, x_{i+1}; \tau)} \\ &= \left\langle \langle \exp(-\tau(\sum_{i=1}^{P}(W^{\mathrm{M}}(x_i) - V(x_{\mathrm{c}})))) \rangle_{T, x_{\mathrm{c}}} \right\rangle_{V(x_{\mathrm{c}})} \end{aligned} \tag{30}$$

where W^{M} is the effective potential according to PA, TI, or CH.

In eq 30 the internal bracket, $\langle \cdots \rangle_{T,x_c}$, is an average over the free-particle distribution which is constrained to the centroid (classical) position, while the external average, $\langle\cdots\rangle_{V(x_c)}$, is over the classical (centroid) potential. In this formulation, which is an extension of the original quantized classical path methods,[9,10] the sampling of the classical centroid coordinate and the quantum PI coordinate may be performed separately. Enhanced sampling may be obtained by using the MC staging algorithm[37] in conjunction with the expression in eq 30. In the case of M = CH, a symmetrized version of the potential must be employed in conjunction with the standard staging algorithm.

2.3. Perturbation Expression for Accurate Isotope Effects.

In principle, one can carry out separate centroid path integral (PI) simulations to make QM corrections to the classical potential of mean force for different isotopes. Then, one can use the free energies for different isotopic reactions to compute the corresponding IEs. However, the statistical errors associated with these separate calculations are at least one order of magnitude greater than the free-energy difference for different isotopic reactions—an error too large to be useful for computing IEs. Thus, a sampling scheme which avoids separate sampling for different isotopes is of great importance. Here we present such a scheme for the staging algorithm.

Assuming we want to sample $P-1$ beads using the staging algorithm, $\{x_2, ..., x_P\}$, between end-points x_1 and x_{P+1}. We define $x_1 = x_{P+1} = 0$ and $\Lambda_m = (2\pi\Omega^2)^{-1/2}$.

Stage 1:

$$\begin{aligned} x_2 &= \frac{x_{P+1} + x_1(P-1)}{P} + \Lambda_m \eta_1 \sqrt{\frac{P-1}{P}} \\ &= \frac{x_{P+1} + x_1(P-1)}{P} + \Lambda_m \theta_1 = \Lambda_m \theta_1 \end{aligned} \tag{31}$$

where $\theta_1 = \eta_1\sqrt{(P-1)/P}$ and η_1 is a random number with normal distribution, zero mean, and unit variance.

Stage 2:

$$\begin{aligned} x_3 &= \frac{x_{P+1} + x_2(P-2)}{P-1} + \Lambda_m \eta_2 \sqrt{\frac{P-2}{P-1}} \\ &= \Lambda_m \theta_1 \frac{P-2}{P-1} + \Lambda_m \theta_2 \end{aligned} \tag{32}$$

where $\theta_2 = \eta_2\sqrt{(P-2)/(P-1)}$

Stage 3:

$$\begin{aligned} x_4 &= \frac{x_{P+1} + x_3(P-3)}{P-2} + \Lambda_m \eta_3 \sqrt{\frac{P-3}{P-2}} \\ &= \Lambda_m \theta_1 \frac{P-3}{P-1} + \Lambda_m \theta_2 \frac{P-3}{P-2} + \Lambda_m \theta_3 \end{aligned} \tag{33}$$

where $\theta_3 = \eta_3\sqrt{(P-3)/(P-2)}$

In general, we may write for stage $k-1$:

$$\begin{aligned} x_k &= \frac{x_{P+1} + x_{k-1}(P-k+1)}{P-k+2} + \Lambda_m \eta_{k-1} \sqrt{\frac{P-k+1}{P-k+2}} \\ &= \frac{x_{k-1}(P-k+1)}{P-k+2} + \Lambda_m \theta_{k-1} \end{aligned} \tag{34}$$

where $\theta_{k-1} = \eta_{k-1}\sqrt{(P-k+1)/(P-k+2)}$ and $x_{k-1} = \Lambda_m \Sigma_{i=2}^{k-1} \theta_i (P-k+1)/(P-i+1)$ and $\theta_i = \eta_i\sqrt{(P-i)/(P-i+1)}$.

We may write x_k in a more compact form:

$$x_k = \Lambda_m \sum_{i=2}^{k} \theta_i \frac{P-k+1}{P-i+1} \tag{35}$$

Thus, we see that the final bead distribution is independent of the initial position and may be written exclusively as a function of mass and random distribution numbers. In practice, we implemented the staging algorithm employing eqs 31−34.

Considering a reaction where the light atom of mass m_L is replaced by a heavier isotope of mass m_H, we use exactly the same sequence of random numbers, that is, displacement numbers $\{\theta_i\}$, to generate the staging PI distribution for both isotopes. Thus, the resulting coordinates of these two bead distributions differ only by the ratio of the corresponding masses, assuming we use an identical random number series for the two isotopes:

$$\frac{x_k^{m_L}}{x_k^{m_H}} = \sqrt{\frac{m_H}{m_L}} = \alpha \tag{36}$$

We thus obtain the following identity for the free particle density matrices of the two isotopes:

$$\begin{aligned} &\Omega_L \cdot \exp\left[-\tau \frac{m_L}{2\tau^2\hbar^2}(x_{i,L} - x_{i+1,L})^2\right] \\ &= \Omega_H \cdot \exp\left[-\tau \frac{m_H}{2\tau^2\hbar^2}(x_{i,H} - x_{i+1,H})^2\right] \end{aligned} \tag{37}$$

This is in accord with our previous work employing the bisection sampling algorithm.[13] We may then write the ratio between the QM partition functions for different isotopes (i.e., IE) as

$$\begin{aligned} \text{IE} &= \frac{Q_L^{QM}}{Q_H^{QM}} \\ &= \frac{\int dx_c \int dx_{1,L}...dx_{P,L} \delta(x_c - \bar{x}) \prod_{i=1}^{P} \rho_T^L(x_{i,L}, x_{i+1,L}; \tau) \rho_V^M(x_{i,L}; \tau)}{\int dx_c \int dx_{1,H}...dx_{P,H} \delta(x_c - \bar{x}) \prod_{i=1}^{P} \rho_T^H(x_{i,H}, x_{i+1,H}; \tau) \rho_V^M(x_{i,H}; \tau)} \\ &= \frac{\int dx_c \int dx_{1,L}...dx_{P,L} \delta(x_c - \bar{x}) \prod_{i=1}^{P} \rho_T^L(x_{i,L}, x_{i+1,L}; \tau) \rho_V^M(x_{i,L}; \tau)}{\alpha^P \int dx_c \int dx_{1,L}...dx_{P,L} \delta(x_c - \bar{x}) \prod_{i=1}^{P} \rho_T^H(\alpha x_{i,L}, \alpha x_{i+1,L}; \tau) \rho_V^M(\alpha x_{i,L}; \tau)} \\ &= \frac{\int dx_c \int dx_{1,L}...dx_{P,L} \delta(x_c - \bar{x}) \prod_{i=1}^{P} \rho_T^L(x_{i,L}, x_{i+1,L}; \tau) \rho_V^M(x_{i,L}; \tau)}{\int dx_c \int dx_{1,L}...dx_{P,L} \delta(x_c - \bar{x}) \prod_{i=1}^{P} \rho_T^L(x_{i,L}, x_{i+1,L}; \tau) \rho_V^M(\alpha x_{i,L}; \tau)} \end{aligned} \tag{38}$$

where we have used the substitution $dx_{i,H} = \alpha dx_{i,L}$. This expression may then be employed to compute IEs (e.g., EIE = IE_{PS}/IE_{RS}).

3. COMPUTATIONAL DETAILS

For the model one-dimensional systems, the numerical integration is performed with the iterative scheme for numerical matrix multiplication (NMM) as it avoids the numerical noise inherent to sampling methods, such as MC integration.[60,61] Additionally, the NMM method allows rapid analysis of the convergence as a function of number of beads. The partition function is obtained by computing the trace of the density matrix (eq 6). The high-temperature density matrix is computed with the PA, TI, and CH approaches. To obtain further insight into the properties of the solutions, we employ density matrix diagonalization (DMD) which gives the eigenvalues and eigenvectors of the density matrix as well as the trace.[62]

The partition function is computed for various well-studied potentials relevant for modeling chemical reactions: harmonic oscillator (HO), Morse oscillator (MO), symmetric double well (SDW), and asymmetric double well (ADW) which possess quantum behavior such as zero point energy and tunneling in the temperature range of 100−500 K.

The HO is given by

$$V_{\mathrm{HO}}(x) = ax^2 \tag{39}$$

where we have employed the mass of a hydrogen atom and a = 309 kcal/mol·Å^2.

The MO is given by

$$V_{\mathrm{MO}}(x) = D_{\mathrm{e}}[1 - e^{-\alpha(x - x_0)}]^2 \tag{40}$$

where D_e = 136.3 kcal/mol, α = 2.2112 Å^{-1}, and x_0 = 0.9166 Å. The values were chosen to resemble those of the HF molecule, and the reduced mass of HF was employed.[66]

The SDW is given by

$$V_{\mathrm{SDW}}(x) = ax^4 + bx^2 + c \tag{41}$$

where we have employed a mass of 1224.259 m_{e} (where m_{e} is the mass of an electron), a = 0.01, b = −0.01, and c = 0.0025 au.

The ADW is given by

$$V_{\mathrm{ADW}}(x) = ax^4 + bx^2 + cx + d \tag{42}$$

where we have employed a mass of 1224.259 m_{e}, a = 0.01, and b = −0.02, c = 0.005, and d = 0.015 au.

For all of the above potentials, the optimal CH value of λ (eq 25) was obtained by varying the parameter t_0 in the range 0 to $(1 - 1/\sqrt{3})/2$. Specifically, 10 values at equal intervals were chosen: 0.0211, 0.0422, 0.0633, 0.0844, 0.1055, 0.1266, 0.1477, 0.1688, 0.1899, and 0.2110. Of these values, t_0 = 0.1899 does not fall in the forward range, as it yields a positive exponent in eq 22. We also attempted to optimize λ and t_0 separately.[58] This was done by initially setting λ = 0 and finding the optimal t_0 value. Subsequently an optimal λ value for this t_0 was sought after. We found that for the potentials examined here we obtain very similar results with both approaches, and we prefer the simplicity of using eq 25.

For simulations employing eq 30, we employed the CHARMM program.[67] Previously, we have implemented the PA method within CHARMM.[11−14] In this work we also implement the TI and CH approaches together with the staging algorithm. Calculations for the water molecule employed the B3LYP functional[68,69] with the 6-31+G(d,p) basis set.[70] In this case, CHARMM was combined with the Gamess-UK electronic structure program.[71] Simulations on the enzyme AlaR employed a hybrid QM/MM potential, where the QM part was described by a specific reaction parameter version of the semiempirical AM1 Hamiltonian.[31,32] Details of the system setup and the classical molecular dynamics simulations have been published previously.[31,32] In the current study we employed eqs 1 and 38 to compute IEs at a temperature of 298 K, as implemented in a development version of CHARMM. In all simulations the bead sampling was performed by simultaneously moving all beads at each PI step using the staging or mass perturbation staging algorithms. The number of classical configurations employed was 5200, while 10 MC PI steps were performed at each classical configuration.

The programs employed for all calculations on model systems are written using the Fortran programming language on a Linux platform with Intel compilers. All mathematical derivations are verified using the Maple 12 software suite (Waterloo Maple Inc.).

4. RESULTS

To demonstrate the performance of the higher-order method, we apply the CH algorithm to compute the partition function (eq 6) for a number of well-studied potentials which model key features of chemical reactions. In particular, we compute the partition function for the HO, MO, SDW, and ADW. The results of CH are compared with those obtained with the PA and TI approaches. All potentials were studied at temperatures of 100, 200, 300, 400, and 500 K. A series of 10 t_0 parameters were tested for the CH algorithm with all the potentials at all the temperatures studied. For the sake of brevity, only the results at 100, 300, and 500 K are presented here. The results are compared with the exact results for the HO and MO as well as for SDW and ADW.[72] In order to assess the performance of the CH methods in simulations, we tested the PA, TI, and CH methods on a water molecule as well as in the enzyme AlaR. The PA, TI, and CH methods are employed for a water molecule treated with DFT and a hybrid QM/MM potential for the enzyme AlaR. Below the following notation will be employed: P refers to the number of discrete points in the PI, while k-level is defined by the integer k as $P = 2^k$.

4.1. Harmonic Oscillator (HO). The HO serves as a simple model for a chemical bond. Key results are shown in Figure 1 and Table 1. In Figure 1 the convergence of the CH algorithm with different t_0 values is presented at temperatures 100, 300, and 500 K. The optimal t_0 values are 0.1055 and 0.1266 for all three temperatures, with errors increasingly greater when deviating from these optimal numbers. Indeed, these values are near the midpoint of the range given in eq 24. At 100 K the parameter t_0 = 0.1899 is clearly an outlier, which is symptomatic of a positive sign in the exponent in eq 21. In Table 1 the partition function is displayed at temperatures 100, 300, and 500 K for the PA, TI, and CH. Inspection of the results at T = 100 K shows that using PA the partition function does not converge to within 1% of the exact partition function value of 3.97 × 10^{-9} with a k-level of up to 6. Indeed, PA converges only at a k-level of 9, corresponding to 2^9 = 512 integrals (results not shown in table). Using TI reduces this to k = 6, corresponding to 2^6 = 64 integrals, whereas with CH the desired accuracy is reached with k = 4 with the optimal t_0, requiring $3 \cdot 2^4$ = 48 integrals. At T = 300 K the exact partition function value is 1.58 × 10^{-3}. PA converges with k = 7, TI requires k = 4, and CH converges with k = 2. At T = 500 K the exact value of the partition function is 2.09 × 10^{-2}. To reach convergence PA requires k = 5, TI requires k = 3, whereas CH needs k = 1. Thus, at all temperatures, the CH method yields a 25% enhancement in performance compared to TI. Moreover, CH reaches the performance of PA at less than 10% of the cost at T = 100 and 300 K, and at 500 K, it reaches the performance of PA at 20% of the cost, where we assume for simplicity that the computation of gradients does not significantly increase the computational cost.

In Table 2 we present the computed IE on the partition functions of hydrogen and deuterium ($Q^{\mathrm{H}}/Q^{\mathrm{D}}$). The exact value is 3.485×10^{-3} at a temperature of 100 K. It is clear from inspection of these results that PA has not yet converged (to within 1% of the exact value) at a *k*-level of 6, whereas TI converges with a *k*-level of 6. On the other hand CH reaches convergence with a *k*-level of 4 and 5 with use of the optimal t_0 value of 0.1266 or the symmetric t_0 value of 1/6, respectively. As expected, the IE converges slightly faster than the absolute partition functions.

4.2. Morse Oscillator (MO). The Morse potential is employed as a simple model for the chemical bond including anharmonicity.[73] Main results are shown in Figure 2 and Table 3. In Figure 2 the convergence of the CH algorithm with different t_0 values is presented at temperatures 100, 300, and 500 K. The optimal t_0 values are 0.1055 and 0.1266 for all three temperatures. Again, the value t_0 = 0.1899 yields slightly greater errors than the other parameter values. In Table 3 the partition function is displayed at temperatures 100, 300, and 500 K for the PA, TI, and CH. Inspection of the results at *T* = 100 K shows that using PA the partition function does not converge to within 1% of the exact partition function value of 2.98×10^{-13} at a *k*-level of 6. Rather a *k*-level of 10 is required using PA, corresponding to 1024 integrals (results not shown in table). Using TI reduces this to *k* = 7, corresponding to 128 integrals (results not shown in table), whereas with CH the desired accuracy is reached with *k* = 5 with the optimal t_0, requiring 96 integrals. At *T* = 300 K the exact partition function value is 6.68×10^{-5}. PA converges with *k* = 8, TI requires *k* = 5, and CH converges with *k* = 3. At *T* = 500 K the exact value of the partition function is 3.12×10^{-3}. To reach convergence PA requires *k* = 6, TI requires *k* = 4, whereas CH needs *k* = 2. Thus, at all temperatures, the CH method yields a 25% enhancement in performance compared to TI. Moreover, CH reaches the performance of PA at less than 10% of the cost at *T* = 100 and 300 K, and at *T* = 500 K, it reaches the performance of PA at 20% of the cost.

4.3. Symmetric Double Well (SDW). The double well potential is a simple model for chemical reactions and hydrogen bonding, such as proton transfer or hydrogen networking in ice.[74] Key results for this potential are shown in Figure 3 and Table 4. In Figure 3 the convergence of the CH algorithm with different t_0 values is presented at temperatures 100, 300, and 500 K. The optimal t_0 values are 0.1055 and 0.1266 for all three temperatures. At a *k*-level of 1, t_0 = 0.1899 failed to converge, something we ascribe to the aforementioned sign problem. In Table 4 the partition function is displayed at temperatures 100, 300, and 500 K for the PA, TI, and CH. Inspection of the results

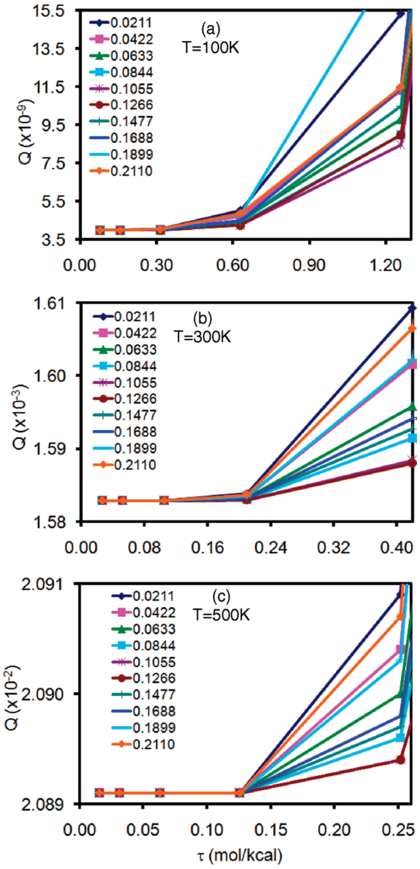


Figure 1. Partition functions for the HO calculated by the CH algorithm at *T* = 100, 300, and 500 K, with varying values of the parameter t_0.

Table 2. Isotope Effect (H/D) for the HO Calculated by the PA, TI, and CH Algorithms at Various *k*-Levels at *T* = 100 K

	Q^H/Q^D (*T* = 100 K)			
$\log_2 P$	PA	TI	CH[a]	CH[b]
1	5.031×10^{-1}	2.584×10^{-1}	2.969×10^{-2}	2.977×10^{-2}
2	2.609×10^{-1}	7.875×10^{-2}	6.384×10^{-3}	6.484×10^{-3}
3	8.363×10^{-2}	1.561×10^{-2}	3.557×10^{-3}	3.688×10^{-3}
4	1.938×10^{-2}	4.925×10^{-3}	3.414×10^{-3}	3.492×10^{-3}
5	6.558×10^{-3}	3.619×10^{-3}	3.460×10^{-3}	3.486×10^{-3}
6	4.176×10^{-3}	3.495×10^{-3}	3.478×10^{-3}	3.486×10^{-3}

[a] The CH method employed t_0 = 1/6. [b] The CH method employed t_0 = 0.1266.

Table 1. Partition Functions for the HO Calculated by the PA, TI, and CH Algorithms at Various *k*-levels at Temperatures *T* = 100, 300, and 500 K

	Q (*T* = 100 K)			Q (*T* = 300 K)			Q (*T* = 500 K)		
$\log_2 P$	PA	TI	CH[a]	PA	TI	CH[a]	PA	TI	CH[a]
1	2.658×10^{-3}	8.303×10^{-5}	1.317×10^{-7}	2.297×10^{-2}	5.330×10^{-3}	1.701×10^{-3}	5.935×10^{-2}	2.811×10^{-2}	2.099×10^{-2}
2	1.096×10^{-4}	1.470×10^{-6}	8.955×10^{-9}	6.597×10^{-3}	2.194×10^{-3}	1.589×10^{-3}	3.236×10^{-2}	2.189×10^{-2}	2.090×10^{-2}
3	2.423×10^{-6}	3.947×10^{-8}	4.244×10^{-9}	2.743×10^{-3}	1.653×10^{-3}	1.584×10^{-3}	2.396×10^{-2}	2.098×10^{-2}	2.089×10^{-2}
4	7.951×10^{-8}	6.271×10^{-9}	3.982×10^{-9}	1.864×10^{-3}	1.589×10^{-3}	1.584×10^{-3}	2.168×10^{-2}	2.090×10^{-2}	2.089×10^{-2}
5	1.100×10^{-8}	4.164×10^{-9}	3.974×10^{-9}	1.653×10^{-3}	1.584×10^{-3}	1.584×10^{-3}	2.109×10^{-2}	2.089×10^{-2}	2.089×10^{-2}
6	5.273×10^{-9}	3.987×10^{-9}	3.973×10^{-9}	1.601×10^{-3}	1.584×10^{-3}	1.584×10^{-3}	2.094×10^{-2}	2.089×10^{-2}	2.089×10^{-2}

[a] The CH algorithm employed t_0 = 0.1266.

at T = 100 K shows that using PA the partition function converges to within 1% of the exact partition function value of 2.57×10^{-3} at a k-level of 7, corresponding to P = 128 integrals (results not shown). Using TI reduces this to k = 5, corresponding to 32 integrals, whereas with CH the desired accuracy is reached with k = 3 with the optimal t_0, requiring 24 integrals. At T = 300 K the exact partition function value is 1.75×10^{-1}. PA converges with k = 5, TI requires k = 3, and CH converges with k = 1. At T = 500K the exact value of the partition function is 4.59×10^{-1}. To reach convergence PA requires k = 4, TI requires k = 2, whereas CH needs k = 0, the latter corresponding to three integrals. Thus, at T = 100 and 300 K, the CH method yields a 25% enhancement in performance compared to TI, while at T = 500 K TI is somewhat more efficient. Moreover, CH reaches the performance of PA at ca. 20% of the cost at T = 100 and 300 K, while at T = 500 K PA is nearly three times as costly as CH.

4.4. Asymmetric Double Well (ADW). Principal results are shown in Figure 4 and Table 5. In Figure 4 the convergence of the CH algorithm with different t_0 values is presented at temperatures 100, 300, and 500 K. The optimal t_0 values are 0.1055 and 0.1266 for all three temperatures. In Table 5 the partition function is displayed at temperatures 100, 300, and 500 K for the PA, TI, and CH. Inspection of the results at T = 100 K shows that in using PA, the partition function converges to within 1% of the exact partition function value of 1.69×10^{-6} at a k-level of 8 (results not shown). Using TI reduces this to k = 6, whereas with CH the desired accuracy is reached with k = 4 with the optimal t_0.

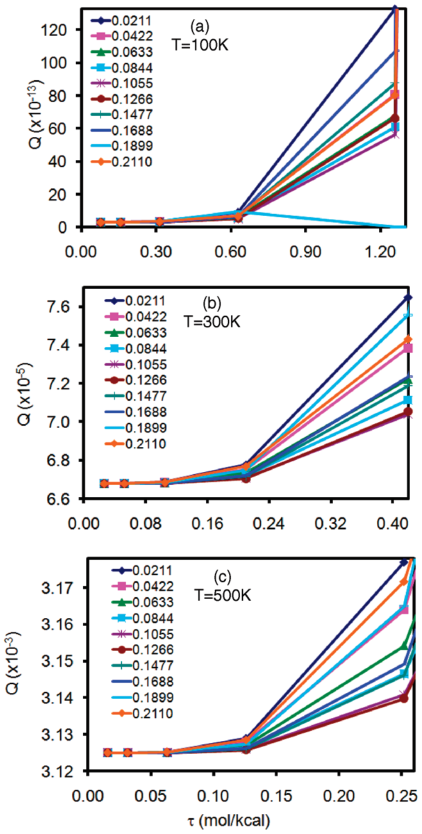


Figure 2. Partition functions for the MO calculated by the CH algorithm at T = 100, 300, and 500 K, with varying values of the parameter t_0.

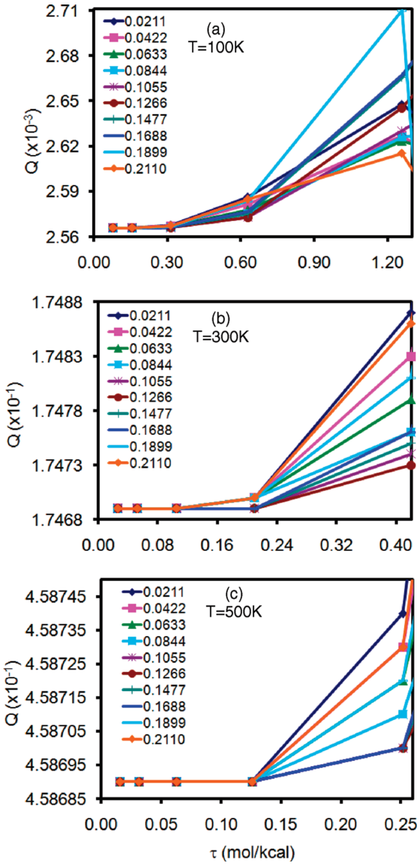


Figure 3. Partition functions for the SDW potential calculated by the CH algorithm at T = 100, 300, and 500 K, with varying values of the parameter t_0.

Table 3. Partition Functions for the MO Calculated by the PA, TI, and CH Algorithms at Various k-Levels at Temperatures T = 100, 300, and 500 K

	Q (T = 100 K)			Q (T = 300 K)			Q (T = 500 K)		
$\log_2 P$	PA	TI	CH[a]	PA	TI	CH[a]	PA	TI	CH[a]
1	1.177×10^{-3}	1.639×10^{-5}	1.491×10^{-9}	1.046×10^{-2}	1.201×10^{-3}	1.019×10^{-4}	2.818×10^{-2}	7.739×10^{-3}	3.317×10^{-3}
2	2.200×10^{-5}	6.363×10^{-8}	6.639×10^{-12}	1.574×10^{-3}	2.046×10^{-4}	7.054×10^{-5}	9.571×10^{-3}	3.963×10^{-3}	3.140×10^{-3}
3	1.133×10^{-7}	1.501×10^{-10}	5.354×10^{-13}	2.938×10^{-4}	8.615×10^{-5}	6.704×10^{-5}	4.741×10^{-3}	3.240×10^{-3}	3.126×10^{-3}
4	3.978×10^{-10}	2.068×10^{-12}	3.156×10^{-13}	1.115×10^{-4}	6.906×10^{-5}	6.681×10^{-5}	3.533×10^{-3}	3.135×10^{-3}	3.125×10^{-3}
5	5.728×10^{-12}	4.235×10^{-13}	2.989×10^{-13}	7.721×10^{-5}	6.698×10^{-5}	6.679×10^{-5}	3.228×10^{-3}	3.126×10^{-3}	3.125×10^{-3}
6	7.519×10^{-13}	3.093×10^{-13}	2.980×10^{-13}	6.935×10^{-5}	6.681×10^{-5}	6.679×10^{-5}	3.151×10^{-3}	3.125×10^{-3}	3.125×10^{-3}

[a] The CH algorithm employed t_0 = 0.1266.

Table 4. Partition Functions for the SDW Potential Calculated by the PA, TI, and CH Algorithms at Various *k*-Levels at Temperatures *T* = 100, 300, and 500 K

	Q (*T* = 100 K)			Q (*T* = 300 K)			Q (*T* = 500 K)		
$\log_2 P$	PA	TI	CH[a]	PA	TI	CH[a]	PA	TI	CH[a]
1	3.490×10^{-2}	4.542×10^{-3}	2.885×10^{-3}	2.524×10^{-1}	1.889×10^{-1}	1.754×10^{-1}	5.225×10^{-1}	4.693×10^{-1}	4.589×10^{-1}
2	1.103×10^{-2}	3.122×10^{-3}	2.649×10^{-3}	2.015×10^{-1}	1.780×10^{-1}	1.747×10^{-1}	4.788×10^{-1}	4.601×10^{-1}	4.587×10^{-1}
3	4.879×10^{-3}	2.838×10^{-3}	2.577×10^{-3}	1.826×10^{-1}	1.751×10^{-1}	1.747×10^{-1}	4.642×10^{-1}	4.588×10^{-1}	4.587×10^{-1}
4	3.208×10^{-3}	2.618×10^{-3}	2.570×10^{-3}	1.768×10^{-1}	1.747×10^{-1}	1.747×10^{-1}	4.601×10^{-1}	4.587×10^{-1}	4.587×10^{-1}
5	2.740×10^{-3}	2.574×10^{-3}	2.570×10^{-3}	1.752×10^{-1}	1.747×10^{-1}	1.747×10^{-1}	4.591×10^{-1}	4.587×10^{-1}	4.587×10^{-1}
6	2.613×10^{-3}	2.570×10^{-3}	2.570×10^{-3}	1.748×10^{-1}	1.747×10^{-1}	1.747×10^{-1}	4.588×10^{-1}	4.587×10^{-1}	4.587×10^{-1}

[a] The CH algorithm employed t_0 = 0.1266.

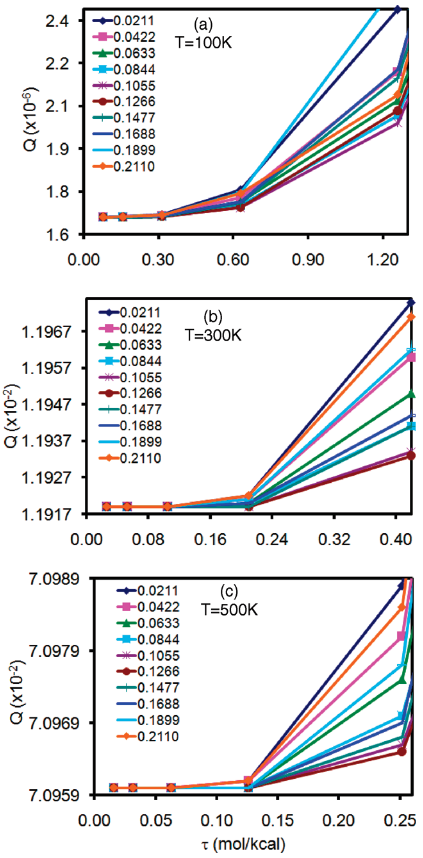


Figure 4. Partition functions for the ADW potential calculated by the CH algorithm at *T* = 100, 300, and 500 K, with varying values of the parameter t_0.

At *T* = 300 K the exact partition function value is 1.19×10^{-2}. PA converges with *k* = 6, TI requires *k* = 4, and CH converges with *k* = 2. At *T* = 500 K the exact value of the partition function is 7.10 $\times 10^{-2}$. To reach convergence PA requires *k* = 5, TI requires *k* = 3, whereas CH needs *k* = 1. Thus, at all temperatures, the CH method yields a 25% enhancement in performance compared to TI. Moreover, CH reaches the performance of PA at ca. 20% of the cost at all temperatures.

4.5. H_2O Molecule. To investigate the application of the various potentials described in this paper, we employed a water molecule at the B3LYP/6-31+G(d,p) level of theory. Specifically, we compute the quantum correction at *T* = 300 K, where the quantum effects are modest. These simulations show that the CH potential is equally applicable to more complex potentials (Figure 5 and Table 6). At this temperature, TI and CH perform similarly well, while PA requires approximately four times as many beads.

4.6. Isotope Effect on Keto–Enol Tautomerism in Alanine Racemase (AlaR). PLP is an essential cofactor for ubiquitous enzyme catalyzed transformations of amines and amino acids, such as racemizations, transaminations, and decarboxylations. A crucial question in all PLP-dependent enzymes is the tautomeric nature of the Schiff-base (Scheme 1), as it may exist in either the iminophenoxide or the enolimine form. The tautomeric form and hence the Schiff-base hydrogen-bond strength is highly sensitive to solvent polarity, and this topic has been addressed experimentally by NMR studies of the hydrogen-bond EIE.[75–77] From a computational perspective, it is therefore important to develop methods which can accurately predict the hydrogen-bond EIE in enzymes. Herein, we employ the PA, TI, and CH methods with the mass-perturbation staging algorithm derived in eqs 31–38.

Initially, to validate the vibrational frequencies of the O–H and N–H stretches in tautomers of the pyridoxal moiety, we performed model calculations on the tautomeric ketoamine NH and enolimine OH forms (Scheme 1). The computed vibrational frequency for the NH-stretch in zwitterionic ketoamine NH tautomer was 3211.7 and 3290.9 cm^{-1} at the target M06/6-31+G(d,p) level[78] and at the AM1-SRP level, respectively (Table 7). The computed vibrational frequency for the OH-stretch in the nonzwitterionic enolimine OH tautomer was 3413.6 and 3435.9 cm^{-1} at the target M06/6-31+G(d,p) level and at the AM1-SRP level, respectively. Thus, the differences between the vibrational frequencies of the two tautomeric forms are 201.9 and 145.1 cm^{-1} at the M06 and semiempirical levels. The computed gas-phase equilibrium IE is 0.89 and 1.06 at the M06 and semiempirical levels, respectively, where we have employed a scaling factor of 0.98 for M06 frequencies.[79]

The enzyme quantum simulations employed the PA, TI, and CH approaches in conjunction with the mass-perturbation staging algorithm. The PLP cofactor in AlaR is presented in Figure 6 with a protonated and deuterated Schiff base, and the numerical results obtained using eqs 1 and 38 are summarized in Tables 8 and 9. We estimate the converged value of the EIE at 298 K as 1.16 ± 0.06. Interestingly, PA shows robust performance with the mass-perturbation staging algorithm for computation of the EIE. Using only 3 or 6 beads, the EIE is estimated to be 1.19. With 12 beads, the result is 1.16, while further increasing the number of beads to 24 or 48 yields 1.15. Surprisingly, the TI and CH show poor performance using 3 or 6 beads. Using TI the

Table 5. Partition functions for the ADW potential calculated by the PA, TI, and CH algorithms at various k-levels at temperatures T = 100, 300, and 500 K

	Q (T = 100 K)			Q (T = 300 K)			Q (T = 500 K)		
$\log_2 P$	PA	TI	CH[a]	PA	TI	CH[a]	PA	TI	CH[a]
1	5.246×10^{-3}	3.091×10^{-4}	4.918×10^{-6}	4.501×10^{-2}	1.775×10^{-2}	1.215×10^{-2}	1.125×10^{-1}	7.711×10^{-2}	7.108×10^{-2}
2	4.274×10^{-4}	1.858×10^{-5}	2.065×10^{-6}	2.162×10^{-2}	1.303×10^{-2}	1.193×10^{-2}	8.375×10^{-2}	7.189×10^{-2}	7.097×10^{-2}
3	3.273×10^{-5}	3.312×10^{-6}	1.725×10^{-6}	1.457×10^{-2}	1.206×10^{-2}	1.192×10^{-2}	7.446×10^{-2}	7.105×10^{-2}	7.096×10^{-2}
4	5.462×10^{-6}	1.916×10^{-6}	1.693×10^{-6}	1.261×10^{-2}	1.193×10^{-2}	1.192×10^{-2}	7.186×10^{-2}	7.097×10^{-2}	7.096×10^{-2}
5	2.427×10^{-6}	1.714×10^{-6}	1.691×10^{-6}	1.209×10^{-2}	1.192×10^{-2}	1.192×10^{-2}	7.119×10^{-2}	7.096×10^{-2}	7.096×10^{-2}
6	1.863×10^{-6}	1.692×10^{-6}	1.691×10^{-6}	1.196×10^{-2}	1.192×10^{-2}	1.192×10^{-2}	7.102×10^{-2}	7.096×10^{-2}	7.096×10^{-2}

[a] The CH algorithm employed t_0 = 0.1266.

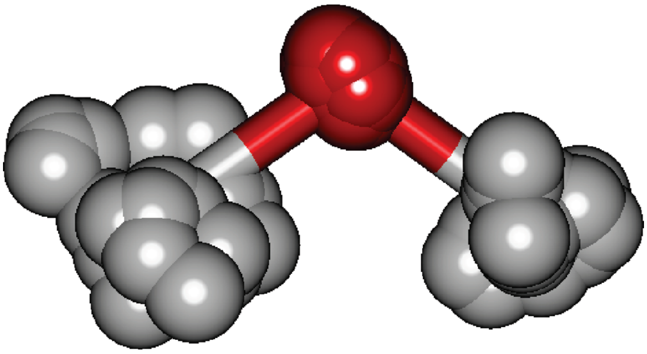


Figure 5. Quantized water molecule treated at the B3LYP/6-31+G(d, p) level with 18 beads (P = 18).

Table 6. Vibrational Quantum → Classical Free Energy Correction[a] (kcal/mol) for a Water Molecule Treated at the B3LYP/6-31+G(d,p) Level Computed Using the Staging Algorithm in Conjunction with the PA, TI, and CH Algorithms with Various Numbers of Beads (P) at Temperature T = 300 K

P	PA	TI	CH
3	4.09 ± 0.06	7.38 ± 0.94	7.25 ± 0.48
6	6.70 ± 0.14	8.89 ± 0.65	8.86 ± 0.46
12	8.42 ± 0.22	9.39 ± 0.31	9.40 ± 0.22
24	9.08 ± 0.17	9.40 ± 0.26	9.27 ± 0.20

[a] The calculations used eq 30 with 100 MC staging algorithm steps per classical point. The total number of classical points was 100. All values are averaged over 10 independent runs.

Table 7. Computed Unscaled Vibrational Frequencies (cm^{-1}) of the Schiff-Base Moiety in Tautomers of a Model Pyridoxal Compound

	M06/6-31+G(d,p)	AM1-SRP
ketoamine NH-form	3211.7	3290.9
enolimine OH-form	3413. 6	3435.9

EIE is estimated to be 1.58 and 1.25 using 3 and 6 beads, respectively. Using CH the EIE is estimated to be 1.44 and 1.19 using 3 and 6 beads, respectively. Further increasing the number of beads to 12, 24, or 48 yields converged values for both TI and CH. At the current simulation temperature and all numbers of beads employed, TI and CH are still expected to perform better than PA (i.e., $O^{\mathrm{TI/CH}}(\tau^4) < O^{\mathrm{PA}}(\tau^2)$). The reason for this seems

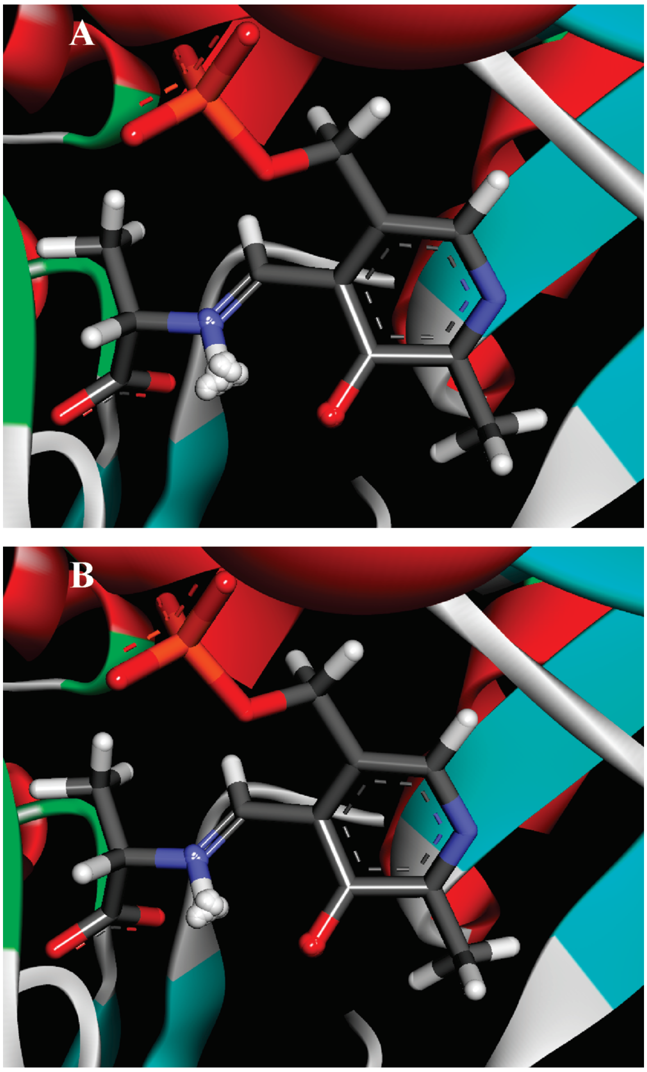


Figure 6. Quantized Schiff base and oxyanion in the PLP cofactor in AlaR with 48 beads and (A) protonated (B) deuterated Schiff base. The bead distributions were computed with the mass-perturbation staging algorithm.

Table 8. RS and PS IE Computed for the Keto−Enol Tautomerism in AlaR at $T = 298$ K[a]

	RS			PS		
P	PA	TI	CH	PA	TI	CH
3	0.227 ± 0.009	0.039 ± 0.005	0.041 ± 0.005	0.270 ± 0.009	0.062 ± 0.005	0.059 ± 0.004
6	0.096 ± 0.004	0.034 ± 0.003	0.035 ± 0.003	0.114 ± 0.004	0.043 ± 0.002	0.041 ± 0.003
12	0.060 ± 0.003	0.042 ± 0.002	0.041 ± 0.002	0.070 ± 0.003	0.048 ± 0.002	0.047 ± 0.002
24	0.051 ± 0.002	0.046 ± 0.002	0.045 ± 0.002	0.059 ± 0.002	0.054 ± 0.002	0.053 ± 0.002
48	0.049 ± 0.002	0.047 ± 0.002	0.047 ± 0.002	0.056 ± 0.002	0.055 ± 0.002	0.054 ± 0.002

[a] The calculations used eq 38 with 10 MC staging algorithm steps per classical point per isotope. The total number of classical points was 10 400. The error is estimated as $\bar{\sigma} = (\Sigma_{i=1}^{N}\sigma_i^2)^{1/2}/N)$, which is the standard deviation in computing eq 38 in the RS or PS wells using N discrete points, and σ_i is the standard deviation in computing IE at a discrete point in the reactant or product well.

Table 9. EIE Computed for the Keto−Enol Tautomerism in AlaR at $T = 298$ K[a]

P	PA	TI	CH
3	1.19 ± 0.06	1.58 ± 0.25	1.44 ± 0.22
6	1.19 ± 0.07	1.25 ± 0.13	1.19 ± 0.13
12	1.16 ± 0.07	1.14 ± 0.08	1.15 ± 0.08
24	1.15 ± 0.06	1.17 ± 0.07	1.16 ± 0.07
48	1.15 ± 0.06	1.15 ± 0.07	1.16 ± 0.06

[a] The calculations used eqs 1 and 38 with 10 MC staging algorithm steps per classical point per isotope. The total number of classical points was 10 400. The error is estimated as $\sigma = ((\bar{\sigma}_{RS}/IE_{RS})^2 + (\bar{\sigma}_{PS}/IE_{PS})^2)^{1/2}\cdot(IE_{PS}/IE_{RS})$, where IE = Q_L/Q_H is the IE in either the RS or PS wells, $\bar{\sigma} = (\Sigma_{i=1}^{N}\sigma_i^2)^{1/2}/N$ is the standard deviation in computing eq 38 in the RS or PS wells using N discrete points, and σ_i is the standard deviation in computing IE at a discrete point in the reactant or product well.

to be the greater uncertainty in the computed EIE using TI and CH with a low number of beads, which is interestingly due to the enhanced accuracy of the methods. Inspection of Table 8 reveals that the simulation error is reduced as the number of beads is increased. Moreover, the standard deviation is similar for all three methods. It is important to note that the standard deviation is not due to the sampling of the kinetic energy term, as this part is sampled exactly by the free-particle mass-perturbation staging algorithm, but rather due to sampling of the potential energy surface. Specifically, both the classical averaging over the potential energy surface as well as the PI sampling of the potential surface contribute to the standard deviation. This is due to the fluctuating nature of the complex potential energy surface in enzymes. Using TI and CH in computing eq 38 in the RS and PS, respectively, with a small number of beads yields fairly converged IE values with respect to number of beads (Table 8). In computing the EIE we need to divide IE in the PS and RS (EIE = IE_{PS}/IE_{RS}), which in the case of TI and CH are small numbers with large error bars, yielding greater errors in the computed EIE (Table 9). On the other hand when using PA, the absolute error in computing the IE is greater due to the small number of beads and lack of higher order terms as in TI and CH. Thus, although PA exhibits greater absolute errors in computing the IE, these errors largely cancel out in the RS and PS, as the error is not in the leading digits of the IE.

Finally, we compare the efficiency of the mass perturbation treatment using the staging and bisection algorithms. Specifically we computed the EIE in AlaR using 8 beads with the two methods, using either 10 or 20 MC steps per classical configuration for a total of 5200 classical configurations. Employing 10 MC steps we obtained 1.19 ± 0.07 and 1.23 ± 0.07 for the staging and bisection algorithms, respectively, while using 20 MC steps we obtained 1.20 ± 0.05 for both the staging and bisection algorithms, respectively. Thus, the two methods give comparable results, and this conclusion is not expected to change when using a greater number of beads. Indeed, both the bisection and staging algorithms sample the kinetic part of the action exactly, and therefore for free particle sampling, their performance will be comparable. Thus, within the framework of eq 30 both sampling schemes may in principle be employed. However, the Chin action requires that the number of beads be a multiple of three (see eq 22) and may be readily achieved with the staging algorithm, which can sample any number of beads. However, the bisection algorithm naturally samples 2^k number of beads in a naïve implementation, where k is the sampling level, and therefore is not generally suitable for the Chin action. Thus, the staging algorithm may be more flexible with respect to number of beads. We note that when the sampling entails not only the kinetic part of the action but also the potential part, the bisection algorithm may be advantageous. Using the bisection algorithm when moving *P* beads, the largest bead move is performed at the first MC step (i.e., the middle bead), and one may reject the collective move of *P* beads based on the move of a single bead (i.e., a single energy and force calculation as opposed to *P* such calculations).

5. DISCUSSION

In this study we initially compare the performance of the PA algorithm and the higher order TI and CH algorithms on four model potentials: HO, MO, SDW, and ADW. We find that the CH algorithm with optimal parameters performs considerably better than the PA when computing the partition function for the model chemical potentials. This conclusion is in accordance with the findings of Sakkos et al for quantum liquids.[58] The use of the CH approach is most beneficial at low temperatures where quantum effects are more pronounced. Nonetheless, we find that the TI approach performs nearly as well as CH, and the main gain is in going beyond the PA. These findings for model systems are of great importance when moving to condensed phase systems, where the addition of numerical noise complicates the performance analysis of the methods.

The parametrized CH algorithm is expected to be of value in condensed phase simulations where the computational bottleneck is the energy evaluation, such as in simulations employing fully QM or hybrid QM/MM potential energy surfaces. In typical uses of such potentials, iterative self-consistent field calculations are required in evaluating the potential energy, and these are computationally expensive. The computation of gradients, on the other hand, requires less effort than the energy evaluation itself. Thus, a PI method which can significantly reduce the number of energy evaluations is of great value. Indeed, the efficiency of the CH factorization in the calculation of nuclear QM effects using a complicated potential energy surface is exemplified in this work by calculations of water treated with DFT. This conclusion is also correct for a considerably more complex hybrid QM/MM potential energy surface such as the one employed here in the case of the enzyme AlaR. However, in computing IEs, numerical noise hampers the performance of both TI and CH with a small number of beads although the quantum effects are treated more accurately than with PA. This is largely due to the simulation noise inherent to any sampling method and not due to inherent properties of TI or CH. This is clear from the model calculations of the IE for the HO, where TI and CH displayed superb performance. Remarkably, PA is highly accurate in computing the EIE on the tautomerism of PLP in AlaR, even when using only three beads, when employing the mass-perturbation staging algorithm.

The approaches employed in this work (eqs 30 and 38) are equally applicable to computing the centroid potential of mean force, and this is currently being investigated in our group. Additionally, PI schemes based on the flux autocorrelation methods which require the calculation of the entire density matrix will benefit from the CH algorithm. In such chemical rate calculations, the PA and TI approaches are expected to be much less efficient. Higher dimensionality derivations of flux autocorrelation methods in conjunction with the CH method are being pursued in our group.

It is interesting to note that the enzyme environment enhances the EIE when compared to the gas-phase results. In the gas-phase, the EIE is computed to be 1.06, while in AlaR it is estimated as 1.16. This is indicative of a weakening of the intramolecular hydrogen bond relative to the gas phase. This is indeed expected as the highly polar active site in AlaR reduces the difference between the zero-point energies of the iminophenoxide or enolimine forms. This is in agreement with experimental work on model PLP systems in solvents of varying degrees of polarity.[77] We believe the current approach will be of great use in the study of the effect of active site polarity on the hydrogen-bond strength in PLP-dependent enzymes as well as other enzymes.

Finally, it may be instructive to compare the current approach for computing IEs to other related approaches. Recently, Wong et al. employed classical TST, PI quantum TST, and the quantum instanton approaches to evaluate the quantized potential of mean force and KIE in malonaldehyde.[80] In this study, the latter two approaches were found to give KIEs in reasonable agreement with each other, although a clear relationship between the two methods has not yet been established. The current mass-perturbation staging approach is in principle similar to the PI quantum TST with thermodynamic integration approach employed by Wong et al. However, the advantage of the current approach is that PI sampling is only required for the light and heavy isotopes, and no sampling of intermediate mass values is required during the perturbation. This one-step perturbation is highly efficient and is possible due to the fact that the current PI sampling is performed using free-particle MC.

6. CONCLUSIONS

Higher-order corrections to the primitive approximation (PA) may considerably enhance the performance of quantum simulation methods. In this report we compare the PA and the higher order Takahashi−Imada (TI) algorithm with the gradient-based forward corrector algorithm due to Chin (CH) on a variety of model potentials. We find a unique parameter for the Chin algorithm which gives a good performance for all model potentials tested. Moreover, the PA, TI, and CH factorizations are employed to compute the quantum correction to a water molecule treated with the B3LYP functional with a 6-31+G(d,p) basis set. Finally, we employ the PA, TI, and CH methods to compute the equilibrium IE on the Schiff base−oxyanion tautomerism in the cofactor pyridoxal-5′-phosphate in the enzyme alanine racemase using a novel mass-perturbation staging algorithm. We find that the Chin algorithm performs well for the complex molecular systems as well, although numerical noise might hamper its performance when computing IEs with a small number of beads.

■ ASSOCIATED CONTENT

S Supporting Information. Figures of model potentials at temperatures 100−500 K. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: majort@mail.biu.ac.il.

■ ACKNOWLEDGMENT

This work has been supported by a start-up grant from Bar-Ilan University, Israel Science Foundation, and the Alon fellowship from the Council for higher education, planning, and budgeting committee. We thank Prof. Siu A. Chin for helpful comments to this manuscript.

■ REFERENCES

(1) Frey, P. A.; Hegeman, A. D. *Enzymatic reaction mechanisms*; Oxford University Press: New York, 2007.

(2) Kohen, A.; Limbach, H. H. *Isotope Effects in Chemistry and Biology*; Taylor and Francis Group, CRC Press: New York, 2006.

(3) Lewis, B. E.; Schramm, V. L. Enzymatic binding isotope effects and the interaction of glucose with hexokinase. In *Isotope Effects in Chemistry and Biology*; Kohen, A., Limbach, H. H., Eds.; Taylor and Francis Group, CRC Press: New York, 2006; pp 1019.

(4) Limbach, H. H.; Denisov, G. S.; Golubev, N. S., Hydrogen bond isotope effects studied by NMR. In *Isotope Effects in Chemistry and Biology*; Kohen, A., Limbach, H. H., Eds.; Taylor and Francis Group, CRC Press: New York, 2006; pp 193.

(5) Voth, G. A. Feynman path integral formulation of quantum mechanical transition state theory. *J. Phys. Chem.* **1993**, *97*, 8365.

(6) Feynman, R. P.; Hibbs, A. R. *Quantum Mechanics and Path Integrals*; McGraw-Hill: New York, 1965.

(7) Berne, B. J.; Thirumalai, D. On the simulation of quantum-systems - path integral methods. *Annu. Rev. Phys. Chem.* **1986**, *37*, 401.

(8) Marx, D.; Parrinello, M. Structural quantum effects and 3-center 2-electron bonding in ch_5+. *Nature* **1995**, *375*, 216.

(9) Hwang, J. K.; Chu, Z. T.; Yadav, A.; Warshel, A. Simulations of quantum-mechanical corrections for rate constants of hydride-transfer reactions in enzymes and solutions. *J. Phys. Chem.* **1991**, *95*, 8445.

(10) Hwang, J. K.; Warshel, A. A quantized classical path approach for calculations of quantum-mechanical rate constants. *J. Phys. Chem.* **1993**, *97*, 10053.

(11) Major, D. T.; Gao, J. L. Implementation of the bisection sampling method in path integral simulations. *J. Mol. Graphics Modell.* **2005**, *24*, 121.

(12) Major, D. T.; Garcia-Viloca, M.; Gao, J. L. Path integral simulations of proton transfer reactions in aqueous solution using combined QM/MM potentials. *J. Chem. Theory Comput.* **2006**, *2*, 236.

(13) Major, D. T.; Gao, J. L. An integrated path integral and free-energy perturbation-umbrella sampling method for computing kinetic isotope effects of chemical reactions in solution and in enzymes. *J. Chem. Theory Comput.* **2007**, *3*, 949.

(14) Gao, J. L.; Wong, K. Y.; Major, D. T. Combined QM/MM and path integral simulations of kinetic isotope effects in the proton transfer reaction between nitroethane and acetate ion in water. *J. Comput. Chem.* **2008**, *29*, 514.

(15) Wang, M. L.; Lu, Z. Y.; Yang, W. T. Nuclear quantum effects on an enzyme-catalyzed reaction with reaction path potential: Proton transfer in triosephosphate isomerase. *J. Chem. Phys.* **2006**, *124*, 124516.

(16) Wang, Q.; Hammes-Schiffer, S. Hybrid quantum/classical path integral approach for simulation of hydrogen transfer reactions in enzymes. *J. Chem. Phys.* **2006**, *125*, 184102.

(17) Marx, D.; Tuckerman, M. E.; Martyna, G. J. Quantum dynamics via adiabatic ab initio centroid molecular dynamics. *Comput. Phys. Commun.* **1999**, *118*, 166.

(18) Tuckerman, M. E.; Marx, D.; Klein, M. L.; Parrinello, M. On the quantum nature of the shared proton in hydrogen bonds. *Science* **1997**, *275*, 817.

(19) Tuckerman, M. E.; Marx, D.; Parrinello, M. The nature and transport mechanism of hydrated hydroxide ions in aqueous solution. *Nature* **2002**, *417*, 925.

(20) Paesani, F.; Iuchi, S.; Voth, G. A. Quantum effects in liquid water from an ab initio-based polarizable force field. *J. Chem. Phys.* **2007**, *127*, 074506.

(21) Ohta, Y.; Ohta, K.; Kinugawa, K. Quantum effect on the internal proton transfer and structural fluctuation in the H-5(+) cluster. *J. Chem. Phys.* **2004**, *121*, 10991.

(22) Hayashi, A.; Shiga, M.; Tachikawa, M. H/D isotope effect on the dihydrogen bond of NH_4^+ center dot center dot center dot BeH_2 by ab initio path integral molecular dynamics simulation. *J. Chem. Phys.* **2006**, *125*, 204310.

(23) Wong, K. Y.; Gao, J. An automated integration-free path-integral method based on Kleinert's variational perturbation theory. *J. Chem. Phys.* **2007**, *127*, 211103.

(24) Wong, K. Y.; Gao, J. Systematic approach for computing zero-point energy, quantum partition function, and tunneling effect based on Kleinert's variational perturbation. *J. Chem. Theory Comput.* **2008**, *4*, 1409.

(25) Alhambra, C.; Corchado, J.; Sanchez, M. L.; Garcia-Viloca, M.; Gao, J.; Truhlar, D. G. Canonical variational theory for enzyme kinetics with the protein mean force and multidimensional quantum mechanical tunneling dynamics. Theory and application to liver alcohol dehydrogenase. *J. Phys. Chem. B* **2001**, *105*, 11326.

(26) Pu, J. Z.; Gao, J. L.; Truhlar, D. G. Multidimensional tunneling, recrossing, and the transmission coefficient for enzymatic reactions. *Chem. Rev.* **2006**, *106*, 3140.

(27) Billeter, S. R.; Webb, S. P.; Agarwal, P. K.; Iordanov, T.; Hammes-Schiffer, S. Hydride transfer in liver alcohol dehydrogenase: Quantum dynamics, kinetic isotope effects, and role of enzyme motion. *J. Am. Chem. Soc.* **2001**, *123*, 11262.

(28) Iyengar, S. S.; Sumner, I.; Jakowski, J. Hydrogen tunneling in an enzyme active site: A quantum wavepacket dynamical perspective. *J. Phys. Chem. B* **2008**, *112*, 7601.

(29) Antoniou, D.; Basner, J.; Nunez, S.; Schwartz, S. D. Computational and theoretical methods to explore the relation between enzyme dynamics and catalysis. *Chem. Rev.* **2006**, *106*, 3170.

(30) Major, D. T.; York, D. M.; Gao, J. L. Solvent polarization and kinetic isotope effects in nitroethane deprotonation and implications to the nitroalkane oxidase reaction. *J. Am. Chem. Soc.* **2005**, *127*, 16374.

(31) Major, D. T.; Nam, K.; Gao, J. L. Transition state stabilization and alpha-amino carbon acidity in alanine racemase. *J. Am. Chem. Soc.* **2006**, *128*, 8114.

(32) Major, D. T.; Gao, J. L. A combined quantum mechanical and molecular mechanical study of the reaction mechanism and alpha-amino acidity in alanine racemase. *J. Am. Chem. Soc.* **2006**, *128*, 16345.

(33) Rubinstein, A.; Major, D. T. Catalyzing Racemizations in the Absence of a Cofactor: The Reaction Mechanism in Proline Racemase. *J. Am. Chem. Soc.* **2009**, *131*, 8513.

(34) Major, D. T.; Heroux, A.; Orville, A. M.; Valley, M. P.; Fitzpatrick, P. F.; Gao, J. Differential quantum mechanical tunneling in the uncatalyzed and in the Nitroalkane Oxidase proton abstraction of nitroethane. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 20734.

(35) Zimmermann, T.; Vanicek, J. Path integral evaluation of equilibrium isotope effects. *J. Chem. Phys.* **2009**, *131*, 024111.

(36) Zimmermann, T.; Vanicek, J. Three applications of path integrals: equilibrium and kinetic isotope effects, and the temperature dependence of the rate constant of the [1,5] sigmatropic hydrogen shift in (Z)-1,3-pentadiene. *J. Mol. Model.* **2010**, *16*, 1779.

(37) Sprik, M.; Klein, M. L.; Chandler, D. Staging - a sampling technique for the Monte Carlo evaluation of path-integrals. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1985**, *31*, 4234.

(38) Pollock, E. L.; Ceperley, D. M. Simulation of quantum many-body systems by path-integral methods. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1984**, *30*, 2555.

(39) Cao, J.; Berne, B. J. A Born-Oppenheimer approximation for path integrals with an application to electron solvation in polarizable fluids. *J. Chem. Phys.* **1993**, *99*, 2902.

(40) Takahashi, M.; Imada, M. Monte Carlo calculation of quantum-systems. 2. Higher-order correction. *J. Phys. Soc. Jpn.* **1984**, *53*, 3765.

(41) Li, X. P.; Broughton, J. Q. High-order correction to the Trotter expansion for use in computer simulation. *J. Chem. Phys.* **1987**, *86*, 5094.

(42) Suzuki, M. Compact exponential product formulas and operator functional derivative. *J. Math. Phys.* **1997**, *38*, 1183.

(43) Suzuki, M. Fractal decomposition of exponential operators with applications to many-body theories and Monte Carlo simulations. *Phys. Lett. A* **1990**, *146*, 319.

(44) Chin, S. A. Symplectic integrators from composite operator factorizations. *Phys. Lett. A* **1997**, *226*, 344.

(45) Chin, S. A.; Chen, C. R. Fourth order gradient symplectic integrator methods for solving the time-dependent Schrodinger equation. *J. Chem. Phys.* **2001**, *114*, 7338.

(46) Chin, S. A.; Chen, C. R. Gradient symplectic algorithms for solving the Schrodinger equation with time-dependent potentials. *J. Chem. Phys.* **2002**, *117*, 1409.

(47) Forbert, H. A.; Chin, S. A. Fourth-order diffusion Monte Carlo algorithms for solving quantum many-body problems. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2001**, *63*, 144518.

(48) De Raedt, H.; De Raedt, B. Applications of the generalized Trotter formula. *Phys. Rev. A: At., Mol., Opt. Phys.* **1983**, *28*, 3575.

(49) Kono, H.; Takasaka, A.; Lin, S. H. Monte Carlo calculation of the quantum partition function via path integral formulations. *J. Chem. Phys.* **1988**, *88*, 6390.

(50) Schwartz, S. D. Accurate quantum mechanics from high order resummed operator expansions. *J. Chem. Phys.* **1994**, *100*, 8795.

(51) Weht, R. O.; Kohanoff, J.; Estrin, D. A.; Chakravarty, C. An ab initio path integral Monte Carlo simulation method for molecules and clusters: Application to Li-4 and Li-5(+). *J. Chem. Phys.* **1998**, *108*, 8848.

(52) Jang, S. J.; Jang, S. M.; Voth, G. A. Applications of higher order composite factorization schemes in imaginary time path integral simulations. *J. Chem. Phys.* **2001**, *115*, 7832.

(53) Omelyan, I. P.; Mryglod, I. M.; Folk, R. Construction of high-order force-gradient algorithms for integration of motion in classical and quantum systems. *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.* **2002**, *66*, 026701.

(54) Yamamoto, T. M. Path-integral virial estimator based on the scaling of fluctuation coordinates: Application to quantum clusters with fourth-order propagators. *J. Chem. Phys.* **2005**, *123*, 104101.

(55) Whitfield, T. W.; Martyna, G. J. Low variance energy estimators for systems of quantum Drude oscillators: Treating harmonic path integrals with large separations of time scales. *J. Chem. Phys.* **2007**, *126*, 074104.

(56) Cuervo, J. E.; Roy, P. N.; Boninsegni, M. Path integral ground state with a fourth-order propagator: Application to condensed helium. *J. Chem. Phys.* **2005**, *122*, 114504.

(57) Chin, S. A. Quantum statistical calculations and symplectic corrector algorithms. *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.* **2004**, *69*, 046118.

(58) Sakkos, K.; Casulleras, J.; Boronat, J. High order Chin actions in path integral Monte Carlo. *J. Chem. Phys.* **2009**, *130*, 204109.

(59) Klemm, A. D.; Storer, R. G. Structure of quantum fluids - helium and neon. *Aust. J. Phys.* **1973**, *26*, 43.

(60) Thirumalai, D.; Berne, B. J. On the calculation of time correlation-functions in quantum-systems - path integral techniques. *J. Chem. Phys.* **1983**, *79*, 5029.

(61) Thirumalai, D.; Bruskin, E. J.; Berne, B. J. An iterative scheme for the evaluation of discretized path-integrals. *J. Chem. Phys.* **1983**, *79*, 5063.

(62) Sethia, A.; Sanyal, S.; Singh, Y. Discretized path integral method and properties of a quantum system. *J. Chem. Phys.* **1990**, *93*, 7268.

(63) Chandler, D.; Wolynes, P. G. Exploiting the isomorphism between quantum theory and classical statistical mechanics of polyatomic fluids. *J. Chem. Phys.* **1981**, *74*, 4078.

(64) Trotter, H. F. On the product of semi-groups of operators. *Proc. Am. Math. Soc.* **1959**, *10*, 545.

(65) Scuro, S. R.; Chin, S. A. Forward symplectic integrators and the long-time phase error in periodic motions. *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.* **2005**, *71*, 056703.

(66) Khristenko, S. V.; Maslov, A. I.; Shevelko, V. P. *Molecules and Their Spectroscopic Properties*; Springer: Berlin, Heidelberg, Germany, 1998.

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

(68) Becke, A. D. Density-functional thermochemistry III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648.

(69) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785.

(70) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; John Wiley & Sons: New York, 1986.

(71) Guest, M. F.; Bush, I. J.; van Dam, H. J. J.; Sherwood, P.; Thomas, J. M. H.; van Lenthe, J. H.; Havenith, R. W. A.; Kendrick, J. The GAMESS-UK electronic structure package: algorithms, developments and applications. *Mol. Phys.* **2005**, *103*, 719.

(72) Mielke, S. L.; Truhlar, D. G. Displaced-points path integral method for including quantum effects in the Monte Carlo evaluation of free energies. *J. Chem. Phys.* **2001**, *115*, 652.

(73) Morse, P. M. Diatomic molecules according to the wave mechanics. II. Vibrational levels. *Phys. Rev.* **1929**, *34*, 57.

(74) Morrone, J. A.; Lin, L.; Car, R. Tunneling and delocalization effects in hydrogen bonded systems: A study in position and momentum space. *J. Chem. Phys.* **2009**, *130*, 204511.

(75) Sharif, S.; Denisov, G. S.; Toney, M. D.; Limbach, H. H. NMR studies of solvent-assisted proton transfer in a biologically relevant Schiff base: Toward a distinction of geometric and equilibrium H-bond isotope effects. *J. Am. Chem. Soc.* **2006**, *128*, 3375.

(76) Sharif, S.; Schagen, D.; Toney, M. D.; Limbach, H. H. Coupling of functional hydrogen bonds in pyridoxal-5′-phosphate-enzyme model systems observed by solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **2007**, *129*, 4440.

(77) Sharif, S.; Denisov, G. S.; Toney, M. D.; Limbach, H. H. NMR studies of coupled low- and high-barrier hydrogen bonds in pyridoxal-5′-phosphate model systems in polar solution. *J. Am. Chem. Soc.* **2007**, *129*, 6313.

(78) Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **2008**, *120*, 215.

(79) Alecu, I. M.; Zheng, J.; Zhao, Y.; Truhlar, D. G. Computational thermochemistry: Scale factor databases and scale factors for vibrational frequencies obtained from electronic model chemistries. *J. Chem. Theory Comput.* **2010**, *6*, 2872.

(80) Wong, K. F.; Sonnenberg, J. L.; Paesani, F.; Yamamoto, T.; Vanicek, J.; Zhang, W.; Schlegel, H. B.; Case, D. A.; Cheatham, T. E., III; Miller, W. H.; Voth, G. A. Proton transfer studied using a combined ab initio reactive potential energy surface with quantum path integral methodology. *J. Chem. Theory. Comp.* **2010**, *6*, 2566.

JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Coupled Cluster Theory on Graphics Processing Units I. The Coupled Cluster Doubles Method

A. Eugene DePrince, III*[,†] and Jeff R. Hammond[‡]

[†]Center for Nanoscale Materials and [‡]Leadership Computing Facility, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, Illinois 60439, United States

ABSTRACT: The coupled cluster (CC) ansatz is generally recognized as providing one of the best wave function-based descriptions of electronic correlation in small- and medium-sized molecules. The fact that the CC equations with double excitations (CCD) may be expressed as a handful of dense matrix−matrix multiplications makes it an ideal method to be ported to graphics processing units (GPUs). We present our implementation of the spin-free CCD equations in which the entire iterative procedure is evaluated on the GPU. The GPU-accelerated algorithm readily achieves a factor of 4−5 speedup relative to the multithreaded CPU algorithm on same-generation hardware. The GPU-accelerated algorithm is approximately 8−12 times faster than Molpro, 17−22 times faster than NWChem, and 21−29 times faster than GAMESS for each CC iteration. Single-precision GPU-accelerated computations are also performed, leading to an additional doubling of performance. Single-precision errors in the energy are typically on the order of 10^{-6} hartrees and can be improved by about an order of magnitude by performing one additional iteration in double precision.

1. INTRODUCTION

The accuracy and extensibility of computational chemistry methods, particularly those which approximately solve the Schrödinger equation, are ultimately limited by the speed at which computer processors can execute floating point and memory operations. Due to fundamental limitations in processor technology, clock speeds are not increasing, and all future increases in computational capability are expected to come from parallelism, which now more than ever can be found within a single processor. Graphics processing units (GPUs) are a type of massively parallel processor in which hundreds of cores can execute many instructions at once, provided they are sufficiently regular. Recently, many groups have demonstrated the incredible power of GPUs for scientific applications when sufficient effort is devoted to programming them to exploit their high degree of instruction-level parallelism.[1] The programmability of GPUs has increased dramatically with the NVIDIA CUDA API[2] and associated SDK including CUBLAS and CUFFT, although these tools as well as the vendor-independent alternative, OpenCL,[3] require more programming effort to realize the same relative performance as CPUs, especially for irregular algorithms.

To date, many computational chemistry methods have been implemented on GPUs (or other accelerators), including classical molecular dynamics,[4−7] atomic integrals,[8−10] density functional[11−14] and Hartree−Fock theory,[15,16] low-order perturbation theory,[17−19] and quantum Monte Carlo (QMC) for both fermions[20,21] and bosons.[22] Related efforts include the development of fast multipole methods for biomolecular electrostatics.[23] Notably lacking is an implementation of a high-accuracy many-body method, such as coupled cluster (CC) or configuration-interaction (CI), for GPUs. Both CC and CI have high floating point cost (N^6 or greater, where N is the number of electrons correlated) and are memory intensive, hence they are ideally suited for GPUs, which have significantly greater floating point capability and memory bandwidth than equivalently priced CPUs.

In this paper, we report the first demonstration of CC executed entirely on GPUs. Specifically, the coupled cluster doubles method (CCD) has been implemented using CUDA and the associated dense linear algebra routines (BLAS). The utility of BLAS, specifically dense matrix−matrix multiplication (MMM) kernels, to achieve high performance in coupled cluster methods is well-known, having been central to the implementation of CC for vector processors in the 1980s[24,25] followed by scalar and superscalar processors in the 1990s.[26] However, simply moving BLAS calls from the CPU to the GPU is not sufficient to achieve good performance. The modest amount of memory available on the GPU relative to the CPU requires the programmer to move data back and forth between the two devices across the PCI bus, which has limited bandwidth compared to that within each device. Simply inlining GPU BLAS calls would generate significantly more data motion than necessary and inhibit performance significantly. Our implementation of CCD minimizes memory motion both by organizing BLAS calls on the GPU in an optimal way and by performing all other computations (e.g., tensor permutations) on the GPU. We utilized both the independence of the different terms in CCD to decompose the calculation as well as by splitting large arrays into tiles when the total input and output required for a particular BLAS call could not fit on the GPU. These two strategies allow us to realize a significant fraction of the theoretical performance possible for CCD and are extensible to more complicated CC theories or other methods relying heavily upon tensor contractions. While the cost of tensor permutations is not significant in CCD, it will become a greater portion of the run time for CCSD

Received: October 13, 2010
Published: April 15, 2011

dx.doi.org/10.1021/ct100584w | *J. Chem. Theory Comput.* 2011, 7, 1287–1295

and higher-order CC methods, hence our code addresses an important issue for implementing more complex theories.

CCD is quite reasonable as a first demonstration of iterative CC on GPUs. The addition of singles provides a method (CCSD) which is not significantly more expensive,[27] but CCD contains all of the same computational bottlenecks, and its simplicity facilitates the design of an optimal algorithm. In addition, accurate energies require an approximate treatment of triples, e.g., CCSD(T), and it is well-known that the non-iterative triples contribution beyond CCSD is a series of very large MMMs even more amenable to GPUs than CCD.[28−30] Hence, the implementation strategy and performance analysis in this paper are immediately applicable to CCSD(T), which is the ultimate goal of our ongoing efforts in this area. Very recently, Kowalski and co-workers demonstrated an 8-fold speedup using GPUs for the computation of perturbative triples corrections.[30]

The performance of our implementation of spin-free CCD is analyzed in two ways. First, the same code is executed on both CPU and GPU; the code runs 4.0−5.2 times faster on a single NVIDIA C2050 GPU than it does on two Intel Xeon CPUs. In both cases, the hardware is fully utilized with vendor-optimized BLAS routines (CUBLAS and Intel MKL), and other routines are parallelized for the GPU using CUDA. Second, we compare to several implementations of CCD found in well-known electronic structure packages, including GAMESS,[31] NWChem,[32] and Molpro.[33] Running on the C2050, our CCD implementation outperforms all other CCD implementations by at least a factor of eight when using double precision throughout. An additional factor of two speed-up can be achieved by performing GPU computations in single precision. Single precision computations give errors on the order of 10^{-6} hartrees; a single iteration in double precision following convergence in single precision reduces the numerical error to 10^{-7} hartrees. While far from the "magic" 100-fold speed-up observed in other applications, both the CPU and GPU implementations use tuned BLAS libraries, which precludes a defective comparison between CPU and GPU implementations of vastly different quality.[34] In fact, the comparison of algorithms dominated by BLAS routines favors the CPU since we find that tuned BLAS libraries for CPUs achieve a higher percentage of theoretical peak than their GPU counterparts, although we find that the gap in implementation quality of BLAS has decreased significantly in CUDA 3.2.

This paper is organized as follows: In Section 2, the equations of CCD are presented, followed by a description of their implementation (Section 3). Performance results and analysis of numerical precision for hydrocarbons with as many as 20 carbon atoms can be found in Section 4. Section 5 contains our conclusions and a discussion of future work.

2. THEORY

A detailed perspective of coupled cluster theory is available in the literature,[26,35,36] so we describe only the equations necessary to understand the specifics of our algorithm. The cluster amplitudes, t_{ij}^{ab}, for the spin-free CC equations with double excitations are determined by the solution of the set of nonlinear equations

$$D_{ij}^{ab}t_{ij}^{ab} = v_{ij}^{ab} + P(ia,jb)\Big[t_{ij}^{ae}I_e^b - t_{im}^{ab}I_j^m + \frac{1}{2}v_{ef}^{ab}t_{ij}^{ef} + \frac{1}{2}t_{mn}^{ab}I_{ij}^{mn} - t_{mj}^{ae}I_{ie}^{mb} - I_{ie}^{ma}t_{mj}^{eb} + (2t_{mi}^{ea} - t_{im}^{ea})I_{ej}^{mb}\Big] \tag{1}$$

Table 1. Dimensions of the Matrix−Matrix Multiplications That Comprise the Spin-Free CCD Equations[a]

	dimension		
no. of occurrences	M	N	K
1	o^2	o^2	v^2
1	o^2	v^2	v^2
1	o^2	v^2	o^2
6	ov	ov	ov
1	v	v	o^2v
1	v	o^2v	v
1	o	o	ov^2
1	o	ov^2	o

[a] The dimensions correspond to C(M × N) = A(M × K)·B(K × N). The symbols o and v refer to the number of occupied and unoccupied orbitals in the reference function.

where we have slightly modified the tensors given by Piecuch et al.:[26]

$$I_b^a = (-2v_{eb}^{mn} + v_{be}^{mn})t_{mn}^{ea} \tag{2}$$

$$I_j^i = (2v_{ef}^{mi} - v_{ef}^{im})t_{mj}^{ef} \tag{3}$$

$$I_{kl}^{ij} = v_{kl}^{ij} + v_{ef}^{ij}t_{kl}^{ef} \tag{4}$$

$$I_{jb}^{ia} = v_{jb}^{ia} - \frac{1}{2}v_{eb}^{im}t_{jm}^{ea} \tag{5}$$

$$I_{bj}^{ia} = v_{bj}^{ia} + v_{be}^{im}\left(t_{mj}^{ea} - \frac{1}{2}t_{mj}^{ae}\right) - \frac{1}{2}v_{be}^{mi}t_{mj}^{ae} \tag{6}$$

$$D_{ij}^{ab} = f_{ii} + f_{jj} - f_{aa} - f_{bb} \tag{7}$$

Here the indices i, j, k, l, m, and n (a, b, c, d, e, and f) represent those orbitals that are occupied (unoccupied) in the reference function. We have used the Einstein summation convention in which repeated upper and lower indices are summed; note, however, that the left-hand side of eq 1 involves no sum. The permutation operator, $P(ia,jb)$, involves a sum of two terms: $P(ia,jb)v_{ij}^{ab} = v_{ij}^{ab} + v_{ji}^{ba}$. The Fock matrix elements are denoted by f_{pq}, and the two-electron integrals are given by

$$v_{ij}^{ab} = \int\int \varphi_a^*(1)\varphi_b^*(2)\frac{1}{r_{12}}\varphi_i(1)\varphi_j(2) \tag{8}$$

where ϕ represents a canonical molecular orbital. In this spin-free representation, eqs 1−6 require 9 MMMs that scale as the sixth power of system size and 4 that scale as the fifth power of system size. The dimensions for these multiplications are given in Table 1.

Equation 1 is solved iteratively, beginning with an MP2 guess for t_{ij}^{ab}. Evaluating the right-hand side of eq 1 yields the updated amplitudes. The algorithm proceeds by simple substitution until the convergence criterion is satisfied, when the norm of the change in t_{ij}^{ab} between iterations falls below 1×10^{-7}. If all integrals and amplitudes are stored on the device, the evaluation of this norm represents the only communication between device and host following the initial copy of the Fock matrix and the two-electron integrals to the GPU. However, if the v_{cd}^{ab} block of integrals is prohibitively large, the associated MMM may be blocked, and the integrals copied to the device as needed.

3. COMPUTATIONAL DETAILS

The CC equations presented in eqs 1−7 were implemented in both single (SP) and double precision (DP) for computations with CPU and GPU hardware. The CPU implementation evaluates the tensor contractions with dense matrix multiplication routines (SGEMM or DGEMM) provided by Intel MKL 10.2. The CPU hardware utilized was a dual socket quad-core 2.67 GHz Intel Xeon X5550 processor with 36 GB of available memory. Both NVIDIA C1060 Tesla and NVIDIA C2050 Tesla graphics processors with total memories of 4 and 2.6 GB (with ECC enabled), respectively, were used to perform GPU computations. See Table 2 for important hardware parameters for both the CPU and GPU processors used in this study. The SP and DP GPU implementations were achieved using an identical algorithm with the MKL BLAS routines substituted with the corresponding routines of the CUBLAS 3.2 library. The required one- and two-electron integrals were generated by the GAMESS electronic structure package on a CPU. For SP GPU computations, the Fock matrix was computed on the host in DP before demoting its elements to SP for use on the device.

Table 3 presents pseudocode for three different algorithms to evaluate one of the more complicated diagrams of CCD using either a CPU or GPU. The naive GPU implementation, labeled GPU 1, copies data to and from the GPU for nearly every MMM and performs all tensor permutations and amplitude updates on the CPU. These operations represent wasted opportunity for acceleration by the enhanced memory bandwidth of the GPU.

Table 2. Summary of Hardware Details for the Processors Used[a]

	CPU	GPU	
	X5550[b]	C1060[c]	C2050[d]
processor speed (MHz)	2660	1300	1150
memory bandwidth (GB/s)	32	102	144
memory speed (MHz)	1066	800	1500
ECC available	yes	no	yes
SP peak (GF)	85.1	933	1030
DP peak (GF)	42.6	78	515
power usage (W)	95[e]	188	238

[a] Note that details are given for a single Intel X5550 processor, but two processors were used for all CPU calculations. [b] http://ark.intel.com/Product.aspx?id=37106. [c] http://www.nvidia.com/object/product_tesla_c1060_us.html. [d] http://www.nvidia.com/object/product_tesla_C2050_C2070_us.html. [e] CPU power usage does not include CPU DRAM, whereas GPU power usage includes GPU DRAM.

Table 3. CPU and GPU Algorithms for Evaluating One Diagram of CCD[a]

CPU	GPU 1	GPU 2
$I_{jb}^{ia} = v_{jb}^{ia} - 1/2 v_{eb}^{im} t_{jm}^{ea}$		
		$\textbf{cudaMemcpy}(t_{\text{gpu}} \leftarrow t(e,a,j,m))$
		$\textbf{cudaMemcpy}(v_{\text{gpu}}^{1} \leftarrow v(i,m,e,b))$
		$\textbf{cudaMemcpy}(v_{\text{gpu}}^{2} \leftarrow v(i,a,j,b))$
$t'(a,j,m,e) \leftarrow t(e,a,j,m)$	$t'(a,j,m,e) \leftarrow t(e,a,j,m)$	$t'_{\text{gpu}}(a,j,m,e) \leftarrow t_{\text{gpu}}(e,a,j,m)$
$v^1(m,e,i,b) \leftarrow v(i,m,e,b)$	$v^1(m,e,i,b) \leftarrow v(i,m,e,b)$	$v'_{\text{gpu}}(m,e,i,b) \leftarrow v_{\text{gpu}}^{1}(i,m,e,b)$
$I(a,j,i,b) = v(i,a,j,b)$	$v^2(a,i,j,b) \leftarrow v(i,a,j,b)$	$I_{\text{gpu}}(a,j,i,b) = v_{\text{gpu}}^{2}(i,a,j,b)$
	$\textbf{cudaMemcpy}(v'_{\text{gpu}} \leftarrow v^1)$	
	$\textbf{cudaMemcpy}(t'_{\text{gpu}} \leftarrow t')$	
	$\textbf{cudaMemcpy}(I_{\text{gpu}} \leftarrow v^2)$	
$I(a,j,i,b) - = 1/2t'(a,j,m,e)$- $v^1(m,e,i,b)$ (**DGEMM**)	$I_{\text{gpu}}(a,j,i,b) - = 1/2t'_{\text{gpu}}(a,j,m,e)$- $v'_{\text{gpu}}(m,e,i,b)$ (**cublasDgemm**)	$I_{\text{gpu}}(a,j,i,b) - = 1/2t'_{\text{gpu}}(a,j,m,e)$- $v'_{\text{gpu}}(m,e,i,b)$ (**cublasDgemm**)
$R_{ij}^{ab} + = -t_{mj}^{ae} I_{ie}^{mb} - I_{ie}^{ma} t_{mj}^{eb}$		
$R'(b,i,a,j) = I(b,i,m,e)[t'(a,j,m,e)]^T$ (**DGEMM**)	$v'_{\text{gpu}}(b,i,a,j) = I_{\text{gpu}}(b,i,m,e)[t'_{\text{gpu}}(a,j,m,e)]^T$ (**cublasDgemm**)	$v'_{\text{gpu}}(b,i,a,j) = I_{\text{gpu}}(b,i,m,e)[t'_{\text{gpu}}(a,j,m,e)]^T$ (**cublasDgemm**)
	$\textbf{cudaMemcpy}(R' \leftarrow v'_{\text{gpu}})$	
$R(a,b,i,j) - = R'(b,i,a,j) + R'(a,j,b,i)$	$R(a,b,i,j) - = R'(b,i,a,j) + R'(a,j,b,i)$	$R_{\text{gpu}}(a,b,i,j) - = v'_{\text{gpu}}(b,i,a,j) + v'_{\text{gpu}}(a,j,b,i)$
$t'(m,e,b,j) \leftarrow t(e,b,m,j)$	$t'(m,e,b,j) \leftarrow t(e,b,m,j)$	$t'_{\text{gpu}}(m,e,b,j) \leftarrow t_{\text{gpu}}(e,b,m,j)$
	$\textbf{cudaMemcpy}(t'_{\text{gpu}} \leftarrow t')$	
$R'(a,i,b,j) = I(a,i,m,e)t'(m,e,b,j)$ (**DGEMM**)	$v'_{\text{gpu}}(a,i,b,j) = I_{\text{gpu}}(a,i,m,e)t'_{\text{gpu}}(m,e,b,j)$ (**cublasDgemm**)	$v'_{\text{gpu}}(a,i,b,j) = I_{\text{gpu}}(a,i,m,e)t'_{\text{gpu}}(m,e,b,j)$ (**cublasDgemm**)
	$\textbf{cudaMemcpy}(R' \leftarrow v'_{\text{gpu}})$	
$R(a,b,i,j) - = R'(a,i,b,j) + R'(b,j,a,i)$	$R(a,b,i,j) - = R'(a,i,b,j) + R'(b,j,a,i)$	$R_{\text{gpu}}(a,b,i,j) - = v'_{\text{gpu}}(a,i,b,j) + v'_{\text{gpu}}(b,j,a,i)$
update t with R	update t with R	update t_{gpu} with R_{gpu}
GPU storage requirements		
0	$3o^2v^2$	$5o^2v^2$

[a] Note that for both GPU implementations, the temporary array v'_{gpu} can be reused. The naive GPU implementation, GPU 1, performs many unnecessary memory transfers and fails to exploit the high memory bandwidth of the GPU for additions and tensor permutations. For GPU implementation 2, all memory transfers can occur before CCD iterations begin, and the amplitudes can be directly updated on the GPU. R denotes the residual of the CCD equations and is defined by the right-hand side of eq 1.

The algorithm labeled GPU 2 is a better implementation that moves data between host and device as seldom as possible and performs all tensor permutations and updates to the CC amplitudes on the GPU. Algorithm GPU 2 is designed such that all memory transfers occur before the CC iterations begin. By monitoring convergence on the device, the amplitudes need never be copied from the GPU. Algorithm GPU 2 has greater memory requirements than GPU 1 for the evaluation of this diagram, but the temporary arrays used therein are useful in the evaulation of the remainder of the CCD equations, and the difference in storage requirements is insignificant when considering the storage of the v_{cd}^{ab} block of integrals. The present GPU implementation is most similar to GPU 2.

For systems with hundreds of active orbitals, the limited memory of GPUs necessitates an algorithm which repeatedly copies data from CPU to GPU memory. In single precision, the global memory of the C1060 GPU can accommodate the v_{cd}^{ab} block of integrals (the largest array required by CCD or CCSD) for no more than 181 virtual orbitals, and the storage requirements for the CC amplitudes and all other blocks of integrals further limit the size of systems which can be treated before such repeated memory transfers become necessary. In our algorithm, all two-electron integrals and CC amplitudes are stored in GPU memory whenever possible. In addition, we allocate several arrays to contain convenient permutations of the amplitudes and smaller blocks of integrals. While on-device storage of the two-electron integrals and CC amplitudes will ultimately limit the size of the applications that may be treated with our code, the lack of significant host/device communication will result in a best case for performance acceleration. To extend the applicability of this implementation, large MMMs can be blocked, and integrals copied to the GPU on-demand. Such tiling of the matrix multiplication involving the v_{cd}^{ab} block of two-electron integrals will alleviate some of these memory limitations but will result in the new overhead of transferring the integrals to the device every iteration. A different approach to large matrix multiplications is the streaming approach employed in LINPACK, wherein the matrix dimensions are much greater than 10 000. We are currently exploring whether this approach is suitable for the evaluation of the CC equations. The two-electron integral memory bottleneck may be entirely avoided through integral-direct techniques, provided one has an atomic integral code that can run on the GPU.

3.1. Comparison to Other Codes. The performance of the multicore CPU and GPU implementations of CCD is compared to a number of implementations in well-known packages. Only the GAMESS package implements the same equations as ours does, but we have also compared to the Molpro and NWChem implementation of CCD. Timings for CCSD are reported for these three packages as well as PSI3[37] to estimate the performance of a GPU CCSD code. It is important to point out if and how each CPU implementation is parallelized, given the essential role of fine-grain parallelism in our implementation. None of the codes tested have threading in their CC codes. In principle, all can take advantage of threading in the BLAS library, but this only improves performance for large matrices (dimension $\gtrsim$500). Both NWChem and Molpro are parallelized using global arrays. It was previously determined that NWChem performance is no worse, and in some cases significantly better, when one core is dedicated to communication, which is an artifact of optimizing interprocess communication within the node.[38] Hence, all NWChem and Molpro jobs used only seven cores for computation. For consistency, all computations are performed in C_1 symmetry.

3.1.1. GAMESS. The CCD algorithm contained in the GAMESS package is detailed in ref 26 and is essentially identical to that presented herein. The algorithm makes heavy use of BLAS DGEMM calls, but these are constrained by design to only use one thread, significantly restricting performance. We note that the present algorithm, when executing on a single CPU core, performs nearly identically to the GAMESS algorithm.

3.1.2. NWChem. NWChem implements CC in two different modules: the first is spin-free and AO-direct in the three- and four-virtual integrals,[39] while the second module (TCE) uses the more expensive spin−orbital CC equations.[40] Due to the much larger memory footprint of the two-electron integrals in spin−orbital form, a comparison could be made to this procedure only for small systems. However, the TCE module also permits of the use spin-free integrals with spin−orbital amplitudes, which provides the generality of the spin−orbital representation with a much-reduced memory footprint.[41]

Neither the TCE implementation of CCD and CCSD nor the partially direct spin-free implementation of CCSD in NWChem are directly comparable to our CCD code. However, as NWChem is parallel throughout, it is capable of utilizing a large number of CPU cores per node, unlike GAMESS.

3.1.3. Molpro. Molpro[33] implements CCD as a special case of spin-free CCSD where single excitations are set to zero. The CCD algorithm is thus not optimal and performs identically to the CCSD algorithm. The spin-free implementation is described in ref 42 .

4. RESULTS

Three different methods were used to evaluate the performance of CCD on CPU and GPU processors. First, we compare both our CPU and GPU implementations of CCD to those found in GAMESS, Molpro, and NWChem. Second, the performance benefit of using single-precision is considered due to the large gap in single- and double-precision performance of some GPUs. Finally, the performance of MMM was measured in both single and double precision on the CPU and GPU to establish an approximate upper bound on the performance of a CC code implemented using BLAS for tensor contractions. The performance of the present implementations is compared to that of the underlying BLAS routines.

The polyacetylene series, C_nH_{n+2}, with n ranging from 8 to 18, the acene series for 2-, 3-, and 4-fused benzene rings, and the smallest fullerene, C_{20}, were used to evaluate the performance of the CC implementations. The 6-31G basis set used throughout was not selected on the basis of chemical considerations but rather to allow us to consider a wide range of system sizes. By using a relatively small basis set, more emphasis is placed upon the performance of tensor contractions involving more than two occupied indices. A very large basis set places almost all the computational work in the evaluation of a single diagram involving four virtual indices. While this may provide a very high flop rate due to the presence of a very large MMM call, it is not particularly useful for evaluating implementation quality.

4.1. Comparison of CPU and GPU Implementations. The per iteration computational costs of our double-precision GPU and CPU algorithms are compared to implementations found in well-known electronic structure packages in Table 4. For our CPU and GPU implementations, the timings correspond to

Table 4. Comparison of CPU and GPU Implementations of CCD[a]

			present implementation			X5550		
molecule	o	v	C1060	C2050	X5550	GAMESS	Molpro	NWChem (TCE)[b]
C_8H_{10}	21	63	0.8	0.3	1.3	6.2	2.3	5.1 (4.7)
$C_{10}H_8$	24	72	1.5	0.5	2.5	12.7	4.8	10.6 (10.1)
$C_{10}H_{12}$	26	78	2.5	0.8	3.5	19.7	7.1	16.2 (15.1)
$C_{12}H_{14}$	31	93	7.1	2.0	10.0	57.7	17.6	42.0
$C_{14}H_{10}$	33	99	10.2	2.7	13.9	78.5	29.9	59.5
$C_{14}H_{16}$	36	108	16.7	4.5	21.6	129.3	41.5	90.2
C_{20}	40	120	29.9	8.8[c]	40.3	238.9	103.0	166.3
$C_{16}H_{18}$	41	123	35.9	10.5[c]	50.2	279.5	83.3	190.8
$C_{18}H_{12}$	42	126	42.2[c]	12.7[c,d]	50.3	329.4	111.8	218.4
$C_{18}H_{20}$	46	138	73.0[c]	20.1[c,d]	86.6	555.5	157.4	372.1

[a] Timings per CC iteration are given in seconds. The symbols o and v represent the number of doubly occupied and virtual orbitals in each system, respectively. [b] Numbers in parentheses are for uncompacted (spin−orbital) two-electron integrals, which runs out of memory for even medium-sized jobs. [c] The MMM involving v_{cd}^{ab} was tiled. [d] Some two-electron integrals were pushed to the GPU every iteration.

Table 5. Comparison of GAMESS, Molpro, NWChem, and PSI3 Implementations of CCD and CCSD[a]

	GAMESS		Molpro		NWChem			PSI3
molecule	CCD	CCSD	CCD	CCSD	CCSD[b]	CCD[c]	CCSD[c]	CCSD
C_8H_{10}	6.2	7.2	2.3	2.4	9.6/3.6	5.1	8.4	7.9
$C_{10}H_8$	12.7	15.3	4.8	5.1	22.8/8.2	10.6	16.8	17.9
$C_{10}H_{12}$	19.7	23.6	7.1	7.2	20.5/11.3	16.2	25.2	23.6
$C_{12}H_{14}$	57.7	65.1	17.6	19.0	53.6/29.4	42.0	64.4	54.2
$C_{14}H_{10}$	78.5	92.9	29.9	31.0	92.7/49.1	59.5	90.7	61.4
$C_{14}H_{16}$	129.3	163.7	41.5	43.1	103.2/65.0	90.2	129.2	103.4
C_{20}	238.9	277.5	103.0	102.0	294.6/175.7	166.3	233.9	162.6
$C_{16}H_{18}$	279.5	345.8	83.3	84.1	169.1/117.5	190.8	267.9	192.4
$C_{18}H_{12}$	329.4	380.0	111.8	116.2	274.2/178.6	218.4	304.5	216.4
$C_{18}H_{20}$	555.5	641.3	157.4	161.4	278.1/216.3	372.1	512.0	306.9

[b] The spin-free CCSD code in NWChem is integral direct for the terms with integrals having 3 or 4 virtual indices. The first number is the first iteration when the stored integrals (those with 0−2 virtual indices) are computed and written to disk; the second number is for subsequent iterations when the stored integrals are read from disk. [c] TCE implementation with spin−orbital CC equations but using compacted (spin-free) integrals. [a] The data given are seconds per CC iteration.

those required to evaluate all tensors given by eqs 2−6 and to update the CC amplitudes according to eq 1. Timings for the GPU version exclude the initial integral push to the device. For standard packages, timings correspond to only the iterative portions of the CCD algorithms; integral generation and sorting were excluded. The four-index transformation and the related I/O and processing are excluded for three reasons: (1) the implementation of these procedures varies greatly between different packages, (2) for larger calculations the time required to generate the integrals is insignificant, as it scales as N^5 while many CC diagrams scale as N^6, and (3) the transfer time from the CPU to GPU will disappear when it is possible to compute all integrals directly on the GPU.

The C2050 GPU-accelerated CCD algorithm outperforms all other implementations on a per iteration basis. As compared to the threaded CPU implementation, we observe application accelerations of 4.0−5.2 for all systems considered. In addition to the acceleration of MMMs, all required transposes are performed on the GPU, which has significantly higher memory bandwidth than the CPU, making these operations much faster. The C1060 GPU-accelerated algorithm performs anywhere from 2.6 to 3.7 times worse than the C2050 algorithm. This generation of hardware provides relatively poor double-precision performance and is thus not optimal for general-purpose scientific computing.

Aside from our CPU algorithm, the best CPU CCD implementation presented here is that found in Molpro. Because CCD is implemented in Molpro as a special case of CCSD, the addition of singles results in a marginal increases in computational cost, and the comparison to the GPU-accelerated CCD code is thus slightly biased in favor of the GPU. Regardless of this deficiency, the Molpro CCD implementation is still the best available, and the GPU-accelerated algorithm consistently outperforms it by a factor of 8−12. For the largest system studied, $C_{18}H_{20}$, a single iteration in Molpro requires about 2.5 min, while a C2050 iteration requires only 20 s.

For comparison, we present in Table 5 the costs for both CCD and CCSD in GAMESS, Molpro, and NWChem. We have also included the CCSD timings for PSI3, which does not implement the CCD method. The addition of single excitations does not significantly increase the cost of CCD, as supported by the relative costs of CCD and CCSD given here. The GAMESS CCD and CCSD timings suggest that the addition of single excitations in our algorithm will increase the cost of the serial CPU algorithm by around 25%. The cost of the threaded CCSD code may be slightly worse than expected due to a less-than-optimal performance of threaded MKL BLAS for some of the very small MMMs that arise in CCSD. Nonetheless, we predict that the CCSD algorithm given in ref 26 would be several times faster than the best available CPU-based CCSD algorithms if implemented properly for GPU hardware.

As stated above, the limited global memory associated with GPUs limits the size of the applications that may be directly treated by eq 1, especially in double precision. For larger systems, the v_{cd}^{ab} block of integrals will not fit into global memory, and it is necessary to break up the associated MMM into multiple calls for subblocks of the input arrays (known as tiling). Those instances in which tiling was necessary are noted accordingly in Table 4. In

Table 6. SP and DP GPU and CPU Timings in Seconds[a]

	time (s)							
	C1060		C2050		X5550		error (μE_h)	
molecule	SP	DP	SP	DP	SP	DP	SP	mixed
C_8H_{10}	0.2	0.8	0.2	0.3	0.7	1.3	0.05	0.01
$C_{10}H_8$	0.4	1.5	0.2	0.5	1.3	2.5	−0.42	−0.04
$C_{10}H_{12}$	0.7	2.5	0.4	0.8	2.0	3.5	−0.13	−0.02
$C_{12}H_{14}$	1.8	7.1	1.0	2.0	5.6	10.0	−0.30	−0.04
$C_{14}H_{10}$	2.6	10.2	1.5	2.7	8.4	13.9	−3.74	−0.35
$C_{14}H_{16}$	4.1	16.7	2.4	4.5	12.1	21.6	−1.00	−0.16
C_{20}	6.7	29.9	4.1	8.8[b]	22.3	40.3	−1.43	0.09
$C_{16}H_{18}$	9.0	35.9	5.0	10.5[b]	28.8	50.2	−2.66	−0.44
$C_{18}H_{12}$	10.1	42.2[b]	5.6	12.7[b,c]	29.4	50.3	−15.03	−1.30
$C_{18}H_{20}$	17.2	73.0[b]	10.1[b]	20.1[b,c]	47.0	86.6	−5.72	−0.91

[a] Errors in the energy for single- and mixed-precision algorithms are presented. The mixed-precision algorithm converges in SP and performs one iteration in DP. All errors are given in units of μE_h (10^{-6} hartrees). [b] The MMM involving v_{cd}^{ab} was tiled. [c] Some two-electron integrals were pushed to the GPU every iteration.

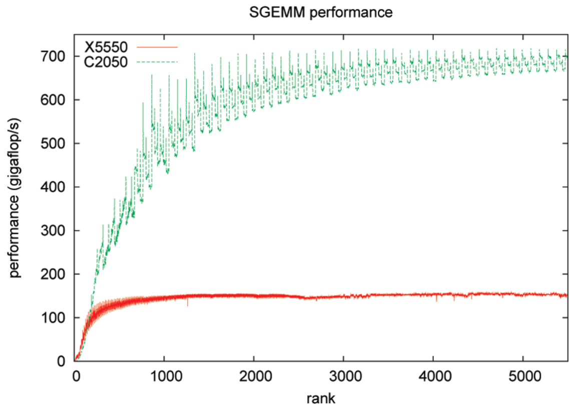


Figure 1. Performance in gigaflop/s (10^9 floating point operations per second) for SGEMM on CPU and GPU devices. The CPU SGEMM implementation utilizes eight threads. The maximum gigaflop/s for the CPU and GPU are 156.2 and 717.6, respectively.

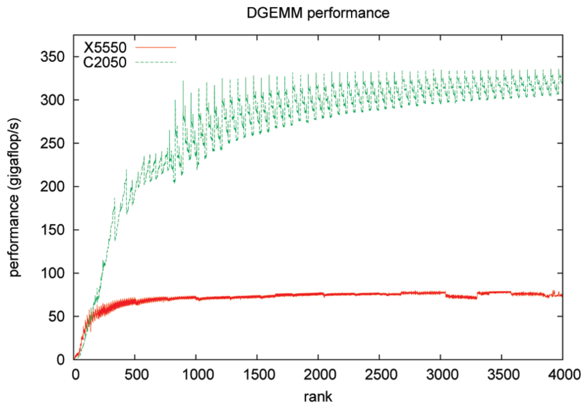


Figure 2. Performance in gigaflop/s (10^9 floating point operations per second) for DGEMM on CPU and GPU devices. The CPU DGEMM implementation utilizes eight threads. The maximum gigaflop/s for the CPU and GPU are 79.2 and 335.6, respectively.

addition, some permutations of integrals that scale as o^2v^2 were, in some cases, pushed to the GPU as needed. The timings presented in Table 4 include all memory transfers occurring during the iterative portions of the algorithm, and the overhead associated with these transfers has minimal implications for performance. In the case of $C_{18}H_{20}$, the v_{cd}^{ab} block of integrals represents 2.7 GB of data. These data, as well as several hundreds of MB of o^2v^2 integrals, are pushed to the GPU every iteration; regardless, the GPU-accelerated algorithm is more than four times faster per iteration than the corresponding CPU implementation and almost eight times faster than the Molpro package. This result suggests that larger systems can be treated with this algorithm provided that the larger MMMs are appropriately tiled. The notion of tiling leads naturally to a framework for many GPU CC; scalable parallelization can be realized by distributing comparably sized tiles among many GPUs.

4.2. Impact of Numerical Precision upon Accuracy and Performance. Many implementations of scientific algorithms on GPU hardware utilize single or mixed precision due to the markedly reduced DP performance of older graphics cards. As accelerator software and hardware mature for HPC, GPUs are becoming increasingly efficient at performing DP operations. On the other hand, most commodity graphics processors cannot support DP. It is important to understand the performance advantages of low-precision computing as weighed against the disadvantages. We have implemented our CCD algorithm in SP on the same CPU and GPU hardware discussed above, and the SP timings are presented in Table 6. In general, we observe the same GPU/CPU accelerations for SP computations as we did for strictly DP computations. On the C2050 processor, SP per iteration costs are 4.2−5.8 times less than that of the SP costs of the CPU-based threaded MKL algorithm, and C2050 SP operations are roughly half the cost of DP operations. The advantages of SP computations are more evident for the older C1060 card, where DP is 3.9−4.4 times more expensive than SP.

Table 6 also lists the energy errors associated with the SP algorithm. For all systems investigated, the observed SP errors are at worst on the order of 10^{-5} E_h. Chemical accuracy is considered to be kcal/mol, which corresponds to roughly 1.6 × 10^{-3} E_h. Clearly, SP CCD on the GPU yields more than acceptable accuracy for single-point energy evaluations. Should the need for higher accuracy arise, it is not difficult to design an algorithm in which we converge to a SP solution, promote the amplitudes to DP, and perform a single iteration in DP. The results in Table 6 labeled "mixed" represent such an algorithm. A single CC iteration in DP on the C2050 graphics card costs about two times as much as a single iteration in SP and can increase the accuracy of the computation by an order of magnitude. Five of the mixed precision errors in Table 6 are below 1×10^{-7} E_h; the largest error decreased from −15.03 μE_h to only −1.30 μE_h. For larger systems, the SP energy errors may be larger, necessitating further DP iterations. Total application speedup is only marginally affected by the final DP iteration.

4.3. Matrix−Matrix Multiplication Performance. The performance of SGEMM and DGEMM with the X5550 using MKL and C2050 using CUBLAS 3.2 are shown in Figures 1 and 2, respectively. MKL SGEMM and DGEMM implementations are threaded and utilize eight threads. The GPU is superior in performance to the CPU for larger matrices (dimension greater than ~600 for SGEMM), and the performance of the GPU is

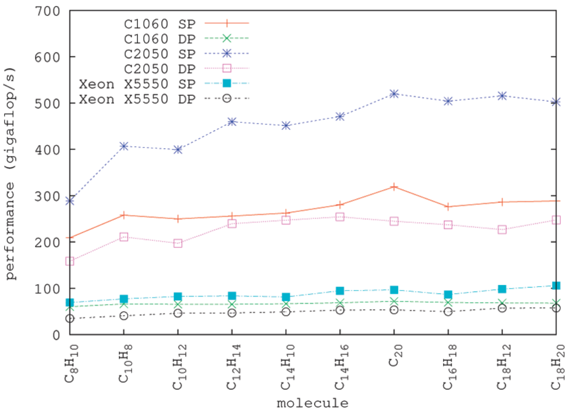


Figure 3. Performance in gigaflop/s (10^9 floating point operations per second) for different implementations of spin-free CCD. Results are given for both CPU and GPU hardware, and CPU BLAS routines utilize eight threads. S SP and DP performance data are given.

more than four times better than that of the CPU in the limit of very large matrices. The best-case performance of the GPU for SGEMM is greater than 700 gigaflop/s (GF), which is approximately 70% of the theoretical peak performance. On the C2050, the performance of DGEMM is approximately half of that of SGEMM, which again is approximately 70% of peak. The 2:1 ratio of single- and double-precision performance is new as of the Fermi GPU architecture; a much larger ratio of single to double precision performance was observed with the Tesla GPU architecture (C1060). The variation in the performance as a function of matrix dimension on the C2050 is more than on the CPU but significantly less than previous generations of NVIDIA hardware and software. The absolute performance of SGEMM and DGEMM also changed significantly upon the release of CUDA 3.2. Prior to this release, algorithm performance was tightly coupled to system size and could be maximized by padding matrix dimensions in an effort to match warp sizes. However, as of CUDA 3.2, padding appears to have been integrated into the implementation of BLAS, and our manual implementation of padding no longer improves performance. Finally, previous generations of CUBLAS achieve only ~50% of peak for very large matrices, whereas the latest version achieves ~70% of peak.

Assuming no overhead for computing transposes and data transfers, the CPU and GPU CCD algorithms should achieve the flop rates of the underlying DGEMM and SGEMM kernels depicted in Figures 1 and 2. Figure 3 depicts the performance of the CPU and GPU implementations of spin-free CCD for all of the systems studied here. The SP CPU implementation utilizes threaded MKL SGEMM calls and achieves 69−105 GF on the Intel Xeon X5550 CPU. The DP implementation achieves only half of that flop rate, 35−57 GF. The C2050 is capable of delivering at or near 500 GF of SP performance for systems larger than and including C_{20}, while DP performance approaches 250 GF. The release of CUDA 3.2 provides enhanced DP capabilities on the C2050 graphics card that is well beyond those of older Tesla products. For the C1060, we observe considerably lower flop rates of only 209−319 (60−72) GF for SP (DP).

5. CONCLUSIONS

We have reported the first implementation of the iterative procedures of any coupled cluster method running entirely on a GPU. We find that the NVIDIA C2050 graphics processor can achieve approximately 500 gigaflop/s performance in single precision (SP) and 250 gigaflop/s in double precision (DP). This performance translates to per iteration accelerations of 4.2−5.8 for SP and 4.0−5.2 for DP as compared to the multithreaded CPU implementation. The quality of both the CPU and GPU implementations are as similar as possible, as both employ the vendor-optimized BLAS libraries provided by Intel and NVIDIA, respectively. To the best of our knowledge, this is the first time such a direct evaluation of hardware performance has been undertaken for any quantum chemistry kernel; previous papers compare an optimized GPU implementation to a standard package written for CPUs or evaluated the impact of using GPUs for a limited set of procedures. In contrast, the entire iterative CC procedure is computed using the GPU, and the overhead of transferring integrals from CPU memory is demonstrated to be nominal. As most CPU packages are neither multithreaded (e.g., with OpenMP) nor optimized for vector floating point instructions (e.g., SSE3 for x86 processors), such a comparison is intrinsically unfair to the CPU. Unless special effort has been devoted to architecture-specific optimization of CPU code, as is done for BLAS calls, it is impossible to make a fair comparison of CPU and GPU hardware.

Our DP implementation running on the C2050 GPU processor is more than an order of magnitude faster than several well-known electronic structure packages for the CC iterative procedure, which dominates the total wall time of a CC calculation (when neglecting the perturbative triples correction). Specifically, the DP algorithm was shown to be 8−12 times faster than Molpro and 17−22 times faster than NWChem when each is executed on two quad-core CPUs. Our DP implementation is 21−29 times faster than the serial implementation of CCD implmented in GAMESS executing on a single CPU core. It is important to point out that none of the software packages tested make efficient use of multicore CPUs using threads. GAMESS is constrained by design to run on a single CPU core and using multiple threads in BLAS did not improve the performance of PSI3 as much as it did ours. For example, our CCD iteration timings for $C_{12}H_{14}$ reduced from 51 to 10 s when the number of threads increased from 1 to 8, whereas we found that the performance of PSI3 improved by less than 50% for the same 8-fold increase in thread utilization. While it was possible for NWChem and Molpro to utilize two quad-core CPUs using global arrays, this process-based parallelism necessarily divides the data into smaller chunks, reducing the efficiency of BLAS calls. In the end, the best predictor of the performance improvement of CC codes using GPUs instead of CPUs is our own CPU code, for which the performance improvement of four to five times is in good agreement with the relative performance of SGEMM and DGEMM we measured.

Because most GPUs have modest DP performance relative to what can be done in SP, computing in SP or some mixture of SP and DP can improve performance, provided the results are still numerically accurate. We demonstrate that CC is amenable to a very simple multiprecision algorithm due to its iterative nature and that we can converge to a standard DP threshold while performing all but one iteration in SP. While the state-of-the art NVIDIA C2050 (Fermi) processor has a similar ratio of SP to DP performance, as is found on CPU hardware, the older NVIDIA C1060 (Tesla) architecture and noncompute-oriented commodity GPUs have a much larger discrepancy between SP and DP. While the Fermi architecture may be more relevant to computational chemists using dedicated high-performance computing resources, which are likely to be equipped with more expensive hardware that is more suitable for scientific computation, the commodity GPU hardware found in laptop and desktop computers is likely to continue to provide substantially more performance in SP than DP. Thus, our mixed precision algorithm will still be relevant in the future. We also note that the mixed precision approach can be used on the CPU as well, but the performance gain will not be more than two times, as all modern CPU architectures we are aware of are optimized for DP floating point computation.

The present implementation of GPU-accelerated CCD attempts to store all two-electron integrals and CC amplitudes in global memory on the GPU device and is thus limited in its applicability to systems with less than 200 spatial orbitals. We have experimented with tiling the multiplication involving the v_{cd}^{ab} block of integrals to allow us to treat systems as large as $C_{18}H_{20}$ in a 6-31G basis in full double precision on the C2050 card, which has only 2.6 GB of global memory. It was shown that this system could be treated roughly four times more efficiently by the C2050 GPU than with the X5550 CPU, despite the fact that 3 GB of integrals were transferred to the GPU every iteration. An algorithm dominated by tiled DGEMM calls is also naturally amenable to parallelization. In the extension of this algorithm to multiple GPUs, the cost of the required MPI collectives will eventually dominate the integral push−pull time. Regardless of these arguments for tiled matrix multiplications, the memory limitation is completely artificial in the sense that it is coupled to current hardware limitations and will therefore change as GPU hardware matures for scientific applications; the NVIDIA C2070 card, which was not used in these experiments, has 6 GB of global memory.

The tremendous performance increase observed for these moderate systems can have profound implications for computations that require multiple energy evaluations. An obvious target of fast CC calculations on GPUs is ab initio molecular dynamics of small molecules. Such calculations might require higher than SP for accurate results, but as has been shown herein, it is trivial to design a mixed precision algorithm that can yield DP accuracy in SP time. Additionally, computations on clusters of GPUs would be ideal for local correlation approximations such as the clusters-in-molecule (CIM) approximation.[43,44] The CIM-CC methods are embarrassingly parallel, and the 1×10^{-6} E_h error for SP GPU algorithms is negligible compared to the corresponding CIM errors, which can be three orders of magnitude larger. Thus, it would be straightforward to utilize many GPUs by performing each cluster simulation on a single GPU, since no data needs to be communicated between subsystem calculations after the original partitioning of the molecule into clusters.

Future work will include the implementation of the full CCSD equations for similar studies in GPU acceleration. Based upon our current results and the performance of the GAMESS implementations of CCD and CCSD, we expect GPU-accelerated CCSD to be several times more efficient than any existing CPU implementation. Additionally, preliminary tests of our implementation of CC for multiple GPUs suggest that distributing independent diagrams allows for the utilization of 5−10 GPUs. Scaling to more than 10 GPUs requires breaking up a single diagram computation across multiple GPUs, which is more difficult due to increased communication but certainly possible for larger calculations. Because CCSD has more diagrams and because these vary greatly in computational cost, more care is required to load balance these calculations to fully utilize all available processor resources. However, we believe that significant speed-ups can be obtained by overlapping CPU and GPU computations in a hybrid CPU-GPU implementation of CCSD.

AUTHOR INFORMATION

Corresponding Author

*E-mail: adeprince@anl.gov.

ACKNOWLEDGMENT

A.E.D. is supported by the Computational Postdoctoral Fellowship program at Argonne National Laboratory. J.R.H. is supported by the Director's Postdoctoral Fellowship program at Argonne National Laboratory and thanks Alán Aspuru-Guzik for interesting discussions and encouragement.

This research used the "Dirac" GPU test bed system of the National Energy Research Scientific Computing Center, which is supported by the Office of Science of the U.S. Department of Energy under contract no. DE-AC02-05CH11231.

We gratefully acknowledge use of "Breadboard," a development computing cluster operated by the Mathematics and Computer Science Division at Argonne National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under contract no. DE-AC02-06CH11357.

We gratefully acknowledge use of "Fusion," a computing cluster operated by the Mathematics and Computer Science Division at Argonne National Laboratory as part of its Laboratory Computing Resource Center, which is supported by the Office of Science of the U.S. Department of Energy under contract no. DE-AC02-06CH11357.

REFERENCES

(1) The line between instruction- and thread-level parallelism in a GPU is blurred relative a CPU, but we consider them to be effectively single-instruction multiple-data (SIMD) processors.

(2) *CUDA Programming Guide*; NVIDIA: Santa Clara, CA; http://developer.download.nvidia.com/compute/cuda/3_2_prod/toolkit/docs/CUDA_Toolkit_Reference_Manual.pdf. Accessed March 10, 2011).

(3) Stone, J. E.; Gohara, D.; Shi, G. *Comput. Sci. Eng.* **2010**, *12*, 66.

(4) Stone, J. E.; Phillips, J. C.; Freddolino, P. L.; Hardy, D. J.; Trabuco, L. G.; Schulten, K. *J. Comput. Chem.* **2007**, *28*, 2618.

(5) Anderson, J. A.; Lorenz, C. D.; Travesset, A. *J. Comput. Phys.* **2008**, *227*, 5342.

(6) Liu, W.; Schmidt, B.; Voss, G.; Møller-Wittig, W. *Comput. Phys. Commun.* **2008**, *179*, 634.

(7) Friedrichs, M. S.; Eastman, P.; Vaidyanathan, V.; Houston, M.; Legrand, S.; Beberg, A. L.; Ensign, D. L.; Bruns, C. M.; Pande, V. S. *J. Comput. Chem.* **2009**, *30*, 864.

(8) Ufimtsev, I. S.; Martnez, T. J. *J. Chem. Theory Comput.* **2008**, *4*, 222.

(9) Asadchev, A.; Allada, V.; Felder, J.; Bode, B. M.; Gordon, M. S.; Windus, T. L. *J. Chem. Theory Comput.* **2010**, *6*, 696.

(10) Titov, A. V.; Kindratenko, V. V.; Ufimtsev, I. S.; Martínez, T. J. In Proceedings of Symposium on Application Accelerators in High-Performance Computing (SAAHPC) ,Knoxville, TN, July 13−15, 2010; National Center for Supercomputing Applications, University of Illinois: Urbana-Champaign, IL, 2010.

(11) Yasuda, K. *J. Chem. Theory Comput.* **2008**, *4*, 1230.

(12) Woods, C. J.; Brown, P.; Manby, F. R. *J. Chem. Theory Comput.* **2009**, *5*, 1776.

(13) Genovese, L.; Ospici, M.; Deutsch, T.; Mehaut, J.-F.; Neelov, A.; Goedecker, S. *J. Chem. Phys.* **2009**, *131*, 034103.

(14) Brown, P.; Woods, C. J.; McIntosh-Smith, S.; Manby, F. R. *J. Comput. Chem.* **2010**, *31*, 2008.

(15) Ufimtsev, I. S.; Martnez, T. J. *J. Chem. Theory Comput.* **2009**, *5*, 1004.

(16) Ufimtsev, I. S.; Martnez, T. J. *J. Chem. Theory Comput.* **2009**, *5*, 2619.

(17) Vogt, L.; Olivares-Amaya, R.; Kermes, S.; Shao, Y.; Amador-Bedolla, C.; Aspuru-Guzik, A. *J. Phys. Chem. A* **2008**, *112*, 2049.

(18) Olivares-Amaya, R.; Watson, M. A.; Edgar, R. G.; Vogt, L.; Shao, Y.; Aspuru-Guzik, A. *J. Chem. Theory Comput.* **2010**, *6*, 135.

(19) Koniges, A.; Preissl, R.; Kim, J.; Eder, D.; Fisher, A.; Masters, N.; Mlaker, V.; Ethier, S.; Wang, W.; Head-Gordon, M. In Proceedings of Cray User Group (CUG), Edinburgh, Scotland, May 24−27, 2010; Cray User Group, Inc.: Corvallis, Oregon, 2010.

(20) Anderson, A. G.; Goddard, W. A., III; Schröder, P. *Comput. Phys. Commun.* **2007**, *177*, 298.

(21) Esler, K.; Kim, J.; Shulenburger, L.; Ceperley, D. Fully accelerating quantum Monte Carlo simulations of real materials on GPU clusters. *Comput. Sci. Eng.*; http://doi.ieeecomputersociety.org/10.1109/MCSE.2010.122. Accessed March 10, **2011**).

(22) Gothandaraman, A.; Peterson, G. D.; Warren, G. L.; Hinde, R. J.; Harrison, R. J. *Parallel Comput.* **2008**, *34*, 278.

(23) Yokota, R.; Hamada, T.; Bardhan, J. P.; Knepley, M. G.; Barba, L. A. Biomolecular electrostatics using a fast multipole BEM on up to 512 GPUs and a billion unknowns. 2011, arXiv:1007.4591v3. arXiv.org ePrint archive. http://arxiv.org/abs/1007.4591. Accessed March 10, 2011).

(24) Kucharski, S. A.; Bartlett, R. J. *Theor. Chim. Acta* **1991**, *80*, 387.

(25) Stanton, J. F.; Gauss, J.; Watts, J. D.; Lauderdale, W. J.; Bartlett, R. J. *Int. J. Quantum Chem.* **1992**, *44*, 879.

(26) Piecuch, P.; Kucharski, S. A.; Kowalski, K; Musial, M. *Comput. Phys. Commun.* **2002**, *149*, 71.

(27) Scuseria, G. E.; Schaefer, H. F., III *J. Chem. Phys.* **1989**, *90*, 3700.

(28) Melicherčk, M.; Demovič, L.; Pitoňák, M.; Neogrády, P. In *Proceedings of the 9th Central European Symposium on Theoretical Chemistry,* 2010.

(29) Ma, W., Krishnamoorthy, S.; Villa, O.; Kowalski, K. In Proceedings of Cluster Computing and the Grid (CCGRID), IEEE/ACM International Symposium, Melbourne, Australia, May 17−20, 2010; IEEE: Piscataway, NJ, 2010.

(30) Ma, W.; Krishnamoorthy, S.; Villa, O.; Kowalski, K. *J. Chem. Theory Comput.* **2011,** in press.

(31) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A., Jr. *J. Comput. Chem.* **1993**, *14*, 1347.

(32) Bylaska, E. J.; de Jong, W. A.; Govind, N.; Kowalski, K.; Straatsma, T. P.; Valiev, M.; Wang, D.; Aprà, E.; Windus, T. L.; Hammond, J.; Autschbach, J.; Nichols, P.; Hirata, S.; Hackler, M. T.; Zhao, Y.; Fan, P.-D.; Harrison, R. J.; Dupuis, M.; Smith, D. M. A.; Nieplocha, J. Tipparaju, V.; Krishnan, M.; Vazquez-Mayagoitia, A.; Wu, Q.; Van Voorhis, T.; Auer, A. A.; Nooijen, M.; Crosby, L. D.; Brown, E.; Cisneros, G.; Fann, G. I.; Früchtl, H.; Garza, J.; Hirao, K.; Kendall, R.; Nichols, J. A.; Tsemekhman, K.; 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.; Pollack, L.; Rosing, M.; Sandrone, G.; Stave, M.; Taylor, H.; Thomas, G.; van Lenthe, J.; Wong, A.; Zhang, Z. *NWChem, A Computational Chemistry Package for Parallel Computers,* version 5.1.1; Environmental Molecular Sciences Laboratory (EMSL): Richland, WA, 2009.

(33) Werner,H.-J.; Knowles, P. J.; Manby, F. R.; Schütz, M.; Celani, P.; Knizia, G.; Korona, T.; Lindh, R.; Mitrushenkov, A.; Rauhut, G.; Adler, T. B.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Goll, E.; Hampel, C.; Hesselmann, A.; Hetzer, G.; Hrenar, T.; Jansen, G.; Köppl, C.; Liu, Y.; Lloyd, A. W.; Mata, R. A.; May, A. J.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklass, A.; Palmieri, P.; Pflüger, K.; Pitzer, R.; Reiher, M; Shiozaki, T.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T.; Wang, M.; Wolf, A. *MOLPRO*, version 2010.1; University College Cardiff Consultants Limited: Wales, U.K., 2010; http://www.molpro.net.

(34) Lee, V. W.; Kim, C.; Chhugani, J.; Deisher, M.; Kim, D.; Nguyen, A. D.; Satish, N.; Smelyanskiy, M.; Chennupaty, S.; Hammarlund, P.; Singhal, R.; Dubey, P. *SIGARCH Comput. Archit. News* **2010**, *38*, 451.

(35) Crawford, T. D.; Schaefer, H. F., III *Rev. Comput. Chem.* **2000**, *14*, 33.

(36) Bartlett, R. J.; Musial, M. *Rev. Modern Phys.* **2007**, *79*, 291.

(37) Crawford, T. D.; Sherrill, C. D.; Valeev, E. F.; Fermann, J. T.; King, R. A.; Leininger, M. L.; Brown, S. T.; Janssen, C. L.; Seidl, E. T.; Kenny, J. P.; Allen, W. D. *J. Comput. Chem.* **2007**, *28*, 1610.

(38) Hammond, J. R.; Krishnamoorthy, S.; Shende, S.; Romero, N. A.; Malony, A. D. *Performance Characterization of Global Address Space Applications: A Case Study with NWChem* **2011**.

(39) Kobayashi, R.; Rendell, A. P. *Chem. Phys. Lett.* **1997**, *265*, 1.

(40) Hirata, S. *J. Phys. Chem. A* **2003**, *107*, 9887.

(41) Kowalski, K.; Hammond, J. R.; de Jong, W. A.; Fan, P.-D.; Valiev, M.; Wang, D.; Govind, N. Coupled Cluster Calculations for Large Molecular and Extended Systems. In *Computational Methods for Large Systems: Electronic Structure Approaches for Biotechnology and Nanotechnology*; Reimers, J. R., Ed.; Wiley: Hoboken, NJ, 2011.

(42) Hampel, C.; Peterson, K. A.; Werner, H.-J. *Chem. Phys. Lett.* **1992**, *190*, 1.

(43) Li, S.; Ma, J.; Jiang, Y. *J. Comput. Chem.* **2002**, *23*, 237.

(44) Li, S.; Shen, J.; Li, W. *J. Chem. Phys.* **2006**, *125*, 074109.

JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Time-Dependent Density-Functional Description of the ^1L_a State in Polycyclic Aromatic Hydrocarbons: Charge-Transfer Character in Disguise?

Ryan M. Richard and John M. Herbert*

Department of Chemistry, The Ohio State University, Columbus, Ohio 43210, United States

S Supporting Information

ABSTRACT: The electronic spectrum of alternant polycyclic aromatic hydrocarbons (PAHs) includes two singlet excited states that are often denoted ^1L_a and ^1L_b. Time-dependent density functional theory (TD-DFT) affords reasonable excitation energies for the ^1L_b state in such molecules, but often severely underestimates ^1L_a excitation energies and fails to reproduce observed trends in the ^1L_a excitation energy as a function of molecular size. Here, we examine the performance of long-range-corrected (LRC) density functionals for the ^1L_a and ^1L_b states of various PAHs. With an appropriate choice for the Coulomb attenuation parameter, we find that LRC functionals avoid the severe underestimation of the ^1L_a excitation energies that afflicts other TD-DFT approaches, while errors in the ^1L_b excitation energies are less sensitive to this parameter. This suggests that the ^1L_a states of certain PAHs exhibit some sort of charge-separated character, consistent with the description of this state within valence-bond theory, but such character proves difficult to identify a priori. We conclude that TD-DFT calculations in medium-size, conjugated organic molecules may involve significant but hard-to-detect errors. Comparison of LRC and non-LRC results is recommended as a qualitative diagnostic.

I. INTRODUCTION

Most contemporary density-functional approximations, including those based on generalized gradient approximations (GGAs) as well as hybrid functionals that do not incorporate full Hartree−Fock (HF) exchange, afford an incorrect asymptotic distance dependence for charge-transfer (CT) excitation energies.[1] In the context of time-dependent density functional theory (TD-DFT), this artifact leads to predictions of spurious, low-energy CT states in large molecules,[2−4] liquids,[5] and clusters.[6,7] One means to mitigate this problem, while retaining the computational simplicity of TD-DFT, is to use long-range-corrected (LRC) density functionals.[8−20] The basic idea behind LRC-DFT is to treat the electron−electron exchange interaction using HF theory at large separation, since HF theory affords the proper distance dependence for CT excitation energies,[1] but to use GGA exchange at short range, in the interest of obtaining an accurate description of dynamical electron correlation. This length-scale separation is accomplished by partitioning the electron−electron Coulomb operator into short- and long-range components.[8,15,21−23]

While conventional TD-DFT's propensity to overstabilize CT states[1−7] and Rydberg states[19,24] is well-known, this method's admirable accuracy for localized, valence excitations in small organic molecules is similarly well-documented.[25,26] For alternant polycyclic aromatic hydrocarbon (PAH) molecules, however, TD-DFT calculations sometimes afford large errors in excitation energies,[27,28] for states that one would not ordinarily associate with CT character.

A particular class of examples is the homologous sequence of linear-condensed acenes (benzene, naphthalene, anthracene, etc.), which exhibit two low-lying $^1\pi\pi^*$ excited states, commonly denoted ^1L_a and ^1L_b.[29−31] The transition densities for these two states are polarized along the short and long axes of the molecule, respectively (see Figure 1), with the ^1L_b transition density exhibiting nodes at the atoms and the ^1L_a transition density displaying nodes at the bond midpoints.[30,31] For the ^1L_a state in the linear acene sequence, errors in TD-DFT excitation energies increase dramatically as a function of the number of aromatic rings, yet errors in the ^1L_b excitation energies appear to be uncorrelated with molecular size.[27,32] (This is perhaps all the more surprising in view of the fact that the ^1L_b state in benzene and naphthalene exhibits substantial double-excitation character, whereas the ^1L_a state does not.[33−35]) Recently, however, certain TD-LRC-DFT have been shown to afford accurate ^1L_a excitation energies for the linear acenes,eliminating the length-dependent trend in the errors.[36,37]

The ^1L_a and ^1L_b states in linear acenes have long been discussed as being "ionic" and "covalent", respectively, in the language of valence-bond (VB) theory.[31] In other words, the ^1L_a wave function is thought to include determinants where both π electrons from a C=C bond are assigned to the same carbon atom. Detailed VB calculations corroborate this conceptual picture,[33−35,38] and this might lead one to suspect that charge separation in the ^1L_a state, which somehow increases as a function of molecular size, could explain the errors observed in TD-DFT excitation energies for the ^1L_a state. This is precisely what was concluded in a recent study,[27] based on a semiempirical charge-decomposition analysis. The goal of the present work is to analyze all-electron TD-DFT and TD-LRC-DFT calculations of ^1L_a and ^1L_b on a more diverse set of PAHs.

Received: October 25, 2010
Published: April 05, 2011

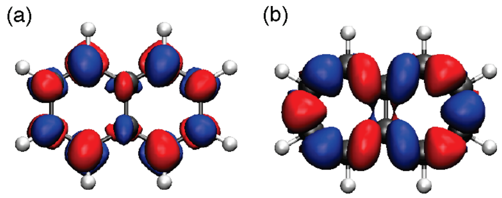


Figure 1. Transition densities for (a) the ^1L_a state and (b) the ^1L_b state of naphthalene, computed at the TD-B3LYP level. The isosurface in either plot encapsulates 90% of the transition density.

II. METHODS

Ground-state geometries were optimized at the B3LYP/6-31G* level, and vertical excitation energies (for singlet states only) were subsequently calculated at the TD-DFT/cc-pVTZ level, using various density functionals. The SG-1 quadrature grid[39] was used for all TD-DFT calculations, as tests using significantly finer grids resulted in changes of less than 0.01 eV in the excitation energies. Except where noted, the commonly used Tamm–Dancoff approximation[40] is *not* employed here. All calculations were performed using a locally modified version of Q-Chem.[41] Cartesian coordinates for the optimized PAH geometries, along with tabulated TD-DFT excitation energies, can be found in the Supporting Information.

A variety of LRC density functionals are examined in this work, including LRC-μBLYP, LRC-μBOP, LRC-ωPBE, and LRC-ωPBEh. The notations "μBLYP" and "μBOP" indicate that the BLYP[42,43] and BOP[44] functionals are used, but with a short-range version of Becke's GGA exchange functional[42] that is constructed according to the prescription developed by Hirao and co-workers.[8] The notation "ωPBE" indicates a short-range version of the PBE exchange functional,[45] constructed according to the procedure of Scuseria and co-workers.[15] (The aforementioned notation is consistent with that used in the Q-Chem program but differs from the nomenclature used in some recent papers.[46]) The LRC-ωPBEh function is a hybrid ("h") that includes 20% HF exchange at short range.[19] All of the LRC functionals examined here include full HF exchange at long range:

$$E_{xc}^{\rm LRC} = E_c + E_x^{\rm GGA,\,SR} + C_{\rm HF}E_x^{\rm HF,\,SR} + E_x^{\rm HF,\,LR} \tag{1}$$

Here, "SR" and "LR" indicate use of the short-range and long-range components of the Coulomb operator, respectively, and $C_{\rm HF}$ is the coefficient of short-range HF exchange.

TD-LRC-DFT excitation energies can be quite sensitive to the value of the Coulomb attenuation parameter (μ or ω),[18,19] especially for CT-type excitations.[49] Values of μ or ω that are optimized using ground-state properties (e.g., atomization energies, ionization potentials, or reaction barrier heights) may afford large errors in TD-DFT excitation energies.[18,19] Previous studies by our group[4,19] have shown that LRC-ωPBE with ω = 0.3 a_0^{-1} and LRC-ωPBEh with ω = 0.2 a_0^{-1} afford the best statistical performance for excitation energies, without degrading ground-state properties. As such, we focus primarily on these two functionals. With the aforementioned parameters, the LRC-ωPBEh functional affords average errors of ~0.3 eV for both localized and CT excitation energies,[19] while the LRC-ωPBE functional performs similarly when ω lies in the range of 0.2–0.3 a_0^{-1}.[4,26]

Table 1. Parameters for the LRC Functionals Employed in This Work

functional	μ or ω/a_0^{-1}	$C_{\rm HF}$	functional	μ or ω/a_0^{-1}	$C_{\rm HF}$
LRC-μBOP	0.47	0.0	LRC-ωPBE	0.30	0.0
LRC-μBLYP[a]	0.17	0.0	LRC-ωPBEh	0.20	0.2
LRC-μBLYP[a]	0.30	0.0			

[a] Two different values of μ are used for LRC-μBLYP.

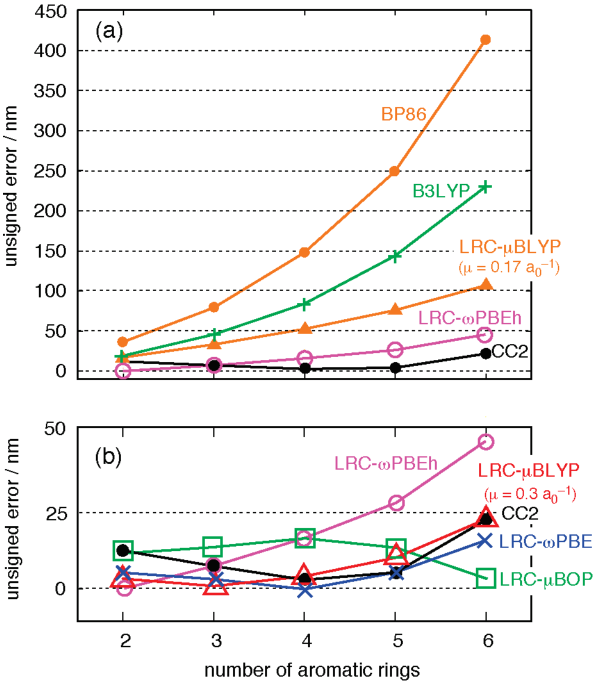


Figure 2. TD-DFT errors in the vertical excitation energies for the ^1L_a state, expressed in wavelength units. Panel (a) illustrates the divergence of the TD-B3LYP and TD-BP86 excitation energies as a function of *n*, while panel (b) shows a close-up view of the errors engendered by several different LRC functionals.

As compared to LRC functionals based upon ωPBE, the functionals μBLYP and μBOP, which utilize the short-range "μB88" functional[46] developed by Hirao and co-workers,[8] have not been studied as extensively in the context of TD-DFT excitation energies. It does appear that the LRC-μPBE and LRC-ωPBE functionals afford comparable excitation energies, at a given value of the Coulomb attenuation parameter (μ or ω),[19] although predicted ground-state properties may be quite different.[18]

In view of these facts, we choose the value μ = 0.3 a_0^{-1} for the LRC-ωPBE and LRC-μBLYP functionals, a choice that is supported by results from a recent TD-LRC-DFT study of linear acenes.[37] At the same time, the value μ = 0.17 a_0^{-1} was found to provide the most accurate excitation energies in a recent TD-LRC-μBLYP study of intramolecular CT states in Coumarin dyes,[50] although the value μ = 0.31 a_0^{-1} performs better for oligothiophenes.[51] Thus, for completeness we will consider LRC-μBLYP with μ = 0.17 a_0^{-1}. Finally, Hirao and co-workers advocate the use of LRC-μBOP with μ = 0.47 a_0^{-1},[10] so we will assess this functional as well, even though our previous work indicates that values of $\mu \gtrsim 0.5$ a_0^{-1} often afford large errors in ground-state properties.[18] Table 1 lists the parameters for each of the LRC functionals used in this work.

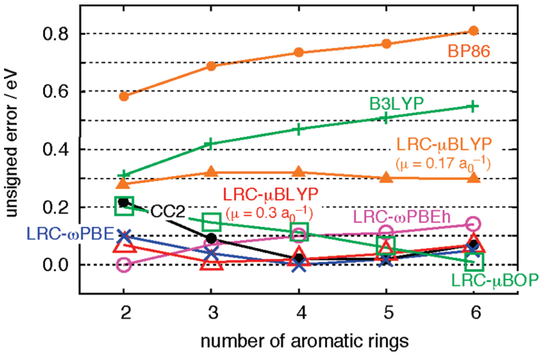


Figure 3. TD-DFT errors in the vertical excitation energies for the ^1L_a state of the linear acene sequence, expressed in energy units.

III. RESULTS

A. Linear-Condensed Acenes. The ^1L_a and ^1L_b states in the linear-condensed acene series are characterized by transition densities that are polarized along the short and long axes of the molecule, respectively,[30,31] as illustrated for naphthalene in Figure 1. Consistent with previous calculations,[27,33−35,37,38,52] we find that the $S_0 \rightarrow {}^1L_a$ excitation is dominated (>90%) by a transition between the highest occupied and lowest unoccupied molecular orbitals (HOMO → LUMO), whereas the $S_0 \rightarrow {}^1L_b$ excitation involves (HOMO − 1) → LUMO and HOMO → (LUMO + 1) transitions, with approximately equal weights.

As noted in previous studies,[27,32,37] errors in the ^1L_a excitation wavelength computed using TD-DFT methods often increase rapidly as a function of the number of aromatic rings, n. Errors in the ^1L_a excitation wavelength are plotted in Figure 2 as a function of n, for the set of functionals considered here. (We note that Wong and Hsieh[37] have recently published similar results, using a slightly different set of LRC functionals.) Also included in Figure 2 are the errors obtained using approximate coupled-cluster theory (CC2), which were obtained from ref 27. Errors are computed on the basis of experimental band maxima that have been corrected to account for excited-state geometry relaxation.[27]

Wong and Hsieh[37] have noted previously that size-dependent errors in the ^1L_a excitation wavelength that are obtained at the TD-BP86 and TD-B3LYP level are greatly reduced using certain TD-LRC-DFT approaches, for which a qualitatively correct distance dependence is obtained. Our results add a caveat, namely, that the erroneous n-dependence of the excitation wavelength remains in LRC-μBLYP calculations performed using $\mu = 0.17\ a_0^{-1}$. This value of μ, which was suggested in two different studies of CT states in Coumarin dyes,[50,53] is the smallest value of the Coulomb attenuation parameter that has been suggested in any benchmark study of LRC-DFT of which we are aware. Other LRC functionals examined here use a Coulomb attenuation parameter of either 0.2 a_0^{-1} or 0.3 a_0^{-1}, and for these functionals the errors in ^1L_a excitation wavelengths for the linear acene series is uncorrelated with molecular size.

At the same time, one should recognize that the length-dependent trends that are evident in the excitation *wavelength* data in Figure 2 amount to relatively small changes in excitation *energies*, at least in comparison to the ∼0.3 eV statistical error bar that is typically ascribed to TD-DFT calculations. Errors in excitation energies for the ^1L_a state of the linear acene sequence are shown in Figure 3. From these data, it is difficult to ascribe any length-dependent trend to the errors obtained using LRC-μBLYP($\mu = 0.17\ a_0^{-1}$); rather, these excitation energies appear to be systematically overestimated by about 0.3 eV. [Mean absolute errors (MAEs) for each method are listed in Table 2.] Excitation energies calculated using LRC-ωPBE ($\omega = 0.3\ a_0^{-1}$) and LRC-μBLYP ($\mu = 0.3\ a_0^{-1}$) are in good agreement with CC2 calculations. As noted by Wong and Hsieh,[37] LRC functionals significantly outperform B3LYP for the ^1L_a excitation energies, but B3LYP affords a smaller MAE for the ^1L_b excitation energies.

Table 2. Mean Absolute Errors (MAEs) in Excitation Energies for the Linear Acene Sequence, n = 2−6

method	MAE [a]/eV	
	^1L_a	^1L_b
CC2[b]	0.08	0.22
TD-BP86	0.72	0.62
TD-B3LYP	0.45	0.15
TD-LRC-μBLYP ($\mu = 0.17\ a_0^{-1}$)	0.30	0.18
TD-LRC-μBLYP ($\mu = 0.30\ a_0^{-1}$)	0.07	0.31
TD-LRC-ωPBE	0.04	0.33
TD-LRC-ωPBEh	0.08	0.35
TD-LRC-μBOP	0.08	0.37

[a] Relative to experimental values corrected for excited-state geometry relaxation (from ref 27). [b] Values taken from ref 27.

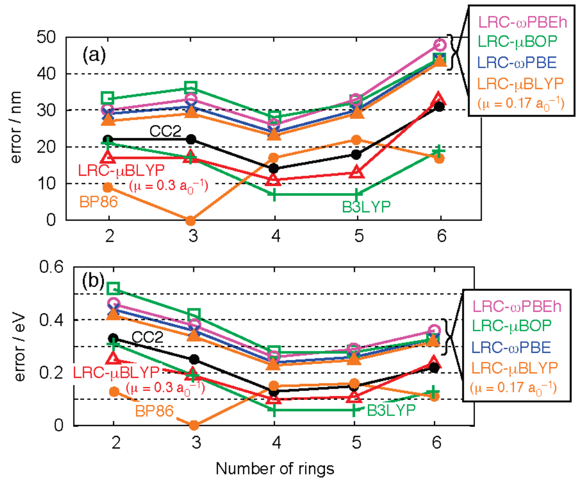


Figure 4. TD-DFT errors in the vertical excitation energies for the ^1L_b state, expressed in (a) wavelength units and (b) energy units.

In contrast to the ^1L_a results, TD-DFT errors for the ^1L_b excitation energies show no clear trend with respect to n, even for the non-LRC functionals (see Figure 4). With the exception of the TD-BP86 calculations, the n-dependence of the TD-DFT errors tracks the CC2 results quite well, albeit with a constant energy offset that varies from one functional to another. This observation, along with the fact that the CC2 MAE is somewhat larger for ^1L_b than for ^1L_a (0.22 eV versus 0.08 eV), suggests that the correction applied to the experimental band maxima in order to obtain an experimental estimate of the vertical excitation energy[27] may be somewhat less accurate for ^1L_b. In any case, most of the TD-DFT MAEs for the ^1L_b state are ≲0.3 eV, which is within the generally accepted accuracy of TD-DFT excitation energies.

Table 3. Comparison of Excitation Energies (in eV) for the ^1L_a State of the Linear Acene Sequence, Computed Using Full TD-DFT and also the Tamm−Dancoff Approximation (TDA)

	LRC-ωPBE		LRC-ωPBEh	
n	TD-DFT	TDA	TD-DFT	TDA
2	4.76	5.01	4.66	4.89
3	3.64	3.92	3.53	3.80
4	2.88	3.19	2.78	3.07
5	2.35	2.69	2.26	2.57
6	1.97	2.31	1.88	2.21

Another point worth noting is the effect of the Tamm−Dancoff approximation (TDA).[40] In benchmark calculations for small molecules, this approximation provides excitation energies within 0.15 eV of full TD-DFT results, at somewhat reduced cost. In larger molecules, however, we have observed that TD-DFT/TDA discrepancies are sometimes more significant. Table 3 summarizes the difference between TDA and full TD-DFT excitation energies for two different density functionals. We find that the TDA systematically increases both the ^1L_a and ^1L_b excitation energies, by about 0.3 eV. In the case of the ^1L_a state, a 0.3 eV shift would bring the TD-LRC-μBLYP ($\mu = 0.17\ a_0^{-1}$) excitation energies into good agreement with experiment, thereby masking errors that appear to indicate a too-small value of μ. In fact, it has previously been suggested that TD-DFT calculations on PAHs *should* invoke the TDA, as more accurate results are obtained (using B3LYP) with the TDA than with full TD-DFT.[52] In our view, this is most likely a fortuitous cancellation of errors, as only full TD-DFT affords the proper linear response of the ground-state density.

It has been determined, experimentally, that the ^1L_a state lies above the ^1L_b state for $n \leq 2$, but that ^1L_b is higher in energy starting at $n = 3$.[54] Both TD-B3LYP and TD-BP86 calculations incorrectly predict that ^1L_b is higher in energy starting at $n = 2$, whereas all of the TD-LRC-DFT methods examined here, with the exception of LRC-μBLYP with $\mu = 0.17\ a_0^{-1}$, place ^1L_a and ^1L_b in the correct energetic order as a function of molecular size, both within the TDA and also at the full TD-DFT level. The failure of TD-B3LYP in this context is potentially a problem in applications beyond PAHs, since indole (and, consequently, tryptophan) also exhibits ^1L_a and ^1L_b states, whose electronic structure is thought to be similar to the corresponding states in naphthalene.[55] TD-B3LYP also fails to predict the correct order of the ^1L_a and ^1L_b states in tryptophan.[56]

B. Nonlinear PAHs. Although LRC-DFT calculations of the linear acenes have been reported previously,[36,37] these methods have not yet been studied for more general, nonlinear PAHs. The TD-B3LYP and TD-BP86 methods *have* been applied to certain larger PAHs, and large errors in the ^1L_a excitation energies are observed in some cases.[28,52] Here, we apply TD-LRC-DFT to a set of nonlinear PAHs, the structures of which are depicted in Figure 5. This data set includes both cata-condensed and peri-condensed examples,[31] ranging in size from three to seven six-membered rings. A numbering scheme for these molecules is introduced in Figure 5; as a rough guideline, larger numbers correspond to larger molecules, although the data set does contain several structural isomers.

For these molecules, we shall restrict our calculations to the functionals B3LYP, LRC-ωPBE, and LRC-ωPBEh. Results presented above and in ref 37 demonstrate that other LRC functionals afford very similar excitation energies for the linear acenes, provided that μ (or ω) is chosen appropriately. In particular, Wong and Hsieh[37] considered several different functionals with $C_{HF} = 0$ and $\mu \approx 0.3\ a_0^{-1}$ and found that MAEs across the linear acene sequence differ by no more than 0.05 eV, for both ^1L_a and ^1L_b. Our results (section III.A) show that μ can

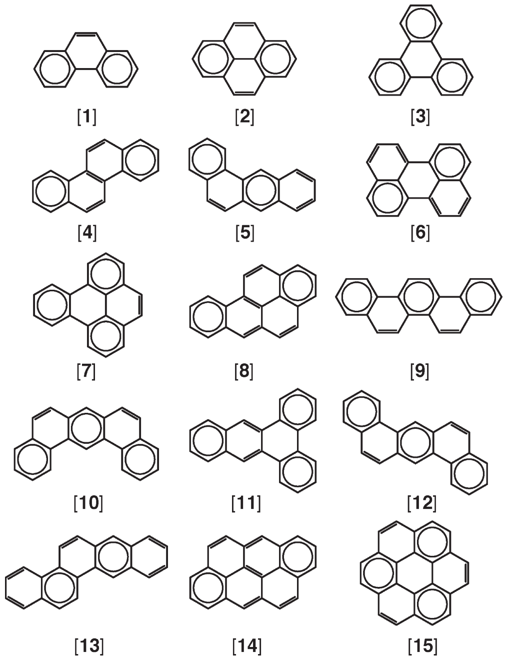


Figure 5. Clar-type resonance structures[57,58] of the nonlinear PAHs considered in this work, along with the numbering scheme that is used to refer to them in the text and figures: phenanthrene (**1**), pyrene (**2**), triphenylene (**3**), chrysene (**4**), benz[*a*]anthracene (**5**), perylene (**6**), benzo[*e*]pyrene (**7**), benzo[*a*]pyrene (**8**), picene (**9**), dibenz[*a*,*j*]anthracene (**10**), dibenz[*a*,*c*]anthracene (**11**), dibenz[*a*,*h*]anthracene (**12**), benzo[*b*]chrysene (**13**), anthanthrene (**14**), and coronene (**15**).

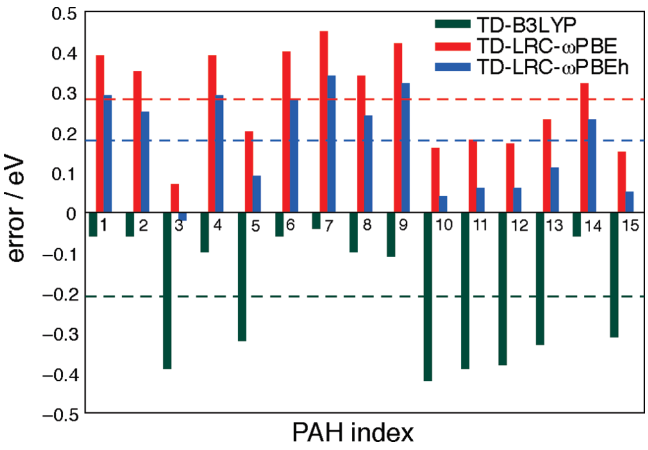


Figure 6. Errors (theory minus experiment) in ^1L_a excitation energies for the PAHs depicted in Figure 5. Dashed horizontal lines represent the average error for each method. Experimental benchmarks are band maxima in nonpolar solvents. The solvent correction suggested in ref 52 would reduce the TD-LRC-DFT errors by 0.11 eV and would make the TD-B3LYP errors more negative by 0.11 eV.

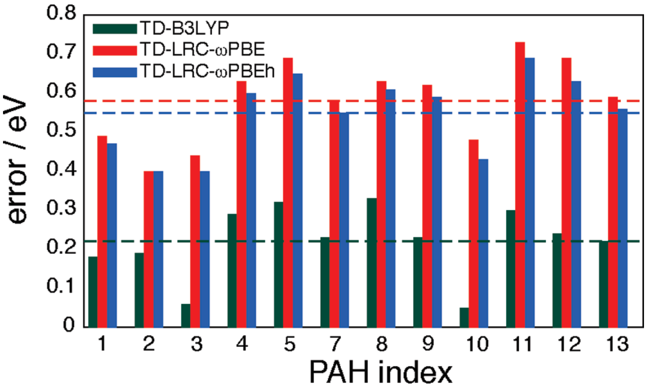


Figure 7. Errors (theory minus experiment) in ^1L_b excitation energies for the PAHs depicted in Figure 5. Dashed horizontal lines represent the average error for each method. Experimental benchmarks are band maxima in nonpolar solvents, and omissions from the data set in Figure 5 correspond to molecules for which no experimental value for the ^1L_b excitation energy is available. The solvent correction suggested in ref 52 would reduce all of the errors by 0.03 eV.

be reduced if short-range HF exchange is introduced, as in LRC-ωPBEh. This conclusion is in accord with previous findings using a more diverse set of molecules and excited states.[19]

Figure 6 shows the errors in the calculated vertical excitation energies for the ^1L_a state of the nonlinear PAHs. As in the case of the linear acenes, B3LYP consistently underestimates the excitation energies, with most of the largest errors associated with the larger PAHs. (Note that the errors in Figure 6 are signed quantities.) The two LRC functionals, on the other hand, consistently overestimate the excitation energies, which was also observed for the linear acene sequence, although the errors are somewhat larger here. Interestingly, the largest errors observed at the TD-B3LYP level seem to correlate with the smallest errors obtained using the LRC functionals.

We should note that the experimental excitation energies used to compute the TD-DFT errors are taken from ref 59 (they are also tabulated in ref 28) and represent band maxima in nonpolar solvents. On the basis of a comparison of solution-phase absorption spectra to gas-phase photoelectron spectra,[60] Wang and Wu[52] suggest that these values should be corrected upward by 0.11 eV to obtain an estimate of the gas-phase $S_0 \rightarrow {}^1L_a$ excitation energy. This correction has *not* been applied in Figure 6. If we were to apply this correction, then the mean error in ^1L_a excitation energies computed at the TD-B3LYP cc-pVTZ level would change from −0.21 eV (the value indicated in Figure 6) to −0.32 eV. Meanwhile, the TD-LRC-DFT values would become more accurate, with corrected mean errors of 0.1−0.2 eV, which is only slightly larger than the mean errors obtained for the linear acenes using these same LRC functionals.

Figure 7 depicts errors in the ^1L_b excitation energies for the nonlinear PAHs. (The solvent correction is also absent from these data, but the value suggested by Wang and Wu[52] is only 0.03 eV for the ^1L_b state.) As in the case of the linear acenes, all three of the TD-DFT methods consistently overestimate the ^1L_b excitation energies, with no clear size-dependent trend, and TD-B3LYP consistently outperforms the LRC functionals. For these molecules, the mean error in TD-LRC-DFT excitations energies (≈ 0.5 eV) is somewhat larger than for the ^1L_b states of the linear acenes and lies outside of the ~0.3 eV accuracy established for these functionals in previous benchmark calculations.[4,19,26]

Table 4. Mean Absolute Errors (eV) in TD-DFT Excitation Energies for Various Subsets of Nonlinear PAHs

	MAE(^1L_a)[b]			MAE(^1L_b)[b]		
subset[a]	B3LYP	LRC-ωPBE	LRC-ωPBEh	B3LYP	LRC-ωPBE	LRC-ωPBEh
full set	0.32	0.17	0.07	0.19	0.55	0.52
cata	0.39	0.14	0.03	0.18	0.57	0.53
peri	0.22	0.23	0.12	0.22	0.51	0.49
1-2	0.36	0.17	0.06	0.19	0.57	0.53
3+	0.29	0.17	0.07	0.19	0.53	0.50

[a] The full data set is shown in Figure 5; see the text for a description of the various subsets. [b] Relative to solution-phase band maxima corrected for solvent effects.[52]

As mentioned above, for the nonlinear PAHs many of the largest TD-B3LYP errors for ^1L_a excitation energies coincide with the largest molecules in this data set, whereas TD-B3LYP errors tend to be smaller for the PAHs that are more condensed (in the sense of possessing more fused rings), and therefore smaller. To analyze this further, we have partitioned the full set of nonlinear PAHs into various subsets that reflect the degree and manner of annulation. In addition to cata-condensed and peri-condensed subsets, we consider a subset "1-2" in which each ring is fused to no more than two other rings, and another subset "3+" in which at least one ring is fused to three other rings.

Table 4 lists separate MAEs for each of these subsets, and these statistics do suggest that the accuracy of TD-B3LYP for the ^1L_a excitation energy is related to the extent of condensation. The cata-condensed and 1−2 subsets have MAEs that are larger than the MAE for the full data set, at the TD-B3LYP level, whereas in the case of the two LRC functionals, the MAE is largely unaffected by how the data set is partitioned. (For ^1L_b excitation energies, the MAE is largely unaffected by the partitioning even in the case of B3LYP.) However, this trend is not *strictly* related to molecular size. For example, the TD-B3LYP error for picene, **9**, is well below the mean, despite having one of the longer end-to-end distances among the nonlinear PAHs considered here.

IV. ANALYSIS AND DISCUSSION

A. Valence-Bond Considerations. One might hypothesize that size-dependent errors in ^1L_a excitation energies are related to well-known size-dependent errors in TD-DFT polarizabilities and hyperpolarizabilities for conjugated molecules,[61,62] problems that are mitigated when LRC functionals are employed.[63] Because the ^1L_b state exhibits no such size-dependent errors, however, we must look elsewhere for an explanation.

Grimme and Parac[27] have previously noted these size-dependent errors for the ^1L_a state and explained them in terms of an excited-state wave function having significant contributions from ionic determinants, to use valence-bond language. In other words, the valence-bond picture is that the ^1L_a wave function exhibits charge separation at the level of individual C−C bonds.[33−35] (The dipole moment of the ^1L_a state is zero, by symmetry, so the $S_0 \rightarrow {}^1L_a$ excitation cannot be associated with any *net* charge separation. In addition, the $S_0 \rightarrow {}^1L_a$ transition is primarily a HOMO → LUMO excitation, and both the HOMO and the LUMO are delocalized over the entire molecule, as required by symmetry.)

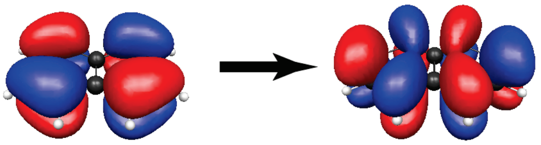


Figure 8. NTOs for the ^1L_a state of naphthalene, computed at the TD-B3LYP/cc-pVTZ level.

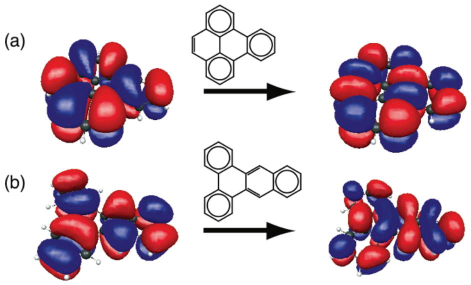


Figure 9. NTOs for two representative PAHs: (a) benzo[*e*]pyrene (**7**) and (b) dibenz[*a*,*c*]anthracene (**11**). The structure of each molecule is also shown. The NTOs shown in (a) accounts for 93% of the transition density, and those in (b) account for 84% of the transition density. Each NTO was computed at the TD-B3LYP/cc-pVTZ level.

Parac and Grimme[28] developed an "ionicity metric" based upon a Mulliken-style atomic partition of time-dependent Pariser—Parr—Pople[64−66] (TD-PPP) transition densities and demonstrated that this metric is strongly correlated with errors in excitation energies computed at the TD-BP86 level. We attempted a similar analysis, using transition densities computed from all-electron TD-DFT calculations, and taking proper account for the nonorthogonality of the atomic orbital (AO) basis. However, we found that the trends obtained from these all-electron calculations were far more muddled and ambiguous than those reported by Parac and Grimme, even when minimal basis sets were used in an effort to avoid well-known problems with Mulliken analysis in extended basis sets.

On the other hand, natural transition orbitals[67] (NTOs) for the ^1L_a state *do* support the notion of charge separation within the C−C bonds. As an example, Figure 8 depicts the most significant pair of NTOs for the ^1L_a state of naphthalene; this pair of NTOs accounts for 88% of the norm of the $S_0 \rightarrow {}^1L_a$ transition density matrix, and the product of these two NTOs is qualitatively similar to the $S_0 \rightarrow {}^1L_a$ transition density (cf. Figure 1a). The same sort of charge separation that is seen in this pair of NTOs might be inferred from the transition density itself, insofar as the latter has nodes centered on the C−C bonds, whereas the $S_0 \rightarrow {}^1L_b$ transition density has nodes located on the carbon atoms. These TD-B3LYP transition densities are consistent with the predictions of a simple particle-on-a-ring model,[30,31] which has long been used as a qualitative model for understanding the electronic structure of the linear acenes.

With the benefit of hindsight and the availability of VB calculations for naphthalene and anthracene,[33−35,38] this analysis of NTOs and transition densities for the linear acenes could be used to rationalize the size-dependence of TD-B3LYP results for the ^1L_a state and the lack of size dependence in TD-B3LYP results for the ^1L_b state. Analysis of the NTOs is more complicated in the case of the nonlinear PAHs, however. Consider two representative examples: benzo[*e*]pyrene (7), for which TD-B3LYP predicts an accurate ^1L_a excitation energy; and dibenz[*a*,*c*]-anthracene (**11**), for which TD-B3LYP significantly underestimates the ^1L_a excitation energy. The NTOs that dominate the $S_0 \rightarrow {}^1L_a$ transition for each of these two PAHs are pictured in Figure 9. In both cases, one could argue that the NTOs show evidence of charge separation within individual C−C bonds.

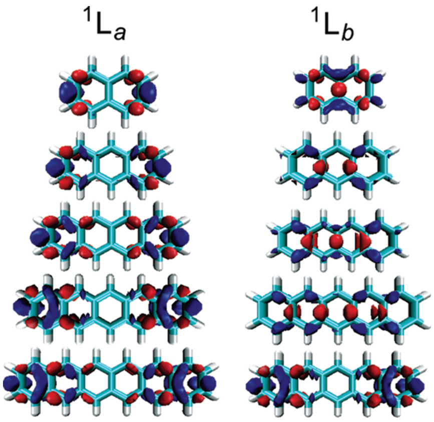


Figure 10. Difference densities, $\Delta\rho$, for the $S_0 \rightarrow {}^1L_a$ and $S_0 \rightarrow {}^1L_b$ excitations of the linear acene sequence, computed at the TD-B3LYP level. The two colored isosurfaces in each plot encapsulate 60% of the positive/negative part of $\Delta\rho$.

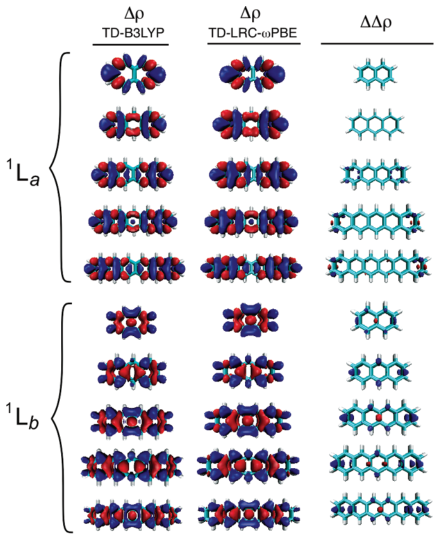


Figure 11. Difference densities, $\Delta\rho$, for the ^1L_a and ^1L_b states of the linear acene sequence computed using two different TD-DFT methods. The two colored isosurfaces in each plot encapsulate 95% of the positive negative part of $\Delta\rho$. The difference between the two difference densities, $\Delta\Delta\rho$, is also plotted, using the same isocontour that is used to plot $\Delta\rho$ at the TD-LRC-ωPBE level.

Detailed VB calculations are not generally available (or even feasible), and in their absence, we must conclude that one cannot unambiguously infer ionic character from NTOs and transition densities alone. Ideally, we would like a predictive means to diagnose errors in TD-DFT calculations. The search for such a diagnostic occupies the remainder of this work.

B. Difference Densities. To this end, we first examine difference densities,

$$\Delta\rho = \rho(\text{excited}) - \rho(\text{ground}) \qquad (2)$$

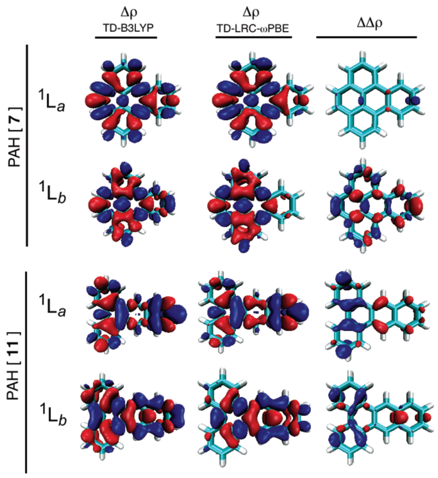


Figure 12. Difference densities, $\Delta\rho$, for the ^1L_a and ^1L_b states of PAHs **7** and **11** computed using two different TD-DFT methods. The two colored isosurfaces in each plot encapsulate 95% of the positive negative part of $\Delta\rho$. The difference between the two difference densities, $\Delta\Delta\rho$, is also plotted, using the same isocontour as the TD-LRC-ωPBE plots.

for the linear acenes. Isosurface representations of $\Delta\rho$ for the ^1L_a and ^1L_b excited states, computed at the TD-B3LYP level, are depicted in Figure 10. In these isosurface representations, we have chosen contour values that encapsulate 60% of the positive and negative lobes of $\Delta\rho$, a value that was selected in order to obtain plots that are qualitatively similar to those published in ref 37, where $\Delta\rho$ was computed at the level of second-order approximate coupled-cluster theory (CC2). Consistent with the CC2 difference densities plotted in ref 37, the TD-B3LYP difference densities in Figure 10 show that the $S_0 \rightarrow {}^1L_a$ excitations are associated with a greater degree of local charge reorganization, as compared to the $S_0 \rightarrow {}^1L_b$ excitations. This fact was previously noted by Wong and Hsieh,[37] as an explanation for improved performance of TD-LRC-DFT for the ^1L_a state.

Unfortunately, the picture becomes a bit more muddled if one plots isosurfaces that contain a larger fraction of $\Delta\rho$, as can be seen from the 95% isocontour surfaces, computed at the TD-B3LYP level, that are depicted on the left side of Figure 11. These isosurface plots fail to provide any clear evidence that the ^1L_a state exhibits a greater degree of charge separation than does the ^1L_b state. Difference densities obtained at the TD-B3LYP level are nearly identical to those obtained at the TD-LRC-ωPBE level, as can be seen by plotting the difference between the difference densities,

$$\Delta\Delta\rho = \Delta\rho(\text{B3LYP}) - \Delta\rho(\text{LRC-}\omega\text{PBE}) \qquad (3)$$

Isosurface representations of $\Delta\Delta\rho$ are similar for both states (see Figure 11). In other words, any sort of charge separation that one might infer on the basis of $\Delta\rho$ for one method is present also in the other method. Analysis of $\Delta\rho$ therefore cannot explain the fact that non-LRC functionals exhibit a qualitatively different size-dependence for the ^1L_a state, as compared to LRC functionals.

Figure 12 presents isosurface representations of $\Delta\rho$ and $\Delta\Delta\rho$ for two different nonlinear PAHs, **7** and **11**. For **7**, where the TD-B3LYP excitation energy for ^1L_a is reasonably accurate, we find almost no difference between $\Delta\rho$ computed at the TD-B3LYP level and $\Delta\rho$ computed at the TD-LRC-ωPBE level; in fact, differences in $\Delta\rho$ between these two functionals are much more significant for the ^1L_b state. In the case of **11**, for which TD-B3LYP error in the ^1L_a excitation energy is large, we do see qualitative differences in $\Delta\rho$ between these two methods. However, these differences are no more significant for the ^1L_a state than they are for the ^1L_b state. (In other words, $\Delta\Delta\rho$ is similar for both states.) Since TD-B3LYP is more accurate for the ^1L_b excitation energy of **11**, while TD-LRC-ωPBE is more accurate for the ^1L_a excitation energy, this cannot explain the origin of the TD-B3LYP errors for ^1L_a .

Table 5. Values of the CT Metric (Equation 4) for the ^1L_a State and the ^1L_b State of the Linear Acene Series, Computed at the TD-B3LYP Level

	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$
$\Lambda(^1L_a)$	0.86	0.83	0.89	0.85	0.90
$\Lambda(^1L_b)$	0.62	0.61	0.63	0.64	0.65

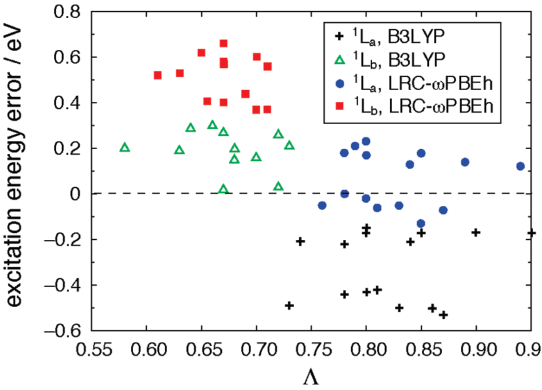


Figure 13. Errors in TD-DFT excitation energies for the nonlinear PAHs, plotted as a function of the CT metric, Λ. The LRC-ωPBE and LRC-ωPBEh functionals afford similar results, so only the latter is shown here.

C. Tozer's CT Metric. Tozer and co-workers[36,47] have proposed a diagnostic test to determine whether a particular TD-DFT excited state is beset by sufficient CT contamination such that the predicted excitation energy may not be reliable. This diagnostic comes in the form of a metric, Λ, given by

$$\Lambda = \frac{\sum_{ia} (X_{ia} + Y_{ia})^2 O_{ia}}{\sum_{jb} (X_{jb} + Y_{jb})^2} \qquad (4)$$

which is defined such that $0 \leq \Lambda \leq 1$. The quantities X_{ia} and Y_{ia} are the TD-DFT transition amplitudes (using standard notation[68]), which determine the transition density matrix, and O_{ia} is the overlap integral between $|\phi_i(\vec{r})|$ and $|\phi_a(\vec{r})|$, where ϕ_i and ϕ_a are occupied and virtual MOs, respectively. When $\Lambda = 0$, the transition in question involves donor and acceptor orbitals with no spatial overlap, and methods such as TD-B3LYP and TD-BP86 will undoubtedly underestimate the excitation energy in such cases, probably by a large amount. On the basis of a set of benchmark tests, Tozer and co-workers suggest that TD-B3LYP excitation energies are unreliable if $\Lambda < 0.3$,[36] although they later reported an example where this metric fails to detect a problematic CT state.[69]

Values of Λ for the linear acene series have been reported in ref 37 and in the Supporting Information for ref 36, but because these data are relevant to the discussion at hand, they are also listed in Table 5. In all cases (naphthalene through hexacene), we find that $\Lambda > 0.8$ for the $S_0 \rightarrow {}^1L_a$ excitation, whereas $\Lambda \approx 0.6$ for the $S_0 \rightarrow {}^1L_b$ excitation. As was pointed out in a previous analysis of the linear acenes,[37] these values are not only above the $\Lambda = 0.3$ threshold established in previous tests, but in fact it is the ${}^1L_b$ state that exhibits the larger value of the CT metric! Furthermore, although TD-DFT errors for the $S_0 \rightarrow {}^1L_a$ excitation energy are clearly correlated with molecular size, Λ exhibits no such size dependence. This is consistent with a transition density comprised of excitations from delocalized π MOs into delocalized π^* MOs.

For the nonlinear PAHs, Figure 13 provides a plot of the excitation energy errors versus Λ; as with the linear acenes, the ${}^1L_a$ state exhibits larger values of Λ than does ${}^1L_b$. The original proposal of Λ as a useful diagnostic was based on an observed correlation between TD-DFT errors and the value of this metric,[36] but no evidence of any such correlation is found in the PAH data. On the other hand, a clear correlation is evident in the data of ref 36 only when $\Lambda \lesssim 0.5$, whereas $\Lambda > 0.55$ for all of the PAHs. Moreover, the range of Λ values that is obtained, for a given excited state, is no different for TD-B3LYP than it is for the TD-LRC-DFT methods. This is consistent with the observed similarity between the difference densities computed using different functionals, and both observations are consistent with the notion that the calculated ${}^1L_a$ (or ${}^1L_b$) electron density is not significantly different among the various functionals. What changes from one functional to the next is the manner in which the excitation energies *depend* on this electron density.

D. Atomic Partition of Particle/Hole Densities. A potentially useful way to analyze the extent of charge separation is to decompose the transition density matrices into "particle" and "hole" components, which can then be analyzed separately. To do this, we define a density matrix $\mathbf{D}^{\mathrm{elec}}$ for the excited electron, whose matrix elements are

$$D^{\mathrm{elec}}_{ab} = \sum_i (\mathbf{X}^\dagger + \mathbf{Y}^\dagger)_{ai}(\mathbf{X} + \mathbf{Y})_{ib} \qquad (5)$$

A density matrix, $\mathbf{D}^{\mathrm{hole}}$, for the hole that is left behind in the occupied space is defined similarly:

$$D^{\mathrm{hole}}_{ij} = \sum_a (\mathbf{X} + \mathbf{Y})_{ia}(\mathbf{X}^\dagger + \mathbf{Y}^\dagger)_{aj} \qquad (6)$$

Note that $\mathbf{D}^{\mathrm{elec}} + \mathbf{D}^{\mathrm{hole}} = \Delta\mathbf{P}$ is the difference between the ground- and excited-state one-electron density matrices. Upon transforming $\mathbf{D}^{\mathrm{elec}}$ and $\mathbf{D}^{\mathrm{hole}}$ into the AO basis, one can write

$$\Delta q = \mathrm{tr}(\mathbf{D}^{\mathrm{elec}}\mathbf{S}) = -\,\mathrm{tr}(\mathbf{D}^{\mathrm{hole}}\mathbf{S}) \qquad (7)$$

where $\mathbf{S}$ is the AO overlap matrix. The quantity Δq is the total charge that is transferred from the occupied space to the virtual space. For TDA calculations, $\Delta q = -1$ (exactly), but deviations from -1 are possible in full TD-DFT calculations. (Typically, however, the Y_{ia} amplitudes are quite small; hence, $\Delta q \approx -1$ even in full TD-DFT calculations.)

Equation (7) immediately suggests that the matrix products $\mathbf{D}^{\mathrm{elec}}\mathbf{S}$ and $\mathbf{D}^{\mathrm{hole}}\mathbf{S}$ are amenable to Mulliken-style population analysis, just as $\mathbf{PS}$ is analyzed in ground-state calculations.[70] In particular, the matrix element $(\mathbf{D}^{\mathrm{elec}}\mathbf{S})_{\nu\nu}$ represents the νth AO's contribution to the excited electron, while $(\mathbf{D}^{\mathrm{hole}}\mathbf{S})_{\nu\nu}$ is a contribution to the hole. The sum of these quantities,

$$\Delta q_\nu = (\mathbf{D}^{\mathrm{elec}}\mathbf{S})_{\nu\nu} + (\mathbf{D}^{\mathrm{hole}}\mathbf{S})_{\nu\nu} \qquad (8)$$

represents the contribution to Δq arising from the νth AO, and

$$\Delta Q_A = \sum_{\nu \in A} \Delta q_\nu \qquad (9)$$

is the change in the Mulliken population of atom A, upon electronic excitation. Generalization to Löwdin-style population analysis[70] is straightforward, and we have implemented these "particle/hole" population analyses into a locally modified version of Q-Chem. Because Mulliken and Löwdin analysis often produce erratic results in extended basis sets, we employ the somewhat more compact 6-31G* basis set for these calculations, rather than the cc-pVTZ basis that is used elsewhere in this work.

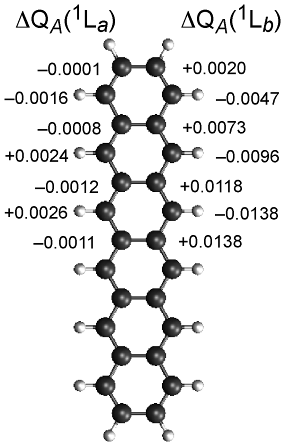


Figure 14. Differences between excited-state and ground-state Mulliken charges [ΔQ_A, from eq (9)] for the carbon atoms in hexacene, computed at the TD-B3LYP/6-31G* level. Only the symmetry-unique carbon atoms have been labeled, with $S_0 \rightarrow {}^1L_a$ charge differences on the left side and $S_0 \rightarrow {}^1L_b$ charge differences on the right side.

The expectation, based on valence-bond considerations, is that the ${}^1L_a$ state should exhibit charge separation on the length scale of C−C bonds. In light of this, it is surprising that both Mulliken and Löwdin population analyses afford an alternating pattern of charges on the carbon atoms *for the* ${}^1L_b$ *state* (see Figure 14). To some extent, the ${}^1L_a$ state exhibits a similar pattern, but in this case there is an additional (albeit quite small) accumulation of negative charge at the ends of the molecule. This suggests that a small amount of charge is pushed to extremities of the molecule in the ${}^1L_a$ state but not in the case of ${}^1L_b$. However, the magnitudes of the charge differences, ΔQ_A, are difficult to reconcile with the valence-bond interpretations of ${}^1L_a$ and ${}^1L_b$; charge differences on individual carbon atoms are ∼100 times larger in the ${}^1L_b$ state than in the ${}^1L_a$ state. (This is true even when we resort to minimal basis sets, in the interest of obtaining more "chemically intuitive" Mulliken charges.) Analysis of the particle and hole contributions to ΔQ_A shows that these contributions, which must have opposite sign, are typically ∼100 times larger than ΔQ_A itself. This is indicative of delocalized NTOs, with a very subtle pattern of net charge separation.

Mulliken charge differences for PAHs **7**, **10**, and **11**, computed at the TD-B3LYP/6-31G* level, are shown in Figure 15. (Charges computed using LRC functionals are quite similar.) As compared to the linear acenes, these examples exhibit far less disparity between the charge differences associated with the $S_0 \rightarrow {}^1L_a$ and $S_0 \rightarrow {}^1L_b$ excitations. The ΔQ_A values in both **7** and **11** suggest some intramolecular charge separation (from the bottom of the molecule to the top of the molecule, as it is shown

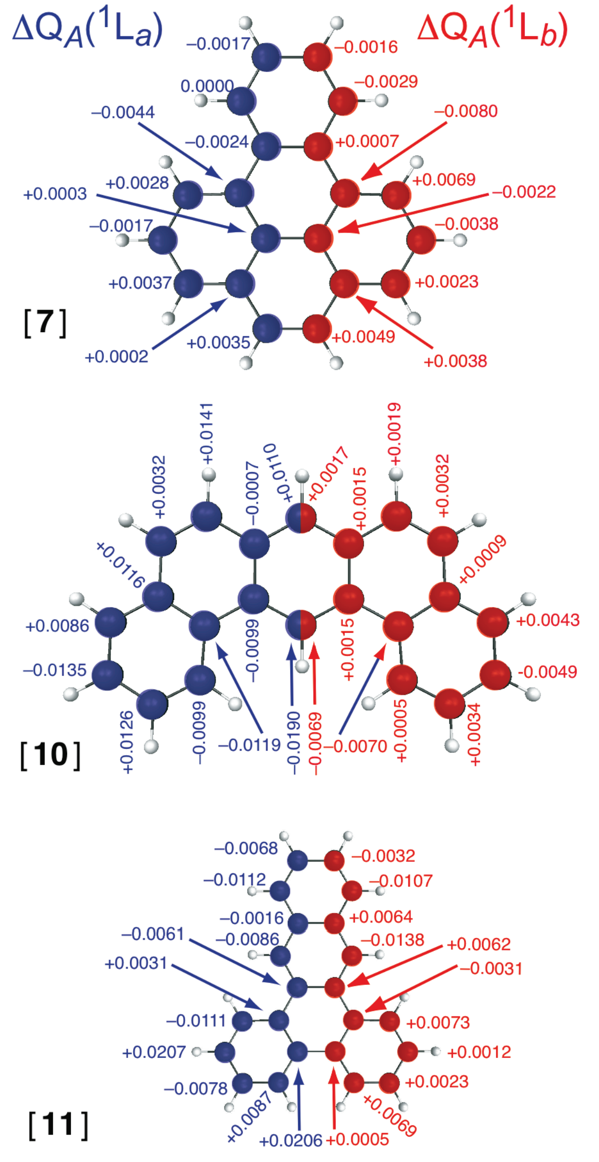


Figure 15. Differences between excited-state and ground-state Mulliken charges for the carbon atoms in PAHs **7**, **10**, and **11**. Charges were computed at the TD-B3LYP/6-31G* level and are listed for the symmetry-unique carbon only; blue labels (left side) correspond to ^1L_a and red labels (right side) correspond to ^1L_b.

in Figure 15), and this effect appears to be more significant in **11** than it is in **7**.

We consider this result in light of the two partitions of the PAH data set that were introduced in section III.B. The ^1L_a excitation energy for **11** is significantly underestimated at the TD-B3LYP level, despite the fact that this molecule falls into the peri-condensed subset that has a somewhat smaller MAE as compared to the cata-condensed subset (see Table 4). PAH **10** falls into the cata-condensed and "3+" subsets, for which the TD-B3LYP MAEs are larger than they are for the full data set, and Figure 15 shows that the magnitudes of the ΔQ_A values for the $S_0 \rightarrow {}^1L_a$ excitation in **10** are comparable to those observed for **11**, even though the former does not exhibit the sort of overall charge separation that is observed in the latter. In both **10** and **11**, the magnitudes of the ΔQ_A values are notably larger than they are in **7**.

This analysis suggests that the magnitude of the TD-B3LYP error in the ^1L_a excitation energy is somehow related to the extent of charge reorganization upon $S_0 \rightarrow {}^1L_a$ excitation. This is certainly not a predictive metric, however, and it is further complicated by examination of the Mulliken charge differences for the $S_0 \rightarrow {}^1L_b$ excitations in Figure 15. In both **7** and **11**, the ΔQ_A values exhibit similar patterns for both $S_0 \rightarrow {}^1L_a$ and $S_0 \rightarrow {}^1L_b$ excitation; namely, the Mulliken charge differences alternate in sign across the carbon backbone. The magnitudes of the ΔQ_A values are also quite similar for both states. Thus, while it appears that Mulliken charge differences may help to explain why the ^1L_a excitation energies in certain PAHs suffer larger TD-B3LYP errors than others, these charge differences are of little help in understanding why these errors are smaller for ^1L_b than for ^1L_a.

E. Summary. In view of these observations, we are left with the following situation. The trends in TD-DFT excitations energies with respect to molecular size strongly suggest that the ^1L_a state in many different PAHs exhibits some sort of CT or charge-separation character that is not present in the ^1L_b state. The fact that TD-LRC-DFT calculations largely mitigate this problem adds to the (circumstantial) evidence for CT character in the ^1L_a state of the linear acene molecules. At the same time, attempts to discern this charge-separated character from the NTOs or transition densities are quite tenuous, and at best these analyses suggest only a very slight concentration of charge at the ends of the molecule. It is essentially impossible to discern any CT character from the MOs or difference density plots, and the TD-DFT charge-overlap metric introduced by Tozer and coworkers[36,47] also fails to raise any warning flags. Mulliken- or Löwdin-style analyses of the transition densities and excited-state atomic charges offer some insight into the nature of the charge separation, but some such charge separation is observed even in the case of excitations where TD-B3LYP predicts the excitation energy accurately.

V. CONCLUSIONS

We have evaluated the performance of TD-DFT and TD-LRC-DFT approaches for calculation of the vertical excitation energies of the ^1L_a and ^1L_b states of various PAHs. While methods such as TD-B3LYP and TD-BP86 provide reasonably accurate values for the ^1L_b excitation energies, ^1L_a excitation energies are consistently underestimated, with errors that increase as the size of the molecule increases. In contrast, TD-LRC-DFT excitation energies are accurate to within ~0.1 eV for the ^1L_a excitation energies. In the linear acene sequence, these methods also correctly predict a crossover point at which the ^1L_a state becomes lower in energy than the ^1L_b state. At the same time, ^1L_b excitation energies are systematically overestimated by LRC functionals (but without any clear size-dependent trend) and are somewhat less accurate than TD-B3LYP results.

The most important result to emerge from this work is an indication, based upon size-dependent trends in excitation energies, that the ^1L_a excited state in many PAHs exhibits some sort of charge-separated character that is not present in the ^1L_b state. This feature causes ^1L_a excitation energies to diverge from experimental values as the size of the molecule increases, when methods such as TD-B3LYP, TD-PBE0, or TD-BP86 are employed. Our hypothesis concerning the charge-separated nature of the ^1L_a state is consistent with the valence-bond language that has long been used to describe the ^1L_a and ^1L_b states, according to which the ^1L_a state is ionic while ^1L_b is covalent.[30−35] However, although the ionic character of ^1L_a emerges cleanly from analysis of TD-PPP calculations,[28] where it correlates well with the error in TD-BP86 excitation energies, analysis of all-electron TD-DFT calculations is much more ambiguous in this respect.

While it is possible, in post hoc analysis, to rationalize the relatively ionic character of ^1L_a by examining TD-DFT transition densities and NTOs, other forms of analysis—including TD-DFT difference density plots and Mulliken population analysis of particle and hole density matrices—do not obviously suggest that ^1L_a exhibits any more CT character than does ^1L_b. A metric specifically designed to detect and quantify CT character in TD-DFT calculations,[36] and which has been successful in this respect, for a variety of molecules,[36,47] also fails to indicate that ^1L_a is more "CT-like" than ^1L_b. This metric is certainly not perfect, and Peach et al.[69] have identified a case where it fails to flag an excitation that (based on examination of the MOs) is clearly a CT state and where the excitation energy is substantially underestimated at the TD-PBE and TD-PBE0 levels. The difference here is that the ^1L_a states in the PAHs are *not* clear examples of CT states.

These observations suggest the possibility that medium- to large-size conjugated organic molecules may exhibit subtle charge-separation effects that are difficult to identify a priori but which cause conventional TD-DFT methods to overstabilize these states, possibly by a significant amount. This is a potentially serious problem in cases where TD-DFT is applied to molecules that are too large to perform any high-level ab initio benchmarks and where reliable experimental data are unavailable. Further analysis is required in order to develop a diagnostic that can automatically detect such states. In the meantime, we recommend performing TD-DFT calculations with both LRC functionals (e.g., LRC-ωPBE or LRC-ωPBEh) and also non-LRC functionals (e.g., B3LYP or PBE0) so that the results may be compared and potentially problematic states may be detected.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information. Cartesian coordinates for the optimized PAH geometries, TD-DFT excitation energies, and values of the Λ metric. This information is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: herbert@chemistry.ohio-state.edu.

■ ACKNOWLEDGMENT

This work was supported by a National Science Foundation CAREER award (Grant CHE-0748448). Calculations were performed at the Ohio Supercomputer Center (Project No. PAS-0291). J.M.H. gratefully acknowledges a fellowship from the Alfred P. Sloan Foundation. The free software programs OpenCubMan,[71] Visual Molecular Dynamics,[72] and MacMolPlt[73] were used to generate some of the figures.

■ REFERENCES

(1) Dreuw, A.; Weisman, J. L.; Head-Gordon, M. *J. Chem. Phys.* **2003**, *119*, 2943–2946.

(2) Dreuw, A.; Head-Gordon, M. *J. Am. Chem. Soc.* **2004**, *126*, 4007–4016.

(3) Magyar, R. J.; Tretiak, S. *J. Chem. Theory Comput.* **2007**, *3*, 976–987.

(4) Lange, A. W.; Herbert, J. M. *J. Am. Chem. Soc.* **2009**, *131*, 3913–3922.

(5) Bernasconi, L.; Sprik, M.; Hutter, J. *J. Chem. Phys.* **2003**, *119*, 12417–12431.

(6) Neugebauer, J.; Louwerse, M. J.; Baerends, E. J.; Wesołowski, T. A. *J. Chem. Phys.* **2005**, *122*, 094115-1–094115-13.

(7) Lange, A.; Herbert, J. M. *J. Chem. Theory Comput.* **2007**, *3*, 1680–1690.

(8) Iikura, H.; Tsuneda, T.; Yanai, T.; Hirao, K. *J. Chem. Phys.* **2001**, *115*, 3540–3544.

(9) Tawada, Y.; Tsuneda, T.; Yanagisawa, S.; Yanai, T.; Hirao, K. *J. Chem. Phys.* **2004**, *120*, 8425–8433.

(10) Song, J.-W.; Hirosawa, T.; Tsuneda, T.; Hirao, K. *J. Chem. Phys.* **2007**, *126*, 154105-1–154105-7.

(11) Sato, T.; Tsuneda, T.; Hirao, K. *J. Chem. Phys.* **2007**, *126*, 234114-1–234114-12.

(12) Yanai, T.; Tew, D. P.; Handy, N. C. *Chem. Phys. Lett.* **2004**, *393*, 51–57.

(13) Peach, M. J. G.; Cohen, A. J.; Tozer, D. J. *Phys. Chem. Chem. Phys.* **2006**, *8*, 4543–4549.

(14) Vydrov, O. A.; Scuseria, G. E. *J. Chem. Phys.* **2006**, *125*, 234109-1–234109-9.

(15) Henderson, T. M.; Janesko, B. G.; Scuseria, G. E. *J. Chem. Phys.* **2008**, *128*, 194105-1–194105-9.

(16) Henderson, T. M.; Janesko, B. G.; Scuseria, G. E. *J. Phys. Chem. A* **2008**, *112*, 12530–12542.

(17) Chai, J.-D.; Head-Gordon, M. *J. Chem. Phys.* **2008**, *128*, 084106-1–084106-15.

(18) Rohrdanz, M. A.; Herbert, J. M. *J. Chem. Phys.* **2008**, *129*, 034107-1–034107-9.

(19) Rohrdanz, M. A.; Martins, K. M.; Herbert, J. M. *J. Chem. Phys.* **2009**, *130*, 054112-1–054112-8.

(20) Baer, R.; Livshits, E.; Salzner, U. *Annu. Rev. Phys. Chem.* **2010**, *61*, 85–109.

(21) Savin, A. In *Recent Advances in Density Functional Methods*, Part I; *Recent Advances in Computational Chemistry*, Vol. 1; Chong, D. P., Ed.; World Scientific: Singapore, 1995; pp 129—153.

(22) Gill, P. M. W.; Adamson, R. D.; Pople, J. A. *Mol. Phys.* **1996**, *88*, 1005–1009.

(23) Adamson, R. D.; Dombroski, J. P.; Gill, P. M. W. *J. Comput. Chem.* **1999**, *20*, 921–927.

(24) Tozer, D. J.; Amos, R. D.; Handy, N. C.; Roos, B. O.; Serrano-Andrés, L. *Mol. Phys.* **1999**, *97*, 859–868.

(25) Silva-Junior, M. R.; Schreiber, M.; Sauer, S. P. A.; Thiel, W. *J. Chem. Phys.* **2008**, *129*, 104103-1–104103-14.

(26) Jacquemin, D.; Wathelet, V.; Perpéte, E. A.; Adamo, C. *J. Chem. Theory Comput.* **2009**, *5*, 2420–2435.

(27) Grimme, S.; Parac, M. *ChemPhysChem* **2003**, *4*, 292–295.

(28) Parac, M.; Grimme, S. *Chem. Phys.* **2003**, *292*, 11–21.

(29) Platt, J. R. *J. Chem. Phys.* **1949**, *17*, 484–495.

(30) Handa, T. *Bull. Chem. Soc. Jpn.* **1963**, *36*, 235–247.

(31) Orchin, M.; Jaffé, H. H. *Symmetry. Orbitals, and Spectra*; Wiley: New York, 1971.

(32) Malloci, G.; Mulas, G.; Cappellini, G.; Joblin, C. *Chem. Phys.* **2007**, *340*, 43–58.

(33) Hashimoto, T.; Nakano, H.; Hirao, K. *J. Chem. Phys.* **1996**, *104*, 6244–6258.

(34) Hirao, K.; Nakano, H.; Nakayama, K.; Dupuis, M. *J. Chem. Phys.* **1996**, *105*, 9227–9239.

(35) Hirao, K.; Nakano, H.; Nakayama, K. *J. Chem. Phys.* **1997**, *107*, 9966–9974.

(36) Peach, M. J. G.; Benfield, P.; Helgaker, T.; Tozer, D. J. *J. Chem. Phys.* **2008**, *128*, 044118-1–044118-8.

(37) Wong, B. M.; Hsieh, T. H. *J. Chem. Theory Comput.* **2010**, *6*, 3704–3712.

(38) Zilberg, S.; Haas, Y.; Shaik, S. *J. Phys. Chem.* **1995**, *99*, 16558–16565.

(39) Gill, P. M. W.; Johnson, B. G.; Pople, J. A. *Chem. Phys. Lett.* **1993**, *209*, 506–512.

(40) Hirata, S.; Head-Gordon, M. *Chem. Phys. Lett.* **1999**, *314*, 291–299.

(41) Shao, Y.; Fusti-Molnar, L.; Jung, Y.; Kussmann, J.; Ochsenfeld, C.; Brown, S. T.; Gilbert, A. T. B.; Slipchenko, L. V.; Levchenko, S. V.; O'Neill, D. P., Jr.; R., A. D.; Lochan, R. C.; Wang, T.; Beran, G. J. O.; Besley, N. A.; Herbert, J. M.; Lin, C. Y.; Van Voorhis, T.; Chien, S. H.; Sodt, A.; Steele, R. P.; Rassolov, V. A.; Maslen, P. E.; Korambath, P. P.; Adamson, R. D.; Austin, B.; Baker, J.; Byrd, E. F. C.; Dachsel, H.; Doerksen, R. J.; Dreuw, A.; Dunietz, B. D.; Dutoi, A. D.; Furlani, T. R.; Gwaltney, S. R.; Heyden, A.; Hirata, S.; Hsu, C.-P.; Kedziora, G.; Khalliulin, R. Z.; Klunzinger, P.; Lee, A. M.; Lee, M. S.; Liang, W.; Lotan, I.; Nair, N.; Peters, B.; Proynov, E. I.; Pieniazek, P. A.; Rhee, Y. M.; Ritchie, J.; Rosta, E.; Sherrill, C. D.; Simmonett, A. C.; Subotnik, J. E.; Woodcock, H. L., III; Zhang, W.; Bell, A. T.; Chakraborty, A. K.; Chipman, D. M.; Keil, F. J.; Warshel, A.; Hehre, W. J.; Schaefer, H. F., III; Kong, J.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172–3191.

(42) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.

(43) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.

(44) Tsuneda, T.; Suzumura, T.; Hirao, K. *J. Chem. Phys.* **1999**, *110*, 10664–10678.

(45) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

(46) We use "μ" to indicate LRC functionals in which the short-range GGA exchange functional is constructed according to the procedure of Iikura et al.[8] Short-range versions of Becke's B88 exchange functional[42] ("μB88") and Perdew—Burke—Ernzerhof exchange[45] ("μPBE") are available in Q-Chem. Henderson et al.[15] have developed an alternative version of short-range PBE exchange that they call ωPBE, and we adopt the same notation. The notation "LC-BOP" has also been used for the functional that we call LRC-μBOP,[10] and "LC-ωPBE" has been used for what we call LRC-ωPBE,[15,26] albeit with different parameters than those used here. The LRC-ωPBE functional has has also been called simply "ωPBE",[47] but we dislike this particular nomenclature, because ωPBE exchange is sometimes used *without* long-range HF exchange.[48]

(47) Plötner, J.; Tozer, D. J.; Dreuw, A. *J. Chem. Theory Comput.* **2010**, *6*, 2315–2324.

(48) Heyd, J.; Scuseria, G. E. *J. Chem. Phys.* **2004**, *121*, 1187–1192.

(49) Lange, A. W.; Rohrdanz, M. A.; Herbert, J. M. *J. Phys. Chem. B* **2008**, *112*, 6304–6308.

(50) Wong, B. M.; Cordaro, J. G. *J. Chem. Phys.* **2008**, *129*, 214703-1–214703-8.

(51) Wong, B. M.; Piacenza, M.; Della Sala, F. *Phys. Chem. Chem. Phys.* **2009**, *11*, 4498–4508.

(52) Wang, Y.-L.; Wu, G.-S. *Int. J. Quantum Chem.* **2007**, *108*, 430–439.

(53) Stein, T.; Kronik, L.; Baer, R. *J. Chem. Phys.* **2009**, *131*, 244119-1–244119-5.

(54) Scher, H.; Montroll, E. W. *Phys. Rev. B* **1975**, *12*, 2455–2477.

(55) Callis, P. R. *Int. J. Quantum Chem.* **1984**, *26*, 579–588.

(56) Rogers, D. M.; Besley, N. A.; O'Shea, P.; Hirst, J. D. *J. Phys. Chem. B* **2005**, *109*, 23061–23069.

(57) Clar, E. *Polycyclic Hydrocarbons*, Vol. 1; Academic Press: London, 1964.

(58) Vijayalakshmi, K. P.; Suresh, C. H. *New J. Chem.* **2010**, *34*, 2132–2138.

(59) Birks, J. B. *Photophysics of Aromatic Molecules*; Wiley: New York, 1970.

(60) Schmidt, W. *J. Chem. Phys.* **1977**, *66*, 828–845.

(61) Champagne, B.; Perpète, E. A.; van Gisbergen, S. J. A.; Baerends, E.-J.; Snijders, J. G.; Soubra-Ghaoui, C.; Robins, K. A.; Kirtman, B. *J. Chem. Phys.* **1998**, *109*, 10489–10498.

(62) Champagne, B.; Perpète, E. A.; Jacquemin, D.; van Gisbergen, S. J. A.; Baerends, E.-J.; Soubra-Ghaoui, C.; Robins, K. A.; Kirtman, B. *J. Phys. Chem. A* **2000**, *104*, 4755–4763.

(63) Sekino, H.; Maeda, Y.; Kamiya, M.; Hirao, K. *J. Chem. Phys.* **2007**, *126*, 014107-1–014107-6.

(64) Pariser, R.; Parr, R. G. *J. Chem. Phys.* **1953**, *21*, 466–471.

(65) Pariser, R.; Parr, R. G. *J. Chem. Phys.* **1953**, *21*, 767–776.

(66) Pople, J. A. *Trans. Faraday Soc.* **1953**, *49*, 1375–1385.

(67) Martin, R. L. *J. Chem. Phys.* **2003**, *118*, 4775–4777.

(68) Dreuw, A.; Head-Gordon, M. *Chem. Rev.* **2005**, *105*, 4009–4037.

(69) Peach, M. J. G.; Le Sueur, C. R.; Ruud, K.; Guillaume, M.; Tozer, D. J. *Phys. Chem. Chem. Phys.* **2009**, *11*, 4465–4470.

(70) Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry*; Collier Macmillan: London, 1982.

(71) Haranczyk, M.; Gutowski, M. *J. Chem. Theory Comput.* **2008**, *4*, 689–693.

(72) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33–38.

(73) Bode, B. M.; Gordon, M. S. *J. Mol. Graphics Modell.* **1998**, *16*, 133–138.

JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Artificial Bee Colony Optimization of Capping Potentials for Hybrid Quantum Mechanical/Molecular Mechanical Calculations

Christoph Schiffmann and Daniel Sebastiani*

Physics Department, Freie Universität Berlin, Arnimallee 14, 14195 Berlin, Germany

S Supporting Information

ABSTRACT: We present an algorithmic extension of a numerical optimization scheme for analytic capping potentials for use in mixed quantum−classical (quantum mechanical/molecular mechanical, QM/MM) ab initio calculations. Our goal is to minimize bond-cleavage-induced perturbations in the electronic structure, measured by means of a suitable penalty functional. The optimization algorithm—a variant of the artificial bee colony (ABC) algorithm, which relies on swarm intelligence—couples deterministic (downhill gradient) and stochastic elements to avoid local minimum trapping. The ABC algorithm outperforms the conventional downhill gradient approach, if the penalty hypersurface exhibits wiggles that prevent a straight minimization pathway. We characterize the optimized capping potentials by computing NMR chemical shifts. This approach will increase the accuracy of QM/MM calculations of complex biomolecules.

1. INTRODUCTION

Accurate simulation of structural and dynamical phenomena of complex biomolecular systems by means of first-principles molecular dynamics simulation techniques is still a challenge for modern physics and chemistry. Despite enormous progress in recent decades, predictive modeling of the interplay of intramolecular and intermolecular interactions is still far from being a routine problem. For determination of structural data in biophysics and biochemistry, the combination of spectroscopic experiments with advanced theoretical predictions and computer simulations is becoming increasingly popular, because this combination often yields a predictive power above the sum of the individual approaches.[1−9] Nevertheless, the first-principles prediction of noncovalent packing effects and the ab initio prediction of experimentally observable spectra is not possible for regular biosystems because of their inherent complexity and, last but not least, their sheer size. Thus, one has to resort either to the modeling of elementary subunits[10−14] or alternatively to hybrid quantum-mechanical + mechanical modeling (QM/MM) approaches.[15−31] One of the difficulties of such a hybrid approach is the interface region between the two different regions. If one of the atoms is located in the quantum (QM) region and the other in the classical (MM) part, then a chemical bond is "broken" as a consequence. This situation is sketched in Figure 1. Similar problems arise when MM atoms are located near a QM region, because the QM and MM descriptions are not genuinely compatible. Thus, a suitable interface has to be used, which can mutually couple the two schemes in a realistic way.

There are several well-established methods to tackle the bond saturation problem, in particular hydrogen[32] or fluorine[33] atoms, precomputed (frozen) atomic orbitals,[34,35] generalized hybrid orbitals,[36−38] quantum capping potentials,[39−42] or designed heptavalent capping potentials.[43] Complementary, effective fragment potentials[44,45] and field-adapted adjustable density matrix assembler[46−48] approach the repartiotioning problem itself. Our approach is conceptually simpler than most of the former ones; we aim at designing a fictitious capping atom to saturate the QM subsystem, which is realized by a regular atomic pseudopotential.

Specifically, we want to improve a method that has been developed recently[49] in view of more complex bond-cleavage situations. This approach is based on analytical effective core potentials (pseudopotentials) of Goedecker−Teter−Hutter (GTH) type,[50,51] in line with previous QM/MM studies.[14,23,52] Our goal is to optimize the pseudopotential parameters in such a way that the change of electronic density in the quantum part of a QM/MM calculation is minimal with respect to a "full-QM" calculation.

In this way, we also ensure that structural parameters and spectroscopic properties in the direct neighborhood of a QM/MM bond cleavage are modeled with a high degree of reliability.

To achieve this aim, we define a penalty functional that quantifies the deviation of the electronic density in a molecular fragment from the corresponding density in the complete molecule, while simultaneously penalizing changes in the equilibrium bond distance and frequency. The penalty functional is minimized iteratively by varying the coefficients of the capping potential placed at the bond-cleavage site.

However, a straightforward mimization approach like steepest descent[53,54] or a simplex method[55] carries the risk of getting stuck in local minima. To avoid this pitfall, we aim for global optimization including stochastic elements by means of a swarm intelligence-based algorithm. In recent years, biology-inspired algorithms[56,57] turned out to be more effective than conventional algorithms.[58]

In this work, we employ a variation of the artificial bee colony algorithm[59−61] (ABC), which mimicks the foraging behavior of honeybees for function minimization. We are especially interested in proving the usablility for optimizations within electronic structure calculations and studying the performance of the algorithm. The optimized capping potentials are intended to

Received: December 10, 2010
Published: April 07, 2011

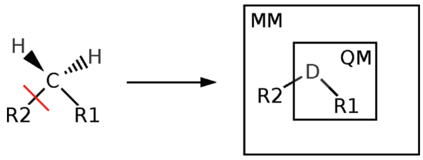


Figure 1. General QM/MM repartioning principle in which the C−R2 bond crosses the QM/MM border and is cleaved. The CH_2 group is replaced by a capping potential associated with a fictitious particle D, saturating the C−R1 bond and hence terminating the QM region.

saturate the quantum region in hybrid QM/MM calculations. In most cases, this saturation affects a single C−C bond and can therefore be done by means of a hydrogenoid atom; however, the properties of this hydrogenoid atom should resemble as much as possible the characteristics of the carbon atom that has been "cut" out from the quantum calculation. Hence, we need a fictitious atom that is monovalent but behaves like a (four-valent) carbon atom in terms of bond distance, potential energy curve(s), and electronic structure.

We further characterize the perturbative effect of bond cleavage by means of NMR chemical shifts, which are known to be particularly sensitive to both intramolecular electronic structure and intermolecular effects such as hydrogen bonding.[62−66] Hence, we can not only gauge the direct perturbing effect of the cleaved bond on the electronic structure of the remaining part of a molecule but also quantitatively describe how strongly its response properties are tainted by the QM/MM bond cleavage.

2. OPTIMIZATION APPROACH

In QM/MM calculations, the dummy atom has to saturate the last covalent bond in the quantum region of the molecule, that is, the bond that is cleaved by QM/MM repartitioning. The true character of the bond, however, cannot be easily reproduced by a simple terminal atom. It is therefore necessary to tune the dummy's properties in a way that the resulting deviation in the quantum region's electronic structure is minimal. To do so, one has to find a capping potential that equips the dummy with the desired properties.

2.1. Definition of Penalty Functional.

Our optimization scheme aims to find a capping potential V_{cap} that gives rise to an electronic density in the quantum region ($\rho[V_{cap}]$) that deviates only in a minimal manner from the reference electron density (ρ^{ref}), that is, the density when the whole molecule is quantum-mechanically treated. Further, we want to preserve the equilibrium bond length and vibrational properties of the bond that is cleaved to allow for an easy coupling of the first classical MM atom and to avoid the need for additional geometric constraints (see Komin and Sebastiani[49] and von Lilienfeld-Toal et al.[67] for a more detailed description).

Therefore, we define a functional that penalizes deviations of these properties from their target values obtained in a full-QM calculation:

$$\mathscr{P}[V_{cap}] = \omega_\rho \sum_{j=1}^{N_{geo}} \int_\Omega d^3r \, |\rho_j{}^{ref}(\mathbf{r}) - \rho_j[V_{cap}](\mathbf{r})|^2 + \omega_f \sum_{j=1}^{N_{geo}} |\mathbf{F}_j{}^{ref} - \mathbf{F}_j[V_{cap}]|^2 + \omega_e \sum_{j=2}^{N_{geo}} |(E_j{}^{ref} - E_1{}^{ref}) - (E_j[V_{cap}] - E_1[V_{cap}])|^2 \tag{1}$$

The integration volume Ω is restricted to an area where penalization is meaningful, that is, the union of spheres around all QM atoms except the dummy with radii $r_{cov}{}^{spc}$, where $r_{cov}{}^{spc}$ is the covalent radius of the atom species (spc). $\mathbf{F}$ denotes the force acting on the dummy (with respect to its uncapped counterpart) and E is the total energy. ω_ρ, ω_f, and ω_e are weighting factors that ensure an adequate relative importance between density, force, and energy penalization. Finally, the penalty is evaluated for $N_{geo} = 3$ molecular geometries, which correspond to variations of the cleaving bond length.

We note at this point that we have replaced a multielectron group (e.g., methyl) with a fictitious monovalent atom, which changes the total number of electrons in the system. Hence, the integration of a direct density difference can never vanish completely, unless the affected regions are entirely excluded from the integration. This also leads to the effect that the penalty functional will in general never reach zero during a capping potential optimization.

2.2. GTH Pseudopotentials.

We assume that an optimal capping potential can be expressed as an analytical GTH potential:[50,51]

$$V_{cap}(\mathbf{r}, \mathbf{r}') = V_{loc}(\mathbf{r}) + \sum_{l=0}^{l_{max}} V_l(\mathbf{r}, \mathbf{r}') \tag{2}$$

consisting of a local component V_{loc}, eq 3, and 0−3 (l_{max}) nonlocal components V_l, eq 4, with the form

$$V_{loc}(\mathbf{r}) = \frac{-Z_{ion}}{r} \operatorname{erf}\left(\frac{r}{\sqrt{2} r_{loc}}\right) + \exp\left[-\frac{1}{2}\left(\frac{r}{r_{loc}}\right)^2\right] \sum_{j=1}^{4} C_j \left(\frac{r}{r_{loc}}\right)^{2(j-1)} \tag{3}$$

$$V_l(\mathbf{r}, \mathbf{r}') = \sum_{i=1}^{3} \sum_{j=1}^{3} \sum_{m=-l}^{+l} Y_{l,m}\left(\frac{\mathbf{r}}{r}\right) p_i{}^l(r) h_{i,j}{}^l p_j{}^l(r') Y_{l,m}\left(\frac{\mathbf{r}'}{r'}\right) \tag{4}$$

$$p_i{}^l(r) = \frac{\sqrt{2} r^{l+2(i-1)} \exp\left(-\frac{r^2}{2r_l{}^2}\right)}{r_l{}^{l+(4i-1)/2} \sqrt{\Gamma\left(l + \frac{4i-1}{2}\right)}} \tag{5}$$

Z_{ion} is a valence charge, $Y_{l,m}$ are spherical harmonics, and $h_{i,j}^l$ are scalars that define the energetic weighting of projectors $p_{i/j}^l$, eq 5, in each angular momentum channel l.

A potential of this type is fully defined via the set of N_σ parameters:

$$\{r_{loc}, C_1, C_2, C_3, C_4, r_0, h^0_{1,1}, h^0_{2,2}, h^0_{3,3}, r_1, h^1_{1,1}, ..., r_2, h^2_{1,1}, ...\} \tag{6}$$

In the following, we use $\{\sigma\}$ as simplified notation for this set.

From physical considerations, we impose an allowed interval for each parameter. Thus, the optimization takes place in an N_σ dimensional orthorhombic manifold in $\mathbb{R}^{N_\sigma}$.

2.3. Artificial Bee Colony (ABC).

The actual optimization algorithm is taken from the field of swarm intelligence and as such is inspired by nature itself. It mimicks the foraging behavior of honeybees to sample a scalar function defined on an N_σ-dimensional unit cube (U_σ) in an efficient manner. We use a set of linear transformations to rescale and shift the allowed parameter space into the unit cube.

We define a *population* as a set of N_{pop} *agents*, each representing a configuration $\{\sigma\}_a \in U_\sigma$ and the corresponding penalty

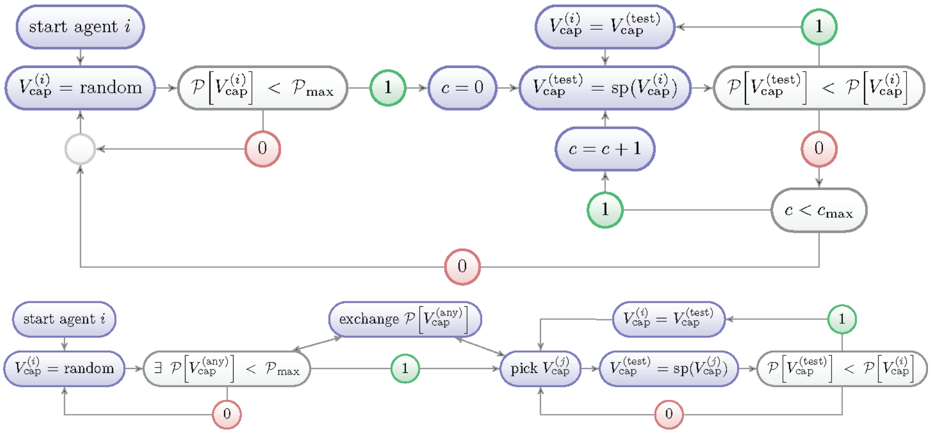

start agent i
$V^{(i)}_{\mathrm{cap}} = \mathrm{random}$
$\mathcal{P}\left[V^{(i)}_{\mathrm{cap}}\right] < \mathcal{P}_{\max}$
1
0
$c = 0$
$V^{(\mathrm{test})}_{\mathrm{cap}} = \mathrm{sp}(V^{(i)}_{\mathrm{cap}})$
$\mathcal{P}\left[V^{(\mathrm{test})}_{\mathrm{cap}}\right] < \mathcal{P}\left[V^{(i)}_{\mathrm{cap}}\right]$
1
$V^{(i)}_{\mathrm{cap}} = V^{(\mathrm{test})}_{\mathrm{cap}}$
0
$c = c + 1$
1
$c < c_{\max}$
0
start agent i
$V^{(i)}_{\mathrm{cap}} = \mathrm{random}$
$\exists\ \mathcal{P}\left[V^{(\mathrm{any})}_{\mathrm{cap}}\right] < \mathcal{P}_{\max}$
0
exchange $\mathcal{P}\left[V^{(\mathrm{any})}_{\mathrm{cap}}\right]$
1
pick $V^{(j)}_{\mathrm{cap}}$
$V^{(\mathrm{test})}_{\mathrm{cap}} = \mathrm{sp}(V^{(j)}_{\mathrm{cap}})$
$\mathcal{P}\left[V^{(\mathrm{test})}_{\mathrm{cap}}\right] < \mathcal{P}\left[V^{(i)}_{\mathrm{cap}}\right]$
1
$V^{(i)}_{\mathrm{cap}} = V^{(\mathrm{test})}_{\mathrm{cap}}$
0

Figure 2. Employee-type (top) and onlooker-type (bottom) agents.

$\mathscr{P}[\{\sigma\}_a]$. This artificial bee colony (ABC) algorithm distiguishes three types of agents which, like honeybees in nature, fulfill simple but different tasks:

- Scout-type agents have the biggest exploration tendency of the three types. A scout chooses in every cycle a random point $\{\sigma\}$ from uniform distribution in U_σ and moves unconditionally to that spot (global sampling).
- Employee-type agents have an additional local sampling component. An employee type starts like a scout but examines in each cycle a random spot $\{\sigma\}'$ from a uniform distribution in a sphere of radius r_s around its present position $\{\sigma\}$. The agent moves only if $\mathscr{P}[\{\sigma\}'] < \mathscr{P}[\{\sigma\}]$. Furthermore, if the agent cannot move for N_a successive cycles, it *abandons* its present position and restarts the search again from a fully random spot in U_σ.
- Onlooker-type agents act, from the point of view of the entire agent population, as feedback and amplify the exploitation of promising areas in U_σ that have been found by other agents. Equipped with knowledge of all agents' positions and penalties, an onlooker type randomly choses another agent's parameter set $\{\sigma\}$, inversely weighted by the penalties. Then it chooses a spot $\{\sigma\}'$ from a uniform distribution in a sphere of radius r_s around $\{\sigma\}$ and moves if $\mathscr{P}[\{\sigma\}']$ is smaller than its original penalty. Thus, this type depends on the other agents' results and ensures that good parameter regions are not lost during an employee-type resetting.

A more detailed description of employee- and onlooker-type agents and their "interactions" is given in the flowcharts in Figure 2.

Optimization starts with *initialization* of the population. In this phase, every agent, regardless of its type, is randomly placed in U_σ. Afterward, the ABC algorithm performs N_{cycle} *cycles*, each a sequence of three steps:

1. Send the employee- and onlooker-type agents to their destinations and evaluate the penalties.
2. Place the scout-type agents and the employee-type agents that abandoned their positions randomly in U_σ.
3. Store the pseudopotential parameter set $\{\sigma\}$ with the lowest penalty in the present population.

Thus, the interaction of the three types of agents, determined by the ABC algorithm, successivly searches for the global minimum $\mathscr{P}[\{\sigma\}_{\min}]$ of the penalty functional.

The pseudocode of the ABC algorithm is given in Chart 1.

2.4. Evolution and Convergence of Optimization.

We define a combined index for the evolution of all agents:

$$\tau := i \cdot N_{\mathrm{pop}} \tag{7}$$

where $0 \leq i \leq N_{\mathrm{cycle}}$ denotes the current optimization cycle. Thus, τ corresponds to the computational cost under the assumption that determination of the penalty $\mathscr{P}[\{\sigma\}]$ has a fixed computational cost for all choices of $\{\sigma\}$. As this assumption cannot be enforced formally, we limit the number of self-consistent-field (SCF) iterations during the wave function optimization for each penalty evaluation to 30. With this combined index we can describe the evolution of the ensemble of agents during one optimization run via

$$\begin{aligned} \mathscr{P}(\tau = 0) \quad &:= \min_{j=1,\ldots,N_{\mathrm{pop}}} \{\mathscr{P}[\{\sigma\}_{j,i=0}]\} \\ \mathscr{P}(\tau > 0) \quad &:= \min\left(\mathscr{P}(\tau - N_{\mathrm{pop}}),\ \min_{j=1,\ldots,N_{\mathrm{pop}}} \{\mathscr{P}[\{\sigma\}_{j,i\neq 0}]\}\right) \end{aligned} \tag{8}$$

where $\{\sigma\}_{j,i}$ are the pseudopotential parameters of agent j in the ith optmization cycle (the case $i = 0$ denotes the initialization phase). Hence, (τ) and $\{\sigma\}(\tau)$ refer to the minimal penalty after an optimization time τ and the corresponding pseudopotential parameter set.

To account for the stochastic nature of the optimization process, we run N_{trial} independent optimizations for each set of control parameters (i.e., the number of employee- and onlooker-type agents as well as the radius of the neighborhood sphere r_s). This enables us to describe the convergence behavior in statistical terms. We distinguish between different optimization runs by a new superscript $1 \leq k \leq N_{\mathrm{trial}}$:

$$\mathscr{P}_{\min}(\tau) = \min_{k=1,\ldots,N_{\mathrm{trial}}} \{\mathscr{P}^k(\tau)\} \tag{9}$$

$$\mathscr{P}_{\max}(\tau) = \max_{k=1,\ldots,N_{\mathrm{trial}}} \{\mathscr{P}^k(\tau)\} \tag{10}$$

Equations 9 and 10 describe a window in which all $\mathscr{P}^k(\tau)$ are located for a fixed setting of control parameters (best- and worst-case scenarios).

2.5. Computational Details.

We perform all calculations within density functional theory[68−70] using the BLYP[71,72] exchange−correlation functional, as implemented in the CPMD package.[73,74]

Chart 1. Pseudocode of the ABC Algorithm

```
initialize population randomly
for i = 1...N_cycles do
        run employee types
        run onlooker types
        run scout types
        abandon solutions
        update best solution
done
```

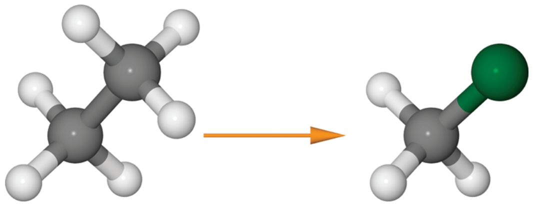


Figure 3. Ethane (C_2H_6) as test system for benchmarking. One methyl group is replaced by a capping potential (green particle), which saturates the remaining methyl group.

We use standard norm-conserving pseudopotentials[50,51] and an energy cutoff of 70 Ry for the plane-wave expansion of the Kohn−Sham orbitals.

Calculation of spectroscopic parameters, for example, NMR chemical shifts, is done within density functional perturbation theory as implemented in the linear response package of CPMD.[75−77]

3. RESULTS

3.1. Stochastic Optimization of Capping Potentials. We have applied the ABC algorithm to the optimization of GTH-type capping potentials V_{cap} for hybrid QM/MM calculations within DFT. In particular, we have examined the influence of control parameters of the ABC algorithm (i.e., the number of employee and onlooker type agents and the radius of the neighborhood sphere r_s) on the optimization process. For this purpose, we benchmarked a series of capping potential optimizations for an isolated ethane molecule (C_2H_6) in which one methyl group is replaced by a capping potential, with respect to a dummy particle D, as shown in Figure 3.

We begin the presentation of our optimization benchmarks with the effect of number of employee- (E) and onlooker- (O) type agents on evolution of the ensemble of agents for different population sizes. For the initial benchmarks, a fixed value of $r_s = 0.2$ is used. N_a is set to 10 and the penalty weights are $\omega_\rho = 1$, $\omega_f = 0.01$, and $\omega_e = 2$ (arbitrary units). We initialize the first agent in each optimization run with the standard carbon GTH pseudopotential. To allow for a higher level of flexibility of the capping potential, we add an angular momentum channel ($l = 1$) with one projector, which leads to a 7-dimensional parameter space.

Figures 4−6 show penalty minimization over an optimization time $0 \leq \tau \leq 800$ for $N_{pop} = 4$, 12, and 20, respectively. Each figure shows the penalty evolution window as described by eqs 9 and 10 for different population setups $E + O = N_{pop}$. The number of trial runs is $N_{trial} = 5$ for each choice of E/O. An employee-type agent abandons its position after $N_a = 10$ unsuccessful cycles.

We observe in Figure 4 ($N_{pop} = 4$) a fast decrease of the lower penalty boundary $\mathscr{P}_{\min}(\tau)$ during $\tau \leq 100$ for all E/O combinations. This boundary remains practically unchanged for the remaining optimization. On the other hand, the upper penalty boundary $\mathscr{P}_{\max}(\tau)$ shows a decay comparable to that of the lower penalty boundary but only for combinations with no or few employee-type agents. For equal numbers of employee/onlooker-type agents or a higher amout of employee types, the upper penalty boundary decreases on a much slower time scale with practically no convergence in the all-employee-type case ($E = N_{pop}$).

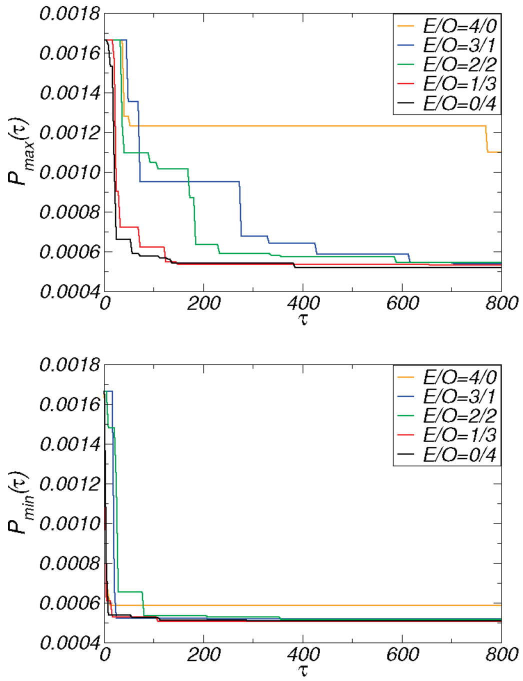


Figure 4. Population size/setup benchmarks: upper penalty boundary $\mathscr{P}_{\max}(\tau)$ and lower penalty boundary $\mathscr{P}_{\min}(\tau)$ with $E + O = 4$, $r_s = 0.2$, and $N_{trial} = 5$.

Very similar behavior of the lower penalty boundary is observed for $N_{pop} = 12$ (Figure 5). The decrease of the upper penalty boundary is similar for combinations from 0 to 6 employee-type agents, and significantly slower for higher numbers of E. The decrease of $\mathscr{P}_{\max}(\tau)$ happens on a slightly longer time scale compared to the optimization benchmarks with $N_{pop} = 4$.

This trend remains valid for $N_{pop} = 20$ (Figure 6). The decrease of the penalty window (i.e., the interval $[\mathscr{P}_{\min}(\tau),\mathscr{P}_{\max}(\tau)]$) is significantly slower compared to smaller populations.

As an extreme case of employee/onlooker combinations, we show in Figure 7 the upper and lower penalty boundaries for a series of $N_{trial} = 5$ optimization runs with a population consisting of only one agent of employee type and $N_a > N_{cycle}$ The lower penalty boundary reaches its minimum after an optimization time of $\tau \approx 230$. The upper penalty boundary needs nearly $\tau \approx 800$ to approach the value of the lower boundary.

The second control parameter of the optimization algorithm is the size of the neighborhood sphere r_s. Again, we run a series of independent optimizations with a fixed population setup of 3 employee and 9 onlooker types, with $N_a = 10$, $\omega_\sigma = 1$, $\omega_f = 0.01$, and $\omega_e = 2$. We show in Figure 8 the upper penalty boundary $\mathscr{P}_{\max}(\tau)$ and lower penalty boundary $\mathscr{P}_{\min}(\tau)$ for different radii r_s, with $N_{trial} = 5$ runs each.

We observe a steady (but slow) minimization behavior for small radii $r_s \leq 0.01$. Intermediate radii ($r_s = 0.1$ and 0.2) lead to a fast decrease of the penalty window ($[\mathscr{P}_{\min}(\tau),\mathscr{P}_{\max}(\tau)]$).

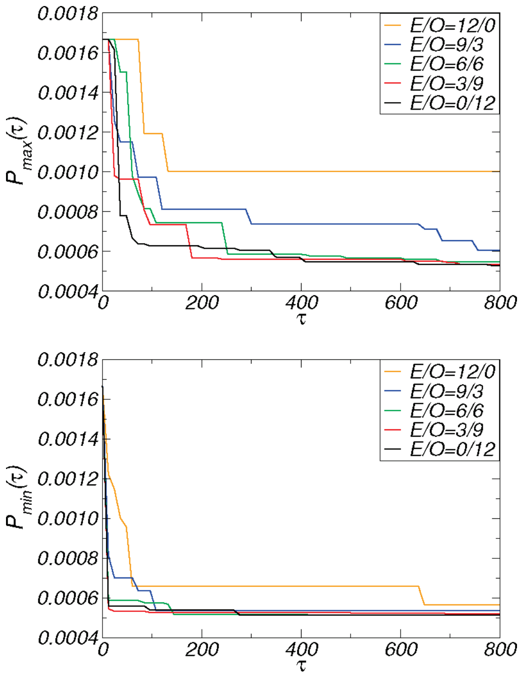


Figure 5. Population size/setup benchmarks: upper penalty boundary $\mathscr{P}_{\mathrm{max}}(\tau)$ and lower penalty boundary $\mathscr{P}_{\mathrm{min}}(\tau)$ with $E + O = 12$, $r_s = 0.2$, and $N_{\mathrm{trial}} = 5$.

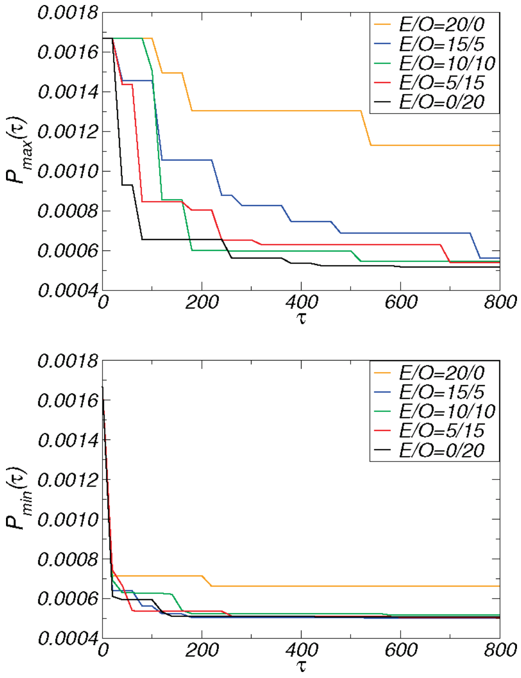


Figure 6. Population size/setup benchmarks: upper penalty boundary $\mathscr{P}_{\mathrm{max}}(\tau)$ and lower penalty boundary $\mathscr{P}_{\mathrm{min}}(\tau)$ with $E + O = 20$, $r_s = 0.2$, and $N_{\mathrm{trial}} = 5$.

Bigger radii lead to a comparable decrease in the lower penalty boundary. The upper penalty boundary, however, decreases quite slowly and unsteadily.

The optimizations presented so far have all used the conventional carbon GTH pseudopotential as starting point. While this appears adequate for the particular situation of homolytic C−C bond capping, we aim at designing capping potentials for more complex settings. In order to test the performance of our ABC algorithm in more difficult circumstances, we repeat the optimization of our C−C capping potential from a starting point with randomized capping parameters.

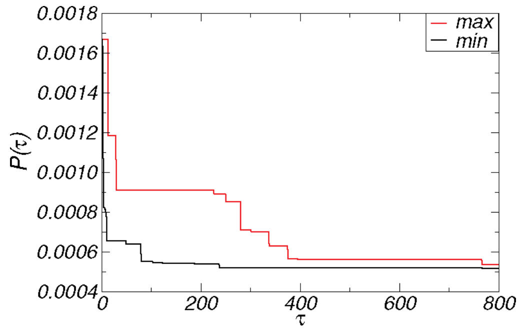


Figure 7. Single employee-type agent: lower penalty boundary $\mathscr{P}_{\mathrm{min}}(\tau)$ and upper penalty boundary $\mathscr{P}_{\mathrm{max}}(\tau)$ with $E = N_{\mathrm{pop}} = 1$, $r_s = 0.2$, and $N_{\mathrm{trial}} = 5$.

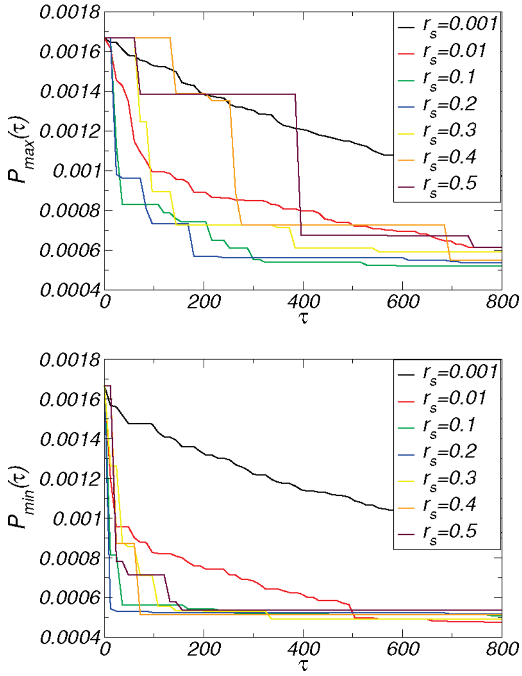


Figure 8. Sphere radius r_s benchmarks: upper penalty boundary $\mathscr{P}_{\mathrm{max}}(\tau)$ and lower penalty boundary $\mathscr{P}_{\mathrm{min}}(\tau)$ with $N_{\mathrm{pop}} = 12$, $E/O = 3/9$, and $N_{\mathrm{trial}} = 5$.

For a population of $E + O = 12$, $N_a = 10$, $r_s = 0.2$, $\omega_\rho = 1$, $\omega_f = 0.01$, and $\omega_e = 2$, we perform $N_{\mathrm{trial}} = 5$ independent optimizations with varying combinations for E/O. The evolution for this unfavorable initialization of agents is shown in Figure 9.

The behavior of the lower penalty boundary is nearly identical for small to intermediate numbers of employee types. It decreases more slowly for a higher amount of employee types. The upper penalty boundary shows a similar pattern for all E/O combinations, but an equal amount of employee and onlooker types shows the best performance in an early stage of the optimization.

Regarding the combination of employee and onlooker agents, it turns out that the optimal ratio depends strongly on whether the present set of agents is "in direct view" of the final minimum,

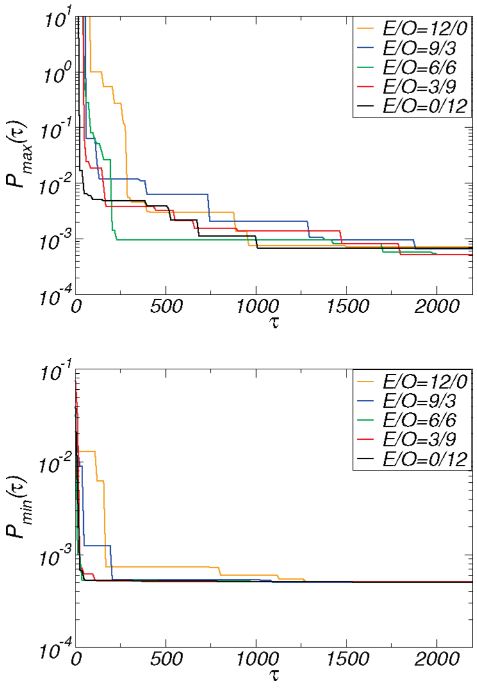


Figure 9. Population setup benchmarks: upper penalty boundary $\mathscr{P}_{\max}(\tau)$ and lower penalty boundary $\mathscr{P}_{\min}(\tau)$ with $E + O = 12$, $r_s = 0.2$, and $N_{\text{trial}} = 5$ with fully random initialization.

that is, in its proximity and without additional barriers on the way. For a good starting point of the optimization, for example, the feedback process inherent to onlooker agents leads to a speedup of the optimization, and zero agents of employee type are best. On the other hand, a nonoptimal starting point (as obtained via the randomized initialization) requires a certain number of agents with explorative character, that is, employee (or scout) type agents. Hence, it might eventually be useful to switch the distribution of employee versus onlooker types during the progress of the optimization. Investigation of this effect, however, exceeds the scope of the present paper.

As for the sphere radius r_s, we find that a small value leads to a slow "speed" of the agents in parameter space. A large value, on the other hand, allows for large moves. However, the plateaus in the evolution of $\mathscr{P}_{\max}$ (Figure 8) for $r_s \geq 0.4$ illustrate that a large neigborhood area leads to a high rejection rate for the proposed moves of the agents. We find that an intermediate choice of $0.1 \leq r_s \leq 0.2$ leads to the fastest decay of the penalty window, due to a trade-off between the "speed" of the agents in parameter space and their rejection rate. We believe that this behavior indicates a rich structure of the penalty surface, even for this simple case of homolytic C−C bond capping.

3.2. Initial Benchmark of Optimized Capping Potentials. To examine the quality of optimized capping potentials obtained during our benchmarks, we compute electronic linear response properties for a linear alkane molecule (hexane) in which the terminal methyl group is replaced by a capping potential. In particular, we compute spectroscopic properties that involve both the occupied and excited manifold of electronic orbitals. These parameters measure the performance of our capping potential beyond the scope that is accessible by the penalty functional (eq 1) because the latter is based only on the ground-state density.

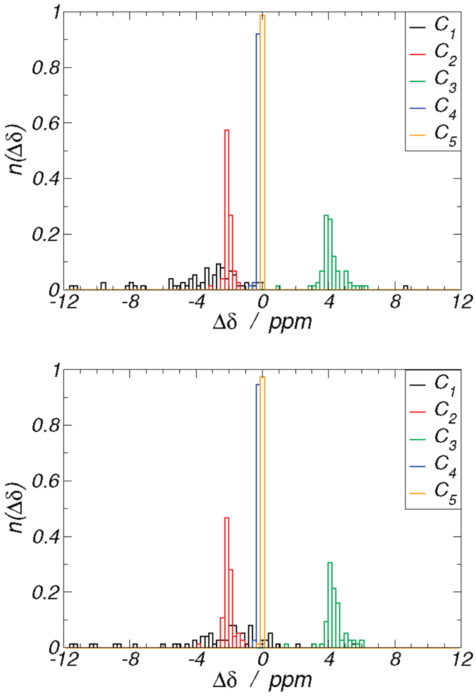


Figure 10. Distribution $n(\Delta\delta)$ of the isotropic NMR chemical shift $\Delta\delta$ of carbon atoms in hexane with V_{cap} bound to C_1 for an ensemble of 75 capping potentials obtained from independent optimizations: (top) hexane geometry; (bottom) optimized geometry for V_{cap}.

We have chosen ^{13}C NMR chemical shifts δ for the characterization of our capping potentials. These chemical shifts are the result of a complex interplay of occupied and excited electronic states. Nevertheless, they are relatively short-sighted, which means that a perturbation in the electronic spectrum reaches no further than a few covalent bonds. Hence, they allow us to monitor the range in which the QM/MM-induced bond cleavage perturbs the electronic subsystem.

Specifically, we compute the distribution of deviations of the trace of the nuclear shielding tensor $\sigma_{\alpha\beta}$ for a capped molecule with respect to a full calculation:

$$\Delta\delta = \text{Tr}(\sigma_{\alpha\beta}[\text{full-QM}] - \sigma_{\alpha\beta}[\text{QM/MM}]) \qquad (11)$$

This is done (1) in the optimized geometry of the full hexane molecule and (2) in a geometry that has been optimized by use of V_{cap}. In both cases, we obtain an ensemble of chemical shift values from the ensemble of independent optimization runs.

Figure 10 shows the distribution of $\Delta\delta$ of carbon atoms C_i, where C_1 is the direct neighbor of V_{cap}. For both geometries, we observe a similar picture: $\Delta\delta$ of the direct neighbor of V_{cap} has a broad distribution with a clustering between −3 and −2 ppm. The distribution for next two carbon atoms (C_2 and C_3) have distinct peaks at −2 and 4 ppm, respectively. As for the last two carbons, we find only minor deviations, far below 1 ppm, from the reference NMR signature.

3.3. Application to Octane. While the main focus in this article is on the algorithmic performance of the ABC algorithm in the QM/MM context, we have nevertheless applied the ABC/capping potential algorithm to C−C bond cleavage in a larger molecule, specifically octane. Here, a butane fragment has been replaced by a capping potential; see Figure 11.

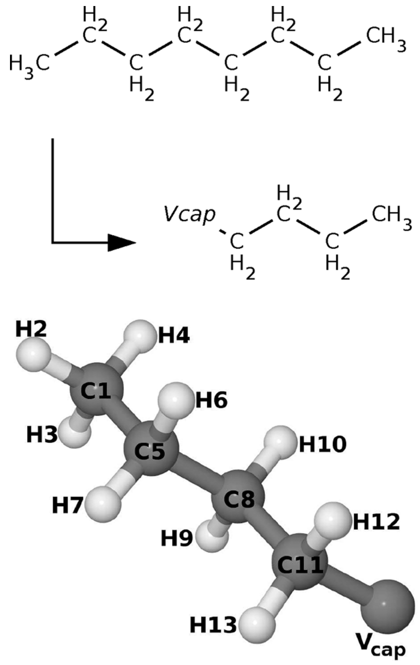


Figure 11. Reference system octane and its capped counterpart: (top) bond-capping scheme and (bottom) atom numbering.

Table 1. GTH Parameters for Regular Carbon and Optimized C−C Capping Potential V_{cap}

	r_{loc}	C_1	C_2	r_0	$h_{1,1}^0$	r_1	$h_{1,1}^1$
regular C	0.3376	−9.1285	1.4251	0.3025	9.6507		
V_{cap}	0.2101	−13.1925	3.4867	0.2416	6.2451	0.3125	9.7340
ref 49	0.7221	9.9086	−2.5466	0.5120	−3.5081	1.4664	0.2316

The optimization is performed over 500 cycles with $E = 4$ and $O = 6$, a neigborhood sphere radius of $r_s = 0.1$, and an integration volume for the density difference consisting of spheres of size 1.5× covalent radius around each atom except V_{cap}. The penalty weighting factors are $\omega_\rho = \omega_f = \omega_e = 1$. See Table 1 for the initial guess (regular carbon) and optimized capping potential parameters (V_{cap}).

All geometric parameters (shown in Table 2) of the capped octane molecule are in excellent agreement with the full octane reference. This agreement holds for our new optimized potential as well as for a previous version,[49] and to some degree even for the simpler hydrogen and fluorine cappings. While the hydrogen termination looks like an accurate way of capping when the H−C bond distance is ignored, it has a strong effect on the properties of the subsequent C−C bond. This is shown in the potential energy curve (Figure 12) of the C_8−C_{11} bond: when hydrogen capping is applied, the equilibrium distance is shortened by about 0.1 Å and its frequency is considerably blue-shifted.

When the NMR chemical shift deviations (shown in Table 3) are examined, a more heterogeneous picture arises. The conventional H- and F-based cappings are only in very rough agreement with the reference system, while both capping potentials yield satisfactory results. When the latter two are compared, it turns out that in our presently optimized capping potential (V_{cap}), the first (C_{11}) and third (C_5) carbon atoms exhibit somewhat larger deviations than the potential from ref 49, while the intermediate carbon (C_8) and most of the hydrogens show better agreement. It is not clear at present what is the specific reason for these deviations in terms of the capping parameters (given numerically in Table 1). It is obvious, however, that the two capping potentials have very different characteristics in terms of the range of their local and nonlocal parts; in particular, the radius of the local part (r_{loc}) differs by a factor of more than 3, as does r_1.

Nevertheless, this result clearly illustrates that the ABC algorithm with its stochastic elements has the very important ability

Table 2. Optimized Bond Lengths, Angles, and Dihedrals of the Octane Reference Molecule and Its Capped Counterpart

	reference	V_{cap}	hydrogen	fluorine	ref 49
C_1−C_5 (Å)	1.54	1.55	1.55	1.54	1.55
C_5−C_8 (Å)	1.55	1.55	1.55	1.55	1.55
C_8−C_{11} (Å)	1.55	1.54	1.55	1.53	1.55
C_{11}−V_{cap} (Å)	1.55	1.62	1.10	1.47	1.55
C_1−C_5−C_8 (deg)	113.4	113.5	113.6	113.1	113.7
C_5−C_8−C_{11}(deg)	113.6	113.5	113.6	111.7	113.4
C_8−C_{11}−V_{cap} (deg)	114.0	113.7	111.3	110.2	116.2
C_1−C_5−C_8−C_{11} (deg)	−179.3	−179.9	179.8	−179.6	−177.2
C_5−C_8−C_{11}−V_{cap} (deg)	−179.5	−178.2	−179.3	179.6	179.7

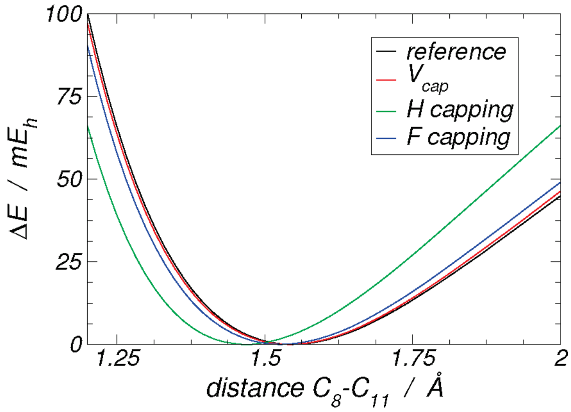


Figure 12. Potential energy curve of the C_8−C_{11} bond.

Table 3. ^{1}H and ^{13}C NMR Chemical Shift Changes of the Capped Octane Molecule with Respect to Its Octane Reference, $\Delta\delta = \sigma^{cap} - \sigma^{ref}$

	chemical shift change (ppm)			
Δδ	V_{cap}	hydrogen	fluorine	ref 49
C_1	0.16	−0.08	0.56	0.56
H_2	0.03	0.01	−0.02	0.02
H_3	0.04	0.01	−0.06	0.03
H_4	0.04	0.02	−0.05	0.02
C_5	−5.03	−1.56	6.77	−1.12
H_6	0.12	0.07	0.14	0.14
H_7	0.13	0.07	0.15	0.22
C_8	−0.02	5.52	0.81	−0.32
H_9	−0.40	−0.07	−0.47	−0.34
H_{10}	−0.41	−0.06	−0.48	−0.41
C_{11}	12.08	18.01	−52.08	−2.44
H_{12}	−0.23	0.37	−3.49	−0.18
H_{13}	−0.24	0.36	−3.47	−0.15

to discover new regions of parameter space, which a downhill algorithm (e.g., conjugate gradients) would never explore. In order to obtain better capping potentials in terms of spectroscopic parameters, the optimization can now be adjusted by means of weighting factors and the exact definition of penalty integration volume. However, this is beyond the scope of the present work and will be highlighted in a forthcoming article.

4. CONCLUSION

In this work, we have presented an algorithmic extension of a numerical optimization scheme for capping potentials that can be used for mixed quantum−classical (QM/MM) ab initio calculations. The new algorithm mixes deterministic (downhill gradient) techniques with stochastic (Monte Carlo-like) moves, which are applied to an analytic potential such that the electronic structure in the quantum region is preserved as well as possible with respect to a reference (full-QM) calculation. Deviations from the ideal electronic (and geometric) structure are characterized by a suitably designed penalty functional, which represents the target quantity that is minimized with respect to the parameters of the capping potential.

Our algorithm is a variant of the artificial bee colony (ABC) approach, which has certain analogies to the foraging behavior of honeybees in nature. From a computational view, it bears similarities to the ideas used in parallel tempering schemes. The stochastic elements that are incorporated into the ABC optimization avoid trapping in local minima of the penalty functional hypersurface. For the benchmark molecule (ethane) used in this work, this surface is still relatively smooth; however, as soon as more complex molecules are targeted, the stochastic components of the ABC algorithm are very important due to the presence of numerous wiggles in this surface. This could be shown by using a randomized starting point for the capping potential optimization. For such more complex situations, several control parameters of the ABC scheme can be adjusted in order to improve the convergence behavior.

The properties of the resulting capping potentials have been characterized in terms of the deviations of carbon NMR chemical shift values with respect to a reference calculation. For our homolytic cleavage of a $C^{sp3}-C^{sp3}$ bond, the properties resemble those of the optimized capping potentials that were obtained previously by the deterministic simplex minimization approach.[49] In turn, our new algorithm has found a considerably different set of values of the capping parameters, mainly because of the use of a slightly different set of weighting parameters within the penalty functional. This illustrates that the penalty surface has indeed a rich substructure, even in a very simple case such as the homolytic capping of ethane.

We believe that the new ABC optimization scheme will help generating better capping potentials for more complex situations in which special care is necessary. In particular, we are presently applying the algorithm to heterolytic bond cleavage (i.e., C−N and C−O bonds), as well as the capping of highly polar and charged groups (i.e., COOH and COO^-), which are of crucial importance for most biophysical QM/MM simulations.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information. Additional text and one figure describing complementary benchmark optimization. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
*E-mail: daniel.sebastiani@fu-berlin.de.

■ ACKNOWLEDGMENT

This work has been supported by the German Research Foundation (DFG) under Grants SE 1008/5 and SE 1008/6. Computing infrastructure was provided by the Northern German Supercomputing Alliance (HLRN) under Grant HLRN/bec00061.

■ REFERENCES

(1) Gascon, J. A.; Sproviero, E. M.; Batista, V. S. *J. Chem. Theory Comput.* **2005**, *1*, 674–685.
(2) Kongsted, J.; Nielsen, C. B.; Mikkelsen, K. V.; Christiansen, O.; Ruud, K. *J. Chem. Phys.* **2007**, *126*, No. 034510.
(3) Sebastiani, D. *Nachr. Chem.* **2009**, *57*, 305.
(4) Schmidt, J.; Hoffmann, A.; Spiess, H. W.; Sebastiani, D. *J. Phys. Chem. B* **2006**, *110*, 23204–23210.
(5) Schmidt, J.; Hutter, J.; Spiess, H. W.; Sebastiani, D. *ChemPhysChem* **2008**, *9*, 2313–2316.
(6) Heller, J.; Elgabarty, H.; Zhuang, B.; Sebastiani, D.; Hinderberger, D. *J. Phys. Chem. B* **2010**, *114*, 7429–7438.
(7) Banyai, D. R.; Murakhtina, T.; Sebastiani, D. *Magn. Reson. Chem.* **2010**, *48*, S56–S60.
(8) Ludueña, G. A.; Wegner, M.; Bjâlie, L.; Sebastiani, D. *ChemPhysChem* **2010**, *11*, 2353–2360.
(9) Hansen, M. R.; Sekharan, S.; Graf, R.; Sebastiani, D. *J. Am. Chem. Soc.* **2009**, 5251–5256.
(10) Gervais, C.; Dupree, R.; Pike, K. J.; Bonhomme, C.; Profeta, M.; Pickard, C. J.; Mauri, F. *J. Phys. Chem. A* **2005**, *109*, 6960–6969.
(11) Yates, J. R.; Dobbins, S. E.; Pickard, C. J.; Mauri, F.; Ghi, P. Y.; Harris, R. K. *Phys. Chem. Chem. Phys.* **2005**, *7*, 1402–1407.
(12) Yates, J. R.; Pham, T. N.; Pickard, C. J.; Mauri, F.; Amado, A. M.; Gil, A. M.; Brown, S. P. *J. Am. Chem. Soc.* **2005**, *127*, 10216–10220.
(13) Murakhtina, T.; Delle Site, L.; Sebastiani, D. *ChemPhysChem* **2006**, *7*, 1215–1219.
(14) Rohrig, U.; Guidoni, L.; Laio, A.; Frank, I.; Rothlisberger, U. *J. Am. Chem. Soc.* **2004**, *126*, 15328–15329.
(15) Deng, R. Z.; Martyna, G. J.; Klein, M. L. *Phys. Rev. Lett.* **1993**, *71*, 267.
(16) Stanton, R. V.; Little, L. R.; Merz, K. M. *J. Phys. Chem.* **1996**, *99*, 11266.
(17) Eichinger, M.; Tavan, P.; Hutter, J.; Parrinello, M. *J. Chem. Phys.* **1999**, *21*, 10452.
(18) Lyne, P.; Hodosceck, M.; Karplus, M. *J. Phys. Chem. A* **1999**, *103*, 3462–3471.
(19) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*, 700–733.
(20) Zhang, Y.; Lee, T.-S.; Yang, W. *J. Phys. Chem.* **1999**, *110*, 46–54.
(21) Brancato, G.; Rega, N.; Barone, V. *J. Chem. Phys.* **2008**, *128*, 144501.
(22) Cui, Q. *J. Chem. Phys.* **2002**, *117*, 4720–4728.
(23) Laio, A.; VandeVondele, J.; Roethlisberger, U. *J. Chem. Phys.* **2002**, *116*, 6941–6947.
(24) Cui, Q.; Karplus, M. *J. Chem. Phys.* **2000**, *112*, 1133.
(25) Laio, A.; VandeVondele, J.; Rothlisberger, U. *J. Phys. Chem. B* **2002**, *106*, 7300–7307.
(26) Bühl, M.; Grigoleit, S.; Kabrede, H.; Mauschick, F. T. *Chem.—Eur. J.* **2006**, *12*, 477–488.
(27) Senn, H. M.; Thiel, W. *Top. Curr. Chem.* **2007**, *268*, 173–290.
(28) Kastner, J.; Thiel, S.; Senn, H. M.; Sherwood, P.; Thiel, W. *J. Chem. Theory Comput.* **2007**, 3, 1064–1072.

(29) Geerke, D. P.; Thiel, S.; Thiel, W.; van Gunsteren, W. F. *Phys. Chem. Chem. Phys.* **2008**, *10*, 297–302.

(30) Benighaus, T.; Thiel, W. *J. Chem. Theory Comput.* **2008**, *4*, 1600–1609.

(31) Komin, S.; Gossens, C.; Tavernelli, I.; Röthlisberger, U.; Sebastiani, D. *J. Phys. Chem. B* **2007**, *111*, 5225–5232.

(32) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1986**, *7*, 718.

(33) Birge, R. R.; Zhang, C.-F. *J. Chem. Phys.* **1990**, *92*, 7178–7195.

(34) Assfeld, X.; Rivail, J.-L. *Chem. Phys. Lett.* **1996**, *263*, 100–106.

(35) Jacob, C. R.; Visscher, L. *J. Chem. Phys.* **2006**, *125*, No. 194104.

(36) Gao, J.; Amara, P.; Alhambra, C.; Field, M. J. *J. Phys. Chem. A* **1998**, *102*, 4714–4721.

(37) Pu, J.; Gao, J.; Truhlar, D. G. *J. Phys. Chem. A* **2004**, *108*, 632–650.

(38) Jung, J.; Choi, C. H.; Sugita, Y.; Ten-no, S. *J. Chem. Phys.* **2007**, *127*, No. 204102.

(39) Jardilliera, N.; Goursot, A. *Chem. Phys. Lett.* **2008**, *454*, 65–69.

(40) Mallik, A.; Taylor, D. E.; Runge, K.; Dufty, J. W. *Int. J. Quantum Chem.* **2004**, *100*, 1019–1025.

(41) DiLabio, G. A.; Wolkow, R. A.; Johnson, E. R. *J. Chem. Phys.* **2005**, *122*, No. 044708.

(42) DiLabio, G. A.; Hurley, M. M.; Christiansen, P. A. *J. Chem. Phys.* **2002**, *116*, 9578–9584.

(43) Xiao, C. Y.; Zhang, Y. K. *J. Chem. Phys.* **2007**, *127*, No. 124102.

(44) Poteau, R.; Ortega, I.; Alary, F.; Solis, A. R.; Barthelat, J.-C.; Daudey, J.-P. *J. Phys. Chem. A* **2001**, *105*, 198–205.

(45) Poteau, R.; Alary, F.; Makarim, H. A. E.; Heully, J.-L.; Barthelat, J.-C.; Daudey, J.-P. *J. Phys. Chem. A* **2001**, *105*, 206–214.

(46) Exner, T. E.; Mezey, P. G. *J. Comput. Chem.* **2003**, *24*, 1980–1986.

(47) Exner, T. E.; Mezey, P. G. *Phys. Chem. Chem. Phys.* **2005**, *24*, 4061–4069.

(48) Eckard, S.; Exner, T. E. *Z. Phys. Chem.* **2006**, *220*, 927–944.

(49) Komin, S.; Sebastiani, D. *J. Chem. Theory Comput.* **2009**, *5*, 1490–1498.

(50) Goedecker, S.; Teter, M.; Hutter, J. *Phys. Rev. B* **1996**, *54*, 1703.

(51) Hartwigsen, C.; Goedecker, S.; Hutter, J. *Phys. Rev. B* **1998**, *58*, 3641.

(52) Rohrig, U. F.; Sebastiani, D. *J. Phys. Chem. B* **2008**, *112*, 1267–1274.

(53) Debye, P. *Math. Ann.* **1909**, *67*, 535–558.

(54) Hestenes, M. R.; Stiefel, E. *J. Res. Natl. Bur. Stand. (U.S.)* **1952**, *49*, 409–436.

(55) Nelder, J. A.; Mead, R. *Comput. J.* **1965**, *7*, 308–313.

(56) Holland, J. H. *Adaptation in natural and artificial systems*; University of Michigan: Ann Arbor, MI, 1975.

(57) Wang, Q. H. *Biol. Cybern.* **1987**, *57*, 95–101.

(58) Yang, X. S. *Lect. Notes Comput. Sci.* **2005**, *3562*, 317–323.

(59) Karaboga, D.; Basturk, B. *J. Global Opt.* **2007**, *39*, 459–471.

(60) Karaboga, D.; Basturk, B. *Appl. Soft Comput.* **2008**, *8*, 687–697.

(61) Karaboga, D.; Akay, B. *Artif. Intell. Rev.* **2009**, *31*, 61–85.

(62) Brown, S. P.; Spiess, H. W. *Chem. Rev.* **2001**, *101*, 4125.

(63) Spiess, H. W. *Macromol. Chem. Phys.* **2003**, *204*, 340–346.

(64) Schulz-Dobrick, M.; Metzroth, T.; Spiess, H. W.; Gauss, J.; Schnell, I. *ChemPhysChem* **2005**, *6*, 315–327.

(65) Ochsenfeld, C.; Brown, S. P.; Schnell, I.; Gauss, J.; Spiess, H. W. *J. Am. Chem. Soc.* **2001**, *123*, 2597–2606.

(66) Bühl, M.; Kabrede, H.; Diss, R.; Wipff, G. *J. Am. Chem. Soc.* **2006**, *128*, 6357–6368.

(67) von Lilienfeld-Toal, A.; Tavernelli, I.; Rothlisberger, U.; Sebastiani, D. *J. Chem. Phys.* **2005**, *122*, No. 014113.

(68) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864.

(69) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133.

(70) Jones, R. O.; Gunnarsson, O. *Rev. Mod. Phys.* **1989**, *61*, 689–746.

(71) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.

(72) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.

(73) Hutter, J.; Curioni, A. *ChemPhysChem* **2005**, *6*, 1788–1793.

(74) Hutter, J. et al. Computer code CPMD, version 3.12.0, Copyright IBM Corp. and MPI-FKF Stuttgart 1990–2007, http://www.cpmd.org.

(75) Putrino, A.; Sebastiani, D.; Parrinello, M. *J. Chem. Phys.* **2000**, *113*, 7102–7109.

(76) Sebastiani, D.; Parrinello, M. *J. Phys. Chem. A* **2001**, *105*, 1951.

(77) Sebastiani, D.; Goward, G.; Schnell, I.; Spiess, H. W. *J. Mol. Struct. (THEOCHEM)* **2003**, *625*, 283–288.

JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

GPU-Based Implementations of the Noniterative Regularized-CCSD(T) Corrections: Applications to Strongly Correlated Systems

Wenjing Ma,[†] Sriram Krishnamoorthy,*[,‡] Oreste Villa,[‡] and Karol Kowalski*[,¶]

[†]Department of Computer Science and Engineering, The Ohio State University, Columbus, Ohio, United States

[‡]Computational Sciences and Mathematics Division, Pacific Northwest National Laboratory, Richland, Washington, United States

[¶]William R. Wiley Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, Washington, United States

ABSTRACT: The details of the graphical processing unit (GPU) implementation of the most computationally intensive (T)-part of the recently introduced regularized CCSD(T) (Reg-CCSD(T)) method [Kowalski, K.; Valiev, M. *J. Chem. Phys.* **2009**, *131*, No. 234107] for calculating electronic energies of strongly correlated systems are discussed. Parallel tests performed for several molecular systems show very good scalability of the triples part of the Reg-CCSD(T) approach. We also discuss the performance of the Reg-CCSD(T) GPU implementation as a function of the parameters defining the partitioning of the spinorbital domain (tiling structure). The accuracy of the Reg-CCSD(T) method is illustrated on three examples: the methyfluoride molecule, dissociation of dodecane, and open-shell Spiro cation (5,5′(4*H*,4*H*′)-spirobi[cyclopenta[*c*]pyrrole] 2,2′,6,6′-tetrahydro cation), which is a frequently used model to study electron transfer processes. It is demonstrated that a simple regularization of the cluster amplitudes used in the noniterative corrections accounting for the effect of triply excited configurations significantly improves the accuracies of ground-state energies in the presence of strong quasidegeneracy effects. For methylfluoride, we compare the Reg-CCSD(T) results with the CR-CC(2,3) and CCSDT energies, whereas for Spiro cation we compare Reg-CCSD(T) results with the energies obtained with completely renormalized CCSD(T) method. Performance tests for the Spiro, dodecane, and uracil molecules are also discussed.

1. INTRODUCTION

The widespread use of highly correlated methods in electronic structure calculations is contingent upon the interplay between advances in the theory and the possibility of utilizing ever-growing computer power of emerging architectures. Due to their accuracy, coupled cluster (CC) methods[1−4] have assumed a special position in high-precision calculations for molecular systems.[5−7] The well established family of iterative approximations CCSD (CC with singles and doubles),[8,9] CCSDT (CC with singles, doubles, and triples),[10,11] etc., provide an increasing level of accuracy of resulting energies. Unfortunately, due to the steep numerical complexity, CCSDT applications are very limited. The development of perturbative methods has played a large part to overcome these difficulties. In the CCSD[T][12] and CCSD(T)[13] approaches, the perturbative corrections are constructed in terms of converged CCSD cluster amplitudes. For equilibrium geometries of closed-shell systems the CCSD(T) approach is capable of providing nearly CCSDT level of accuracy. For problems that require knowledge of ground-state potential energy surfaces (PESs) and energies for stretched internuclear geometries, many approaches have been devised to alleviate the problems caused by the divergent nature of the perturbative expansion. There are two main groups of these approaches. The first class of methods is related to alternative perturbative expansions for similarity transformed Hamiltonian,[14−28] while the other class of methods is deeply rooted in the Method of Moments of Coupled Cluster equations.[29−35] Regardless of the origin, all these noniterative methods lead to very accurate results for processes where a single bond is broken. These CCSD(T)-like approaches (and approaches of even higher order) can be also used in highly accurate thermochemistry calculations.[36,37]

Given the importance of CCSD(T)-like methods for high-precision calculations, significant progress has been made toward the development of scalable CC codes[38−47] enabling calculations on large-scale molecular systems. The NWChem[48] implementation of the (T)-part of the CCSD(T) approach has been shown to scale across 250 000 cores,[49] which should be attributed to the natural parallelism of noniterative approaches. Other components of the whole CCSD(T) calculation, Hartree−Fock, 4-index transformation, and CCSD implementations, because of the much smaller task pool, scale across much smaller number of cores.

Equally important to the development of new theoretical algorithms are the implementations on emerging computer architectures. The emergence of general purpose graphic processing units (GPGPUs) has revolutionized computational science by making available an unprecedented amount of computing capability. There are several examples of successful development of GPU-based software in computational chemistry.[50−64] While GPUs have been employed in accelerating scientific calculations in the past, programming them required mapping the application to the graphic processing pipeline, a challenging task. The advent of higher-level programming support, such as through CUDA and OpenCL, has made it easier to exploit their potential. The

Received: December 16, 2010
Published: April 11, 2011

widespread availability of GPU-accelerated systems, ranging from workstations to supercomputers, stresses the need to develop algorithms to effectively utilize them, and the work's impact on researchers with access to varying computing resources. Recently, we discussed the first GPU implementation of the CCSD(T) method,[65] which due to its large flop count is ideally suited for this type of computer architectures. In this paper, we will discuss accuracies and numerical performance of the GPU implementation of the regularized CCSD(T) approach (Reg-CCSD(T)),[66] which can be used in calculations for strongly correlated systems. In analogy to the CCSD(T) method, the Reg-CCSD(T) approach is characterized by the same $n_o^3 n_u^4$ (n_o and n_u designate the number of occupied and unoccupied orbitals, respectively) numerical complexity. This high numerical overhead is associated with calculating triples correction or (T)-part in short, which is especially challenging for systems with large number of correlated electrons and large virtual space (assuming that two (T) calculations were performed with the same number N of correlated orbitals, $N = n_o + n_u = 1000$, but for two different values of n_o, n_o = 20, and n_o = 400, the latter calculations is more than 3 orders of magnitude more expensive compared to the n_o = 20 case). Our discussion will be based on the calculations for several challenging systems: C−F bond elongation in the methylfluoride, dissociation of dodecane, $C_{12}H_{26}$, into $C_{11}H_{23}$ and CH_3, and the mixed valence system 5,5′(4*H*,4*H*′)-spirobi[cyclopenta[*c*]pyrrole]2,2′,6,6′-tetrahydro cation (or Spiro cation for short).[67] While the methylfluoride and dodecane molecules epitomize common problems the CCSD(T) approach stumble into when both static and dynamic correlation effects play an equally important role, the Spiro molecule is frequently used in fundamental studies of charge transport processes and poses a significant challenge even for multireference perturbative methods.[67−70] The paper is organized as follows: in section 2, we give a brief overview of regularization techniques for CC theory, in section 3 details of the GPU implementation of the most expensive (T)-part are described. In section 4 we discuss the quality of the potential energy surfaces (PESs) for methylfluoride, $C_{12}H_{26}$ dissociation, and Spiro cation and illustrate the parallel performance on the example of calculations for Spiro cation and uracil.

2. THEORY

In this section, we present only the salient features of the regularized methods derived from the generating functional expansion of ref 71, where it was demonstrated that the exact energy (E) for the ground electronic state can be expanded as

$$E = E^{(A)} + \sum_{J;J\neq 0} \overline{M}_J^{(A)} \left[\frac{\partial}{\partial S_J} W(\Sigma, S) \right] \Bigg|_{S^{(A)} = T^{(A)}; S^{(R)} = 0} \qquad (1)$$

Here the energy $E^{(A)}$ is an approximate energy obtained in approximate CC calculations (CC-A) defined by the approximate cluster operator $T^{(A)}$

$$T^{(A)} = \sum_{n=1}^{m_A} T_n \qquad (2)$$

$$T_n = \sum_{i_1 < \ldots < i_n; a_1 < \ldots < a_n} t_{a_1 \ldots a_n}^{i_1 \ldots i_n} X_{a_1}^{+} \ldots X_{a_n}^{+} X_{i_n} \ldots X_{i_1} \qquad (3)$$

The i_1, ..., i_n (a_1, ..., a_n) indices refer to the occupied (unoccupied) spinorbitals in the reference function $|\Phi\rangle$ and the X_p^+ (X_p) operators are creation (annihilation) operators for electrons in p-th single particle state. The cluster operator includes excitations of rank equal or lower than m_A. In practice $m_A \ll N$ (N stands for the total number of correlated electrons). In expansion (eq 1) the Σ operator corresponds to the exact cluster operator ($\Sigma = \Sigma_{n=1}^{N} \Sigma_n$, where the n-tuply excited many-body component Σ_n of Σ, in analogy to (eq 2) is defined by the $\Sigma_{a_1 \ldots a_n}^{i_1 \ldots i_n}$ amplitudes). The auxiliary cluster operator S ($S = \Sigma_{n=1}^{N} S_n$), introduced in refs 66 and 71 is chosen in such a way that the auxiliary wave function $e^S|\Phi\rangle$ is in a close vicinity of the approximate wave function $e^{T^{(A)}}|\Phi\rangle$. We also assume that the exact wave function falls into the same vicinity (these assumptions are critical from the point of view of convergence properties of the connected form of the generating functional). The $S^{(A)}$ part of the S operator in eq 1 is defined by the excitations used to define the $T^{(A)}$ operator, while the $S^{(R)}$ part of the S operator contains higher excitations. The reference function $|\Phi\rangle$ is usually represented by the Hartree−Fock (HF) determinant. In eq 1 the quantities $\overline{M}_J^{(A)}$ correspond to the matrix elements of the moments operator (see ref 71 for details) and are defined as

$$\overline{M}_J^{(A)} = \langle \Phi_{i_1 \ldots i_n}^{a_1 \ldots a_n} | \overline{H}^{(A)} | \Phi \rangle \qquad (4)$$

where the string convention of ref 32 is invoked, that is, the nonzero string J corresponds to the excitation designated by the ordered set of occupied/unoccupied indices $\{i_1 < ... < i_n; a_1 < ... < a_n\}$. The similarity transformed Hamiltonian, $\bar{H}^{(A)}$, is defined as $\bar{H}^{(A)} = e^{-T^{(A)}} H e^{T^{(A)}}$, where the H operator represents the electronic Hamiltonian. The $\bar{H}^{(A)}$ operator contains connected diagrams only. The central role in eq 1 is played by the so-called generating functional $W(\Sigma,S)$, which is defined as a connected part of the overlap between the exact CC wave function and auxiliary wave function defined by the auxiliary cluster operator, that is,

$$\begin{aligned} W(\Sigma, S) &= \langle \Phi | (e^{\Sigma^+} e^S)_C | \Phi \rangle = \ln(\langle \Phi | e^{\Sigma^+} e^S | \Phi \rangle) \\ &= \ln(1 + \gamma(\Sigma, S)) \end{aligned} \qquad (5)$$

where subscript "C" designates connected part of a given expression and the $\gamma(\Sigma,S)$ function corresponds to the correlated part of the overlap between auxiliary and exact CC functions, that is, $\gamma(\Sigma,S) = \langle \Phi | (e^{\Sigma^+} - 1)(e^S - 1) | \Phi \rangle$, where Σ^+ is a Hermitian conjugate of the Σ operator. The features of the generating functional fully determine the basic features of the expansion in eq 1. In particular, the connectedness of $W(\Sigma,S)$ implicates the connectedness of expansion eq 1 (assuming the connectedness of cluster amplitudes). The formula (eq 1) was derived using the Taylor expansion for $\ln(1 + \gamma(\Sigma,S))$, which is only valid when the condition

$$|\gamma(\Sigma, S)| < 1 \qquad (6)$$

is satisfied. This fact limits the applicability of the expansion in (eq 1) to weakly correlated systems. To maintain the form of the expansion in eq 1 in the strong interaction regime one has to artificially redefine certain parameters of the generating functional expansion in order to decrease the value of $|\gamma(\Sigma,S)|$. This type of procedure is commonly referred to as the regularization procedure. There is great flexibility as far as the choice of the regularization procedure is concerned. For the sake of simplicity, we pursue perhaps the most obvious choice corresponding to the regularization of the Σ operator in the generating functional $W(\Sigma,S)$.[71] To address this issue we adopted ideas similar to the

Tikhonov regularization, which have recently been explored by Taube and Bartlett[72] in the context of ill defined linear CC equations for quasidegenerate systems. In our procedure, described in ref 66, the regularized Σ operator ($\Sigma_{\rm reg}$) is obtained by solving modified CC equations, which use the regularized form of the Hamiltonian ($H^{\rm reg}$) defined as

$$H^{\rm reg} = H + \omega^2 N_u \tag{7}$$

where the N_u operator is defined as $N_u = \Sigma_a X_a^+ X_a$, and represents particle number operator for particles in virtual states. The presence of this operator in the equations for regularized cluster operators $\Sigma_{\rm reg}$ is similar to the level shift techniques.

The described regularization scheme was used to define the due-to-triples corrections to energies obtained with the CCSD approach ($m_A = 2$, $T^{(A)} = T_1 + T_2$, $E^{(A)} = E^{\rm CCSD}$). For this purpose, we used the following form of the generating functional:

$$\bar{W}(\Sigma_{\rm reg}, S_3) \simeq \left\langle \Phi \left| \left\{ \left(\Sigma_{\rm reg,\,3} + \Sigma_{\rm reg,\,1}\Sigma_{\rm reg,\,2} + \frac{1}{6}(\Sigma_{\rm reg,\,1})^3 \right)^\dagger S_3 \right\}_C \right| \Phi \right\rangle \tag{8}$$

where the contributions containing only S_1 and S_2 were neglected because singly (${M_{J_1}}^{\rm CCSD}$) and doubly excited moments (${M_{J_2}}^{\rm CCSD}$) are zeroed in the process of solving CCSD equations. By substituting eq 8 into eq 1, we can derive the so-called Reg-GF(T) approximation (see ref 66)

$$E^{\text{Reg-GF(T)}} = E^{\rm CCSD} + \left\langle \Phi \left| \left\{ \left(\Sigma_{\rm reg,\,3} + \Sigma_{\rm reg,\,1}\Sigma_{\rm reg,\,2} + \frac{1}{6}(\Sigma_{\rm reg,\,1})^3 \right)^\dagger M_3^{\rm CCSD} \right\}_C \right| \Phi \right\rangle \tag{9}$$

where $\Sigma_{\rm reg,1}$, $\Sigma_{\rm reg,2}$, and $\Sigma_{\rm reg,3}$ are the regularized Σ amplitudes. While the $\Sigma_{\rm reg,1}$ and $\Sigma_{\rm reg,2}$ amplitudes are approximated by singly and doubly excited cluster amplitudes obtained by solving CCSD equations with regularized form of the electronic Hamiltonian $H^{\rm reg}$, the $\Sigma_{\rm reg,3}$ amplitudes ($\tilde{\Sigma}^{i_1 i_2 i_3}_{a_1 a_2 a_3}$) are obtained in a perturbative manner

$$\tilde{\Sigma}^{i_1 i_2 i_3}_{a_1 a_2 a_3} = \frac{\overline{M}^{i_1 i_2 i_3}_{a_1 a_2 a_3}(\Sigma_{\rm reg,\,1}, \Sigma_{\rm reg,\,2})}{\varepsilon_{i_1} + \varepsilon_{i_2} + \varepsilon_{i_3} - \varepsilon_{a_1} - \varepsilon_{a_2} - \varepsilon_{a_3} - 3\omega^2} \tag{10}$$

where $\bar{M}^{i_1 i_2 i_3}_{a_1 a_2 a_3}(\Sigma_{\rm reg,1},\Sigma_{\rm reg,2})$ are triply excited moments for regularized CCSD equations. In ref 66, the regularized version of the CCSD(T) approach (Reg-CCSD(T)) was introduced

$$E^{\text{Reg-CCSD(T)}} = E^{\rm CCSD} + \langle \Phi | (V_N \Sigma_{\rm reg,\,2} + V_N \Sigma_{\rm reg,\,1})^+ R_0^{(3)}(\omega^2)(V_N T_2) | \Phi \rangle \tag{11}$$

where V_N is two-body part of the electronic Hamiltonian in normal product form and $R_0^{(3)}(\omega^2)$ is the ω^2-dependent resolvent defined by eq 33 of ref 66. The Reg-CCSD(T) approach can be easily implemented using existing CCSD(T) implementations. In the current form both Reg-GF(T) and Reg-CCSD(T) approaches as ω^2-dependent methods should be classified as semiempirical approaches. Despite its simplicity, the Reg-CCSD(T) method offers considerable improvements upon the CCSD(T) results especially for strongly correlated systems at the same $n_{\rm o}^3 n_{\rm u}^4$ numerical cost as the genuine CCSD(T) method. We believe that efficient parallel GPU implementation of the triples part of the regularized CCSD(T) method has potential to evolve into a tool that is capable of providing credible predictions for strongly correlated systems. In the forthcoming sections, we will give details of our (T) implementations and discuss their parallel performance.

3. IMPLEMENTATION DETAILS

In this section, we present the evaluation of the noniterative triples correction on GPUs. It involves a code-generation based approach to generating CUDA code from a high-level specification of tensor contractions. Several optimizations are identified in mapping the tensor contractions to the resources in a GPU. We then develop a hybrid implementation that effectively utilizes the cores and GPU accelerators available in a cluster of SMP nodes with GPUs.

To calculate triples correction two general type quantities, which appear on the right and on the left of the $R_0^{(3)}(\omega^2)$ resolvent, need to be calculated:

$$\begin{aligned}\langle \Phi^{abc}_{ijk} | V_N T_2 | \Phi \rangle = {} & v^{ij}_{ma} t^{mk}_{bc} - v^{ij}_{mb} t^{mk}_{ac} + v^{ij}_{mc} t^{mk}_{ab} \\ & - v^{ik}_{ma} t^{mj}_{bc} + v^{ik}_{mb} t^{mj}_{ac} - v^{ik}_{mc} t^{mj}_{ab} + v^{jk}_{ma} t^{mi}_{bc} - v^{jk}_{mb} t^{mi}_{ac} + v^{jk}_{mc} t^{mi}_{ab} \\ & - v^{ei}_{ab} t^{jk}_{ec} + v^{ei}_{ac} t^{jk}_{eb} - v^{ei}_{bc} t^{jk}_{ea} + v^{ej}_{ab} t^{ik}_{ec} - v^{ej}_{ac} t^{ik}_{eb} + v^{ej}_{bc} t^{ik}_{ea} \\ & - v^{ek}_{ab} t^{ij}_{ec} + v^{ek}_{ac} t^{ij}_{eb} - v^{ek}_{bc} t^{ij}_{ea}, \; (i < j < k, a < b < c)\end{aligned} \tag{12}$$

and

$$\begin{aligned}\langle \Phi^{abc}_{ijk} | V_N T_1 | \Phi \rangle = {} & v^{ij}_{ab} t^{k}_{c} - v^{ij}_{ac} t^{k}_{b} + v^{ij}_{bc} t^{k}_{a} - v^{ik}_{ab} t^{j}_{c} + v^{ik}_{ac} t^{j}_{b} \\ & - v^{ik}_{bc} t^{j}_{a} + v^{jk}_{ab} t^{i}_{c} - v^{jk}_{ac} t^{i}_{b} + v^{jk}_{bc} t^{i}_{a}, \; (i < j < k, a < b < c)\end{aligned} \tag{13}$$

where T_1 and T_2 refer either to genuine or regularized CCSD amplitudes. The $i, j, k, l, m, n, \ldots$ ($a, b, c, d, e, \ldots$) indices designate occupied (unoccupied) spinorbitals. Of the two terms described by the above equations, the first one (section 3) contributes to the $n_{\rm o}^3 n_{\rm u}^4$ scaling. Tensors corresponding to 2-electron integrals and doubly excited amplitudes are assumed to be antisymmetric in all pairs of lower and upper indices. In order to provide granularity for the parallel Tensor Contraction Engine[38] generated codes, the whole spinrobital domain is partitioned into smaller pieces called *tiles* which contain several spinorbitals of the same spatial- and spin-symmetry. The maximum number of elements in the tile is often referred to as the *tilesize*. This partitioning induces partitioning or block-structure of all tensors used in the CC calculations, including: amplitudes, recursive intermediates, integrals, and residual vectors. In the parallel implementation of the (T)-part each core takes care of different set of projections defined by tiles: $[i]$, $[j]$, $[k]$, $[a]$, $[b]$, $[c]$, i.e., each core generates on-the-fly the set of $\langle \Phi^{abc}_{ijk} | V_N T_2 | \Phi \rangle$ and $\langle \Phi^{abc}_{ijk} | V_N T_1 | \Phi \rangle$ projections with $i \in [i]$, $j \in [j]$, $k \in [k]$, $a \in [a]$, $b \in [b]$, $c \in [c]$. These projections are stored on the six-dimensional matrices $P3$ ($\langle \Phi^{abc}_{ijk} | V_N T_2 | \Phi \rangle$)and $R3$ ($\langle \Phi^{abc}_{ijk} | V_N T_1 | \Phi \rangle$). In our implementation, this condition is replaced by the do -loop structure for each tile corresponding to ($[i] \le [j] \le [k]$) and ($[a] \le [b] \le [c]$). This incurs a small amount of redundancy at the boundary of the conditionals (when the equality is satisfied). This has been shown in practice to be very small as compared to the total work done and is correctly incorporated through the use of appropriate constant coefficients. For example, $P3$ is defined as the following matrix

$$P3 \equiv P3(\dim[a], \dim[b], \dim[c], \dim[i], \dim[j], \dim[k]) \tag{14}$$

where $\dim[i], \ldots, \dim[c]$ are the dimensions of the corresponding tiles (they will be denoted *id, jd, kd, ad, bd, cd*). Therefore, the local

Table 1. Architectural Comparison of Tesla T10 and T20 Series

	Tesla T10	Tesla T20
num. multiprocessors (SM)	30	14
num. cores per SM	8	32
SM clock frequency	1.3 GHz	1.15 GHz
single precision peak GFLOPS	933	1030
double precision peak GFLOPS	78	515
memory frequency	800 MHz	1.5 GHz
memory bandwidth	102 GB/sec	144 GB/sec
memory interface	512 bit	384 bit
shared memory	16 KB	16 KB/48 KB
L1 cache		48 KB/16 KB
L2 cache		768 KB

memory requirement for storing $P3$ and $R3$ matrices is defined by *tilesize*^{6}. If *tilesize* equals 20 this is equivalent to 0.48 GB (in recently developed algorithm for (T)-part of TCE these tensors can be "sliced" along the first two dimensions, which lead to a less intensive use of the local memory and effectively larger *tilesize* can be used in the (T) calculations). Because the total flop count on each core associated with forming $P3$ and $R3$ tensors is proportional to *tilesize*^{6} $*n_u$, where n_u stands here for the total number of correlated virtual spinorbitals, this type of calculation is ideally suited to take advantage of GPU accelerators. The whole process is split into number of smaller tasks, where the summation goes over indices from a single tile, for example

$$P3(a,b,c,i,j,k) - = \sum_{e\in[e]} V(a,b,e,i)*T2(j,k,e,c)$$

$$(i\in[i], j\in[j], k\in[k], a\in[a], b\in[b], c\in[c]) \qquad (15)$$

where $V(a,b,e,i)$ and $T2(j,k,e,c)$ are 2-electron integrals and doubly excited amplitudes tensors. For this elementary task the flop count is equal to *tilesize*^{7}, which for *tilesize* = 20 corresponds to 1.2 GF. We expect, that the utilization of the GPU accelerators should lead to considerable speedups in the case of large numerical load, which is created by the use of larger tiles.

3.1. GPU Architecture and Execution Model. CUDA[73] is a language extension to C developed by NVIDIA to program GPGPUs. GPU devices that support the CUDA programming environment consist of several multiprocessors (SMs), each with a fast shared memory, a constant cache, a single instruction unit, and multiple processor cores.

The CUDA programming model views the GPU as an accelerator to which parts of the computation, referred to as kernels, are offloaded by the host CPU. A kernel consists of a grid of thread blocks, with each thread block consisting of multiple threads. All threads in a kernel invocation can access the global memory, which is persistent across kernel invocations. A thread-block is mapped to a multiprocessor (MP). The shared memory associated with the MP is only accessible to threads in a thread block, and is not persistent across thread blocks. Given that each MP consists of a single instruction unit, effective utilization requires that all threads execute the same instruction whenever possible. Conditional branches that diverge among the threads can greatly reduce achieved performance. The execution in a thread block in which few threads are performing work, referred to as thread block under-utilization, also inhibits performance. Each thread can identify its position in the thread block, and in the grid of thread blocks through implicit variables. These are the only distinguishing identifiers of a given thread, and are used to encode all thread-specific computation.

Table 1 summarizes the specification of the two GPU architectures that were the target of our optimizations. Tesla T10 contains the GF100 GPU processor, while Tesla T20, known as "Fermi", is the latest series of graphic card products by NVIDIA. The peak double precision performance has improved from previous generations by a factor of 6.6 (a peak of ~480 GFlops). However, the memory bandwidth and the clock speed for device memory accesses, and the data transfer rates between the host and the device, have improved by a much smaller factor. Thus, we can expect that accesses to device memory and data transfers from host memory can be even more of a bottleneck with T20 cards. Fermi has a much larger shared memory with respect to previous generations, that can also act as a Level 1 data cache. It can be configured as 48 KB shared memory and 16 KB L1 cache, or 48 KB L1 cache and 16KB shared memory. Shared memory is still allocated per thread block. The register file is also much larger. Unlike previous cards, a Level 2 cache is also available.

3.2. Multi-Dimensional Tensor Contractions. We are interested in a direct implementation of optimized tensor contractions on CUDA-programmable GPU devices. A tensor contraction can be viewed as a generalized multidimensional matrix multiplication. Often, it is transformed into a regular matrix multiplication operation through transposition operations. However, such an approach might not most effectively utilize the available hardware resources.

While matrix multiplication can be optimized assuming large dimensions lending themselves to tiling for every level of the memory hierarchy to ensure that the operation is computation-bound, the large dimensionality of the tensors could often result in each dimension being relatively small. The small size of each dimension interferes with achieving good locality. The encoding of the indices into the thread coordinates incurs high index computation overhead. The small size of the common dimension results in the equivalent of highly rectangular matrix multiplication operations, which are harder to optimize than the square versions typically employed in benchmark studies.

3.3. CUDA-Targeted Tensor Contraction Implementation. In this section, we present the various optimizations performed, with illustrative code snippets, in generating an efficient CUDA implementation of a given tensor contraction. We begin with optimizations that are generally applicable to CUDA-capable devices, in particular both T10 and T20 NVIDIA GPU cards. This discussion is followed by a presentation of the optimizations specifically targeted at the T20 cards with their latest generation Fermi GPUs. We consider a typical tensor contraction specification given by the following example:

$$\mathtt{P3(b,c,a,i,k,j)} - = \sum_{l\in[l]} \mathtt{V2(l,a,i,j)*T2(l,k,c,b)}$$

which is taken from the TCE generated code for the noniterative (T) correction (slightly different convention is used in eq 15. `P3`, `V2`, and `T2` are tensors, and `a`, `b`, `c`, `i`, `k`, `j`, and `l` are particle or hole indices as appropriate). Note that the GPU implementation faithfully reproduces the computation structure on the CPU, ensuring correctness of the results produced and avoiding any potential redundant computation, other than those in the CPU version.

3.3.1. Memory Management. A typical computation in the application involves numerous calls to the sequential tensor contraction. Allocating and deallocating CUDA memory, both on the host and on the device, for every kernel invocation would be expensive. We implemented a memory manager that serves allocation requests from previously allocated memory that is current not being used. The memory allocation for the arrays is as shown. The getGPUmem calls are the wrappers implemented for more efficient memory management

```
  double  *P3_d=getGPUmem(size_of_P3);
  double  *T2_d=getGPUmem(size_of_T2);
  double *V2_d=getGPUmem(size_of_V2);
```

In the sample code above, size_of_P3 denote the size of the array P3. It reuses the previously freed GPU memory if it is enough for P3.

3.3.2. Kernel Arguments. Accessing portions of multidimensional arrays in GPU memory requires computation of the strides along the different dimensions. This scalar arithmetic is redundantly executed by each thread and cannot be overlapped with memory or floating-point operations on these systems. We therefore compute the strides in the host CPU, as shown below, and pass them as arguments to the kernel. Note that P3 is stored as a one-dimensional array. The offsets for other arrays are calculated in a similar fashion.

```
  /* parameters calculated in host: */
  P3_offset_c = bd;
  P3_offset_a = bd*cd;
  ...
 /* offset computation in GPU kernel: */
 int offset_P3(b,c,a,i,k,j) {
   return b + P3_offset_c*c +
    P3_offset_a*a + ··· +
  P3_offset_j*j;
  }
  /* offset_T2(1,k,c,b)
  and offset_V2(1,a,i,j) are
   similarly defined */
```

3.3.3. Encoding Thread-Block Specific Arguments. The indices to be operated upon by a given thread are decoded from its coordinates in the thread block and thread block grid by modulo and division operations. These operations are not very efficient on GPUs as they require the use of Special Function Units. To reduce such operations, the indices of the output matrix are mapped to a two dimensional thread grid, according to the two input matrices, which means the indices from the first input matrix are mapped to the y dimension of the thread grid, and indices from the second input matrix are mapped to the *x* dimension, resulting in the configuration below.

```
  dim2 dimGrid(ceil(kd/BLOCK_DIM_Y)*cd
  *bd,ceil(ad/BLOCK_DIM_X)*id*jd);
```

3.3.4. Index Combining. In some tensor contractions, a sequence of indices might occur in the same order in every occurrence. These sequences of indices can be replaced by a combined index. This optimization reduces index computation overhead while improving thread block utilization. While combining of indices might not be possible for all tensor contractions, there is no negative side-effect and we employ this optimization where possible. For the illustrative tensor contraction, by using index combining, we set *x* as the multiplication of *a* and *i*, therefore it results in the following:

```
    P3(b,c,x,k,j) –  = T2(1,k,c,b)*V2(1,x,j)
```

3.3.5. Dimension Flattening. Tiling of the loops in a tensor contraction enables different threads to cooperate in data movement, and enables maximal reuse of the transferred data to minimize data transfer per floating point operation. The tensor contraction could be implemented as a sequence of matrix multiplication on possibly strided data. In the example that we have been using through this section, all threads in one thread block work on elements P3(b,c,aT:aT+16,i,kT:kT+16,j), assuming one block has 16 × 16 threads. This works well when the tiled dimensions are large or match the thread block configuration. The application tile sizes encountered at runtime often do not match the fixed thread block size chosen at compile time. The result is the execution of numerous threadblocks with fewer threads than available, with ensuing poor utilization of threads blocks and poor performance. For example, a dimension size of 17 for a thread block size of 16 results in two thread blocks, with total only 17 of total 32 threads being used.

To solve this problem, we optimize the code by *dimension flattening*. This is illustrated in Figure 1. In the dimension-flattened version, we group the indices of the output array into two dimensions, according to the origin of the indices. In the example we have been using, the two groups are (k,c,b) and (a,i,j). Tiling is done on the grouped indices, instead of a single index as in the original algorithm. Thus, different threads in one thread block could compute output elements that have more than two different indices. The order in which the indices are flattened has an impact on the data access costs for the arrays. We considered flattening orders that favors inputs and outputs and found the order favoring the inputs to be most effective. This is the order used in the experiments.

3.3.6. Pipelined Execution. In a parallel execution of coupled cluster calculations, the data is transferred from remote to local memory for processing by the host CPU. This data needs to be transferred to the GPU memory prior to kernel invocation and the result transferred back to host memory. CUDA enables overlap of this data transfer with kernel invocation through the use of streams. To ensure efficient data transfer, we use the outermost dimension of the output tensor as the streaming index. The number of kernel invocations is the same as the value of the streaming index. dimension.

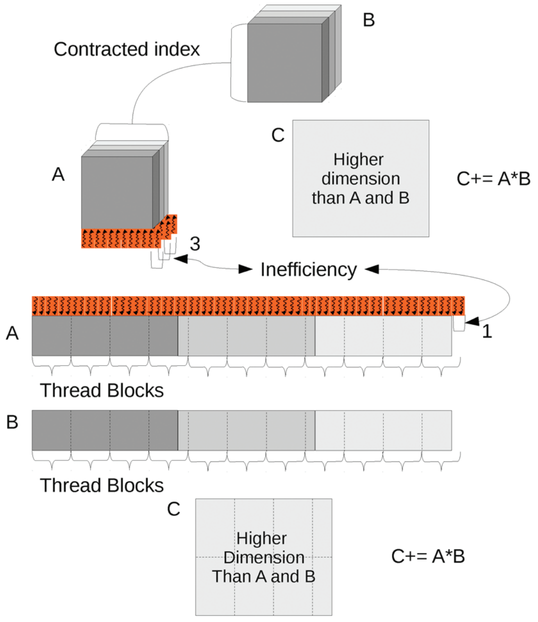


Figure 1. Illustration of dimension flattening. The solid lines correspond to different two-dimensional regions. The dotted lines correspond to the mapping of the data blocks to the thread blocks. A and B matrices are implicitly flattened into two-dimensional arrays. This implicit flattening is used to map the work to be performed to the thread blocks. As shown in the figure, this results in most thread-blocks being fully utilized.

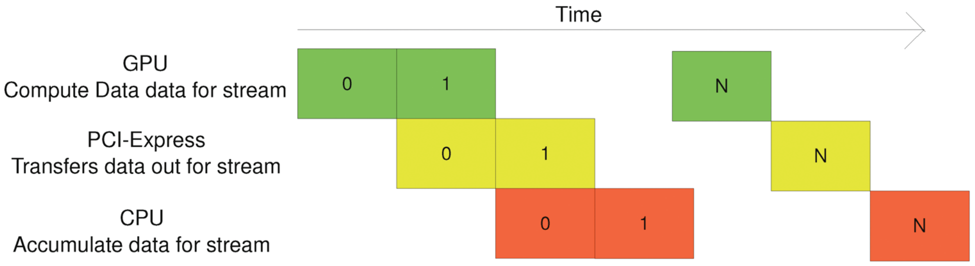

Time
GPU
Compute Data data for stream
PCI-Express
Transfers data out for stream
CPU
Accumulate data for stream
0
1
N
0
1
N
0
1
N

Figure 2. Three-stage pipelined execution with CPU and GPU concurrently participating in the contraction.

Building upon the streaming strategy we also implement a pipelined approach with the host participating in the contraction by performing the accumulation and thus avoiding the initial copy of the large tensors in the GPU memory. With this approach we traded off data transfer with computation, allowing a better utilization of the available resources and reducing the amount of traffic on the PCI-Express bus. This three-stage pipelined design ensures that the different components of the architecture: the CPU, the GPU, and the PCI-express bus are all utilized to minimize the execution time. Figure 2 shows the concept above-discussed, assuming the total amount of streams to transfer equal to *N*.

3.3.7. Hybrid Execution. The above optimization, with the GPU and CPU collaborating in performing a single contraction, is a clear instance of hybrid execution. However, general high-performance SMP nodes (as the one used in our experiments) have more CPU computing cores than GPUs. As each GPU in an SMP node requires a separate core that drives the kernel execution and participates in its pipelined processing, the "spare" cores can be used to compute serially other contractions. This can be seen as a second level of hybrid execution. We implemented this approach introducing a high level load balancer to distribute the contractions. This design maximally utilizes the computational resources available to reduce the time to solution.

3.3.8. Optimizations Specific to the Fermi (T20) GPU Architecture. The CUDA code generator had to be modified to adapt to the different architectural trade-offs presented by the Fermi GPU architecture. Here we identify the specific architectural differences and the adaptations we undertook to account for them.

3.3.8.1. Elimination of PCI Express Data Transfer. As discussed above, the data transfer across the PCI express can be effectively overlapped with the kernel execution on the GPU, through different pipelining approaches. The greater factor of improvement in the computation rate on the Fermi GPU architecture as compared to the improvement in memory bandwidth results in the execution being bound by data transfer, even with the pipelining approaches. To achieve the best execution time, we modified the application to keep the intermediate in memory, across tensor contractions. The inputs need to be transferred into the GPU. The intermediate is computed and translated into a scalar contribution to the energy value, which is transferred back to the host memory. The compute structure is shown below:

```
double *P3_s = getGPUmem(size_inter-
mediate);
/*intermediate tile */
double *P3_d = getGPUmem(sizeof(
double));
/*energy scalar*/
ccsd_t_single(P3_s);
/*Moves inputs to GPU memory;*/
/*Result updated in GPU memory*/
ccsd_t_doubles_1((P3_s); /*as above*/
ccsd_t_doubles_2(P3_s); /* -do- */
compute_energy(P3_s, P3_d); /* -do- */
transfer_to_cpu(P3_d); /*one element */
```

3.3.8.2. Register Tiling. Fermi includes a wider register file and larger shared memory enabling register tiling, which improves data reuse and reduces the data transfer with the GPU memory hierarchy. We modified our implementation such that each thread contributes to 16 output elements, rather than 1 as in the original algorithm. All contributions are stored in 16 double precision registers and finally written back to GPU memory.

3.3.8.3. Scalar Optimizations. While the double precision performance is improved by a factor of 6.6 in Fermi, the scalar instructions do not see a corresponding improvement in performance. However, register tiling results in each thread performing more computation enabling scalar optimizations across calculations for the 16 output elements. Each write to the GPU memory is enclosed in a boundary check to ensure that thread has a valid contribution to make, in cases when the tile being executed is smalled the thread block size. We coalesce the condition checks to minimize them in the common case. The original boundary checking order for four output elements is as shown below:

```
if(thread_y<total_y)
  p3[offset_1]+=tlocal1
if(thread_y+16<total_y)
  p3[offset_2]+=tlocal2;
if(thread_y+16*2<total_y)
  p3[offset_3]+=tlocal3;
if(thread_y+16*3<total_y)
  p3[offset_4]+=tlocal4;
```

The modified order we employ is shown below:

```
if(thread_y+16*3<total_y)
  {p3[offset_1]+=tlocal1;
  p3[offset_2]+=tlocal2;
  p3[offset_3]+=tlocal3;
  p3[offset_4]+=tlocal4;
  }
 else if(thread_y+16*2<total_y)
  {p3[offset_1]+=tlocal1;
  p3[offset_2]+=tlocal2;
  p3[offset_3]+=tlocal3;
  }
 else if(thread_y+16<total_y)
  {p3[offset_1]+=tlocal1;
  p3[offset_2]+=tlocal2;
  p3[offset_3]+=tlocal3;
  }
 else if(thread_y<total_y)
  {p3[offset_1]+=tlocal1;
  }
```

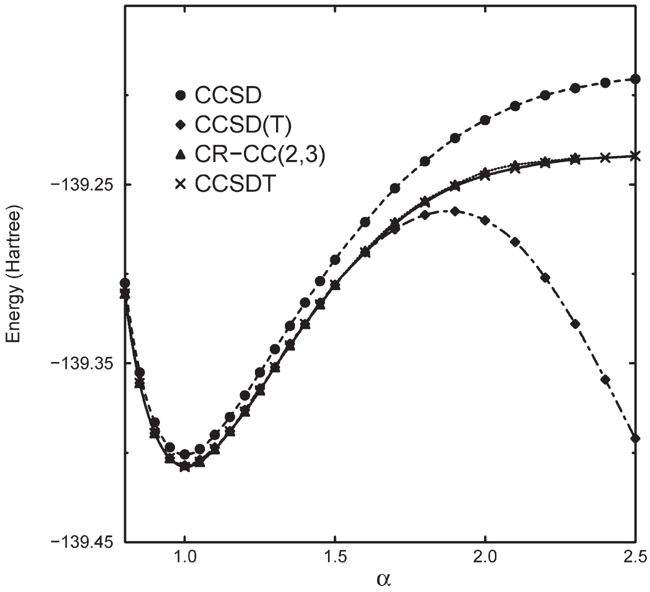


Figure 3. CC energies for methylfluoride system as a function of C−F elongation obtained with the cc-pVDZ basis set (see text for details).

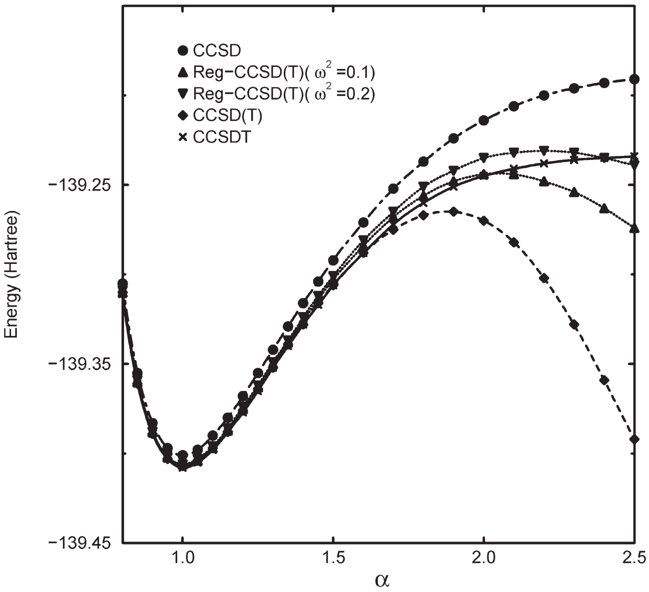


Figure 4. Regularized CC energies for methylfluoride system as a function of C−F elongation obtained with the cc-pVDZ basis set (see text for details).

Note that the above optimizations do not have an impact on the older GPU architecture, while being crucial to maximize performance on Fermi.

4. RESULTS AND DISCUSSION

In this section, we discuss the quality of the Reg-CCSD(T) method when applied to several challenging molecular systems and provide several illustrative performance examples of our GPU implementation of the noniterative triples Reg-CCSD(T) correction.

For the H_3CF system we used the cc-pVDZ basis set,[74] for the dodecane we employed 6-311G basis set,[75] whereas for Spiro cation we used Sadlej's basis set (POL1),[76] which is composed of 486 atomic basis set functions. The geometry of the methylfuoride molecule was optimized with the B3LYP method[77] using cc-pVTZ basis set.[74] The geometry of Spiro cation is as discussed in ref 78. The geometry of dodecane was optimized with the B3LYP approach using cc-pVTZ basis set. In all calculations reported here core electrons were not correlated.

It is well-known that various renormalized CC approaches can provide correct description of processes involving single bond breaking. The first two examples are to illustrate the performance of the Reg-CCSD(T) methods for these processes. We start our discussion from the ground-state singlet potential energy surface of the methylfluoride molecule H_3CF as a function of the C−F bond elongation. The B3LYP/cc-pVTZ equilibrium values of the $H-C$ ($R_{H-C}(eq)$) and $C-F$ ($R_{C-F}(eq)$) bond lengths are equal to 1.09021 and 1.38655 Å, respectively, whereas the equilibrium H−C−H angle is equal to 109.878 deg. The geometry of the system is defined by a single parameter α (see Figure 3 and Figure 4), which defines the C−F distance (R_{C-F} as $R_{C-F} = \alpha R_{C-F}(eq)$). In the present studies we consider the set of geometries with corresponding α's falling into the 0.8,2.5 interval.

It has been already shown[79] that for larger internuclear distances the restricted Hartree−Fock (RHF) determinant is a rather poor choice of the reference, which results in divergent behavior of perturbative triples estimates of the CCSD(T) method. These problems are eliminated by the iterative inclusion of connected triply excited clusters which is sufficinet to obtain nearly MRCI(Q)[80,81] level of accuracy for geometries considered here (see ref 79). Since the approximate CCSDT method (the CCSDT-1b approach[12]) has been considered in ref 79, we should expect that the full CCSDT results provide further improvement of the CCSDT-1b energies, which can in turn be used to calibrate the accuracy of the regularized CCSD(T) method.

In Figure 3, we show the CCSD, CCSD(T), CR-CC(2,3),[33−35] and CCSDT energies. The CR-CC(2,3) calculations were perfromed using the GAMESS implementation.[82,83] While the CCSD(T) method provides a good approximation of the CCSDT results at the equilibrium region (the CCSD(T) energy error with respect to the CCSDT energy is as small as 0.5 milliHartree) it fails at larger internuclear distances. For example the CCSD(T) error for α = 2.5 is equal to −158 milliHartree. This is a direct consequence of the poor choice of the RHF reference. The absolute values of the largest T_1 and T_2 amplitudes: 0.60 and 0.96 for α = 2.5, provide a good illustration of these problems. The very efficient CR-CC(2,3) approach is capable of removing all deficiences characterizing the CCSD(T) in this case. For example, the errors of the CR-CC(2,3) method for α = 1.0, 1.5, 2.0, and 2.3 are equal to 0.02, 0.5, 1.8, and 0.6 milliHartree, respectively (the CR-CC(2,3) results are reported up to α = 2.3 only; for larger distances NWChem and GAMESS were converging to different RHF solutions).

It is interesting to analyze the extent to which the regularization procedure can offset the problems plaguing the CCSD(T) approach. In Figure 4 we compare the Reg-CCSD(T) results with the CCSDT ones for two choices of the ω^2 parameter: 0.1 and 0.2. The impact of the choice of ω^2 on the Reg-CCSD(T) accuracies was a subject of discussion in ref 66. Using the example of the HF molecule we concluded that for strong quasidegeneracy the choice of larger ω^2 may be beneficial. Indeed, for the HF system the best agreement between CCSDT and Reg-CCSD(T) for stretched geometries was obtained for larger values of ω^2 (see ref 66 for details). This can also be observed using the H_3CF

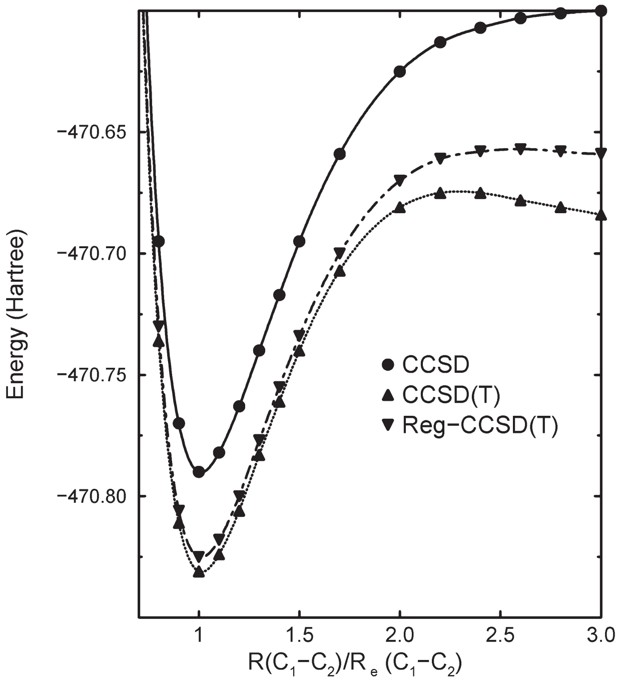


Figure 5. CCSD, CCSD(T), Reg-CCSD(T)(ω^2 = 0.1) energies as functions of the C_1-C_2 bond stretch in dodecane.

example. Despite of the fact that the Reg-CCSD(T)(ω^2 = 0.1) approach is capable of reducing the −158 milliHartree CCSD(T) error at α = 2.5 almost 4-fold, still large error of −40 milliHartree remains. The increase of the ω^2 value results in further reduction of the Reg-CCSD(T)(ω^2 = 0.1) error down to −5 milliHartrre obtained with the Reg-CCSD(T)(ω^2 = 0.2) approach. At the same time the overall error of the Reg-CCSD(T) (ω^2 = 0.2) approach do not exceed 10 milliHartree for all geometries considered here (the Reg-CCSD(T)(ω^2 = 0.2) errors are 1.8, 5.1, 9.96, and −4.89 milliHartree for α = 1.0, 1.5, 2.0, and 2.5, respectively). This clearly demonstrate the advantages of using stronger regularization for quasidegenerate systems and also demonstrates that more flexible regularization methods should be developed in order to minimize the errors for the equilibrium and stretched geometries.

Recently, the state-of-the-art CR-CCSD(2,3) method[33−35] was used to describe ground-state PES corresponding to the $C_{12}H_{26}$ dissociation into $C_{11}H_{23}$ and CH_3.[84] It was shown that the CR-CC(2,3) curve along bond breaking coordinates (corresponding to varied C_1-C_2 separation, see Figure.3 of ref 84) smoothly approaches dissociation limit without any unphysical barriers, which are often observed in the methods employing many-body perturbation theory. With Reg-CCSD(T) method we performed similar studies using 6-311G basis set. The results of our calculations are shown in Figure 5. One can notice that the CCSD(T) curve discloses (in analogy to the methyfluoride) typical symptoms of perturbative breakdown. As in the previous example (H_3CF), for large internuclear distance, the doubly excited amplitudes assume the largest values (the largest amplitude assumes −0.8 value). In contrast to methylfluoride no big T_1 amplitudes have been observed in the CCSD calculations for dodecane. The unphysical hump of the CCSD(T) method is located around 2.2 $R_e(C_1-C_2)$. The CCSD(T) energy for 3$R_e(C_1-C_2)$ is located around 9 milliHartree below the CCSD(T) energy calculated for the "hump" geometry. The Reg-CCSD(T) approach to a large extent eliminates this pathological behavior: the analogous difference is reduced to 1.8 milliHartree. At the same time the Reg-CCSD(T) method yields energy of the CCSD(T) quality at the equilibrium geometry.

The main goal of studying the Spiro cation is to characterize the lowest doublet state of A_2 symmetry along a defined reaction pathway, which corresponds to the electron transfer from one π to other π moiety (see refs 67−70 and 78 for details). The Spiro cation has two equivalent C_{2v} minima, with a D_{2d} intermediate geometry (the neutral ground state is D_{2d}). In our studies of electron transfer we used the geometric change parameter ζ from the work of ref.,[70] which defines a simple linear mixing of the two mirror image C_{2v} minima (Q_A and Q_B)

$$Q(\zeta) = \left(\frac{1}{2}-\zeta\right)Q_A + \left(\frac{1}{2}+\zeta\right)Q_B \qquad (16)$$

The barrier region corresponds to ζ = 0.0 value.

Recently, the Spiro cation was the subject of intensive studies (see refs 68−70 and 78) using high-level theory including multireference perturbative (MRPT) approaches such as CASPT2,[85] various orders of the NEVPT method,[86] and second order of multiconfigurational quasi-degenerate perturbation theory (MCQDPT2)[87]). Using the approximate pathway (eq 16) it was demonstrated that the CASPT2, NEVPT2, and MCQDPT2 formalisms experience a serious problem (an unphysical minimum) in the description of correlation effects in the vicinity of the avoided crossing between the 1^2A_2 and 2^2A_2 states. This unphysical minimum on the ground-state PES in the vicinity of the avoided crossing region can be removed either by invoking higher orders of theory (NEVPT3) or by averaging the orbital energies of two charge-localized one particle states.[70] In ref 78, we showed that the single reference CC methods are capable of providing a satisfactory description of the ground-state PES as a function of the ζ parameter, avoiding to a large extent the problems plaguing multireference methods. In this paper we compare the Reg-CCSD(T) results with those obtained with the CCSD(T) and CR-CCSD(T) (version CR-CCSD(T),IA of ref 88) approaches. The CCSD(T), Reg-CCSD(T) (ω^2 = 0.1), and CR-EOMCCSD(T) energies for the 1^2A_2 state are shown in Table 2 and Figure 6. One should notice that the CCSD(T) corrections to the CCSD energies are in excess of 100 milliHartree for all geometries discussed here. Moreover, the presence of the overlap denominator in the CR-CCSD(T) correction makes the CR-CCSD(T) corrections two times smaller than the CCSD(T) ones. It is interesting to notice that the Reg-CCSD(T) results are invariably between the CCSD(T) and CR-CCSD(T) energies. The same observation is valid for the barrier heights which are equal to 0.078, 0.065, 0.069, and 0.077 eV for the CCSD, CCSD(T), Reg-CCSD(T), and CR-CCSD(T) methods, respectively. Unlike higher orders of MRPT theory, the single reference formulations suffer from cuspy behavior at the transition state geometry (ζ = 0.0). We believe that this feature can be eliminated by employing MCSCF orbitals. Another unresolved issue is related to a "kink" at all levels of CC theory in the vicinity of ζ = ± 0.25. We expect that this may be associated with the reference change or ROHF instability at that region. Unfortunately, using even very small ζ increments and continuing the ROHF solution from the equilibrium geometry we could not find an alternative solution.

We now focus on evaluating the performance of our implementation. The experiments were performed on two clusters,

Table 2. Comparison of the CC Energies As Functions of ζ-Parameter for the Spiro Moleclue[a]

geom.	CCSD	CCSD(T)	Reg-CCSD(T)	CR-CCSD(T)
ζ= 1.50	−611.36232	−611.46506	−611.45046	−611.40588
ζ= 1.40	−611.36396	−611.46675	−611.45214	−611.40755
ζ= 1.30	−611.36540	−611.46826	−611.45363	−611.40902
ζ= 1.20	−611.36666	−611.46959	−611.45494	−611.41030
ζ= 1.10	−611.36773	−611.47072	−611.45606	−611.41139
ζ= 1.00	−611.36861	−611.47167	−611.45699	−611.41228
ζ= 0.90	−611.36930	−611.47243	−611.45773	−611.41298
ζ= 0.80	−611.36980	−611.47300	−611.45829	−611.41351
ζ= 0.70	−611.37013	−611.47340	−611.45867	−611.41384
ζ= 0.60	−611.37026	−611.47361	−611.45885	−611.41398
ζ= 0.50	−611.37021	−611.47362	−611.45885	−611.41394
ζ= 0.40	−611.36997	−611.47346	−611.45866	−611.41371
ζ= 0.30	−611.36954	−611.47312	−611.45829	−611.41330
ζ= 0.20	−611.36921	−611.47289	−611.45803	−611.41298
ζ= 0.10	−611.36840	−611.47217	−611.45728	−611.41216
ζ= 0.05	−611.36793	−611.47175	−611.45684	−611.41168
ζ= 0.00	−611.36740	−611.47125	−611.45634	−611.41117

[a] In all calculations Sadlej's TZ basis set was used, and core electrons were not correlated.

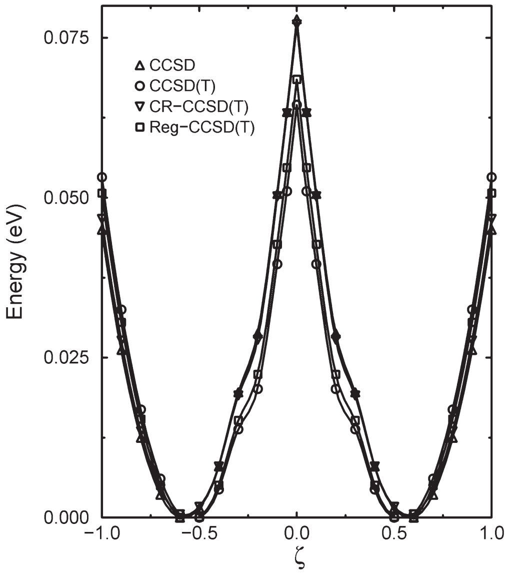


Figure 6. CC 1^2A_2 PESs obtained for the Spiro molecule described using POL1 basis set.[76] The lowest point of a given theory has been shifted to zero.

Table 3. Performance of Different Algorithms (In Milliseconds) on Microbenchmark

problem size	cublasDgemm	baseline	combining	flattening	pipelining
(16,16,16,16,16,16,16)	464	109	101	105	50
(17,17,17,16,16,16,16)	336	233	216	131	59
(10,10,10,19,19,18,19)	114	65	60	44	21

one with Tesla T10 GPUs, and the other with the Fermi cards. The T10 cluster consists of 64 nodes while the T20 cluster contains 16 nodes. Each node on the cluster with Tesla T10 GPUs has two Quad-Core Intel Xeon X5560 CPUs, with a frequency of 2.80 GHz, and 8 MB L2 cache. Two nodes share one Tesla S1070 box, implying that every node has two Tesla T10 GPUs. Each node on the Fermi cluster is equipped with two Quad-Core Intel Xeon E5520 CPUs, with the frequency of 2.27 GHz. Each node has a single GPU. PCI Express 2.0 is used for I/O between the host and the device on both systems. GNU 4.1.2 and NVCC 2.3 compilers are used for compilation, and CUBLAS 2.3 was used for the cublas-dgemm based algorithm. We begin with an evaluation of the individual block contributions, which constitute a parallel tensor contraction, on single GPUs, followed by the full parallel execution on the two clusters of GPUs.

In Table 3, we show the effects of our optimizations. These experiments are done on a micro benchmark which only includes one block contribution, on 3 problem sizes, using a single process. The comparison is among 5 versions, explained as below.

cublasDgemm. This is based on the original algorithm, in which each tensor contraction is implemented with index permutation and dgemm operations. We replaced the dgemm function with the Fortran wrapper and CUBLAS call, a library function provided for dgemm in CUDA.

Baseline. The basic CUDA algorithm described in the previous section.

Combining. Baseline. version with index combination.

Flattening. Combining version with flattened index.

Pipelining. Flattening version with pipelining to overlap data movement and kernel execution.

Among the 3 problem sizes, the first one has all dimensions set as 16, which fits the thread block configuration perfectly; the second problem size has the first dimensions as 17, resulting in big difference between thread block configuration and matrix index size; the third one is a random problem size picked from the NWChem trace. From the table, we can see that the baseline version is already doing better than the cublas version, because of reduced permutation operations. Pipelining is playing an important role in improving performance. Index combining improves performs further. Index flattening provides significant improvement when the problem size does not fit in the thread block configuration.

In the following paragraphs, we present our experiments on the full Reg-CCSD(T) execution. We demonstrate the scalability and performance improvements due to GPU execution using three systems: the Spiro, uracil, and dodecane molecules.

We did a comparison between GPU and CPU versions by running the two versions of code for the Spiro molecule. Since each node has only two GPUs, we ran 2 processes on each node, each driving a GPU. With 32 nodes, the time for CPU version is 42776 s, and the time of GPU version is 6950 s. So the GPU version yields a speedup of more than 6. When using 7 processes on each node, the GPU version has 2 processes using GPU and 5 processes using CPU only. This mixed version also has a speedup of about 3 over the version of using 7 CPU processes on each node. This demonstrates the speed-ups achieved by the GPU-based approach as compared to the CPU-only implementation. To provide a perspective on these times with respect to the overall execution time, we measured the times involved in the different modules executed. For the mixed version utilizing 2 GPUs and 5 CPUs on 32 nodes, for a total of 224 processes, the Hartee-Fock routine consumed 170 s, the four-index tranformation procedure consumed 2061 s, and the iterative CCSD

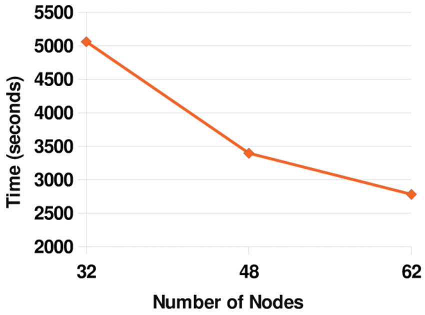


Figure 7. Running time for the (T) correction for Spiro molecule.

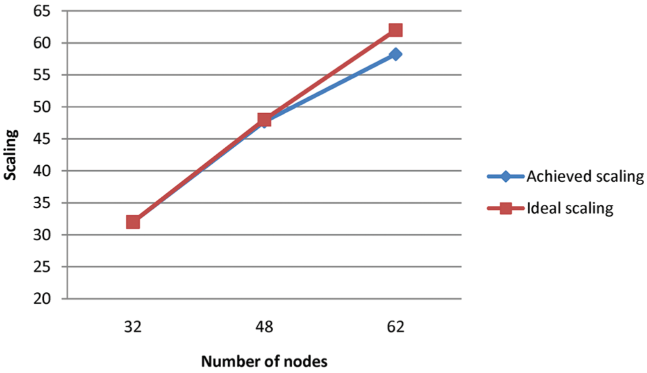


Figure 8. Scaling of the (*T*) correction for Spiro molecule.

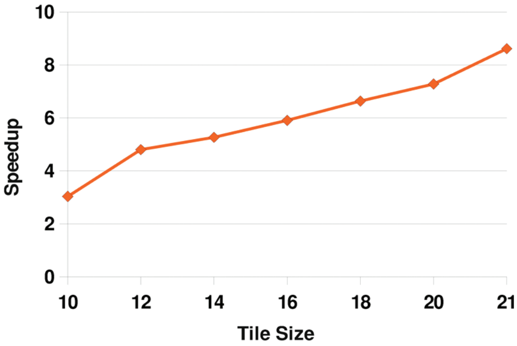


Figure 9. Speedup of GPU for the uracil molecule.

routines 167 s. In comparison, 5079 s were consumed by the noniterative triples correction.

The scalability of the GPU implementation is evaluated on 32, 48, and 62 nodes of the T10-based cluster. The execution times are shown in Figure 7 and scalability in Figure 8. The scalability is plotted with a baseline scaling of 32 on 32 nodes. We observe that the execution time scales almost linearly, achieving good scalability in addition to its speed-up over the CPU-only implementation.

The Spiro molecule represents systems characterized by relatively high symmetry (all calculations were performed using C_{2v} symmetry). This fact results in tiles which may be too small to provide optimum flop count for efficient use of GPU cores (tiles correspond to the partitioning of the spinorbital domain, which in turns implicates the block structure of all multidimensional tensors used in CC calculations: cluster amplitudes, recursive intermediates, and residual vectors; the length of tiles will be commonly referred to as the tilesize).

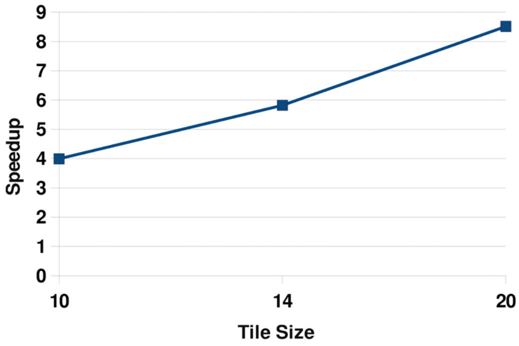


Figure 10. Speedup of GPU for the dodecane molecule.

In order to test the impact of tilesize on the GPU performance we tested GPU speedup on the uracil molecule in 6-31G* basis set where the spatial symmetry was not invoked. This situation commonly occurs in calculations for large systems without symmetry where tilesize can be sufficiently large. Of special importance is to understand the impact of tilesize on the GPU speedup.

The experiments are run on 30 nodes, with 2 processes on each node. The speedup of GPU over CPU version is shown in Figure 9. It can be seen that with larger tile size, which implies more FLOPS per process (proportional to $(tilesize)^7$), the speedup of GPU is more obvious. While for small tilesizes (tilesize = 10) the GPU speedup is rather modest (around 3), for larger tiles (tilesize = 21) the speedup is much better (around 8.75). From data shown in Figure 9 we should expect that the further increase in the tilesize should result in a further improvement in the GPU speedup.

To verify the generality of these observation, we evaluated the impact of tile sizes on the dodecane molecule. Shown in Figure 10, we observe speedups improving with tile sizes, reaching more than a factor of 8 with a tile size of 20.

5. CONCLUSIONS

We demonstrated that the Reg-CCSD(T) approach can improve the accuracies of the CCSD(T)method in studies of strongly correlated closed- and open-shell systems. We showed that for the methylfluoride, the regularization procedure can to a large extent eliminate the problems characteristic for the standard CCSD(T) approach. Our studies on the dodecane dissociation clearly indicate that the Reg-CCSD(T) approach provides significant improvements of the CCSD(T) results for the stretched geometries. We hope that this observation is generally valid for processes involving single bond breaking. For dodecane we showed that $\omega^2 = 0.1$ regularization is sufficient to obtain reliable shape of potential energy surface. The Reg-CCSD(T) results obtained for the methylfluoride suggest that the regularization procedures for singly and doubly excited clusters may require various ω^2 values, especially for states with large T_1. We also demonstrated that the Reg-CCSD(T) energies for the Spiro cation are free of the problems plaguing multireference approaches. We believe that for strong quasidegeneracy effects the use of larger values of ω^2 leads to more reliable results. It should be also stressed that the cost of triples part of Reg-CCSD(T) approach is exactly the same as its CCSD(T) analog. This problem will be pursued in a separate paper. The new GPU implementation of the Reg-CCSD(T) codes showed a great promise as far as the parallel performance of the most computationally N^7 part is concerned. We hope that the further advances in GPU-based technology will enable reliable CC calculations for medium (large) molecular systems on desktop (massively parallel) GPU computers.

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: sriram@pnl.gov (S.K.); karol.kowalski@pnl.gov (K.K.).

■ ACKNOWLEDGMENT

The work related to the development of the GPU Reg-CCSD(T) implementation was supported by the Extreme Scale Computing Initiative (SK,OV,KK), a Laboratory Directed Research and Development Program at Pacific Northwest National Laboratory. The calculations were performed using the Barracuda cluster in EMSL, a national scientific user facility sponsored by the Department of Energy's Office of Biological and Environmental Research and located at Pacific Northwest National Laboratory. The Pacific Northwest National Laboratory is operated for the U.S. Department of Energy by the Battelle Memorial Institute under Contract DE-AC06-76RLO-1830. The Barracuda cluster was purchased with funds received under the American Recovery and Reinvestment Act.

■ REFERENCES

(1) Coester, F. *Nucl. Phys.* **1958**, *7*, 421–424.
(2) Coester, F.; Kümmel, H. *Nucl. Phys.* **1960**, *17*, 477–485.
(3) Čížek, J. *J. Chem. Phys.* **1966**, *45*, 4256–4266.
(4) Paldus, J.; Shavitt, I.; Čížek, J. *Phys. Rev. A* **1972**, *5*, 50–67.
(5) Crawford, T. D.; Schaefer, H. F. *Rev. Comput. Chem.* **2000**, *14*, 33–136.
(6) Piecuch, P.; Kowalski, K.; Pimienta, I. S. O.; McGuire, M. J. *Int. Rev. Phys. Chem.* **2002**, *21*, 527–655.
(7) Bartlett, R. J.; Musiał, M. *Rev. Mod. Phys.* **2007**, *79*, 291–352.
(8) Purvis, G. D.; Bartlett, R. J. *J. Chem. Phys.* **1982**, *76*, 1910–1918.
(9) Cullen, J. M.; Zerner, M. C. *J. Chem. Phys.* **1982**, *77*, 4088–4109.
(10) Noga, J.; Bartlett, R. J. *J. Chem. Phys.* **1987**, *86*, 7041–7050.
(11) Noga, J. *J. Chem. Phys.* **1988**, *89*, 3401–3401.
(12) Urban, M.; Noga, J.; Cole, S. J.; Bartlett, R. J. *J. Chem. Phys.* **1985**, *83*, 4041–4046.
(13) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479–483.
(14) Stanton, J. F. *Chem. Phys. Lett.* **1997**, *281*, 130–134.
(15) Stanton, J. F.; Gauss, J. *J. Chem. Phys.* **1995**, *103*, 1064–1076.
(16) Stanton, J. F.; Gauss, J. *Theor. Chim. Acta* **1996**, *93*, 303–313.
(17) Kucharski, S. A.; Bartlett, R. J. *J. Chem. Phys.* **1998**, *108*, 5243–5254.
(18) Kucharski, S. A.; Bartlett, R. J. *J. Chem. Phys.* **1998**, *108*, 5255–5264.
(19) Crawford, T. D.; Stanton, J. F. *Int. J. Quantum Chem.* **1998**, *70*, 601–611.
(20) Gwaltney, S. R.; Head-Gordon, M. *Chem. Phys. Lett.* **2000**, *323*, 21–28.
(21) Gwaltney, S. R.; Head-Gordon, M. *J. Chem. Phys.* **2001**, *115*, 2014–2021.
(22) Gwaltney, S. R.; Byrd, E. F. C.; Van Voorhis, T.; Head-Gordon, M. *Chem. Phys. Lett.* **2002**, *353*, 359–367.
(23) Hirata, S.; Nooijen, M.; Grabowski, I.; Bartlett, R. J. *J. Chem. Phys.* **2001**, *114*, 3919–3928.
(24) Hirata, S.; Nooijen, M.; Grabowski, I.; Bartlett, R. J. *J. Chem. Phys.* **2001**, *115*, 3967–3968.
(25) Bomble, Y. J.; Stanton, J. F.; Kállay, M.; Gauss, J. *J. Chem. Phys.* **2005**, *123*, 8.
(26) Kállay, M.; Gauss, J. *J. Chem. Phys.* **2005**, *123*, 13.
(27) Taube, A. G.; Bartlett, R. J. *J. Chem. Phys.* **2008**, *128*, 13.
(28) Taube, A. G.; Bartlett, R. J. *J. Chem. Phys.* **2008**, *128*, 9.
(29) Kowalski, K.; Piecuch, P. *J. Chem. Phys.* **2000**, *113*, 18–35.
(30) Kowalski, K.; Piecuch, P. *J. Chem. Phys.* **2000**, *113*, 5644–5652.
(31) McGuire, M. J.; Piecuch, P.; Kowalski, K.; Kucharski, S. A.; Musiał, M. *J. Phys. Chem. A* **2004**, *108*, 8878–8893.
(32) Kowalski, K.; Piecuch, P. *J. Chem. Phys.* **2005**, *122*, 12.
(33) Piecuch, P.; Włoch, M. *J. Chem. Phys.* **2005**, *123*, 10.
(34) Piecuch, P.; Włoch, M.; Gour, J. R.; Kinal, A. *Chem. Phys. Lett.* **2006**, *418*, 467–474.
(35) Włoch, M.; Gour, J. R.; Piecuch, P. *J. Phys. Chem. A* **2007**, *111*, 11359–11382.
(36) Tajti, A.; Szalay, P. G.; Császár, A. G.; Kállay, M.; Gauss, J.; Valeev, E. F.; Flowers, B. A.; Vázquez, J.; Stanton, J. F. *J. Chem. Phys.* **2004**, *121*, 11599–11613.
(37) Bomble, Y. J.; Vázquez, J.; Kállay, M.; Michauk, C.; Szalay, P. G.; Császár, A. G.; Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **2006**, *125*, 8.
(38) Hirata, S. *J. Phys. Chem. A* **2003**, *107*, 9887–9897.
(39) Lotrich, V.; Flocke, N.; Ponton, M.; Yau, A. D.; Perera, A.; Deumens, E.; Bartlett, R. J. *J. Chem. Phys.* **2008**, *128*, 15.
(40) Kuś, T.; Lotrich, V. F.; Bartlett, R. J. *J. Chem. Phys.* **2009**, *130*, 7.
(41) Janowski, T.; Ford, A. R.; Pulay, P. *J. Chem. Theory Comput.* **2007**, *3*, 1368–1377.
(42) Janowski, T.; Pulay, P. *Chem. Phys. Lett.* **2007**, *447*, 27–32.
(43) Janowski, T.; Ford, A. R.; Pulay, P. *Mol. Phys.* **2010**, *108*, 249–257.
(44) Bentz, J. L.; Olson, R. M.; Gordon, M. S.; Schmidt, M. W.; Kendall, R. A. *Comput. Phys. Commun.* **2007**, *176*, 589–600.
(45) de Jong, W. A.; Bylaska, E.; Govind, N.; Janssen, C. L.; Kowalski, K.; Muller, T.; Nielsen, I. M. B.; van Dam, H. J. J.; Veryazov, V.; Lindh, R. *Phys. Chem. Chem. Phys.* **2010**, *12*, 6896–6920.
(46) Kowalski, K.; Krishnamoorthy, S.; Villa, O.; Hammond, J. R.; Govind, N. *J. Chem. Phys.* **2010**, *132*, 154103.
(47) Yoo, S.; Apra, E.; Zeng, X. C.; Xantheas, S. S. *J. Phys. Chem. Lett.* **2010**, *1*, 3122–3127.
(48) Valiev, M.; Bylaska, E. J.; Govind, N.; Kowalski, K.; Straatsma, T. P.; Van Dam, H. J. J.; Wang, D.; Nieplocha, J.; Apra, E.; Windus, T. L.; de Jong, W. A. *Comput. Phys. Commun.* **2010**, *181*, 1477–1489.
(49) Apra, E.; Harrison, R.; de Jong. W. A.,; Rendell, A.; Tipparaju, V.; Xantheas, S.; Olsen, R. *Proc. of the ACM/IEEE Supercomp. 2009 Conf.* 2009; pp 66:1−66:7.
(50) Anderson, A. G.; Goddard, W. A.; Schröder, P. *Comput. Phys. Commun.* **2007**, *177*, 298–306.
(51) Owens, J. D.; Luebke, D.; Govindaraju, N.; Harris, M.; Kruger, J.; Lefohn, A. E.; Purcell, T. J. *Comput. Graphics Forum* **2007**, *26*, 80–113.
(52) Stone, J. E.; Phillips, J. C.; Freddolino, P. L.; Hardy, D. J.; Trabuco, L. G.; Schulten, K. *J. Comput. Chem.* **2007**, *28*, 2618–2640.
(53) Hardy, D. J.; Stone, J. E.; Schulten, K. *Parallel Comput.* **2009**, *35*, 164–177.
(54) Stone, J. E.; Hardy, D. J.; Ufimtsev, I. S.; Schulten, K. *J. Mol. Graph. Model.* **2010**, *29*, 116–125.
(55) Yasuda, K. *J. Comput. Chem.* **2008**, *29*, 334–342.
(56) Yasuda, K. *J. Chem. Theory Comput.* **2008**, *4*, 1230–1236.
(57) Ufimtsev, I. S.; Martinez, T. J. *J. Chem. Theory Comput.* **2008**, *4*, 222–231.
(58) Ufimtsev, I. S.; Martinez, T. J. *J. Chem. Theory Comput.* **2009**, *5*, 1004–1015.
(59) Ufimtsev, I. S.; Martinez, T. J. *J. Chem. Theory Comput.* **2009**, *5*, 2619–2628.
(60) Anderson, J. A.; Lorenz, C. D.; Travesset, A. *J. Chem. Phys.* **2008**, *227*, 5342–5359.
(61) Vogt, L.; Olivares-Amaya, R.; Kermes, S.; Shao, Y.; Amador-Bedolla, C.; Aspuru-Guzik, A. *J. Phys. Chem. A* **2008**, *112*, 2049–2057.

(62) Friedrichs, M. S.; Eastman, P.; Vaidyanathan, V.; Houston, M.; Legrand, S.; Beberg, A. L.; Ensign, D. L.; Bruns, C. M.; Pande, V. S. *J. Comput. Chem.* **2009**, *30*, 864–872.

(63) van Meel, J. A.; Arnold, A.; Frenkel, D.; Zwart, S. F. P.; Belleman, R. G. *Mol. Simul.* **2008**, *34*, 259–266.

(64) Eastman, P.; Pande, V. S. *J. Comput. Chem.* **2010**, *31*, 1268–1272.

(65) Ma, W.; Krishnamoorthy, S.; Villa, O.; Kowalski, K. *IEEE Intl. Conf. Cluster Comp.* **2010**, 207–216.

(66) Kowalski, K.; Valiev, M. *J. Chem. Phys.* **2009**, *131*, 12.

(67) Farazdel, A.; Dupuis, M.; Clementi, E.; Aviram, A. *J. Am. Chem. Soc.* **1990**, *112*, 4206–4214.

(68) Pastore, M.; Helal, W.; Isti, S. E.; Leininger, T.; Malrieu, J. P.; Maynau, D.; Angeli, C.; Cimiraglia, R. *J. Chem. Phys.* **2008**, *128*, 9.

(69) Helal, W.; Evangelisti, S.; Leininger, T.; Maynau, D. *J. Comput. Chem.* **2009**, *30*, 83–92.

(70) Pastore, M.; Helal, W.; Angeli, C.; Evangelisti, S.; Leininger, T.; Cimiraglia, R. *THEOCHEM* **2009**, *896*, 12–17.

(71) Kowalski, K.; Fan, P. D. *J. Chem. Phys.* **2009**, *130*, 11.

(72) Taube, A. G.; Bartlett, R. J. *J. Chem. Phys.* **2009**, *130*, 14.

(73) NVIDIA, NVIDIA CUDA C Programming Guide; http://developer.download.nvidia.com/compute/cuda/3_2_prod/toolkit/docs/CUDA_C_Programming_Guide.pdf (accessed 11/9/2010).

(74) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(75) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650–654.

(76) Sadlej, A. J. *Collect. Czech. Chem. C.* **1988**, *53*, 1995–2016.

(77) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623–11627.

(78) Glaesemann, K. R.; Govind, N.; Krishnamoorthy, S.; Kowalski, K. *J. Phys. Chem.* **2010**, *114*, 8764–8771.

(79) Schütz, M. *J. Chem. Phys.* **2002**, *116*, 8772–8785.

(80) Knowles, P. J.; Werner, H. *J. Chem. Phys. Lett.* **1988**, *145*, 514–522.

(81) Werner, H. J.; Knowles, P. J. *J. Chem. Phys.* **1988**, *89*, 5803–5814.

(82) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347–1363.

(83) Piecuch, P.; Kucharski, S. A.; Kowalski, K.; Musiał, M. *Comput. Phys. Commun.* **2002**, *149*, 71–96.

(84) Li, W.; Piecuch, P.; Gour, J. R.; Li, S. H. *J. Chem. Phys.* **2009**, *131*, 30.

(85) Andersson, K.; Malmqvist, P. A.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. *J. Phys. Chem.* **1990**, *94*, 5483–5488.

(86) Angeli, C.; Cimiraglia, R.; Evangelisti, S.; Leininger, T.; Malrieu, J. P. *J. Chem. Phys.* **2001**, *114*, 10252–10264.

(87) Nakano, H. *J. Chem. Phys.* **1993**, *99*, 7983–7992.

(88) Kowalski, K.; Piecuch, P. *J. Chem. Phys.* **2004**, *120*, 1715–1738.

JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

An Extension of the Hirshfeld Method to Open Shell Systems Using Fractional Occupations

D. Geldof, A. Krishtal, F. Blockhuys, and C. Van Alsenoy*

Department of Chemistry, University of Antwerp, Universiteitsplein 1, B2610 Antwerp, Belgium

ABSTRACT: In this work, a new partitioning method is presented which allows one to calculate properties of radicals, in particular, atomic spin populations. The method can be seen as an extension of the Hirshfeld-I method [Bultinck, P. et al. *J. Chem. Phys.* **2007**, *126*, 144111], in which the atomic weight functions, defining the atoms-in-molecules, are constructed by means of an iterative scheme in which the charges of the atoms-in-molecules are altered but the spin remains fixed. The Hirshfeld-I method is therefore not suitable for the calculation of atomic spin populations of open-shell systems. The new fractional occupation Hirshfeld-I (FOHI) uses an iterative scheme in which both the atomic charge and spin are optimized, resulting in a self-consistent method for the calculation of atomic spin populations. The results obtained with the FOHI method are compared with experimental results obtained using polarized neutron diffraction, thus serving as a validation of the FOHI method as well as the Hirshfeld definition of atoms-in-molecules in general.

1. INTRODUCTION

Atomic charges are one of the quantities most frequently addressed by chemists when rationalizing structure and reactivity of molecules. Organic chemists tend, for example, to use partial charges when interpreting reaction mechanisms,[1] where they can be used to calculate interaction energies between molecules and to construct potential energy surfaces. Since no unique definition of the charge of an atom can be formulated, different methods have been developed to calculate this property. These methods can be divided into two categories. The first category uses LCAO coefficients of the basis functions, which are used to represent the wave function. The first method developed along these lines, and until now the most widely used due to its simplicity, is the Mulliken population analysis.[2] The main drawback of the method is its strong basis set dependence, which becomes particularly problematic when diffuse functions are used, leading to results which have no physical meaning.[3] The natural population analysis[4,5] is at this moment the most elaborate population-analysis-based method, being less sensitive to the choice of the basis set but challenging for extensions to other properties than charges.[6] The other category is based on the electron density in real space. In these methods, the electron density $\rho(\vec{r})$ is divided into atomic densities $\rho_A(\vec{r})$ by making use of atomic weight functions $w_A(\vec{r})$

$$\rho_A(\vec{r}) = w_A(\vec{r})\rho(\vec{r}) \tag{1}$$

Different approaches are possible to define the weight function. In Bader's quantum chemical topology,[7] the weight function is discrete and can only be equal to one or zero. At every point, the molecular density is thus assigned to one single atom. This atom is enclosed in a basin, separated from the rest of the atoms by a surface constructed using the zero density flux condition. Another example is the method by Becke, which is an overlapping atoms-in-molecules method,[8] in which the atoms are separated from each other by means of Voronoi polyhedra, and the Voronoi deformation density (VDD),[9] which also uses Voronoi polyhedra, but in combination with the deformation density rather than the full density. The Hirshfeld method[10] also uses diffuse boundaries where the weight function of an atom A can be in principle nonequal to zero at any point $\vec{r}$ of the space. The "share" of each atom at point $\vec{r}$ is relative to the share of the free atom in the promolecular density.

$$w_A^H(\vec{r}) = \frac{\rho_A^0(\vec{r})}{\sum_B \rho_B^0(\vec{r})} \tag{2}$$

In economics terms, an atom can thus be seen as owning a share of the molecule's stock of electrons, wherefrom the original name of this procedure, "stockholder method", was derived. The promolecular density is defined as the sum of the densities of the isolated atoms, positioned at the same coordinates as the atomic nuclei in the real molecule. Integration of the atomic density leads to the population of every atom:

$$N_A = \int \rho_A(\vec{r})\, d\vec{r} = \int w_A(\vec{r})\, \rho(\vec{r})\, d\vec{r} \tag{3}$$

In this first version of the Hirshfeld method, the isolated atoms are usually chosen as neutral atoms. The arbitrary character of the choice of the isolated atomic densities to build up the promolecular density is one of the major criticisms on this method.[11] One actually assumes that the atoms-in-molecules resemble best the free spherically symmetric neutral atoms. This may be the case for molecules consisting of covalently bonded atoms with no significant electronegativity differences but questionable for molecules such as LiF. For example, the atomic charges in the molecule LiF will depend strongly on whether the promolecular density is composed of Li^0 and F^0, Li^+ and F^-, or Li^- and F^+, the second being evidently the more chemically reasonable choice.

Received: December 23, 2010
Published: April 20, 2011

But since making a chemically sensible guess is not always as straightforward in more extended systems, one would prefer a method which does not require any prior knowledge about the chemical properties of the substrate. Moreover, this scheme is clearly problematic when partitioning charged systems, since in that case the actual system and the promolecule contain a different number of electrons.

These problems have recently been solved for the atomic charges in the Hirshfeld-I method[11] (HI), by making use of an iterative scheme. In this method, the promolecule in each iteration is constructed from atomic densities obtained in the previous iteration, thus allowing the atomic populations of the atoms-in-molecules to change. As a result, the resulting promolecular density and atomic weight functions no longer depend on the first guess: in the case of LiF, identical atomic charges are obtained when the iterative procedure is started from any of the three possibilities mentioned above. This procedure brings the Hirshfeld method more in line with information theory, on the basis of which expression 3 can be derived.[12] The validity of this improved definition of the Hirshfeld weight-function was confirmed also for other quantities such as polarizabilities,[13,14] electrostatic potentials,[15] and dispersion interactions.[16,17]

In principle, atomic spin populations, which are useful in the field of molecular magnetism, can be calculated with the partitioning methods mentioned above.[18,19] However, straightforward application of these methods to the spin density leads to results which suffer from the same inconsistenties as the charges in the classic Hirshfeld method, namely, the arbitrairiness of the choice of spin. In the case of the Hirshfeld-I method, the total number of electrons in each atom in the promolecule is allowed to vary until convergence, but no constraints are laid upon the spin of the atoms. As a result, the Hirshfeld-I method does not seem suitable for open-shell systems when atomic spin populations need to be calculated. In this article, we propose a new method (fractional occupation Hirshfeld-I, FOHI) which extends the Hirshfeld-I method so that the atomic spin populations are also calculated in an iterative way.

The outline of this paper is as follows. Section 2 contains the details of the new method, followed by computational details in section 3. In section 4, the atomic spin populations of both methods will be compared for a set of radicals using different levels of theory. In order to validate the new method, experimental results will be used to compare different partitioning methods. Finally, section 5 contains a brief summary and the conclusions of this paper.

2. METHOD

In order to solve the problem of the fixed isolated atomic densities in the original Hirshfeld method,[12] the HI method[11] used the following interpolation formula:

$$\rho_A^{N_A}(r) = \rho_A^{\mathrm{lint}(N_A)}(r)[\mathrm{uint}(N_A) - N_A] + \rho_A^{\mathrm{uint}(N_A)}(r)[N_A - \mathrm{lint}(N_A)] \qquad (4)$$

for the calculation of the atomic densities (N_A) needed to construct the promolecular density during the different iterations of their procedure. In eq 4, $\mathrm{lint}(x)$ represents the integer part of x (i.e., lower integer), while $\mathrm{uint}(x) = \mathrm{lint}(x) + 1$ (i.e., upper integer). This function interpolates the atomic density between two atoms with an integer population to compute the atomic density of an atom with an noninteger number of electrons. On the basis of these atomic densities, a new promolecule is constructed, and a new weight function is calculated. The atomic densities of every atom ranging from integer charges -2 up to $+2$ in their respective spectroscopic ground state are calculated and stored beforehand. Finally, integration leads to new atomic populations, and this process is repeated until the atomic populations are converged for all atoms.

For open-shell systems, atomic spin populations can be obtained by integrating the molecular spin density with the atomic weight functions.

$$N_A^{\mathrm{spin}} = \int w_A(\vec{r})(\rho^{\alpha}(\vec{r}) - \rho^{\beta}(\vec{r}))\,\mathrm{d}\vec{r} \qquad (5)$$

In the HI method, the weight function $w_A(\vec{r})$ obtained after convergence of the atomic charges is used to calculate the atomic spin populations. However, during the iterations which determine the weight function, the molecular spin density is not used.

The weight function in the HI method was shown to minimize the loss of information entropy by equalizing the number of electrons in the atoms building up the promolecule with the number of electrons in the actual atoms-in-molecules.[11,12] For example, in a diatomic molecule AB, the loss of information due to the partitioning of the density $\rho(\vec{r})$ into a sum of the approximate atomic densities $\rho_A(\vec{r})$ and $\rho_B(\vec{r})$, instead of the "true" (but unknown) atomic densities $\rho_A^0(\vec{r})$ and $\rho_B^0(\vec{r})$, is given by the following function:

$$I = I_A + I_B = N_A \int \sigma_A(\vec{r}) \ln\left(\frac{\sigma_A(\vec{r})}{\sigma_A^0(\vec{r})}\right) \mathrm{d}\vec{r} + N_B \int \sigma_B(\vec{r}) \ln\left(\frac{\sigma_B(\vec{r})}{\sigma_B^0(\vec{r})}\right) \mathrm{d}\vec{r} + N_A \ln\left(\frac{N_A}{N_A^0}\right) + N_B \ln\left(\frac{N_B}{N_B^0}\right) \qquad (6)$$

where N_A and N_B are the atomic populations of the atoms constituting the promolecule, N_A^0 and N_B^0 are the "true" atomic populations of the atoms in the molecule, and $\sigma(\vec{r})$ is the shape function, defined as[20]

$$\sigma_A(\vec{r}) = \frac{\rho_A(\vec{r})}{N_A} \qquad (7)$$

One can see that eq 6 is minimized if the normalization constraints $N_A = N_A^0$ and $N_B = N_B^0$ are fullfilled, which is achieved by means of the iterative procedure in the HI method.

In open shell molecules, where the electrons of opposite spin are described by separate densities $\rho^{\alpha}(\vec{r})$ and $\rho^{\beta}(\vec{r})$, additional constraints must be added, namely, $N_{\mathrm{a}}^{\alpha} = N_A^{0,\alpha}$, $N_A^{\beta} = N_A^{0,\beta}$, $N_B^{\alpha} = N_B^{0,\alpha}$, and $N_B^{\beta} = N_B^{0,\beta}$. The information loss is then described for each atom as

$$I_A^{\text{open-shell}} = N_A^{\alpha} \int \sigma_A^{\alpha}(\vec{r}) \ln\left(\frac{\sigma_A^{\alpha}(\vec{r})}{\sigma_A^{0,\alpha}(\vec{r})}\right) \mathrm{d}\vec{r} + N_A^{\beta} \int \sigma_A^{\beta}(\vec{r}) \ln\left(\frac{\sigma_A^{\beta}(\vec{r})}{\sigma_A^{0,\beta}(\vec{r})}\right) \mathrm{d}\vec{r} + N_A^{\alpha} \ln\left(\frac{N_A^{\alpha}}{N_A^{0,\alpha}}\right) + N_A^{\beta} \ln\left(\frac{N_A^{\beta}}{N_A^{0,\beta}}\right) \qquad (8)$$

where the atomic shape function is now defined for α and β densities separately

$$\sigma_A^\alpha(\vec{r}) = \frac{\rho_A^\alpha(\vec{r})}{N_A^\alpha} \tag{9}$$

In the HI method, these constraints are not met, so there is no guarantee that the atoms in the promolecule have the same spin as the atoms in the molecule. In order to minimize the information loss, the α and β densities must be iterated separately. An additional complication is that a certain inconsistency is present in the interpolation formula (eq 4), even for closed shell molecules. The atomic densities are usually interpolated between atoms of different charge, but calculated at the Hartree−Fock level and at their spectroscopic ground-state, which is not necessarily a singlet. This means that for closed-shell molecules, the atoms in the promolecule do not have zero spin, as one assumes they should.

A solution for the problems outlined above is proposed by extending the HI procedure in such a way that both charge and spin are calculated every iteration. However, extending this procedure by using the interpolation formula eq 4 requires a 2D interpolation scheme of both charge and spin in every iteration, which in our view is not practical. Inspired by ref 6, we propose to use another approach to build up the promolecule. The orbitals for a spherically symmetric atom with a given charge and multiplicity are calculated in the unrestricted approach using fractional occupations for degenerate valence orbitals while all other orbitals lower in energy are fully occupied. A calculation of this type is performed for every atom in the molecular system at the same level of theory (DFT functional and basisset) as the molecular system in every step of the iterative procedure. For example, in a given iteration, for a carbon atom with a charge of −0.1 and a spin (α − β) of +0.3, the occupation of the 1s, 2s, $2p_x$, $2p_y$, and $2p_z$ orbitals would be 1, 1, 0.4, 0.4, and 0.4 for the α electrons and 1, 1, 0.3, 0.3, and 0.3 for the β electrons, respectively. In every iteration, the promolecular density is constructed using the atomic densities calculated using the description described above, for every atom in the system. The weight function of the HI method based on eq 2 cannot be used, because this weight function is only based on the total density $\rho(\vec{r}) = \rho^\alpha(\vec{r}) + \rho^\beta(\vec{r})$. Two new weight functions are defined, one based on the α density and one based on the β density. The weight function for the α density is defined as:

$$w_A^\alpha(\vec{r}) = \frac{\rho_A^\alpha(\vec{r})}{\sum\limits_B \rho_B^\alpha(\vec{r})} \tag{10}$$

The molecular α density is converted to atomic α densities by integration:

$$N_A^\alpha = \int \rho_A^\alpha(\vec{r})\, \mathrm{d}\vec{r} = \int w_A^\alpha(\vec{r})\, \rho^\alpha(\vec{r})\, \mathrm{d}\vec{r} \tag{11}$$

Equivalent formulas are used for the β density.

The new method can be summarized as follows:

- On the basis of the atomic charge and spin, atomic SCF calculations are performed on the basis of fractional occupations to compute the corresponding densities for every atom in the system. In the first step, both the charge and spin of every atom are set equal to zero.
- Two promolecular densities are constructed: one based on the atomic α density and one based on the atomic β density.
- A Hirshfeld partitioning of the α and β molecular density is performed using the respective promolecular densities constructed in the previous step.
- The atomic charge and spin of every atom is compared with the corresponding values in the previous iteration. If convergence is not reached, the previous steps are repeated.

Table 1. Atomic Charges for a Set of Closed-Shell Systems Obtained Using the HI and FOHI Methods Obtained with the RB3LYP Functional

		HI		FOHI	
molecule	atom	6-31G	6-31++G**	6-31G	6-31++G**
C_2H_2	C	−0.211	−0.217	−0.215	−0.219
	H	0.211	0.217	0.215	0.219
CH_4	C	−0.680	−0.597	−0.716	−0.571
	H	0.170	0.149	0.179	0.142
CO_2	C	0.838	0.950	0.832	0.942
	O	−0.419	−0.475	−0.416	−0.470
H_2CO	C	0.276	0.304	0.270	0.262
	O	−0.326	−0.380	−0.320	−0.365
	H	0.025	0.038	0.025	0.052
H_2O	O	−0.841	−0.920	−0.830	−0.900
	H	0.421	0.459	0.415	0.450
H_2S	S	−0.264	−0.276	−0.245	−0.255
	O	0.132	0.138	0.123	0.128
HCN	N	−0.235	−0.283	−0.239	−0.289
	C	0.023	0.066	0.022	0.067
	H	0.212	0.217	0.217	0.223
HF	F	−0.472	−0.513	−0.464	−0.511
	H	0.472	0.513	0.464	0.511
NH_3	N	−1.070	−1.065	−1.066	−1.001
	H	0.357	0.355	0.356	0.334
$N^{(1)}N^{(2)}O$	$N^{(1)}$	−0.259	−0.271	−0.254	−0.258
	$N^{(2)}$	0.567	0.595	0.552	0.567
	O	−0.308	−0.325	−0.298	−0.309
PH_3	P	−0.117	−0.062	−0.076	−0.009
	H	0.039	0.021	0.025	0.003
SO_2	S	1.015	1.072	0.995	1.079
	O	−0.507	−0.536	−0.497	−0.539

Whether or not the level of theory (method/basisset) of the atomic self-consistent field calculations is the same as that with which the molecular density is calculated is not a prerequisite of the partitioning procedure. In section 4, self-consistent atomic densities obtained with the UB3LYP level of theory will be used to partition molecular densities obtained at the same level of theory, as well as densities obtained using the UMP2 and the UCCSD levels of theory.

3. COMPUTATIONAL DETAILS

For the molecules used in this article, the geometry of every structure was optimized using the Gaussian 03 program[22] with the UB3LYP functional[23] using the 6-31++G**[24] basis set. Additional single points were also performed at the UMP2 and UCCSD levels using the aug-cc-pVTZ[25] basis set and at the DFT/UB3LYP level using the 6-31G,[26] 6-311++G**,[27] and aug-cc-pVTZ basis sets. Mulliken partitioning of the spin density

Table 2. Atomic Spin Populations for a Set of Radicals Based on HI and FOHI Using the aug-cc-pVTZ Basis Set[a]

		HI			FOHI		
molecule doublet	atom	UB3LYP	UMP2	UCCSD	UB3LYP	UMP2	UCCSD
CN	C	0.790	0.512	0.877	1.003	0.589	1.140
	N	0.209	0.488	0.123	−0.003	0.411	−0.140
CF_3	C	0.654	0.681	0.681	1.007	1.058	1.058
	F	0.115	0.106	0.106	−0.002	−0.019	−0.019
COH	C	0.562	0.509	0.589	0.655	0.567	0.704
	O	0.288	0.344	0.264	0.187	0.266	0.151
	H	0.150	0.147	0.146	0.159	0.167	0.145
H_2NCO	C	0.512	0.421	0.544	0.711	0.559	0.768
	O	0.221	0.343	0.203	0.103	0.266	0.070
	N	0.255	0.231	0.243	0.227	0.225	0.204
	H	0.006	0.003	0.005	−0.020	−0.025	−0.021
$C^{(1)}H_3C^{(2)}ONH$	$C^{(1)}$	0.006	0.018	0.004	0.025	−0.074	0.026
	H	0.004	0.010	0.001	0.005	0.026	0.002
	$C^{(2)}$	−0.006	0.228	−0.015	−0.157	0.314	−0.176
	O	0.277	−0.136	0.230	0.321	−0.210	0.272
	N	0.702	0.841	0.762	0.872	0.962	0.950
	H	0.019	0.028	0.020	−0.057	−0.053	−0.065
C_3H_5	$C^{(1)}$	0.466	0.458	0.477	0.788	0.764	0.821
	$C^{(2)}$	−0.032	−0.018	−0.052	−0.350	−0.315	−0.398
	$H^{(1)}$	0.024	0.026	0.023	−0.067	−0.060	−0.072
	$H^{(2)}$	0.028	0.030	0.027	0.026	0.011	0.029
triplet							
HCN	N	0.382	0.388	0.285	−0.004	0.009	−0.191
	C	1.539	1.555	1.646	1.897	1.929	2.076
	H	0.080	0.058	0.069	0.107	0.063	0.115
NF	N	1.709	1.742	1.752	1.862	1.905	1.915
	F	0.291	0.258	0.248	0.138	0.096	0.085
singlet							
C_6H_4	$C^{(1)}$	0.757	0.734		1.166	1.070	
	$C^{(2)}$	0.010	0.107		−0.258	−0.036	
	$C^{(3)}$	−0.010	−0.107		0.258	0.036	
	$C^{(4)}$	−0.757	−0.734		−1.166	−1.070	
	$C^{(5)}$	−0.010	−0.107		0.258	0.036	
	$C^{(6)}$	0.010	0.107		−0.258	−0.036	

[a] The UB3LYP functional is compared with the UMP2 and UCCSD level of theory.

was performed using the Gaussian 03 program.[22] Bader analysis[28] was performed using Gaussian cube files. For the fractional occupation Hirshfeld-I partitioning (FOHI), the atomic densities were calculated at every iteration using the BRABO package[29] with the UB3LYP method and with the same basis set used for the molecule. These SCF calculations were performed using fractional occupations as described above. Both HI and FOHI charge and spin populations were evaluated by using the STOCK program.[30] In the HI method, convergence is reached when abs($N_A^i - N_A^{i-1}$) < 0.001. Convergence in the FOHI method is reached when two conditions are met: abs($N_A^i - N_A^{i-1}$) < 0.001 and abs($S_A^i - S_A^{i-1}$) < 0.001, where S stands for the spin population of an atom.

4. RESULTS AND DISCUSSION

4.1. Charges of Closed-Shell Systems. It is possible to calculate the atomic densities necessary to build up the promolecule using two methods: the interpolation method (HI), where the atomic density of an atom containing a noninteger number of electrons N_A is interpolated between two atomic densities with an integer number of electrons according to eq 4, and the fractional occupation method described above (FOHI), in which the density is obtained directly through an SCF calculation using fractional occupation numbers. For closed-shell systems, the spin density is zero in every point of the space. Strictly speaking, this does not mean that the atomic spin density of each atom at each point in space should be zero, but only their sum at each point in space. However, since we are working with spherically symmetric promolecular atoms, this condition can only be met by restricting the spin density of these atoms to zero. Both methods were applied for a set of closed-shell molecules: C_2H_2, CH_4, CO_2, H_2CO, H_2O, H_2S, HCN, HF, NH_3, N_2O, PH_3, and SO_2. Table 1 compares the atomic charges calculated with both methods, using the 6-31G and 6-31++G** basis sets. The very high regression coefficients (R^2 = 0.9996 and R^2 = 0.9993 for the basis

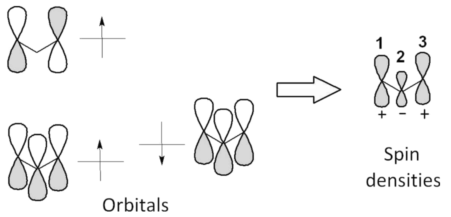


Figure 1. The second highest occupied orbitals and the SOMO for the allyl radical.

sets 6-31G and 6-31++G**, respectively) clearly show that the choice of method has, for closed shell systems, little influence on the charge of the atom. Indeed, according to the information theory, the constraint of charge is fullfilled equally well in the HI method (eq 6) as in the FOHI method (eq 8), regardless of whether we are dealing with closed-shell or open-shell systems. The main difference between the two methods, when applied to charges, comes from the manner of obtaining the atomic densities, namely, through interpolation or through a fractional occupations calculation. This high correlation between both methods is not surprising, since the spectroscopic state of the atom has a negligible effect on the atomic density.[30] This finding serves as a validation of the fact that the interpolation method is a good approximation for the calculation of atomic charges of closed shell systems. This high correlation between the results is also in agreement with a previous study by Ayers,[31] who has shown that the density of a system with an irrational number of electrons is equivalent with the density acquired through interpolation (eq 4). Finally, one can also see that the basis set dependence of both methods is small, as was to be expected from a previous study.[32]

4.2. Spin Densities of Open-Shell Systems. In order to further compare both methods, now for open-shell systems, the spin populations of a set of small radicals were calculated using both the HI and the FOHI methods. The list of radicals and their corresponding spin populations can be found in Table 2 for a number of doublets, triplets, and a singlet diradical. As mentioned above, the results for the charges are very similar for both methods and are therefore not present in the table. The spin populations were calculated using the UB3LYP, UMP2, and UCCSD methods and the aug-cc-pVTZ[25] basis set. The unrestricted formalism is used throughout this work since it has been shown that the restricted open-shell methods cannot reproduce spin polarization effects, which are of importance for spin densities.[33] As mentioned above, it is known that the spin density is influenced by two major effects: spin delocalization and spin polarization.[34] The first effect describes the spreading out of the spin over the molecule, which is a direct consequence of the delocalization of the SOMO (singly occupied molecular orbital). Spin polarization is a consequence of the minimization of electron−electron repulsion of two electrons of parallel spin sharing the same space, according to the Pauli principle. Consequently, one may expect induced negative spin density on the nodal atoms of the SOMO if there is an underlying π orbital which is localized on the same atoms as the SOMO.[34] An example is given for the allyl radical in Figure 1, which can also be found in ref 6, where the upper part of the scheme represents the SOMO orbital and the lower part of the scheme represents the α and β π orbitals in the unrestricted case. One can see that if the SOMO orbital has a node, as is the case for the second carbon atom, the polarization of the π-orbital due to the presence of the radical leads to a negative spin density on the nodal atom. In the RODFT formalism, no polarization of the π orbital is possible, and no negative spin densities are possible. Unlike negative electronic populations, negative spin populations have in fact physical meaning and can be observed experimentally.

The HI and FOHI methods are compared in Table 2 for the three levels of theory. First, we will discuss the HI and FOHI methods based on the DFT level of theory. For the CN and CF$_3$ radicals, the FOHI method localizes the unpaired electron on the carbon atom, whereas the HI method spreads the spin density over several atoms. As we have mentioned for the allyl radical, a negative spin density is expected to be found on the central carbon atom. Both HI and FOHI indeed show a negative spin density for this carbon atom, although this is more pronounced in the FOHI method. With the exception of the UMP2 method, the amide radical has a large negative spin density on the carbon atom in the FOHI method, which is again very small in the HI method. Also for the triplet molecules, one can see that the HI method has more difficulties localizing the spin populations in comparison with the FOHI method. For example, in the HCN molecule in the triplet state, the HI method localizes one unpaired electron on the C atom while the other is spread out over the C and the N atoms, whereas in the FOHI method both unpaired electrons are localized on the C atom. The singlet diradical 1,4-didehydrobenzene has two unpaired electrons in the para position from each other with opposite spin. The position of the two unpaired electrons in the para position having opposite spins is found in both methods, but the alternation of the spin density over the carbon atoms due to spin polarization on the ortho- and meta-carbons is only showing in the FOHI method, while the HI method divides the molecule in a spin-positive and spin-negative part.

Both methods were further compared in partitioning molecular densities obtained at the UMP2 and UCCSD level of theory. For both methods, one can state that the DFT and UCCSD levels are in agreement for the doublet radicals. The UMP2 level shows significant differences for doublet radicals for both methods. The spin density on the CN radical is spread over both atoms. For the H$_2$NCO radical, the spin density on the oxygen atom is much larger compared to DFT and UCCSD. Finally, for the amide radical, the UMP2 level is the only one which localizes a negative spin density on the oxygen atom. The results for the molecules in the triplet state are alike between the different levels of theory in the HI method. For the FOHI method, the UCCSD level shows a significant negative spin density on the nitrogen atom for the HCN triplet. For the singlet diradical, the spin polarization is much more pronounced with the DFT level of theory than with the UMP2 level of theory in the FOHI method.

4.3. Comparing Hirshfeld with Other AIM Methods and Experimental Results. As could be concluded from the previous section, the DFT level of theory gives satisfactory results when compared to higher levels of theory. In this section, we will compare FOHI with other AIM methods for a set of larger systems for which we will use the UB3LYP functional and 6-311++G** basis set. The first molecules we compared are based on a phenalenyl system[35] and are stable radicals with a highly delocalized spin density. The structures of both molecules can be found in Figure 2. The results for the different AIM methods are summarized in Table 3. Although all three methods seem to account for the spin polarization, the negative spin density is more pronounced in the Mulliken and FOHI method, while the Bader analysis shows rather small values of negative spin density.

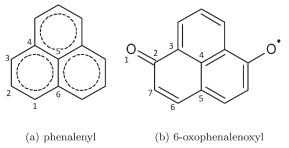


Figure 2. Lewis structures of phenalenyl systems.

Table 3. Comparison of Atomic Spin Populations of the Phenalenyl Structures, TCNE and Nitph, Based on Different AIM Methods Using the UB3LYP Functional and 6-311++G Basis Set**[a]

molecule	atom	Mulliken	Bader	FOHI	experiment
phenalenyl	$C^{(1)}$	0.312	0.224	0.338	
	$C^{(2)}$	−0.137	−0.070	−0.169	
	$C^{(3)}$	0.312	0.225	0.338	
	$C^{(4)}$	−0.142	−0.060	−0.159	
	$C^{(5)}$	0.062	0.032	0.085	
	$C^{(6)}$	−0.142	−0.060	−0.159	
6-oxophenalenoxyl	$C^{(1)}$	0.213	0.176	0.213	
	$C^{(2)}$	−0.070	0.003	−0.086	
	$C^{(3)}$	0.083	0.061	0.109	
	$C^{(4)}$	−0.089	−0.041	−0.118	
	$C^{(5)}$	0.399	0.281	0.421	
	$C^{(6)}$	−0.186	−0.093	−0.211	
	$C^{(7)}$	0.335	0.245	0.365	
TCNE	$N^{(1)}$	0.165	0.117	0.168	0.12
	$C^{(2)}$(sp)	−0.107	0.008	−0.081	−0.05
	$C^{(3)}$(sp^2)	0.385	0.252	0.326	0.33
	$C^{(4)}$(sp)	−0.107	0.008	−0.081	−0.03
	$N^{(5)}$	0.165	0.117	0.168	0.13
	$C^{(6)}$(sp^2)	0.385	0.252	0.326	0.33
	$C^{(7)}$(sp)	−0.107	0.008	−0.081	−0.04
	$N^{(8)}$	0.165	0.117	0.168	0.12
	$C^{(9)}$(sp)	−0.107	0.008	−0.081	−0.08
	$N^{(10)}$	0.165	0.117	0.168	0.16
NitPh	$O^{(1)}$	0.357	0.327	0.320	0.277
	$N^{(2)}$	0.263	0.219	0.315	0.278
	$C^{(3)}$	−0.169	−0.065	−0.245	−0.121
	$N^{(4)}$	0.263	0.219	0.314	0.278
	$O^{(5)}$	0.357	0.327	0.320	0.247
	$C^{(6)}$	−0.009	−0.032	0.063	0.024
	$C^{(7)}$	−0.025	0.013	−0.053	0.000
	$C^{(8)}$	0.022	0.000	0.029	0.025
	$C^{(9)}$	−0.045	0.000	−0.050	−0.016
	$C^{(10)}$	0.022	0.000	0.028	0.011
	$C^{(11)}$	−0.025	−0.032	−0.053	−0.037

[a] Experimental values are listed for the TCNE and NitPh radicals.

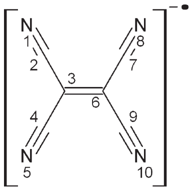


Figure 3. Lewis structure of TCNE.

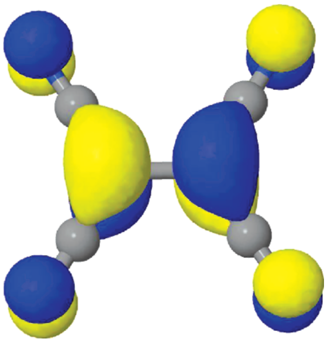


Figure 4. Singly occupied molecular orbital (SOMO) of the radical anion TCNE. The SOMO shows nodes on the sp carbon atoms.

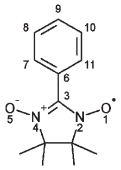


Figure 5. Lewis structure of NitPh.

The spin populations obtained using different partitioning methods were also compared with experimental results obtained with polarized neutron diffraction. This technique is able to quantify the spin density on every atom, which gives the unique opportunity to compare values of atoms-in-molecules, which are generally unobservable, with experimental data. Two molecules were examined, the first one being tetracyanoethylene (TCNE). The structure of this radical anion can be found in Figure 3. The atomic spin populations based on the different partitioning methods and the experimental results can be found in Table 3. Experimental results[18] show that the density is mostly localized on the sp^2 carbon atoms. However, due to spin delocalization and spin polarization, there is a considerable amount of spin density on the nitrogen atom and a negative spin density on the sp carbon atoms. The unequal experimental values on the sp carbon atoms and the nitrogen atoms are due to the reduced symmetry of the monoclinic unit cell.[18] The negative spin density is in agreement with the theory, since the SOMO of TCNE clearly shows nodes at the sp carbon atoms (Figure 4). The negative values are also reproduced by the FOHI method and the Mulliken method but not by the Bader method.

The second molecule examined is the nitronyl nitroxide 2-phenyl-4,4,5,5-tetramethyl-4,5-dihydro-1H-imidazole-1-oxyl-3-oxide (NitPh), the structure of which can be found in Figure 5. Comparison of the values obtained by means of the partitioning methods and the experimental results can be found in Table 3. Experimental results[36] show that most of the spin density is concentrated on the NO groups, where it seems to be equally shared between the nitrogen and oxygen atoms. A large negative spin population is observed on the carbon atom connecting the two NO groups. Also, here, these results are in agreement with the theory, as can be seen from the visualization of the SOMO orbital of NitPh in Figure 6. The spin density in the phenyl ring is very low, although alternation of the spin on the carbon atoms is observed. This has also been confirmed by ^1H and ^{13}C NMR

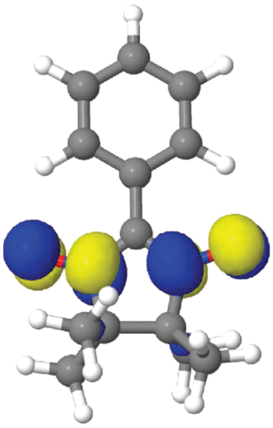


Figure 6. Singly occupied molecular orbital (SOMO) of the nitronyl nitroxide NitPh. The SOMO shows a node on the carbon connecting the two NO groups.

experiments.[37,38] The almost equal distribution of the spin density between the nitrogen and oxygen atoms, observed in the experimental results, is reproduced in the results obtained by the FOHI method and the Mulliken method but not in the Bader method, where the oxygen has higher spin populations. The alternation of the spin on the phenyl group is only observed in the FOHI method. Although the Bader method seems to underestimate the negative spin density on the carbon atom connecting the two NO groups (C_3), both of the other methods overestimate this negative spin density.

5. CONCLUSION

A new partitioning method, inspired by the iterative procedure in Hirshfeld-I (HI),[11] was introduced in order to calculate properties of open-shell systems, in particular, atomic spin populations. The iterative procedure has been expanded so that both the spin and the charge of the atoms are altered during the iterations. This is achieved by performing in each iteration an SCF calculation for each of the atoms using fractional occupation numbers. Because the atomic SCF calculations have to be repeated every iteration for every atom, this can become time-consuming for larger systems. However, the execution time can be reduced drastically, first, by using converged densities of a previous iteration as an initial guess together with fully exploiting the spherical symmetry of the atoms and eventually skipping atoms for which the results are already converged, and second, since the atomic calculations are completely independent, by processing these calculations in parallel, the CPU time can be drastically decreased.

The Hirshfeld-I method is compared with the new fractional occupation Hirshfeld-I method (FOHI). It is found that for properties where only the sum of the α and β density is of importance, such as charges, both methods lead to similar results. Whereas for properties where the difference between the α and β density is of importance, such as spin populations, the FOHI method appears superior. In particular, the ability to reproduce negative spin populations, observed in experimentation, in a consistent manner is a major advantage of the FOHI method. This is due to the fact that the HI method does not make use of the molecular spin density during its iterations and the spin of the atomic densities is arbitrary chosen.

The results obtained by different AIM methods are compared with experimental results obtained with polarized neutron diffraction. The Bader analysis seems to underestimate the spin polarization effect. Therefore, certain atoms have a positive spin density, although experimental results clearly show a negative spin density due to spin polarization. The Mulliken method accounts for the spin polarization, but the results are strongly basis set dependent. The FOHI method is in agreement with the experimental results, but the effect of spin polarization may be overestimated.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: kris.vanalsenoy@ua.ac.be.

■ ACKNOWLEDGMENT

The authors are thankful to the University of Antwerp for access to the university's CalcUA Supercomputer Cluster. Financial support by the University of Antwerp under Grant GOA-2404 is gratefully acknowledged. A.K. acknowledges Research Foundation—Flanders (FWO) for a postdoctoral position.

■ REFERENCES

(1) Smith, B. S.; March, J. *March's Advanced Organic Chemistry: Reactions, Mechanisms and Structure*, 6th ed.; Wiley-Interscience: New York, 2001; pp 16−25.

(2) Mulliken, R. S. *J. Chem. Phys.* **1962**, *36*, 3428.

(3) Jensen, F. *Introduction to Computational Chemistry*; Wiley: Chichester, U. K., 1999; pp 293−296.

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

(5) Reed, A. E.; Weinhold, F.; Curtiss, L. A. *Chem. Rev.* **1988**, *88*, 899.

(6) Bohmann, J. A.; Weinhold, F.; Farrar, T. C. *J. Chem. Phys.* **1997**, *107*, 1173.

(7) Bader, R. F. W. *Chem. Rev.* **1991**, *91*, 893.

(8) Becke, A. D. *J. Chem. Phys.* **1988**, *88*, 2547.

(9) Fonseca Guerra, C.; Handgraaf, J.; Baerends, E. J.; Bickelhaupt, F. M. *J. Comput. Chem.* **2003**, *25*, 2.

(10) Hirshfeld, F. L. *Theor. Chem. Acta* **1977**, *44*, 129.

(11) Bultinck, P.; Van Alsenoy, C.; Ayers, P. W.; Carbo-Dorca, R. *J. Chem. Phys.* **2007**, *126*, 144111.

(12) Nalewajski, R. F.; Broniatowska, E. *Int. J. Quantum Chem.* **2005**, *101*, 349.

(13) Krishtal, A.; Senet, P.; Van Alsenoy, C. *J. Chem. Theory Comput.* **2008**, *4*, 2122.

(14) Krishtal, A.; Senet, P.; Van Alsenoy, C. *J. Chem. Phys.* **2010**, *133*, 154310.

(15) Van Damme, S.; Bultinck, P.; Fias, S. *J. Chem. Theory Comput.* **2009**, *5*, 334.

(16) Krishtal, A.; Vannomeslaeghe, K.; Olasz, A.; Veszpremi, T.; Van Alsenoy, C.; Geerlings, P. *J. Chem. Phys.* **2009**, *130*, 174101.

(17) Steinmann, S. N.; Corminboeuf, C. *J. Chem. Theory Comput.* **2010**, *6*, 1990.

(18) Ressouche, E.; Schweizer, J. *Monatsch. Chem.* **2003**, *134*, 235.

(19) Harrison, J. F. *J. Chem. Phys.* **2009**, *131*, 044117.

(20) Parr, R. G.; Ayers, P. W.; Nalewajski, R. F. *J. Phys. Chem. A* **2005**, *109*, 3957.

(21) Leenaerts, O.; Partoens, B.; Peeters, M. *Appl. Phys. Lett.* **2008**, *92*, 243125.

(22) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.;

Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, Revision B.05; Gaussian, Inc.: Wallingford, CT, 2004.

(23) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648. Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785. Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200. Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.

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

(25) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007.

(26) Ditchfield, R.; Hehre, W. J.; Pople, J. A. *J. Chem. Phys.* **1971**, *54*, 724.

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

(28) Henkelman, G.; Arnaldsson, A.; Jónsson, H. *Comput. Mater. Sci.* **2006**, *36*, 254.

(29) Van Alsenoy, C.; Peeters, A. *THEOCHEM* **1993**, *286*, 19.

(30) Rousseau, B.; Peeters, A.; Van Alsenoy, C. *Chem. Phys. Lett.* **2000**, *324*, 189.

(31) Ayers, P. W. *J. Math. Chem.* **2006**, *43*, 1.

(32) Bultinck, P.; Ayers, P. W.; Fias, S.; Tiels, K.; Van Alsenoy, C. *Chem. Phys. Lett.* **2007**, *444*, 205.

(33) Lim, M. H.; Worthington, S. E.; Dulles, F. J; Cramer, C. J.; Density Functional Calculations of Radicals and Diradicals In *Chemical Applications of Density Functional Theory*; ACS Symposium Series, American Chemical Society: Washington, DC, 1996.

(34) Öhrström, L. *C. R. Chim.* **2005**, *8*, 1374.

(35) Morita, Y.; Nishida, S.; Kawai, J.; Takui, T.; Nakasuji, K. *Pure Appl. Chem.* **2008**, *3*, 507.

(36) Zheludev, A.; Barone, V.; Bonnet, M.; Delley, B.; Grand, A.; Ressouche, E.; Rey, P.; Subra, R.; Schweizer, J. *J. Am. Chem. Soc.* **1994**, *116*, 2019.

(37) Davis, M. S.; Morokuma, K.; Kreilick, R. W. *J. Am. Chem. Soc.* **1972**, *94*, 5588.

(38) Neely, J. W.; Hatch, G. F.; Kreilick, R. W. *J. Am. Chem. Soc.* **1974**, *96*, 652.

JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Molecules-in-Molecules: An Extrapolated Fragment-Based Approach for Accurate Calculations on Large Molecules and Materials

Nicholas J. Mayhall and Krishnan Raghavachari*

Department of Chemistry, Indiana University, Bloomington, Indiana 47405, United States

ABSTRACT: We present a new extrapolated fragment-based approach, termed molecules-in-molecules (MIM), for accurate energy calculations on large molecules. In this method, we use a multilevel partitioning approach coupled with electronic structure studies at multiple levels of theory to provide a hierarchical strategy for systematically improving the computed results. In particular, we use a generalized hybrid energy expression, similar in spirit to that in the popular ONIOM methodology, that can be combined easily with any fragmentation procedure. In the current work, we explore a MIM scheme which first partitions a molecule into nonoverlapping fragments and then recombines the interacting fragments to form overlapping subsystems. By including all interactions with a cheaper level of theory, the MIM approach is shown to significantly reduce the errors arising from a single level fragmentation procedure. We report the implementation of energies and gradients and the initial assessment of the MIM method using both biological and materials systems as test cases.

I. INTRODUCTION

Although electronic structure theory can now routinely yield chemically accurate results for small molecules,[1−17] the steep computational scaling of the most accurate methods precludes their direct application to large molecular systems. For sufficiently local chemical processes, hybrid-energy methods[18−20] provide a strategy for using accurate (and expensive) computational methods on only those parts of the molecule that comprise the chemically active region. Morokuma and co-workers have developed a particularly useful hybrid energy method called ONIOM (our own N-layer integrated molecular orbital molecular mechanics),[21−27] in which the hybrid energy for a system divided into two regions (I and II) is given as

$$E_{\mathrm{ONIOM}} = E(\mathrm{I} + \mathrm{II})_{\text{Low-Level}} + E(\mathrm{I})_{\text{High-Level}} - E(\mathrm{I})_{\text{Low-Level}} \tag{1}$$

Although hybrid methods such as QM/MM (quantum mechanics/molecular mechanics) and ONIOM have been incredibly useful in a variety of situations, the necessity to identify a chemically active region prevents them from being truly black-box methods. Moreover, it is clearly important to develop theoretical methods that call for a uniformly accurate treatment of the entire molecule without any bias.

To develop a strategy for performing high level of theory computations on large systems, one must also overcome the $O(N^5)-O(N^8)$ computational scaling of the most accurate electronic structure methods. In this direction, several researchers have focused their efforts on developing linear scaling methods which exploit the local nature of a molecule's electronic structure in a single calculation.[28−52] In the alternative fragmentation-based approach, significant scaling reductions can be achieved by dividing a molecule into smaller fragments, performing electronic structure calculations on each independently and then summing up the results appropriately. An additional, and possibly equally important, benefit of the fragmentation approach is the inherent ease in constructing a massively parallel implementation.

One of the earliest fragment-based techniques developed is the fragment molecular orbital method (FMO) of Kitaura and co-workers, in which the fragment subsystems are embedded in a Coulomb bath.[53−59] Gadre and co-workers have developed the cardinality-guided molecular tailoring approach (CG-MTA) for energies, gradients, and properties.[60−66] Other fragmentation methods include the molecular fractionation with conjugate caps (MFCC) method of Zhang and co-workers,[67,68] the systematic molecular fragmentation (SMF) method of Collins and co-workers,[69−73] which was later coupled with the effective fragment potential (EFP) of Gordon and co-workers,[74−76] Truhlar et al.'s electrostatically embedded many-body (EE-MB) method,[77−82] the generalized energy-based fragmentation (GEBF) method of Li and co-workers,[83−85] the kernel energy method (KEM) of Karle and co-workers,[86,87] the hybrid many-body interaction (HMBI) model of Beran and Nanda,[88,89] and the multilevel fragment-based approach (MFBA) of Řezàč and Salahub.[90]

The various existing fragment-based methods use a variety of different approaches to implement and to perform the computations. However, it is convenient to discuss them in terms of three main components common to most fragmentation methods: (1) partitioning, (2) subsystem formation, and (3) energy summation.

A. Partitioning. First, the system (large molecule) is partitioned into defined subunits or fragments, most often by cleaving single bonds between adjacent non-hydrogen atoms. This would, as an example, divide hexane into six fragments: 2($-CH_3$) and 4($-CH_2-$). Less general but possibly more useful fragmentation schemes may be employed if one is well versed in the chemical composition of the target system. For example, as the

Received: January 17, 2011
Published: April 05, 2011

MFCC and MFBA methods have been developed for modeling polypeptide systems, the fragmentation scheme in these two methods involves only the cleavage of specific bonds in the peptide backbone. Since the fragments typically have dangling bonds, direct electronic structure calculations on the fragments are not possible without further manipulation in most cases. The fragments are analogous to "atoms" and make up the fundamental building blocks of the system.

B. Subsystem Formation. After partitioning the molecule into fragments, the interacting fragments are combined to form subsystems that can be used in individual electronic structure calculations. The subsystems must be constructed in a way which provides the appropriate balance between accuracy and computational efficiency. A key step in subsystem formation involves the capping of truncated bonds to form well-defined units (i.e., small molecules) on which direct electronic structure calculations can be performed. Most methods (i.e., MFCC, MFBA, GEBF) use hydrogen atoms to cap the dangling bonds, whereas the FMO methods use a more sophisticated potential to satisfy valencies. For the formation of subsystems involving a systematic inclusion of interfragment interactions, many methods (FMO, EE-MB, MFCC, MFBA, SMF, KEM, HMBI) employ the many-body expansion described by Xantheas.[91] This provides a simple way to systematically improve one's results, albeit at greatly increasing computational cost for high expansion orders. Linear scaling can be recovered by including only the *n*-body interactions which fall within a specified interfragment distance. Typically, however, most such calculations include only two-body (and three-body, if possible) interactions.

In an alternative but equally powerful approach, one may form overlapping subsystems comprised of a central fragment surrounded by proximal fragments (either in connectivity or spatial distance). The overlap denotes regions where the interfragment interactions are overcounted and must be appropriately subtracted out in accordance with the Inclusion−Exclusion Principle:

$$|A_1 \cdots \cup A_n| = \sum_i |A_i| - \sum_{i<j} |A_i \cap A_j| + \sum_{i<j<k} |A_i \cap A_j \cap A_k| \ldots + (-1)^{n-1} |A_i \cap \ldots \cap A_n| \tag{2}$$

This is the approach taken by the GEBF and CG-MTA methods.

C. Energy Summation. The final part involves a summation of the energies from the individual subsystems (small molecules) to yield the total energy for the entire system (large molecule). The overall energy expression is clearly dependent on the manner in which the individual subsystems are formed. Methods which use a many-body expansion are then summed according to the form of the expansion. The energy of methods which use overlapping subsystems must be assembled more carefully, as the energy for each subsystem must be summed along with its appropriate coefficient. This will be discussed in more detail in the Methodology section.

In this contribution, we report the development and initial assessment of a new hybrid energy method which we have termed molecules-in-molecules (MIM). In this method, we use multiple levels of theory to extrapolate the fragment-based energy to obtain better convergence of the total energy of the large molecule. Our method is similar in spirit to the popular ONIOM methodology (*vide infra*) and allows one to couple very accurate electronic structure methods, performed at a modest level of fragmentation (small subsystems), with cheaper methods performed at more aggressive levels of fragmentation (larger subsystems). We use features from many of the existing approaches and generalize other features.

II. METHODOLOGY

Most previous treatments consider a single level of partitioning to create the fragments and treat the resulting subsystems (interacting fragments including capping) with one or more levels of theory. The central point of the MIM method is that it is a multilevel fragment/subsystem approach, in which cheaper ab initio or semiempirical methods can be used to describe longer range interactions. In the most straightforward approach, after the initial partitioning, subsystems of different sizes can be generated by using different cutoff distances to describe the interactions between fragments. Alternatively, different partitioning schemes can be used to create fragments of different sizes. As an example, in a three-level MIM scheme (MIM3), the entire molecule can be treated at a low level of theory (e.g., HF), medium subsystems can be treated at a medium level of theory (e.g., MP2), and small subsystems can be treated at a high level of theory (e.g., CCSD). For a biomolecule containing hundreds of atoms, in a two-level MIM scheme (MIM2), B3LYP can be used to treat the subsystems while the entire molecule can be treated with PM6 to provide a correction for long-range interactions. This is very similar in spirit to the popular ONIOM approach by Morokuma, though the subsystems are not centered on the active site (as in ONIOM) but are individually centered throughout the large molecule (in MIM). Thus, the energy expressions can also be written in a manner similar to that in the ONIOM approach. However, unlike in ONIOM we are approximating the high level of theory on the whole molecule. Therefore, the error associated with the fragmentation method is

$$\text{Error} = E^r_{\text{High}} - E_{\text{High}} \tag{3}$$

In our method, we extrapolate the fragmentation energy by approximating the error of the fragmentation method with a more efficient level of theory as shown here:

$$\begin{aligned} E_{\text{High}} &= E^r_{High} - (E^r_{\text{High}} - E_{\text{High}}) \\ &\approx E^r_{\text{High}} - (E^r_{\text{Low}} - E_{\text{Low}}) \end{aligned} \tag{4}$$

Use of the approximation shown in eq 4 sets up a general hierarchy for an arbitrary number of extrapolations.

$$E^{\text{MIM1}} = E^r_{\text{High}} \tag{5}$$

$$E^{\text{MIM2}} = E^r_{\text{High}} - (E^r_{\text{Low}} - E^{\infty}_{\text{Low}}) \tag{6}$$

$$E^{\text{MIM3}} = E^r_{\text{High}} - (E^r_{\text{Med}} - E^{r'}_{\text{Med}}) - (E^{r'}_{\text{Low}} - E^{\infty}_{\text{Low}}) \tag{7}$$

where the superscript *r* indicates the accuracy threshold parameter used in that fragmentation level (i.e., distance/number cutoff or the order of a many-body expansion), and the Low, Med, and High refer to the low level, medium level, and high level of theory, respectively. It should be noted, however, that E^r_{High} and E^r_{Low} each represent *a composite energy for the entire system assembled from a number of subsystem calculations* at the fragmentation level denoted by the parameter r. Since the energy of the entire system (large molecule) is assembled from the individual energies of the subsystems (small molecules), we label our method as "molecules in molecules" (MIM).

In eq 7, the MIM3 expression now has two defined parameters, both r and r', where $r > r'$. The method can also be generalized to more than three levels (multilevel MIM) in more complex systems in an analogous manner. Although we have used the full system calculation in the above equations, E^{∞}_{Low}, we do not necessarily have to extrapolate to the full calculation, which may be intractable for very large systems. We can alternatively decide to extrapolate only to a more accurate method of fragmentation, such as in

$$E^{\text{MIM2}} = E^{r}_{\text{High}} - (E^{r}_{\text{Low}} - E^{r'}_{\text{Low}}) \qquad (8)$$

where $r' \gg r$. However, in this paper, we will only be considering the case in which $r' = \infty$.

The energy expressions given in eqs 5−7 are generic, and independent of the actual procedure used for generating the subsystems. Therefore, it may be used with any fragmentation procedure such as the GEBF or the FMO methods, allowing one to couple various fragmentation methods together. Furthermore, as the MIM approach uses an ONIOM framework, it exhibits several attractive features: (1) When the two (or three) levels of theory become identical, the exact energy (E^{∞}_{High}) is recovered. (2) When the distance parameter r becomes sufficiently large, the exact energy (E^{∞}_{High}) is recovered. (3) As all individual subcalculations are performed on well-defined molecular systems, any electronic structure theory method may be used, which allows this approach to take advantage of the most recent advances in semiempirical methods or electron correlation methods. (4) Finally, an efficient parallel implementation is easy since the individual subsystem calculations can be carried out on different processors. The basic methodology of the MIM method is displayed in Figure 1 with a comparison to the ONIOM methodology.

In the initial implementation of the MIM method reported in this paper, we use a fragmentation scheme similar to the GEBF method.[85] In the GEBF approach, all important interactions must be included in the subsystem calculations. However, since we are using a multilevel approach, we should be able to relax our conditions for assembling subsystems to permit smaller calculations, while picking up the additional long-range interactions with the cheaper methods.

For a MIM calculation to be performed, each level of calculation is carried out in the following manner:

1 *Defining Fragments.* To form the initial fragments, a connectivity table is processed, and all single bonds between heavy atoms are cleaved. As noted earlier, fragments are the most fundamental units considered and are never broken up in the subcalculations. Therefore, a manual modification of the connectivity table provides one with a simple mechanism for controlling the composition of the subsystems generated by the automated program. Customized fragment lists may also be used for generating the subsystems in the next step.

2 *Primary Subsystem Formation.* Subsystems which are centered on a particular fragment with nearby fragments appended are called *primary subsystems*.[108] Primary subsystems may be assembled either by appending all fragments which are within some distance cutoff r (distance-based cutoff) or by appending a given number of fragments η which are closest to the central fragment (number-based cutoff).[85] We assemble the primary subsystems according to the following pseudocode:

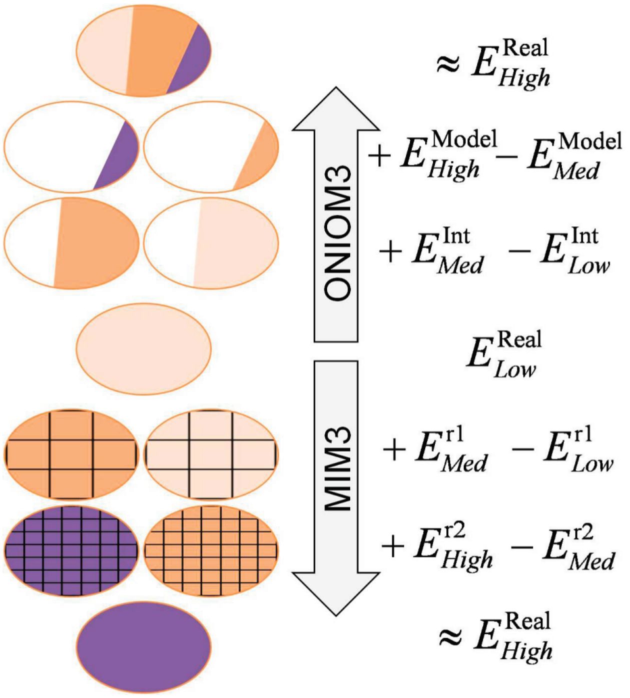

≈ E^{Real}_{High}
ONIOM3
$+ E^{Model}_{High} - E^{Model}_{Med}$
$+ E^{Int}_{Med} - E^{Int}_{Low}$
E^{Real}_{Low}
MIM3
$+ E^{r1}_{Med} - E^{r1}_{Low}$
$+ E^{r2}_{High} - E^{r2}_{Med}$
≈ E^{Real}_{High}

Figure 1. Schematic representation of the MIM methodology. The faint orange areas represent regions computed with the low level of theory. The strong orange areas represent regions computed with the medium level of theory. The purple areas represent regions computed with the high level of theory.

```
1 Define $r$, for distance-based cutoff, or $\eta$, for number based cutoff.
2 For each fragment (f)
    a Initiate a new primary subsystem (p).
    b If distance-based cutoff:
        - Append all fragments to the primary subsystem that are less than distance $r$ away from the central fragment (f)
    c If number-based cutoff:
        - Append $\eta - 1$ fragments to p which are nearest to the central fragment (f). p now contains $\eta$ fragments.
        - In case of systems with high symmetry, include any additional fragments if they are at the same distance as the farthest included fragment.
    d Are there any atoms (a) not in p which are also covalently connected to two or more atoms in p?
        - If yes:
            -- Append a to p.
            -- Go to step d again.
3 Reduce list of primary subsystems to only unique primary subsystems, which are not subsets of other primary subsystems. Here, step d is done to ensure that the same center is not replaced twice by link atoms, as would be the case when only five atoms of a six-membered ring are part of a primary subsystem.
```

3 *Derivative Subsystem Formation*. A consequence of using overlapping primary subsystems is that much of the molecule is overcounted. The primary subsystems overlap with one another, and we must therefore construct subsystems which cancel the overlapping regions of the primary subsystems. These are called *derivative subsystems*[85] and are generated in accordance with eq 2. This can become a time-consuming process for large values of r, though our algorithm as implemented above is sufficient to get quite accurate results for reasonably large molecules (*vide infra*). In the future, we plan to investigate other methods of subsystem determination such as the algorithm put forth by Bettens and Lee.[92]

While we have used a GEBF-like fragmentation procedure in this initial MIM implementation, we have major differences. The first key difference between the current fragmentation scheme and the GEBF method is that we do not use the "extension rules" employed in constructing the GEBF primary subsystems, as our aim is to use a cheaper electronic structure method for the longer range interactions. This permits us to get accurate results employing significantly smaller high-level subsystem calculations (*vide infra*). Since we are currently not including any high-level electronic embedding in our calculations, our energy gradients are exact.[85] However, we are working on implementing electronic embedding in a formalism similar to our earlier ONIOM-EE work,[93-95] in which all of the small subsystem calculations are performed in the presence of the charge distribution obtained in the $E_{\mathrm{Low}}^{\infty}$ level in a MIM calculation. Electronic embedding done this way will remove the necessity to iteratively obtain charge distributions.

In essence, the MIM method provides a general hierarchy for coupling different electronic structure methods to achieve accurate energies and properties for large molecules. This is, to the best of our knowledge, the first *fully general, extrapolated ONIOM-like methodology for a multilevel fragmentation energy approach* (*vide infra*). The method allows transparent coupling of many different methods including semiempirical schemes, easy implementation of analytical gradients for the exploration of potential energy surfaces or dynamics, and efficient parallelization across many platforms including massively parallel architectures. An additional advantage is that our energy extrapolation procedure is generic and is easily coupled with many possible fragmentation schemes. We have already implemented schemes based on bond-space cutoff, real space (distance-based) cutoff, and number-based cutoff treatments. Moreover, the seamless inclusion of link atoms (when needed) makes it possible to treat bonded systems (such as peptides) or nonbonded systems (such as water clusters) on an equal footing. Another major advantage of our methodology is that multilevel MIM calculations do *include all interactions*, albeit, at the low-level of theory. Finally, unlike most previous schemes, our use of the ONIOM-like extrapolation scheme results in our method having the formal property that it yields exact energies when the different levels of theory become identical.

Of the many previous fragment-based approaches developed by other groups, several share some of the advantageous details featured here. The ability to obtain system—system interactions by using overlapping subsystems has been demonstrated previously by Li et al. (GEBF[83-85]) and Gadre et al. (CG-MTA[60-66]). However, both of these methods have thus far used only a single level of fragmentation that attempts to incorporate all significant interactions into the fragmented energy calculations, e.g., via electrostatic embedding of the subsystems. As an alternative, the MIM approach uses a lower level of theory to compute the energy of the full molecule (or simply a less fragmented system) which does include all interactions. Several previous methods (MC-QM:QM,[96-100] HMBI,[88,89] EE-MB-CE,[78] MFBA,[90]) make use of multiple levels of theory, as we do in this work. Typically, many-body energy expansions are used in these approaches, where the lower levels of theory are employed for evaluations of the higher order many-body terms. However, the MIM method is defined in a general and flexible manner, and thus it can couple both overlapping subsystem approaches and many-body expansion approaches (though the latter has not yet been fully implemented in our code). The more recent XO method[101] not only uses overlapping subsystems for describing fragment-fragment interactions but also couples multiple levels of theory, as we do in our current approach. However, the XO method has not yet included a systematic and programmable algorithm for the fragmentation/subsystem assembly procedure as we have done in this contribution. Overall, while MIM is clearly related to many previous fragment-based methods, its broad definition and generality make it an accurate and applicable approach for investigating a wide range of problems for both bonded and nonbonded systems.

In our implementation, Gaussian 09 is used for both the electronic structure calculations and the geometry optimizer.[102]

III. ASSESSMENT OF MIM

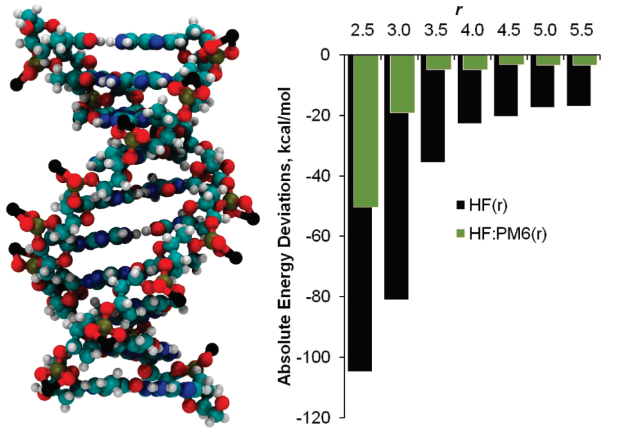
A. Absolute Energies—DNA. For the MIM method to provide accurate results, the low level of theory must be capable of recovering the long-range interactions lost in the high-level subsystem calculations. In this section, we couple a semiempirical method (PM6) with an ab initio method (HF/6-31G) to compute the MIM energy of a large molecular system to test the energy convergence with increasing r both with and without the PM6 level. This allows us to gauge the ability to capture long-range effects with only a low-level of theory. As a test molecule, we have chosen the DNA poly(dA·dT) decamer taken from the nucleic acid database,[103] a system considered previously with the GEBF method.[85] However, as we are currently not including any solvation effects, the presence of 18 excess electrons may make the results difficult to interpret. We therefore cap each of the phosphate groups with a sodium ion to neutralize the highly anionic systems as we have done in a previous study.[104] Although electron correlation effects are known to be important in the description of $\pi-\pi$ interactions, we have used the HF method to permit comparison to previously reported results on a similar system.[85] The modified system is shown in Figure 2a.

Figure 2b illustrates the deviations in absolute energy between the unfragmented HF/6-31G results and two fragmented schemes: HF/6-31G(r) (black columns) and HF/6-31G:PM6-(r) (green columns). The benefits of including the PM6 correction for long-range interactions are immediately seen. The PM6 correction greatly reduces the deviations at all levels of fragmentation.

B. Relative Energies—2NPV. While reproducing the absolute energies of the full unfragmented method is clearly a sufficient condition for accuracy, it is not necessary, as quantum chemistry is largely concerned with energy differences. Therefore, while systematic errors in a fragmentation method may prevent rapid convergence of the absolute energies, these errors may cancel when computing relative energies.

Here, we assess the ability of the MIM method to reproduce the relative conformer energies of a cyclic lipopeptide surfactin (PDB entry 2NPV).[105] We have chosen 2NPV as a test molecule based both on its small size (which allows full, unfragmented calculations to be performed) and on the wide range of conformations sampled by the fatty acid group (as shown in Figure 3a). This molecule has a well-defined, saddle-shaped polypeptide backbone with a long alkyl chain extension. In this example, we use the first 10 conformers given in the PDB file.

Using both distance-based cutoffs (r) and number-based cutoffs (η) for the primary subsystem assembly, we calculated the energy of the 10 conformers with both 1-level [MP2/6-31G*(r or η)] and 2-level [MP2/6-31G*:B3LYP/6-31G*(r or η)] fragmented methods and compared the results to the full, unfragmented MP2/6-31G* (1262 basis functions) relative energies. This allows us to examine the effect of the low-level correction on relative energies. For the distance-based cutoff, we use an r value of 3.0 Å,[109] and for the number-based cutoff, we use a value of η = 9. As can be seen in Figure 3c, the low-level (B3LYP/6-31G*) correction greatly improves the results by closely reproducing the unfragmented MP2/6-31G* relative energies. In fact, for all but conformer 4, there is essentially no discernible difference between the MP2:B3LYP(η = 9) results (dashed line) and the MP2 results (solid line), despite significant deviations for the single level fragmented MP2(η = 9) results (dotted line).

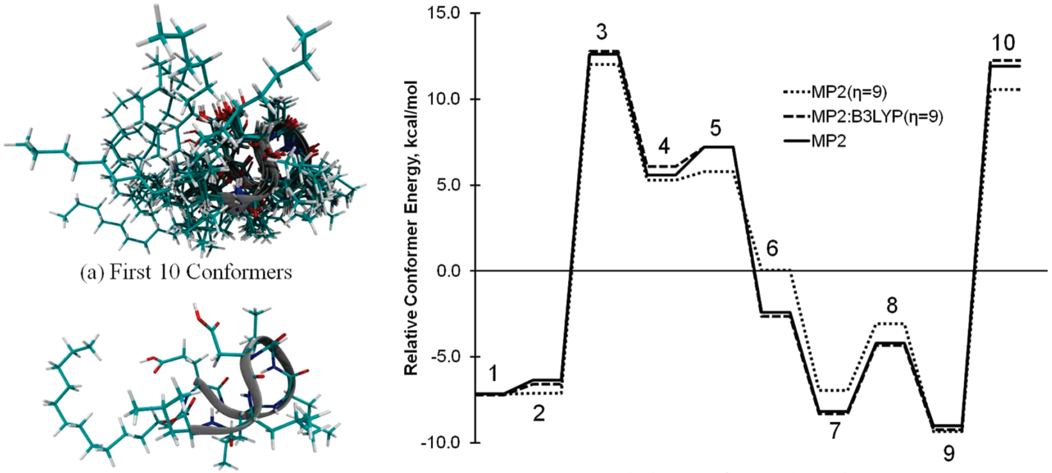


(a) DNA (b) Convergence of r

Figure 2. DNA: (a) DNA decamer used in this study. Na^+ ions are shown as black centers. (b) Convergence of absolute energy with r. Deviations between unfragmented and fragmented calculations reported as HF(r)−HF or HF:PM6(r)−HF.

The results, summarized in Table 1, show that while the low-level correction does indeed decrease both the RMS and maximum absolute errors, the relative energies are improved more dramatically. A careful analysis of the distance- and number-based cutoff results illustrates the differences in using the different procedures of subsystem assembly. We can make a preliminary observation from this study that the η = 9 results appear to be more accurate and have smaller individual MP2 calculations while the r = 3.0 results require fewer subsystem calculations. These results are clearly impressive and underscore the extent of error cancellation in the fragmented methods. While the errors in absolute energies are still rather large (RMS = 3.27: max = 3.9 kcal/mol for MP2:B3LYP(r = 3.0) and RMS = 2.44: max = 2.94 kcal/mol for MP2:B3LYP(η = 9)), the errors in relative energies are much smaller; both the RMS and maximum errors (for both r and η) are less than 1 kcal/mol. Our best results are obtained for the number-based cutoff method where the RMS deviation in relative energies falls from 1.18 kcal/mol to an impressive 0.25 kcal/mol on going from MIM1 to MIM2.

C. Geometry Optimizations—Si(100)-2×1. The theoretical study of semiconductor surface chemistry has benefited greatly from the development of efficient and accurate cluster models.[106] However, the computational demands for studying processes such as a chemical line growth across a surface are greatly increased due to the large cluster models required to capture the lateral movement of the adsorbed molecules. Here, we demonstrate the ability to use the MIM method to optimize the geometry of a large silicon cluster with an adsorbed allylic mercaptan molecule.[107] The structure is shown in Figure 4.



(a) First 10 Conformers

(b) Conformer 9

(c) Relative Conformer Energies

Figure 3. 2NPV: (a) Superposition of the first 10 conformers. (b) Structure of conformer 9 given in the PDB file. (c) Relative energies for the first 10 conformers of 2NPV centered at zero. Dotted line: MP2(η = 9) results. Dashed line: MP2:B3LYP(η = 9) results. Solid line: unfragmented MP2 results. All calculations use the 6-31G* basis set. Relative energies for each method are shown centered at zero.

Table 1. 2NPV—Differences between the Unfragmented MP2 Results and the Fragmented Results[a]

	MIM1		MIM2	
RMS(MAX)	MP2(r = 3.0)	MP2(η = 9)	MP2:B3LYP(r = 3.0)	MP2:B3LYP(η = 9)
relative, kcal/mol	1.31(2.51)	1.18(2.44)	0.48(0.90)	0.25(0.51)
absolute, kcal/mol	4.69(7.02)	4.02(6.29)	3.27(3.91)	2.44(2.94)

[a] Both distance-based cutoffs [MP2(r = 3.0) and MP2:B3LYP(r = 3.0)] and number-based cutoffs [MP2(η = 9) and MP2:B3LYP(η = 9)] were used. Energy results in kcal/mol. Values reported are RMS, with MAX values in parentheses. The 6-31G* basis set used in all calculations.

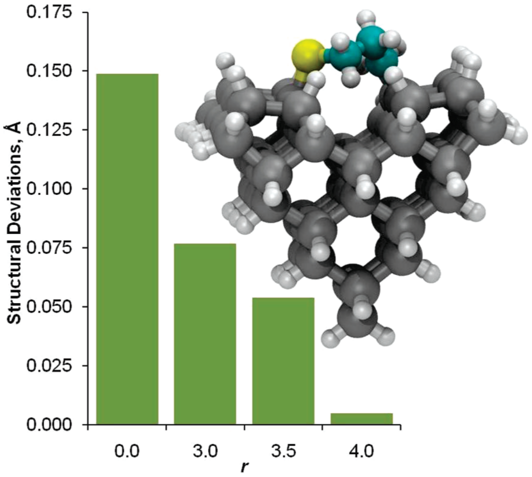


Figure 4. Si(100)—RMS deviations between the B3LYP/6-31G*:AM1(r) optimized geometries and the unfragmented B3LYP/6- 31G* optimized geometries. Units in Å.

As the MIM gradient expressions look just like the energy expressions, we can easily obtain MIM gradients from gradient evaluations of the individual subcalculations. Link atom force projections are performed as outlined for the ONIOM methodology. In Figure 4, we show the convergence of the optimized geometries with increasing r. The columns represent RMS deviations (in Å) between the B3LYP/6-31G*:AM1(r) optimized geometry and the unfragmented B3LYP/6-31G* optimized geometry. Using a value of r = 3.0, the RMS is nearly halved with respect to the pure AM1 results (r = 0.0). Further improvements to the geometry can be made by increasing r, with r = 4.0 giving nearly exact results.

IV. CONCLUSIONS

In this paper, we present a new hybrid energy fragmentation method called molecules-in-molecules (MIM), which treats individual fragmentation calculations as levels in an ONIOM-like hybrid energy framework.

As an initial assessment of the approach, we have studied the convergence of the total energy for a large DNA molecule. By comparing to the single-level fragmentation approach, we observe significant improvements to the total energies upon inclusion of a low-level correction.

By calculating the relative conformer energies of a surfactin molecule, we report improvements to the reproduction of MP2 relative energies, with an RMS deviation from the unfragmented MP2 calculations of only 0.25 kcal/mol.

We have also demonstrated the use of the MIM method for the optimization of molecular geometries for a silicon cluster model system, and our results indicate that accurate molecular geometries can be obtained using the MIM methodology.

Overall, our work suggests that our multilevel MIM approach can be employed for a wide range of large systems with high accuracy.

Finally, a cautionary note on the applicability of MIM is warranted. Central to our method is the assumption that the electronic characteristics are relatively local or near-sighted. Therefore, classes of molecular systems, such as metallic or highly conjugated systems, which are characterized by long-range electron delocalization are not expected to be treatable with the MIM methodology. Further, as we are currently using only hydrogen atoms for capping truncated bonds, our current implementation is only expected to be successful for those systems which are dominated by covalent bonds. Fragmentation of dative or ionic bonds is not recommended. However, we are currently investigating alternative bond-capping approaches (such as the use of pseudoatoms) for various bonding scenarios.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: kraghava@indiana.edu.

■ ACKNOWLEDGMENT

This work was supported by a National Science Foundation grant, CHE-0911454, at Indiana University. The authors would like to thank Professor Balaji Raghavachari for insightful discussions facilitating the development of this work.

■ REFERENCES

(1) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. *J. Chem. Phys.* **2007**, *126*, 084108–084119.

(2) DeYonker, N. J.; Cundari, T. R.; Wilson, A. K. *J. Chem. Phys.* **2006**, *124*, 114104.

(3) Petersson, G. A.; Bennett, A.; Tensfeldt, T. G.; Allaham, M. A.; Shirley, W. A.; Mantzaris, J. *J. Chem. Phys.* **1988**, *89*, 2193–2218.

(4) Ochterski, J. W.; Petersson, G. A.; Montgomery, J. A. *J. Chem. Phys.* **1996**, *104*, 2598–2619.

(5) Karton, A.; Rabinovich, E.; Martin, J. M.; Ruscic, B. *J. Chem. Phys.* **2006**, *125*, 144108.

(6) Tajti, A.; Szalay, P.; Csaszar, A.; Kallay, M.; Gauss, J.; Valeev, E.; Flowers, B.; Vazquez, J.; Stanton, J. *J. Chem. Phys.* **2004**, *121*, 11599.

(7) Bomble, Y. J.; Vazquez, J.; Kallay, M.; Michauk, C.; Szalay, P. G.; Csaszar, A. G.; Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **2006**, *125*, 064108.

(8) Harding, M. E.; Vazquez, J.; Ruscic, B.; Wilson, A. K.; Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **2008**, *128*, 114111.

(9) Kowalski, K.; Piecuch, P. *J. Chem. Phys.* **2000**, *113*, 18–35.

(10) Piecuch, P.; Kucharski, S. A.; Kowalski, K.; Musial, M. *Comput. Phys. Commun.* **2002**, *149*, 71–96.

(11) Martin, J. M. L.; de Oliveira, G. *J. Chem. Phys.* **1999**, *111* 1843–1856.

(12) Boese, A. D.; Oren, M.; Atasoylu, O.; Martin, J. M. L.; Kallay, M.; Gauss, J. *J. Chem. Phys.* **2004**, *120*, 4129–4141.

(13) Karton, A.; Taylor, P. R.; Martin, J. M. L. *J. Chem. Phys.* **2007**, *127*, 064104.

(14) Karton, A.; Parthiban, S.; Martin, J. M. L. *J. Phys. Chem. A* **2009**, *113*, 4802–4816.

(15) Karton, A.; Martin, J. M. L. *J. Chem. Phys.* **2010**, *133*, 144102.

(16) Feller, D.; Dixon, D. A. *J. Phys. Chem. A* **2000**, *104*, 3048–3056.

(17) Barnes, E. C.; Petersson, G. A.; Montgomery, J. A.; Frisch, M. J.; Martin, J. M. L. *J. Chem. Theory Comput.* **2009**, *5*, 2687–2693.

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

(19) Maseras, F.; Morokuma, K. *J. Comput. Chem.* **1995**, *16*, 1170–1179.

(20) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1986**, *7*, 718–730.

(21) Humbel, S.; Sieber, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 1959–1967.

(22) Svensson, M.; Humbel, S.; Froese, R.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357–19363.

(23) Karadakov, P. B.; Morokuma, K. *Chem. Phys. Lett.* **2000**, *317* 589–596.

(24) Vreven, T.; Morokuma, K. *J. Comput. Chem.* **2000**, *21*, 1419–1432.

(25) Vreven, T.; Mennucci, B.; da Silva, C.; Morokuma, K.; Tomasi, J. *J. Chem. Phys.* **2001**, *115*, 62–72.

(26) Vreven, T.; Morokuma, K. *Theor. Chem. Acc.* **2003**, *109*, 125–132.

(27) Rega, N.; Iyengar, S.; Voth, G.; Schlegel, H.; Vreven, T.; Frisch, M. *J. Phys. Chem. B* **2004**, *108*, 4210–4220.

(28) He, X.; Merz, K. M. *J. Chem. Theory Comput.* **2010**, *6*, 405–411.

(29) Kobayashi, M.; Nakai, H. *J. Chem. Phys.* **2009**, *131*, 114108.

(30) Kobayashi, M.; Imamura, Y.; Nakai, H. *J. Chem. Phys.* **2007**, *127*, 074103.

(31) Monard, G.; Bernal-Uruchurtu, M. I.; van der Vaart, A.; Merz, K. M.; Ruiz-Lopez, M. F. *J. Phys. Chem. A* **2005**, *109*, 3425–3432.

(32) der Vaart, A. V.; Gogonea, V.; Dixon, S. L.; Merz, K. M. *J. Comput. Chem.* **2000**, *21*, 1494–1504.

(33) Yang, W. T. *Phys. Rev. Lett.* **1991**, *66*, 1438–1441.

(34) Makowski, M.; Korchowiec, J.; Gu, F. L.; Aoki, Y. *J. Comput. Chem.* **2010**, *31*, 1733–1740.

(35) Imamura, A.; Aoki, Y.; Maekawa, K. *J. Chem. Phys.* **1991**, *95*, 5419.

(36) Saebo, S.; Pulay, P. *J. Chem. Phys.* **1987**, *86*, 914–922.

(37) Maslen, P.; Head-Gordon, M. *Chem. Phys. Lett.* **1998**, *283*, 102–108.

(38) Lee, M. S.; Maslen, P. E.; Head-Gordon, M. *J. Chem. Phys.* **2000**, *112*, 3592–3601.

(39) Subotnik, J. E.; Head-Gordon, M. *J. Chem. Phys.* **2005**, *123* 064108.

(40) Pisani, C.; Maschio, L.; Casassa, S.; Halo, M.; Schtz, M.; Usvyat, D. *J. Comput. Chem.* **2008**, *29*, 2113.

(41) Hampel, C.; Werner, H.-J. *J. Chem. Phys.* **1996**, *104*, 6286–6297.

(42) Adler, T. B.; Werner, H.-J. *J. Chem. Phys.* **2009**, *130*, 241101.

(43) Flocke, N.; Bartlett, R. J. *J. Chem. Phys.* **2004**, *121*, 10935–10944.

(44) Stratmann, R. E.; Scuseria, G. E.; Frisch, M. J. *Chem. Phys. Lett.* **1996**, *257*, 213–223.

(45) Strain, M. C.; Scuseria, G. E.; Frisch, M. J. *Science* **1996**, *271*, 51–53.

(46) Scuseria, G. E.; Ayala, P. Y. *J. Chem. Phys.* **1999**, *111*, 8330–8343.

(47) Scuseria, G. E. *J. Phys. Chem. A* **1999**, *103*, 4782–4790.

(48) Millam, J. M.; Scuseria, G. E. *J. Chem. Phys.* **1997**, *106*, 5569–5577.

(49) Kudin, K. N.; Scuseria, G. E. *Phys. Rev. B* **2000**, *61*, 16440–16453.

(50) Daniels, A. D.; Scuseria, G. E. *J. Chem. Phys.* **1999**, *110*, 1321–1328.

(51) Burant, J. C.; Scuseria, G. E.; Frisch, M. J. *J. Chem. Phys.* **1996**, *105*, 8969–8972.

(52) Ayala, P. Y.; Scuseria, G. E. *J. Chem. Phys.* **1999**, *110*, 3660–3671.

(53) Fujimoto, H.; Koga, N.; Fukui, K. *J. Am. Chem. Soc.* **1981**, *103* 7452–7457.

(54) Fedorov, D. G.; Ishida, T.; Kitaura, K. *J. Phys. Chem. A* **2005**, *109*, 2638–2646.

(55) Fedorov, D. G.; Ishida, T.; Uebayasi, M.; Kitaura, K. *J. Phys. Chem. A* **2007**, *111*, 2722–2732.

(56) Fedorov, D. G.; Kitaura, K. *J. Phys. Chem. A* **2007**, *111*, 6904–6914.

(57) Kitaura, K.; Ikeo, E.; Asada, T.; Nakano, T.; Uebayasi, M. *Chem. Phys. Lett.* **1999**, *313*, 701–706.

(58) Fedorov, D. G.; Slipchenko, L. V.; Kitaura, K. *J. Phys. Chem. A* **2010**, *114*, 8742–8753.

(59) Pruitt, S. R.; Fedorov, D. G.; Kitaura, K.; Gordon, M. S. *J. Chem. Theory Comput.* **2010**, *6*, 1–5.

(60) Gadre, S. R.; Shirsat, R. N.; Limaye, A. C. *J. Phys. Chem.* **1994**, *98*, 9165–9169.

(61) Gadre, S. R.; Ganesh, V. *J. Theor. Comput. Chem.* **2006**, *5* 835–855.

(62) Ganesh, V.; Dongare, R. K.; Balanarayan, P.; Gadre, S. R. *J. Chem. Phys.* **2006**, *125*, 104109.

(63) Kavathekar, R.; Khire, S.; Ganesh, V.; Rahalkar, A. P.; Gadre, S. R. *J. Comput. Chem.* **2009**, *30*, 1167–1173.

(64) Rahalkar, A. P.; Ganesh, V.; Gadre, S. R. *J. Chem. Phys.* **2008**, *129*, 234101.

(65) Rahalkar, A. P.; Katouda, M.; Gadre, S. R.; Nagase, S. *J. Comput. Chem.* **2010**, *31*, 2405–2418.

(66) Yeole, S. D.; Gadre, S. R. *J. Chem. Phys.* **2010**, *132*, 094102.

(67) Mei, Y.; Ji, C. G.; Zhang, J. Z. H. *J. Chem. Phys.* **2006**, *125*, 094906.

(68) Zhang, D. W.; Zhang, J. Z. H. *J. Chem. Phys.* **2003**, *119*, 3599–3605.

(69) Addicoat, M. A.; Collins, M. A. *J. Chem. Phys.* **2009**, *131*, 104103.

(70) Collins, M. A.; Deev, V. A. *J. Chem. Phys.* **2006**, *125*, 104104.

(71) Collins, M. A. *J. Chem. Phys.* **2007**, *127*, 024104.

(72) Deev, V.; Collins, M. A. *J. Chem. Phys.* **2005**, *122*, 154102.

(73) Netzloff, H. M.; Collins, M. A. *J. Chem. Phys.* **2007**, *127*, 134113.

(74) Gordon, M. S.; Freitag, M. A.; Bandyopadhyay, P.; Jensen, J. H.; Kairys, V.; Stevens, W. J. *J. Phys. Chem. A* **2001**, *105*, 293–307.

(75) Mullin, J. M.; Roskop, L. B.; Pruitt, S. R.; Collins, M. A.; Gordon, M. S. *J. Phys. Chem. A* **2009**, *113*, 10040–10049.

(76) Slipchenko, L. V.; Gordon, M. S. *J. Comput. Chem.* **2007**, *28* 276–291.

(77) Dahlke, E. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2007**, *3*, 46–53.

(78) Dahlke, E. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2007**, *3*, 1342–1348.

(79) Dahlke, E. E.; Leverentz, H. R.; Truhlar, D. G. *J. Chem. Theory Comput.* **2008**, *4*, 33–41.

(80) Dahlke, E. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2008**, *4*, 1–6.

(81) Leverentz, H. R.; Truhlar, D. G. *J. Chem. Theory Comput.* **2009**, *5*, 1573–1584.

(82) Sorkin, A.; Dahlke, E. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2008**, *4*, 683–688.

(83) Li, W.; Li, S.; Jiang, Y. *J. Phys. Chem. A* **2007**, *111*, 2193–2199.

(84) Hua, W. J.; Fang, T.; Li, W.; Yu, J. G.; Li, S. H. *J. Phys. Chem. A* **2008**, *112*, 10864–10872.

(85) Hua, S. G.; Hua, W. J.; Li, S. H. *J. Phys. Chem. A* **2010**, *114* 8126–8134.

(86) Huang, L.; Massa, L.; Karle, J. *Proc. Natl. Acad. Sci.* **2006**, *103*, 1233.

(87) Huang, L.; Massa, L.; Karle, J. *Int. J. Quantum Chem.* **2005**, *103*, 808.

(88) Beran, G. J. O. *J. Chem. Phys.* **2009**, *130*, 164115.

(89) Beran, G. J. O.; Nanda, K. *J. Phys. Chem. Lett.* **2010**, *1*, 3480.

(90) Řezàč, J.; Salahub, D. R. *J. Chem. Theory Comput.* **2010**, *6*, 91–99.

(91) Xantheas, S. S. *J. Chem. Phys.* **1994**, *100*, 7523.

(92) Bettens, R. P. A.; Lee, A. M. *J. Phys. Chem. A* **2006**, *110*, 8777–8785.

(93) Hratchian, H. P.; Parandekar, P. V.; Raghavachari, K.; Frisch, M. J.; Vreven, T. *J. Chem. Phys.* **2008**, *128*, 034107.

(94) Mayhall, N. J.; Raghavachari, K.; Hratchian, H. P. *J. Chem. Phys.* **2010**, *132*, 114107.

(95) Parandekar, P. V.; Hratchian, H. P.; Raghavachari, K. *J. Chem. Phys.* **2008**, *129*, 145101.

(96) Elsohly, A. M.; Shaw, C. L.; Guice, M. E.; Smith, B. D.; Tschumper, G. S. *Mol. Phys.* **2007**, *105*, 2777–2782.

(97) Hopkins, B. W.; Tschumper, G. S. *Chem. Phys. Lett.* **2005**, *407* 362–367.

(98) Hopkins, B. W.; Tschumper, G. S. *Mol. Phys.* **2005**, *103*, 309–315.

(99) Hopkins, B. W.; Tschumper, G. S. *J. Comput. Chem.* **2003**, *24*, 1563–1568.

(100) Tschumper, G. S. *Chem. Phys. Lett.* **2006**, *427*, 185–191.

(101) Guo, W.; Wua, A.; Xu, X. *Chem. Phys. Lett.* **2010**, *498*, 203–208.

(102) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, A ; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, revision A.1; Gaussian Inc.: Wallingford, CT, 2009.

(103) http://ndbserver.rutgers.edu (accessed: 12/2010).

(104) Herbert, H. E.; Halls, M. D.; Hratchian, H. P.; Raghavachari, K. *J. Phys. Chem. B* **2006**, *110*, 3336.

(105) Tsan, P.; Volpon, L.; Besson, F.; Lancelin, J.-M. *J. Am. Chem. Soc.* **2007**, *129*, 1968–1977.

(106) Raghavachari, K.; Halls, M. D. *Mol. Phys.* **2004**, *102*, 381–393.

(107) Ferguson, G. A.; Than, C. T.-L.; Raghavachari, K. *J. Phys. Chem. Lett.* **2010**, *1*, 679–685.

(108) In the GEBF work, these subsystems are referred to as primitives.[85]

(109) To obtain this value of r, we first performed a series of calculations with a much cheaper B3LYP:PM6(r) model to calibrate the convergence of r. For $r = 30$ Å, both the RMS and the MAX relative conformer energy deviations between the fragmented B3LYP:PM6(3.0) and the unfragmented B3LYP methods were less than 1 kcal/mol (RMS = 0.46 kcal/mol; MAX = 0.92 kcal/mol).

(110) The largest number of calculations required for a MIM(r) energy was 311, whereas for a MIM(η) energy, it was 507. The largest MP2 subcalculation for MIM(r) had 580 basis functions, whereas MIM(η) had only 402 basis functions. A more detailed comparison of the different types of cutoffs (r or η) will be the focus of future work.

JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Modeling Fast Electron Dynamics with Real-Time Time-Dependent Density Functional Theory: Application to Small Molecules and Chromophores

Kenneth Lopata* and Niranjan Govind*

William R. Wiley Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory

ABSTRACT: The response of matter to external fields forms the basis for a vast wealth of fundamental physical processes ranging from light harvesting to nanoscale electron transport. Accurately modeling ultrafast electron dynamics in excited systems thus offers unparalleled insight but requires an inherently nonlinear time-resolved approach. To this end, an efficient and massively parallel real-time real-space time-dependent density functional theory (RT-TDDFT) implementation in NWChem is presented. The implementation is first validated against linear-response TDDFT and experimental results for a series of molecules subjected to small electric field perturbations. Second, nonlinear excitation of green fluorescent protein is studied, which shows a blue-shift in the spectrum with increasing perturbation, as well as a saturation in absorption. Next, the charge dynamics of optically excited zinc porphyrin is presented in real time and real space, with relevance to charge injection in photovoltaic devices. Finally, intermolecular excitation in an adenine-thymine base pair is studied using the BNL range separated functional [Baer, R.; Neuhauser, D. *Phys. Rev. Lett.* **2005**, *94*, No. 043002], demonstrating the utility of a real-time approach in capturing charge transfer processes.

1. INTRODUCTION

The time-dependent response of molecules under external fields forms the basis of a host of fundamental physical processes including light harvesting, photodissociation, electron transport, and higher harmonic emission. Broadly, radiation−molecule interactions can be classified as either weak or strong. When the interaction with the field is much smaller than the intramolecular interactions, the excitation is weak and the field induces only a small perturbation from the ground state. Perturbation theories such as linear-response time-dependent density functional theory (LR-TDDFT)[1−3] are excellent at modeling weak excitations and can accurately predict properties such as the absorption spectra of molecules and materials.[4] In the general case, however, matter−radiation interactions require going beyond linear response.

A fundamental understanding of nonlinear excited state dynamics at the femto- and subfemtosecond time regimes offers unparalleled insight into unsolved problems such as nonlinear spectra of single molecules, the nature of photoabsorption and exciton dynamics in photovoltaic devices, transport through molecules, and many others. Modeling nonlinear dynamics at molecular length scales, however, is a challenge requiring a combination of careful theoretical formulation and considerable computational effort. Unlike the weak excitation limit, where frequency domain perturbative approaches suffice, the strong excitation regime involves a complex interplay of electronic and nuclear dynamics and is best captured with a real-time, real-space approach. Here, the electron density, and in some cases nuclear motion, is monitored in time and space, which sheds light directly on the fundamental mechanisms of the excitation. Moreover, fully nonlinear (beyond perturbation theory) spectral information is readily obtainable from a real-time simulation via Fourier transform of time-dependent expectation values, such as the dipole moment.

All this comes at a cost, and studying real-time dynamics in molecules and materials is a daunting task. In particular, evolving a system in time requires calculating these potentials at every time step, which is extremely time-consuming. Additionally, care must be taken in evolving the system in time, and propagators must strike a balance between accuracy, stability, and speed. Finally, since excited states tend to be quite delocalized, dynamics simulations typically require the use of larger basis sets compared to traditional ground state calculations.

Despite the challenges, many successful approaches have been developed to study real-time electron dynamics in realistic systems. Within the Born−Oppenheimer approximation, these include: direct integration of the Schrödinger equation for very small systems (e.g., H_2);[5] real-time configuration interaction singles (CIS);[6,7] real-time orbital-free/Thomas-Fermi;[8,9] real-time, time-dependent density functional theory (RT-TDDFT) (discussed below); real-time Hartree−Fock (RT-TDHF);[10,11] and time-dependent semiempirical methods (e.g, TD-PM3).[12] DFT in particular offers a good trade-off between accuracy and efficiency for both ground and excited states, which has motivated extensive interest in TDDFT for real-time modeling, of which we present a representative sampling below.

Real-space (grid-based), time-dependent local density approximation RT-TDDFT, which was first developed by Theilhaber,[13] and pioneered by Yabana and Bertsch,[14] has been applied to systems ranging from aluminum dimers[15] to quantum dots in magnetic fields.[16] The Octopus[17] real-time TDDFT package derives from this lineage. Real-time TD-LDA using a planewave basis has similarly been applied to aluminum dimers,[18] modeling enhanced absorption of a nanoshell,[19] and conduction through a molecular junction.[20] RT-TDDFT has also been performed with

Received: February 23, 2011
Published: April 19, 2011

numerical orbitals using Siesta,[21] applied to atomic clusters[22] and higher harmonic generation in chromophores,[23] and extended to include ionic motion.[24] Another approach is to use a tight binding Hamiltonian for RT-TDDFT,[25] for which linear scaling implementations have been used to study absorption spectra.[26] Finally, RT-TDDFT with an atom-centered Gaussian basis has been used to study molecular conductance,[27,28] excited states at metal surfaces,[29] absorption properties of silicon clusters,[30] and double excitations[31] and singlet−triplet transitions.[32]

There has also been extensive work in developing schemes which go beyond the Born−Oppenheimer approximation to explicitly treat nuclear motion such as Ehrenfest dynamics,[33−35] Liouville−von Neumann molecular dynamics with real-time tight binding,[36] using surface hopping[37] to emulate nonadiabatic switching between adiabatic states,[38−40] and correlated electron−ion dynamics.[41]

In this paper, we present a massively parallel RT-TDDFT implementation in NWChem[42] geared towards simulating large systems while still maintaining generality and flexibility (e.g, various basis sets and functionals) and use it to explore the linear and nonlinear response properties of a series of molecules. The remainder of the paper takes the following form: The overall methodology is outlined in section 2.1. The structure of the time-dependent complex Fock matrix in discussed in section 2.2. The propagation scheme is detailed in section 2.3, and section 2.4 highlights some of the computation considerations. Next, the scheme is validated against LR-TDDFT for a few small molecules in section 3.1. The nonlinear absorption properties of the green fluorescent protein (GFP) are explored in section 3.2. Real-time, real-space visualization of resonant excitation in zinc porphyrin is presented in section 3.3, and finally, intramolecular charge transfer excitation in the adenine−thymine base pair is studied in section 3.4 using the BNL range-separated functional.[43,44]

2. METHODOLOGY

2.1. Overview of Real-Time TDDFT. Time-dependent density functional theory casts the time-dependent Schrödinger equation into a fictitious system of noninteracting electrons that satisfy the effective single particle time-dependent Kohn−Sham (TDKS) equations with an effective potential $v_{\mathrm{KS}}(\mathbf{r},t)$ uniquely described by the time-dependent charge density $\rho(\mathbf{r},t)$,[1] which in atomic units is

$$\begin{aligned} i\frac{\partial\psi_i(\mathbf{r},t)}{\partial t} &= \left[-\frac{1}{2}\nabla^2 + v_{\mathrm{KS}}[\rho](\mathbf{r},t)\right]\psi_i(t) \\ &= \left[-\frac{1}{2}\nabla^2 + v_{\mathrm{ext}}(\mathbf{r},t) + v_{\mathrm{H}}(\mathbf{r},t) + v_{\mathrm{XC}}[\rho](\mathbf{r},t)\right]\psi_i(t) \end{aligned} \tag{1}$$

Here, the charge density is the sum over all orbitals

$$\rho(\mathbf{r},t) = \sum_i^{\mathrm{occ}} |\psi_i(\mathbf{r},t)|^2 \tag{2}$$

and $v_{\mathrm{ext}}(\mathbf{r},t)$ contains the nuclear−electron and applied field potentials and $v_{\mathrm{H}}(\mathbf{r},t)$ is the Hartree (electron−electron) mean-field potential. Note that all potentials are explicit functions of time. Moreover, the exchange-correlation potential $v_{\mathrm{XC}}[\rho](\mathbf{r},t)$ is nonlocal in both space and time and is formally a functional of the initial wave functions and the entire history of the charge density $\rho(\mathbf{r},t)$. However, all practical implementations use the adiabatic approximation, which assumes locality in time (see discussion by Baer[45]). Real-time TDDFT involves explicitly propagating the coupled one-particle KS wave functions via eq 1. This is in contrast to the traditional linear-response approach, which is not actually a time-resolved method but instead solves eq 1 in the frequency domain for the excitation energies of a system subject to a small perturbation;[3] there are also real-time linear-response TDDFT approaches.[46,47]

In practical applications, the KS molecular orbitals are either solved in real space (e.g., finite element approaches) or expanded in a set of basis functions. In the case of an orthogonal basis (e.g., plane waves), time evolution consists of propagating the time-dependent coefficients of each of the mutually orthogonal basis functions. Localized basis functions, on the other hand, offer a good compromise between speed and flexibility, and in the case of a Gaussian basis set, they offer the added ability to use hybrid nonlocal exchange-correlation functionals in a seamless manner. In a Gaussian basis, it is most natural to use the single particle reduced density matrix

$$\mathbf{P}'_{\mu\nu} = \sum_i^{N_{\mathrm{MO}}} \mathbf{C}^*_{\mu i}(t)\,\mathbf{C}_{i\nu}(t) \tag{3}$$

where we have introduced the time-dependent molecular orbital coefficient matrix $\mathbf{C}(t)$, which describes the occupations of the molecular orbitals:

$$\psi_i(\mathbf{r},t) = \sum_{\mu=1}^{N_{\mathrm{AO}}} \mathbf{C}_{\mu i}(t)\,\phi_\mu(\mathbf{r}) \tag{4}$$

where $\{\phi(\mathbf{r})\}$ are the atomic orbitals, N_{AO} is the number of atomic orbitals, and N_{MO} is the number of molecular orbitals. From here on, we use primes to denote matrices in the molecular orbital (MO) basis and no primes to denote matrices in the atomic orbital (AO) basis. In the MO (orthonormal) basis, the time evolution of the density matrix is governed by the von Neumann equation

$$i\frac{\partial\mathbf{P}'}{\partial t} = [\mathbf{F}'(t), \mathbf{P}'(t)] \tag{5}$$

where $\mathbf{F}'(t)$ is the time-dependent Fock matrix in the MO basis, which depends on the density matrix at that time. Evolving the system in time reduces to computing the Fock matrix and stepping $\mathbf{P}'(t)$ forward using eq 5. For large systems with diffuse basis sets, linear dependencies in the basis become unavoidable; see Appendix A for a detailed discussion on AO↔MO transformations and how to deal with linear dependencies in RT-TDDFT.

Unlike ground state DFT calculations where the density matrix and Fock matrix are purely real (at least in the case of a real basis set like Gaussians), both become complex quantities in real time due to the i in eq 5. Moreover, whereas they are symmetric in the ground state, they must remain Hermitian under time propagation, with the additional constraint that the density matrix remains idempotent and trace invariant ($\mathbf{P}'\mathbf{P}' = \mathbf{P}'$, $Tr[\mathbf{P}'] = N_{\mathrm{e}}$), where N_{e} is the total number of electrons in the system.

This approach for evolving the density matrix in time is intuitive and easy to implement with a variety of time propagators (see section 2.3). Moreover, it can be readily extended to use matrix and current functionals, and phenomenological damping can be introduced via the off-diagonal elements of $\mathbf{P}'(t)$ or through friction functionals.[48] The dominant computational burden is in computing the Fock matrix $\mathbf{F}'(t)$ at every time step, which is outlined in the following section.

2.2. Time-Dependent Fock Matrix.

The proper choice of exchange-correlation functional is critical, as the inclusion of nonlocal exchange has been shown to be essential in a broad range of cases.[49−51] LR-TDDFT with global hybrid functionals (e.g, B3LYP, PBE0, and others) has been highly successful in predicting excitation energies for a range of systems.[4,52] More recently, range-separated functionals have been shown to capture charge transfer states successfully.[44,53−56] As such, we consider a composite time-dependent complex Fock matrix, which in general contains a blend of DFT exchange-correlation and Hartree−Fock exchange. This can be written in a general way as follows:

$$\begin{aligned}\mathbf{F}_{\mu\nu}[\mathbf{P}(t)] &= \mathbf{H}_{\mu\nu}^{\text{core}} + \mathbf{G}_{\mu\nu}^{J}[\mathbf{P}(t)] + \alpha\mathbf{G}_{\mu\nu}^{K}[\mathbf{P}(t)] + \beta\mathbf{G}_{\mu\nu}^{\text{X-DFT}}[\rho(\mathbf{r},t)]\\ &\quad + \gamma\mathbf{G}_{\mu\nu}^{\text{C-DFT}}[\rho(\mathbf{r},t)] + \mathbf{V}^{\text{app}}(t)\end{aligned} \tag{6}$$

where μ and ν index the atomic orbitals. Here, $\mathbf{H}_{\mu\nu}^{\text{core}}$ is the time-independent one-electron part, $\mathbf{G}_{\mu\nu}^{J}(t)$ is two-electron Coulomb interaction between the electrons, $\mathbf{G}_{\mu\nu}^{K}(t)$ is the two-electron exact exchange, $\mathbf{G}_{\mu\nu}^{\text{X-DFT}}(t)$ is the DFT exchange part, and $\mathbf{G}_{\mu\nu}^{\text{C-DFT}}$ is the DFT correlation. The α, β, and γ coefficients quantify the mixing of DFT and HF (exact exchange), e.g., $\alpha = 1, \beta = \gamma = 0$ for pure HF; $\alpha = 0, \beta = \gamma = 1$ yields for DFT; and intermediate values for global hybrid functionals. $\mathbf{V}^{\text{app}}(t)$ includes any external perturbation to the system, such as an applied electric field.

The Coulomb and DFT exchange-correlation contributions are all purely real symmetric matrices that depend only on the real part of the density matrix. The exact exchange matrix $\mathbf{G}_{\mu\nu}^{K}(t)$, however, is complex Hermitian and depends on both the real and imaginary parts of $\mathbf{P}(t)$; see Appendix B for a derivation and discussion of these symmetries. As a consequence, the imaginary part of $\mathbf{P}(t)$ only enters into the Fock matrix if there is exact exchange (i.e., either pure Hartree−Fock or hybrid functionals), and despite the complex phase introduced into the density matrix via eq 5, the Fock matrix remains purely real in pure DFT calculations.

2.3. Magnus Propagator.

The final component of the real-time scheme involves integrating eq 5 to get the time-dependent density matrix. Nonsymplectic integrators, such as Euler or Runge−Kutta methods, are unsuitable for large scale simulations as they become increasingly unstable with increased simulation size and require a very small time step to maintain the idempotency constraint of the density matrix. A better choice for von Neumann dynamics is the Magnus expansion, which steps $\mathbf{P}'$ forward in time using a unitary propagator which conserves the indempotency. We briefly summarize the method below, without derivation. For a general overview of time propagation schemes, see the review by Kosloff.[57] For additional details concerning propagators for the time-dependent Kohn−Sham equations, see refs 58 and 36 and also the excellent discussion by Castro and co-workers.[59]

The exact unitary propagator for eq 5 is given by

$$\mathbf{U}(t+\Delta t, t) = T\exp\{-i\int_{t}^{t+\Delta t}\mathbf{F}'(\tau)\,\mathrm{d}\tau\} \tag{7}$$

such that

$$\mathbf{P}'(t+\Delta t) = \mathbf{U}(t+\Delta t, t)\,\mathbf{P}'(t)\,\mathbf{U}^{\dagger}(t+\Delta t, t) \tag{8}$$

where T is the time-ordering operator which orders operators from those associated with later times to earlier times. The explicit time dependence of $\mathbf{F}'(t)$ makes it impossible to evaluate this propagator directly. Instead, a convenient solution to eq 7 is given by a Magnus expansion[60]

$$T\exp\{-i\int_{t}^{t+\Delta t}\mathbf{F}'(\tau)\,\mathrm{d}\tau\} = \mathrm{e}^{\Omega_1+\Omega_2+\ldots} \tag{9}$$

where the $\{\Omega_i\}$ are a series of nested commutator integrals:

$$\Omega_1(t+\Delta t, t) = -i\int_{t}^{t+\Delta t}\mathbf{F}'(\tau)\,\mathrm{d}\tau \tag{10}$$

$$\Omega_2(t+\Delta t, t) = -i\int_{t}^{t+\Delta t}\mathrm{d}\tau_1\int_{t}^{\tau_1}[\mathbf{F}'(\tau_2), \mathbf{F}'(\tau_2)] \tag{11}$$

$$\vdots \tag{12}$$

The resulting approximation is valid to order Δt^{2M}, where M is the number of Magnus terms. The integrals in eq 10 can be evaluated using quadrature. For example, for $M = 1$, we have

$$\mathbf{U}(t+\Delta t, t) \simeq \mathrm{e}^{\Omega_1} \tag{13}$$

$$\Omega_1 \simeq -i\mathbf{F}'(t+\Delta t/2) \tag{14}$$

The results presented here all used a $M = 1$ Magnus expansion; increasing to $M = 2$ would allow larger time steps, at the cost of more Fock builds per time step.

The main difficulty in using a Magnus scheme arises from the fact that the propagation (e.g., eq 14) requires knowledge of the Fock matrix at a future time, which is unknown. In the case of a second order ($M = 1$) Magnus propagator, the obvious solution is to form a guess for $\mathbf{F}'(t+\Delta t/2)$ from a linear extrapolation of $\mathbf{F}'$ at previous times. Unfortunately, crude predictors such as this inevitably fail for larger time steps. The most accurate method is to extrapolate $\mathbf{F}'$, propagate $\mathbf{P}'$ forward, interpolate to find a better $\mathbf{F}'$, and repeat until converged. This approach is costly, however, as you must rebuild the Fock matrix every convergence step. Instead, we adopted a two step predictor−corrector scheme proposed by Van Voorhis and co-workers,[27] whereby you predict $\mathbf{F}'(t+\Delta t/4)$ by linear extrapolation from previous values and use this to step $\mathbf{P}'$ forward by $\Delta t/2$ using eq 7. Overall, the predictor−corrector scheme was found to be sufficiently accurate and stable for a wide variety of systems. Predictor schemes such as this, however, fail to conserve the time-reversibility of eq 5. One alternate approach is the modified midpoint unitary transformation (MMUT) method developed by Li and co-workers;[11] the MMUT approach will be implemented in the future.

Finally, the exponentiation of the Ω matrices can be performed using a range of methods such as diagonalization, power series, Lanczos, etc. Note, however, that $\mathbf{U}^{\dagger} = \mathbf{U}^{-1}$ (unitary), so eq 7 is of the form

$$\mathbf{P}'(t+\Delta t) = \mathrm{e}^{\mathbf{W}}\mathbf{P}'(t)\,\mathrm{e}^{-\mathbf{W}} \tag{15}$$

where $\mathbf{W}(t+\Delta t,t) = \Omega_1(t+\Delta t,t) + \Omega_2(t+\Delta t,t) + \ldots$ Thus, we can apply the Baker−Campbell−Hausdorff (BCH) formula

$$\begin{aligned}\mathbf{P}'(t+\Delta t) &= \mathbf{P}' + \frac{1}{1!}[\mathbf{W}, \mathbf{P}'(t)] + \frac{1}{2!}[\mathbf{W}, [\mathbf{W}, \mathbf{P}'(t)]]\\ &\quad + \frac{1}{3!}[\mathbf{W}, [\mathbf{W}, [\mathbf{W}, \mathbf{P}'(t)]]] + \ldots\end{aligned} \tag{16}$$

For clarity, we have dropped the explicit time dependence of the $\mathbf{W}(t+\Delta t,t)$ matrix. The BCH expansion converges much better than a simple power series expansion, and the primary advantage of the BCH expansion over diagonalization is that eq 16 consists entirely of matrix multiplications, which are operations that parallelize extremely well. Although diagonalization formally scales as $O(N^3)$ and is thus more efficient than a series of matrix multiplications (each of which takes exactly N^3 effort in the absence of sparsity), in practical applications, diagonalization is hard to parallelize well. Due to inefficiencies, diagonalization becomes a serious bottleneck in quantum chemistry simulations run on more than a few hundred processors, and thus for large-scale simulations, diagonalization-free approaches are better suited (see ref 61 and references therein). In practice, it usually took on the order of tens of terms to converge the BCH expansion to 10^{-8} accuracy.

2.4. Computational Considerations. Accurate electron dynamics simulations of realistic systems can easily involve thousands of electrons and basis functions, propagated for long times. Clearly, an efficient implementation is necessary to make such simulations feasible. Since the vast majority of the computational work comes from building the Fock matrix at each time step, effort should be taken to either increase Δt or speed up construction of the Fock matrix.

Higher order Magnus propagators allow for larger time steps, for example, but require added Fock builds at each step; in practice, however, this is a system-specific trade-off. On the other hand, the Fock matrix construction can be sped up by using smaller basis sets or pure DFT functionals (e.g., ALDA), in conjunction with approaches such as charge density fitting.[62] Semiempirical Hamiltonians are also an alternative[12] but need to be properly parametrized and carefully validated.

The von Neumann dynamics approach as formulated is straightforward to implement in any quantum chemistry suite, as it can be built on fundamental routines. Constructing the time-dependent Fock matrix (eq 6) is akin to the building the ground state $\mathbf{F}$ in standard SCF (self-consistent field) schemes, save for the imaginary part due to exact exchange. Provided parallelization bottlenecks like diagonalization are avoided (as in the BCH exponentiation approach), RT-TDDFT will scale as well as standard Gaussian orbital-based SCF DFT with no loss of generality. We have implemented this approach, in NWChem which allows us to take advantage of the efficient parallelization capabilities offered by the code. This in turn allows us to tackle large systems with a high level of accuracy.

3. RESULTS

In this section, we first validate the real-time TDDFT approach against linear response TDDFT for series of small molecules, then move on to study the response of two chromophores to weak and strong perturbations, and finally study charge transfer across a DNA base pair using a long-range corrected, or range-separated, functional. Throughout, we mostly use atomic units (au), but for convenience we also present values in more customary units: 1 au length = 0.5292 Å; 1 au energy = 27.21 eV; 1 au time = 0.02419 fs; 1 au dipole moment = 2.542 D; 1 au electric field = 514.2 V/nm. Unless noted otherwise, all basis sets used in this study were obtained from the EMSL Basis Set Exchange.[63]

3.1. Validation on Small Molecules. To validate that the RT-TDDFT approach correctly reduces to linear response in the small perturbation limit, we studied the lowest excitation energies of various molecules and compared the results to standard LR-TDDFT and experimental results. An optical absorption spectrum can be obtained from a real-time simulation via Fourier transform of the time-dependent dipole moment resulting from a small δ-function-like electric field "kick". Starting from the ground state, we perturb the system with a narrow, transient, linearly polarized uniform Gaussian electric field:

$$\mathbf{E}(t) = \kappa \exp[-(t-t_0)^2/2w^2]\hat{d} \tag{17}$$

where t_0 is the center of pulse, w is the pulse width (typically a few time steps), which has dimensions of time, $\hat{d} = \hat{x}, \hat{y}, \hat{z}$ is the polarization of the pulse, and κ is the maximum field strength (dimensions of electric field). Note that the total energy added is therefore dependent on the time step Δt and the pulse width w; alternatively a normalized pulse can be used. The applied field excites the system through a dipole coupling term added to the Fock matrix (in the AO basis)

$$\mathbf{V}_{\mu\nu}^{\text{app}}(t) = -\mathbf{D}_{\mu\nu}\cdot\mathbf{E}(t) \tag{18}$$

where $\mathbf{D}$ is the transition dipole tensor of the system, e.g,

$$\mathbf{D}_{\mu\nu}^{x} = \int \phi_{\mu}^{*}(\mathbf{r})\, x\phi_{\nu}(\mathbf{r})\, d\mathbf{r} \tag{19}$$

A Gaussian-type electric field was chosen instead of a δ-function to avoid introducing nonphysical artifacts or instabilities due to a sudden change in potential. Despite the finite width, the pulse essentially excites all electronic frequencies simultaneously, save perhaps very high frequencies. The system is allowed to evolve in time, and the dipole moment is computed in the AO basis according to

$$\mu(t) = Tr[\mathbf{DP}(t)] \tag{20}$$

Likewise, the time-dependent occupation of the kth molecular orbital is computed by projecting the density matrix onto the ground state orbitals

$$n_k(t) = \mathbf{C}_k^{\prime\dagger}\, \mathbf{P}'(t)\, C_{k'}^{\prime} \tag{21}$$

where $\mathbf{C}_k'$ is the kth eigenvector of the ground state Fock matrix. To get the absorption spectrum from RT-TDDFT, a simulation is performed for each polarization of kick (symmetries in the system may alleviate this need), and the complex polarizability tensor is constructed from the Fourier transforms of the dipole signals

$$\alpha_{d,j}(\omega) = \frac{1}{\kappa}\tilde{\mu}_{d,j}(\omega) \tag{22}$$

where d is the index for the kick direction and j is the index for the measurement direction. The absorption cross-section tensor is obtained from $\alpha(\omega)$ via

$$\sigma(\omega) = \frac{4\pi\omega}{c}\,\text{Im}[\alpha(\omega)] \tag{23}$$

and finally the dipole strength function (absorption spectrum) is

$$S(\omega) = \frac{1}{3}Tr[\sigma(\omega)] \tag{24}$$

Figure 1 shows real-time data for benzene described using the 6-31G* basis set, subjected to a small x kick with $\kappa = 2 \times 10^{-5}$ au = 10 mV/nm, $w = 0.2$ au = 4.8×10^{-3} fs, and $t_0 = 3$ au = 0.07 fs. Before perturbing the system, the nuclear geometry was optimized

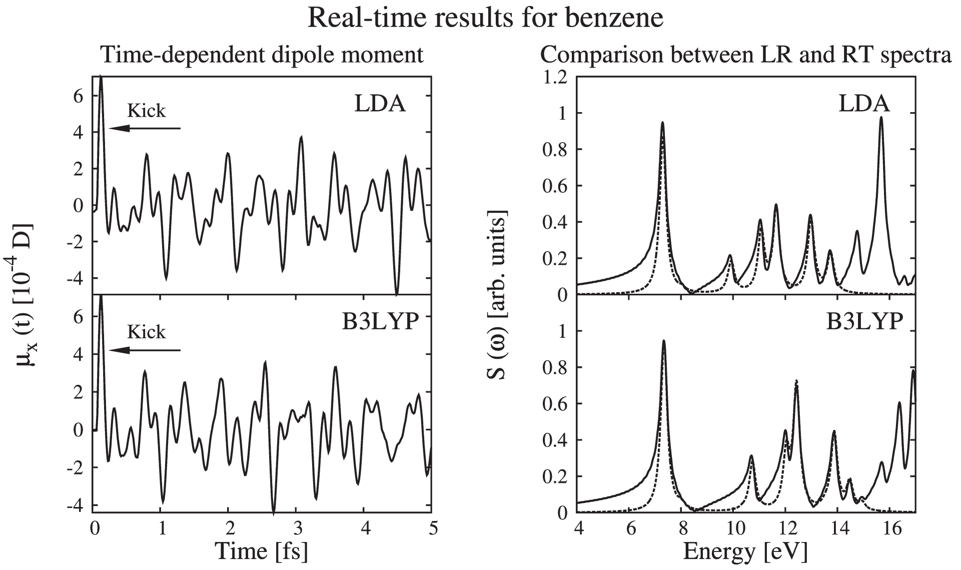


Figure 1. Real-time (RT) results for benzene in the perturbation limit, described using the 6-31G* basis and LDA (top) and B3LYP (bottom) functionals. The left panel shows the *x* dipole moment after a narrow Gaussian electric field kick. The resulting absorption spectra are shown on the right (solid lines), with the corresponding linear-response (LR) spectra shown for comparison (fine dashed lines). The peaks were artificially broadened and normalized (see text).

using the same basis and functional, and the ground state density converged via standard DFT. The narrow electric field pulse simultaneously excites all electronic modes, and the full dipole response (left panel) shows essentially dipolar oscillations composed of multiple frequencies.

A time step of $\Delta t = 0.5$ au = 0.012 fs was used, and the system was evolved for 1000 au = 24 fs. Two functionals were used, LDA and B3LYP, which yielded qualitatively similar results. For speed, in the LDA case, the Coulomb part of the Fock matrix was evaluated using charge density fitting instead of the explicit two-electron integrals;[62] this significantly reduced the computational cost. We checked that the charge density fitting approach yielded very similar results to explicit calculation of the Coulomb integrals. As an additional check, we confirmed that after the pulse had passed, the total system energy remained constant over time. The resulting absorption spectra are shown in the right panel, where the peaks have been broadened by artificially damping the time signal by $e^{-t/\tau}$, $\tau = 250$ au = 6 fs before taking the Fourier transform.

To validate these results, the corresponding linear response TDDFT spectra were compared with those obtained using the linear response TDDFT module in NWChem. The dashed lines show the LR spectrum, artificially broadened with Lorenzians of width 0.01 au = 0.3 eV. Additionally, both RT and LR spectra were normalized for clarity. The two spectra are essentially identical. For each of the finite LR signals (100 total roots computed for each simulation yielding six signals with appreciable oscillator strength), the RT signal agrees perfectly. Since the RT result effectively samples all excitations, rather than a finite number of roots as in LR, the RT spectra contain higher frequency signals beyond those computed in the LR simulation.

As an aside, the spectral resolution of the RT approach is limited by the time step; i.e., if a molecule has a spectral bandwidth of $\omega_{\max}$, the maximum time step is $\Delta t_{\max} = \pi/\omega_{\max}$. As an extreme example, to resolve the spectrum of a molecule with a maximum excitation frequency of 2 au (54 eV), one requires a time step of $\Delta t = 1.57$ au = 0.034 fs, or smaller. In practice, however, the time step is limited by the stability of the propagator (e.g, Magnus) rather than the bandwidth.

As further validation, a similar analysis was conducted for a range of small molecules, basis sets, and functionals, using the same kick parameters as described above. As before, geometry optimizations and convergence of the ground state densities were performed using the same basis and functionals as each of the real-time simulations. Table 1 shows a comparison between the linear-response, real-time, and gas-phase experimental lowest excitation energies for dihydrogen, methane, carbon monoxide, and benzene. Overall, there is excellent agreement between the linear-response (LR) and real-time (RT) values, as well as with experimental results. At worst, the RT energy deviated from the LR result by ~1% and from the experimental one by ~10%, with agreement generally improving with size of the basis set.

We note that a RT-TDDFT kick-type simulation yields the full electronic spectrum, up to the cutoff energy due to the finite time step. Thus, unlike a LR-TDDFT approach which requires a large number of roots or a windowed solver to compute higher energy transitions, a kick approach yields all roots in one single simulation, or at most three simulations (x, y, z kicks) in the absence of symmetries. Indeed, if one is interested in computing many excitation energies for a very large system, the RT approach is actually more efficient than a LR approach, which requires $\simeq O(N^4)$ effort for each root.[64] The one caveat is that, as formulated, a RT-TDDFT simulation can only probe excitations with a nonzero oscillator strength and thus cannot be effectively used to study "dark" excitations, which are typically measured experimentally via emission.

3.2. Linear and Nonlinear Excitation of Green Fluorescent Protein. Green fluorescent protein (GFP), which is responsible for the bioluminescence of some species of jellyfish, is nearly ubiquitous in biotechnology, with technological applications ranging from visualizing tagged proteins using fluorescence microscopy to developing transgenic fluorescent organisms.[65] Depending on the variant, GFP absorbs light in the blue or

Table 1. Comparison of Real-Time (RT) TDDFT, Linear Response (LR) TDDFT, and Experimental Lowest Excitation Energies (in eV) for a Selection of Molecules, Basis Sets, and Exchange-Correlation Functionals

	6-311G/LDA		6-311G/B3LYP		cc-pVTZ/LDA		cc-pVTZ/B3LYP		
	LR	RT	LR	RT	LR	RT	LR	RT	expt
H_2	12.52	12.49	13.09	13.12	12.32	12.31	12.88	12.90	11.19
CH_4	10.67	10.67	11.10	11.13	10.29	10.29	10.72	10.75	9.70
CO	8.00	8.03	8.16	8.22	8.29	8.28	8.52	8.55	8.55
C_6H_6	7.31	7.35	7.35	7.36	7.09	7.10	7.15	7.18	6.90

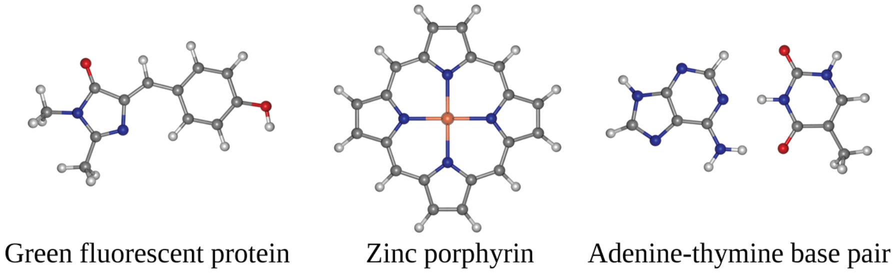


Figure 2. Structures of the molecules studied.

ultraviolet and fluoresces in the green. The actual GFP chromophore, which is a small molecule embedded in the larger overall protein, has a strong absorption due to a single excitation, which makes it an ideal candidate for TDDFT studies. The particular chromophore variant we studied (see Figure 2) has an intense experimental absorption at 3.51 eV, which corresponds to 354 nm light.[66] In the weak-field regime, linear response TDDFT calculations (3.32 eV B3LYP/POL1) agree well with both coupled cluster (3.60 eV CR-EOMCCSD(T)/POL1) and the experimental values for the absorption (see ref 67). Modeling the full response of GFP to strong fields beyond the weak perturbation limit, however, requires a real-time approach, and in this section, we use RT-TDDFT to study the nonlinear excitation of GFP subject to a range of perturbations.

To explore the nonlinear absorption properties, we subjected the GFP chromophore to a series of kicks (as in section 3.1) with field maximum κ ranging from 8×10^{-4} au = 0.41 V/nm (weak perturbation) to 0.24 au = 123 V/nm (strong perturbation). These narrow pulses ($w = 4.8 \times 10^{-3}$ fs) are nonphysical fields that simultaneously excite all electronic modes; this results in a nonphysical dipole moment which is a convolution of all excitations. Correspondingly, these simulations describe the immediate absorption properties of the molecule (i.e., how the light is absorbed and excites the density), and although the larger values of κ correspond to extremely strong electric fields, these simulations do not capture photoionization, which is difficult to describe using TDDFT in an atom-centered basis. Estimating ionization probabilities over a range of frequencies from a single kick-type nonphysical excitation is not straightforward, as the energy is distributed among all electrons in the system (i.e., core to valence); this will be quantified in future studies.

To ensure that there were sufficient basis functions to capture the diffuse excited states for the highly excited cases, we used the POL1 basis set.[68] In total, there were 114 electrons and 492 basis functions. To rule out unphysical confinement of the charge density due to the finite basis, we also tried using the smaller 6-31G* basis set; the results were essentially identical. We used the B3LYP exchange-correlation functional for this study. To ensure we start at an energy minimum, we used the same basis set and functional for geometry optimization and ground state density convergence before starting the time-dependent simulation. The density matrix was propagated for 1300 au = 31 fs with a time step of Δt = 0.1 au = 0.0024 fs.

Figure 3 shows the dipole moment and absorption spectrum (artificially broadened via damping by $e^{-t/\tau}$; τ = 400 au = 9.7 fs) for this range of perturbations. For illustrative purposes, relative values are presented—the dipole moment is scaled by the kick height such that any results in the weak excitation regime will be identical. Alternatively, in the linear response regime, if you double the kick, the dipole moment will double; this does not hold true in the strong excitation regime, which provides a simple graphical indicator of nonlinearity.

Once the perturbation is strong enough (i.e., $\kappa \gtrsim 40$ V/nm in Figure 3), two interesting effects emerge. First, the scaled responses (heights of the dipole moment and absorption peaks) decrease in magnitude; this is a saturation effect. Physically, once one goes beyond linear response, the highly excited molecule cannot absorb subsequent radiation, and the absorption, which is due to a single excitation, saturates. Nonlinear saturation effects like this, which are neglected by linear response, are critical for realistic modeling of spectroscopy under intense fields and transport in nanosystems where LR transport calculations tend to drastically overestimate certain effects.

Second, in the nonlinear regime, higher frequency signals begin to dominate. The inset in the left panel of Figure 3 shows how in the strong perturbation regime the time signal is no longer dominated by the main slow (low frequency) excitation, but instead by higher frequency modes. The corresponding spectrum shows how the absorption is likewise spread out over a wider range of frequencies. Moreover, in the nonlinear regime, the secondary absorption around 6.8 eV is increasingly blue-shifted with stronger fields.

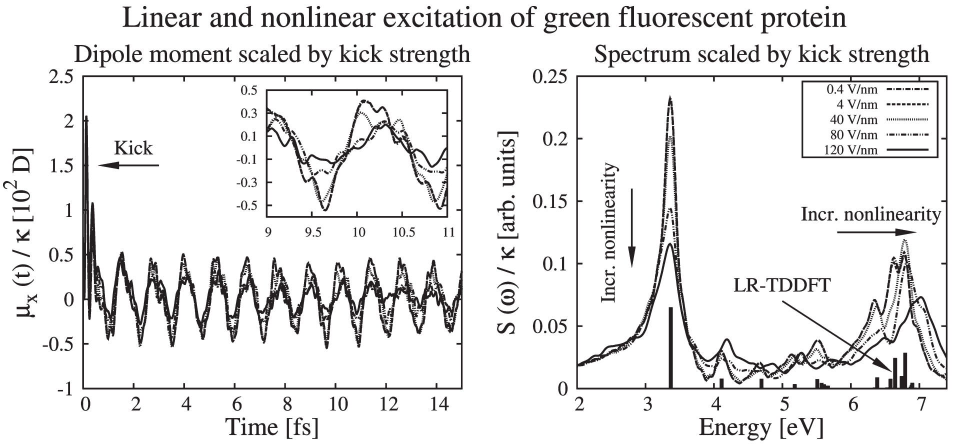


Figure 3. The dipole moment (left) and spectrum (right) of the green fluorescent protein chromophore subjected to a series of kicks ranging from the weak to the highly nonlinear regimes. For comparison, the values are scaled by the kick strength and the spectrum plot also shows the corresponding linear response TDDFT roots. In the nonlinear regime ($\kappa \gtrsim 40$ V/nm), the molecule's excitation is saturated, and higher frequency signal modes begin to dominate the time-resolved dipole moment. The strength of the main absorption in the spectrum decreases with nonlinearity, and the secondary higher energy absorption is increasingly blue-shifted.

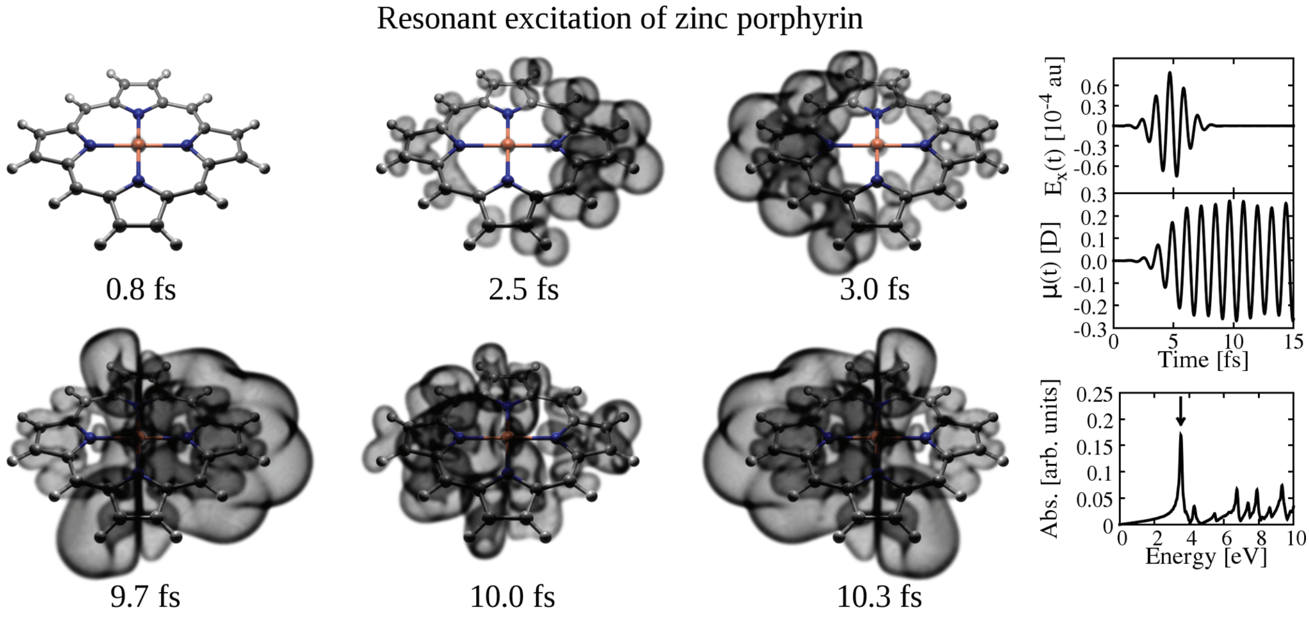


Figure 4. Isosurface snapshots of the difference between the excited and ground state charge densities, $\rho(\mathbf{r},t) - \rho(\mathbf{r},0) = 7 \times 10^{-7}$ Å^{-3}, for zinc porphyrin described using B3LYP and 6-31G* for H, C, and N and the Stuttgart RSC 1997 effective core potential for Zn. The system was excited at its resonance of 3.53 eV (see bottom inset spectrum) with a transient x-polarized laser pulse (top inset), resulting in charge oscillation back and forth along the molecule's conjugated π backbone, with each complete oscillation taking ~1.2 fs. The charge density was plotted using Blender.[75]

A simple physical interpretation is that the highly displaced charge density experiences a nonlinear restoring force which is stronger than in the linear regime, just like a highly displaced spring.

3.3. Resonant Excitation of Zinc Porphyrin. Porphyrin forms the structural basis for the various chlorophyll molecules and has been exploited as the functional unit in light harvesting systems.[69,70] The absorption properties of zinc and free-base porphyrin have been studied extensively using LR-TDDFT and coupled cluster theory,[54,67,71] but from a photovoltaic device point of view, the mechanism of charge injection (i.e., from the light harvesting porphyrin into a nearby substrate) has many unanswered questions. RT-TDDFT is an excellent tool to probe these charge dynamics in real time and real space.

As a first step in this direction, we model the response of zinc porphyrin (Figure 2) to a transient laser pulse tuned to its strongest absorption. We used B3LYP for the exchange correlation; for the basis set, we used 6-31G* for the hydrogen, carbon, and nitrogen atoms and the Stuttgart RSC 1997 effective core potential (ECP) for the zinc center, which replaces 10 of the Zn electrons. Since optical absorption does not typically involve core electrons, ECPs offer a simple way to boost the speed of real-time calculations without a loss of accuracy; we confirmed that the use of an ECP in this case did not alter the results. The total propagation time was 1500 au = 36 fs with a time step of Δt = 0.1 au = 0.0024 fs, which was chosen to ensure a smooth density animation rather than being limited by the propagator.

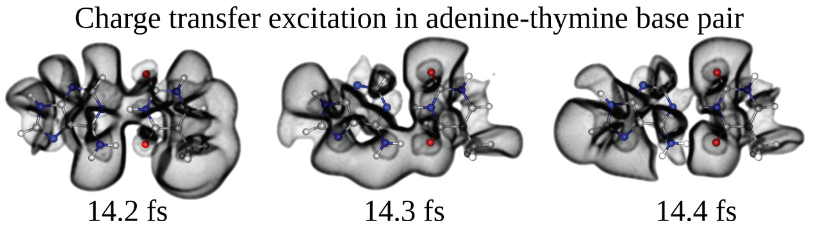


Figure 5. Snapshots of $\rho(\mathbf{r},t) - \rho(\mathbf{r},0) = 10^{-7}$ Å^{-3} for the adenine (A, left molecule)/thymine (T, right molecule) base pair excited at the 6.36 eV resonance, corresponding to an intermolecular charge transfer state. The snapshots correspond to times long after the transient laser pulse has passed, leaving the system in an excited state. Initially, both molecules are polarized in the $+x$ direction, and there is an excess of charge on T. Next, charge is transferred A $\leftarrow$ T via the NH2···O═C bridge, resulting in a net $-x$ overall polarization and excess charge on A. The overall charge transfer happens in approximately 0.3 fs. The charge density was plotted using Blender.[75]

The enveloped monochromatic laser pulse took the form

$$\begin{aligned} E_x(t) &= k\exp[-(t-t_0)^2/2w^2]\cos(\omega_0 t), \\ E_y &= E_z = 0 \end{aligned} \tag{25}$$

The frequency of the pulse was taken to be $\omega_0 = 0.1296$ au = 3.53 eV, and the Gaussian had a half-width of $w = 50$ au = 1.2 fs. The maximum field strength was $\kappa = 8.0 \times 10^{-5}$ au = 41 mV/nm, which is within the linear response regime. The ZnP resonance frequency was determined from a kick-type simulation (see eq 17 in section 3.1) and artificially broadened. The resulting spectrum is shown in the bottom inset of Figure 4, with an arrow denoting the laser pulse frequency. The top inset shows the profile for the homogeneous x-polarized monochromatic electric field pulse used in this resonant excitation (overall the pulse lasts approximately 7 fs). The middle plot shows the resulting dipole moment which is essentially monochromatic, and as there is no damping in the system, the oscillation continues indefinitely.

The six panels in Figure 4 show snapshots of the deviation of the charge density from the ground state, $\rho(\mathbf{r},t) - \rho(\mathbf{r},0)$ during and after resonant illumination. The snapshots show the positive 7×10^{-7} Å^{-3} isosurface, i.e., the smooth surface where $\rho(\mathbf{r},t) - \rho(\mathbf{r},0) = 7 \times 10^{-7}$ Å^{-3}; the corresponding negative deviation was omitted for clarity. The top three slices depict the response of the charge density while being driven by the laser pulse. $\rho(\mathbf{r},t)$ starts essentially in the ground state (first slice). The second slice occurs just at the first significant peak in the dipole moment, which occurs just after the peak in the driving field, as it takes time for the density to respond. Here, the extra charge density is beginning to populate the space above and below the carbon backbone on the $+x$ (right) side of the molecule, which corresponds to a $\pi \rightarrow \pi^*$ transition. The third slice, which is the first significant trough in the total dipole moment, demonstrates that the charge density now populates the π^* orbitals on the $-x$ (left) side of the molecule. The bottom three slices show how the charge density of the excited ZnP molecule evolves after the driving field has passed. The charge density sloshes along the delocalized π^* orbitals from the right to the left, which takes ~24 au = 0.6 fs, which is in agreement with the time-resolved dipole moment (middle inset).

Using RT-TDDFT to directly visualize the electron dynamics offers insight into the fundamental nature of the excitation, not just concerning which molecular orbitals are at play but also how they are being populated in time, and where in space the charge is concentrated. In Figure 4, not only is the $\pi \rightarrow \pi^*$ transition obviously visible, but additionally, as expected, there is a clear buildup of charge at the x end of the molecule during the oscillations. It is easy to see that bonding a porphyrin to a substrate will then allow the excited charge density, which has high momentum, to hop from the end of the molecule to the surface in an ultrafast injection process; future RT-TDDFT studies will explore this effect further.

3.4. Long-Range Charge Transfer in DNA Base-Pair Fragments.

As a final example, we demonstrate how long-range corrected functionals can be used in a RT-TDDFT framework to correctly capture charge transfer excitations. The charge transfer between adenine (A) and thymine (T) is a classic example where local exchange-correlation functionals (LDA) and even global hybrids (e.g, B3LYP) underestimate the energy of the A$\pi \rightarrow$Tπ^* intermolecular transition, to the point that they incorrectly predict it will be the lowest excitation.[50,55] The source of this error is the incorrect asymptotic behavior of the exchange term, which should go as r_{12}^{-1} but goes as $0.2r_{12}^{-1}$ in B3LYP, for example. Recently developed long-range corrected functionals have shown great promise in addressing this shortcoming.[44,53−56,72] These functionals split the exchange into a short-range part and a long-range piece which converges to the correct Hartree−Fock asymptote:

$$\frac{1}{r_{12}} = \frac{1 - \mathrm{erf}(\mu r_{12})}{r_{12}} + \frac{\mathrm{erf}(\mu r_{12})}{r_{12}} \tag{26}$$

Here, μ is a tuning parameter for partitioning the exchange, where $\mu \rightarrow 0$ tends to the pure DFT limit and $\mu \rightarrow \infty$ tends to the pure HF limit.

The real-time response of the A−T pair (see Figure 2) was modeling using the 6-31G* basis set and the BNL range separated functional[43] with $\mu = 0.3$. The Coulomb part of the Fock matrix was computed using charge density fitting with the Ahlrichs Coulomb fitting basis set.[62] The system was excited with a transient laser pulse tuned to the charge transfer excitation in the linear response regime: $\omega_0 = 0.234$ au = 6.36 eV, $w = 30$ au = 0.7 fs, and $\kappa = 1.2 \times 10^{-4}$ au = 62 mV/nm (see eq 25). The resonant frequency was found via kick-type simulation as described previously; it compares well with the value of 6.25 eV from ref 55 computed using LR-TDDFT with BNL ($\mu = 0.3$) and the aug-cc-pVTZ basis set.

Figure 5 shows three snapshots of the $\rho(\mathbf{r},t) - \rho(\mathbf{r},0) = 10^{-7}$ Å^{-3} density deviation isosurface long after the exciting laser pulse has passed. In the first frame ($t = 14.2$ fs), both molecules are polarized in the $+x$ direction, and there is shared electron density in the central N···H bridge. In the second frame (0.1 fs later), excess charge passes from the thymine molecule through the NH_2···O═C bridge, finally resulting in a charge buildup on adenine another 0.1 fs later and a net $-x$ polarization for both molecules. These simulations show that the complete A $\leftarrow$ T charge transfer process occurs in approximately 0.3 fs. Note that these dynamics are much faster than the experimentally observed 100 fs decay time for this excitation (likely due to internal conversion to $n\pi^*$).[73]

4. CONCLUSIONS

We have presented Gaussian basis set-based real-time, time-dependent density functional theory simulations using NWChem and have shown that the calculated spectra for a range of small molecules correctly reduce to the linear response TDDFT spectra in the small perturbation limit. Going beyond linear response, we studied the optical response of the green fluorescent chromophore to a series of perturbations of increasing strength. In the strong perturbation regime, the main absorption saturates and the higher energy absorption becomes blue-shifted with increasing field strengths; this has implications for strong field studies of molecules and electron transport in nanosystems. Next, we studied the resonant excitation of the light-harvesting molecule zinc porphyrin. Direct visualization of the charge density in time and space shows that the excitation, which corresponds to a $\pi \rightarrow \pi^*$ transition, induces delocalized charge oscillations across the carbon backbone, with a buildup of charge near the ends of the molecule. Real-time, real-space studies of this kind offer powerful insight into electron dynamics and are uniquely well-suited to modeling fast electron processes in a variety of devices, such as photovoltaics. Finally, we visualized the adenine $\pi \rightarrow$thymine π^* transition, which shows that the charge transfer happens through the oxygen−amine bridge on the order of 0.3 fs. From a computational point of view, the implementation is massively parallel and is scalable with system size; as there is no diagonalization, the main burden is construction of the Fock matrix, which is easily distributed across many processors. Further improvements to the implementation are planned, which will be presented in future publications.

■ APPENDIX A: CANONICAL ORTHOGONALIZATION TRANSFORMS

The real-time TDDFT scheme requires working in both the atomic orbital (AO) and molecular orbital (MO) representations. The propagation is done entirely in the MO basis via the von Neumann equation, eq 5, whereas the Fock matrix is built in the AO basis, eq 6. The time-dependent dipole moment is computed in the AO basis, eq 20, and the time-dependent orbital occupations are computed in the MO basis, eq 21. It is therefore useful to outline how to perform AO↔MO transformations.

For a given overlap matrix $\mathbf{S}_{\mu\nu} = \langle \phi_\nu | \phi_\nu \rangle$, there may be linear dependencies in the eigenvectors which necessitates truncating the number of molecular orbitals using canonical orthogonalization. Although not typically a problem in smaller systems, as the system size increases or many diffuse atomic orbitals are used (which is necessary to capture diffuse excited states), linear dependencies become unavoidable. The well-known transformation matrix for converting from the AO to a truncated MO basis is[74]

$$\mathbf{X} = \mathbf{U}\mathbf{s}^{-1/2} \tag{27}$$

where $\mathbf{U}$ is the matrix with eigenvectors of $\mathbf{S}$ as columns, and $\mathbf{s}$ is the diagonal matrix of eigenvalues of $\mathbf{S}$. If we have N atomic orbitals and d linear dependencies, $\mathbf{X}$ becomes a rectangular matrix of dimensions $N \times M$, where $M = N - d$ is the number of molecular orbitals. Converting the Fock matrix from the AO basis to the MO basis is then straightforward:

$$\mathbf{F}' = \mathbf{X}^\dagger \mathbf{F} \mathbf{X} \tag{28}$$

Note that $\mathbf{F}$ (AO basis) is an $N \times N$ matrix, whereas $\mathbf{F}'$ (MO basis) is a smaller $M \times M$ matrix. Converting the density matrix from the MO to the AO basis is likewise very simple:

$$\mathbf{P} = \mathbf{X}\mathbf{P}'\mathbf{X}^\dagger \tag{29}$$

where, as before, $\mathbf{P}$ is $N \times N$ and $\mathbf{P}'$ is $M \times M$. It is slightly more complicated to convert $\mathbf{P} \rightarrow \mathbf{P}'$, which is necessary when converting the ground state density matrix, which is computed in the AO basis in an SCF approach, to the MO basis for subsequent von Neumann propagation. Simple inversion of eq 29 is complicated by the fact that $\mathbf{X}$ is not square and cannot be easily inverted.

The simplest solution is to use left and right inverses. The left inverse of $\mathbf{X}$ is given by

$$\mathbf{X}_\mathrm{L}^{-1} = (\mathbf{X}^\dagger \mathbf{X})^{-1} \mathbf{X}^\dagger \tag{30}$$

while the right inverse of $\mathbf{X}^\dagger$ is given by

$$(\mathbf{X}^\dagger)_\mathrm{R}^{-1} = \mathbf{X}(\mathbf{X}^\dagger \mathbf{X})^{-1} \tag{31}$$

We know these inverses exist because all zero (or near zero) eigenvectors have been removed. From eq 29, we get

$$\mathbf{X}_\mathrm{L}^{-1} \mathbf{P} (\mathbf{X}^\dagger)_\mathrm{R}^{-1} = \mathbf{P}' \tag{32}$$

which means

$$\mathbf{P}' = (\mathbf{X}^\dagger \mathbf{X})^{-1} \mathbf{X}^\dagger \mathbf{P} \mathbf{X} (\mathbf{X}^\dagger \mathbf{X})^{-1} \tag{33}$$

From eq 27, we know

$$\mathbf{X}^\dagger = \mathbf{s}^{-1/2} \mathbf{U}^\dagger \tag{34}$$

Although $\mathbf{U}$ is not strictly unitary (as it is not square), we know that $\mathbf{U}^\dagger \mathbf{U} = I_m$, and thus $(\mathbf{X}^\dagger \mathbf{X})^{-1} = \mathbf{s}$. The transformation from the density matrix in the AO basis to the MO basis then becomes

$$\mathbf{P}' = \mathbf{s}\mathbf{X}^\dagger \mathbf{P} \mathbf{X} \mathbf{s} \tag{35}$$

which in a more compact form is simply

$$\mathbf{P}' = \mathbf{Y}^\dagger \mathbf{P} \mathbf{Y} \tag{36}$$

where $\mathbf{Y} \equiv \mathbf{X}\mathbf{s} = \mathbf{U}\mathbf{s}^{1/2}$ is an $N \times M$ transformation matrix.

■ APPENDIX B: SYMMETRIES IN THE COMPLEX FOCK MATRIX

In this section, we prove that, for a basis set of purely real functions, in pure RT-TDDFT (without Hartree−Fock exchange), the Fock matrix is purely real and symmetric and depends only on the real part of the complex density matrix. In the case of hybrid RT-TDDFT, however, the HF exchange term of the Fock matrix is complex Hermitian and depends on the full complex density matrix. The derivation presented is similar to that given in ref 36. For simplicity, we assume a closed shell system, but the results are identical for an open shell system.

Recall that in hybrid DFT-HF, the elements of the Fock matrix take the general form

$$\begin{aligned} \mathbf{F}_{\mu\nu}[\mathbf{P}(t)] = {} & \mathbf{H}_{\mu\nu}^\mathrm{core} + \mathbf{G}_{\mu\nu}^J(t) + \alpha \mathbf{G}_{\mu\nu}^K(t) + \beta \mathbf{G}_{\mu\nu}^\text{X-DFT}(t) \\ & + \gamma \mathbf{G}_{\mu\nu}^\text{C-DFT}(t) + \mathbf{V}^\mathrm{app}(t) \end{aligned} \tag{37}$$

where μ and ν are indexes for the atomic orbitals, $\mathbf{H}^{\text{core}}$ is the time-independent one-electron part, $\mathbf{G}^{J}(t)$ and $\mathbf{G}^{K}(t)$ are the time-dependent Coulomb and exchange terms, $\mathbf{G}^{\text{X-DFT}}$ and $\mathbf{G}^{\text{C-DFT}}$ are the DFT exchange and correlation terms, and $\mathbf{V}^{\text{app}}(t)$ is the potential due to an external perturbation (e.g, electric field). In RT-TDDFT, the Fock matrix $\mathbf{F}(t)$ and the density matrix $\mathbf{P}$ are in general complex and Hermitian. We will discuss the symmetries in eq 37 term by term.

First, we note that the applied potential $\mathbf{V}^{\text{app}}(t)$ is independent of the density matrix, and for all physical potentials, it is purely real. This is not true in the case of nonphysical potentials such as complex absorbing boundary conditions, but in such situations, the Fock matrix ceases to be Hermitian and the total system charge is not conserved in time, which requires a careful reformulation of TDDFT.

The time-independent one-electron part $\mathbf{H}^{\text{core}}$ includes kinetic and electron–nuclear terms

$$\mathbf{H}^{\text{core}}_{\mu\nu} = \mathbf{T}_{\mu\nu} + \mathbf{V}^{\text{eN}}_{\mu\nu} \tag{38}$$

$$= \int \mathrm{d}\mathbf{r}_1\ \phi_\mu(\mathbf{r}_1)\left[-\frac{1}{2}\nabla_1^2\right]\phi_\nu(\mathbf{r}_1)$$
$$+\int \mathrm{d}\mathbf{r}_1\ \phi_\nu(\mathbf{r}_1)\left[-\sum_A \frac{Z_A}{|\mathbf{r}_1 - \mathbf{R}_A|}\right]\phi_\nu(\mathbf{r}_1) \tag{39}$$

where $\{\phi(\mathbf{r})\}$ are the atomic orbitals which we henceforth assume are real. Since this expression is independent of the density matrix, and the integrals in eq 39 are symmetric with respect to the exchange of μ and ν, *the core term is pure real and symmetric.*

Next, in adiabatic RT-TDDFT, the DFT exchange and correlation terms are all functionals uniquely determined by the instantaneous charge density (and possibly its gradients):

$$\mathbf{G}^{\text{X-DFT}}_{\mu\nu} = \mathbf{G}^{\text{X-DFT}}_{\mu\nu}[\rho(\mathbf{r},t)] \tag{40}$$

$$\mathbf{G}^{\text{C-DFT}}_{\mu\nu}(t) = \mathbf{G}^{\text{C-DFT}}_{\mu\nu}[\rho(\mathbf{r},t)] \tag{41}$$

The charge density $\rho(\mathbf{r},t)$ is dependent only on the real part of the density matrix

$$\rho(\mathbf{r},t) = \sum_\mu\sum_\nu \mathrm{Re}[\mathbf{P}_{\mu\nu}(t)]\phi_\mu(t)\phi_\nu(\mathbf{r}) \tag{42}$$

and therefore *the DFT XC terms are both real and symmetric and depend only on the real part of* $\mathbf{P}_{\mu\nu}(t)$.

The Coulomb term takes the form

$$\mathbf{G}^{J}_{\mu\nu}(t) = \sum_{\lambda\sigma} \mathbf{P}_{\lambda\sigma}(t)(\mu\nu|\sigma\lambda) \tag{43}$$

where $(\mu\nu|\sigma\lambda)$ are the standard two-electron integrals

$$(\mu\nu|\sigma\lambda) \equiv \int \phi_\mu(\mathbf{r}_1)\ \phi_\nu(\mathbf{r}_1)\frac{1}{r_{12}}\phi_\sigma(\mathbf{r}_2)\ \phi_\lambda(\mathbf{r}_2)\ \mathrm{d}\mathbf{r}_1\ \mathrm{d}\mathbf{r}_2 \tag{44}$$

Note that since the basis functions are real, these two-electron integrals are symmetric to permutation of λ and σ:

$$(\mu\nu|\sigma\lambda) = (\mu\nu|\lambda\sigma) = (\nu\mu|\lambda\sigma) = (\nu\mu|\sigma\lambda) \tag{45}$$

The double sum in eq 43 can be split into three parts

$$\mathbf{G}^{J}_{\mu\nu}(t) = \sum_\lambda\sum_{\sigma<\lambda} \mathbf{P}_{\lambda\sigma}(t)(\mu\nu|\sigma\lambda) + \sum_\lambda \mathbf{P}_{\lambda\lambda}(t)(\mu\nu|\lambda\lambda)$$
$$+\sum_\lambda\sum_{\sigma>\lambda} \mathbf{P}_{\lambda\sigma}(t)(\mu\nu|\sigma\lambda) \tag{46}$$

Swapping the indices of summation for the third term gives

$$\mathbf{G}^{J}_{\mu\nu}(t) = \sum_\lambda\sum_{\sigma<\lambda} \mathbf{P}_{\lambda\sigma}(t)(\mu\nu|\sigma\lambda) + \sum_\lambda \mathbf{P}_{\lambda\lambda}(t)(\mu\nu|\lambda\lambda)$$
$$+\sum_\sigma\sum_{\sigma>\lambda} \mathbf{P}_{\sigma\lambda}(t)(\mu\nu|\lambda\sigma) \tag{47}$$

and using the symmetry of the two-electron integrals (eq 45), the Coulomb matrix elements become

$$\mathbf{G}^{J}_{\mu\nu}(t) = \sum_{\sigma<\lambda} [\mathbf{P}_{\lambda\sigma}(t) + \mathbf{P}_{\sigma\lambda}(t)](\mu\nu|\sigma\lambda) + \sum_\lambda \mathbf{P}_{\lambda\lambda}(t)(\mu\nu|\lambda\lambda) \tag{48}$$

The real part is

$$\mathrm{Re}[\mathbf{G}^{J}_{\mu\nu}(t)] = \sum_{\sigma<\lambda} \{\mathrm{Re}[\mathbf{P}_{\lambda\sigma}(t)] + \mathrm{Re}[\mathbf{P}_{\sigma\lambda}(t)]\}(\mu\nu|\sigma\lambda)$$
$$+\sum_\lambda \mathrm{Re}[\mathbf{P}_{\lambda\lambda}(t)](\mu\nu|\lambda\lambda) \tag{49}$$

but since $\mathbf{P}$ is Hermitian, the real part is symmetric, which gives

$$\mathrm{Re}[\mathbf{G}^{J}_{\mu\nu}(t)] = 2\sum_{\sigma<\lambda} \mathrm{Re}[\mathbf{P}_{\lambda\sigma}(t)](\mu\nu|\sigma\lambda) + \sum_\lambda \mathrm{Re}[\mathbf{P}_{\lambda\lambda}(t)](\mu\nu|\lambda\lambda) \tag{50}$$

Equation 50 is symmetric to the exchange of μ and ν; thus $\mathrm{Re}[\mathbf{G}^{J}_{\mu\nu}(t)] = \mathrm{Re}[\mathbf{G}^{J}_{\nu\mu}(t)]$, and the real part of the Coulomb term is symmetric. The imaginary part is

$$\mathrm{Im}[\mathbf{G}^{J}_{\mu\nu}(t)] = \sum_{\sigma<\lambda} \mathrm{Im}\{\mathbf{P}_{\lambda\sigma}(t) + \mathrm{Im}[P_{\sigma\lambda}(t)]\}(\mu\nu|\sigma\lambda)$$
$$+\sum_\lambda \mathrm{Im}[\mathbf{P}_{\lambda\lambda}(t)](\mu\nu|\lambda\lambda) \tag{51}$$

but here due to Hermicity the imaginary part of the density matrix is antisymmetric with on-diagonal elements of zero; thus

$$\mathrm{Im}[\mathbf{P}_{\lambda\sigma}(t)] + \mathrm{Im}[P_{\sigma\lambda}(t)] = 0 \tag{52}$$

$$\mathrm{Im}[\mathbf{P}_{\lambda\lambda}(t)] = 0 \tag{53}$$

and the imaginary part of the Coulomb matrix vanishes

$$\mathrm{Im}[\mathbf{G}^{J}_{\mu\nu}(t)] = 0 \tag{54}$$

Thus, *the Coulomb term is a real-valued symmetric matrix which only depends on the real part of the complex density matrix.*

A similar analysis can be done for the exchange matrix,

$$\mathbf{G}^{K}_{\mu\nu}(t) = \sum_{\lambda\sigma} \mathbf{P}_{\lambda\sigma}(t)(\mu\lambda|\sigma\nu) \tag{55}$$

(note the different two electron integrals from Coulomb part), which after expanding into three terms and swapping summation in the third term gives

$$\mathbf{G}^{K}_{\mu\nu}(t) = \sum_\lambda\sum_{\sigma<\lambda} \mathbf{P}_{\lambda\sigma}(t)(\mu\lambda|\sigma\nu) + \sum_\lambda \mathbf{P}_{\lambda\lambda}(t)(\mu\lambda|\sigma\nu)$$
$$+\sum_\sigma\sum_{\sigma<\lambda} \mathbf{P}_{\sigma\lambda}(t)(\mu\lambda|\sigma\nu) \tag{56}$$

The real part is

$$\begin{aligned}\mathrm{Re}[\mathbf{G}^K_{\mu\nu}(t)] = & \sum_{\sigma<\lambda} \mathrm{Re}[\mathbf{P}_{\lambda\sigma}(t)](\mu\lambda|\sigma\nu) + \sum_{\lambda} \mathrm{Re}[\mathbf{P}_{\lambda\lambda}(t)](\mu\lambda|\lambda\nu) \\ & + \sum_{\sigma<\lambda} \mathrm{Re}[\mathbf{P}_{\sigma\lambda}(t)](\mu\sigma|\lambda\nu) \qquad (57)\end{aligned}$$

$$\begin{aligned}= & \sum_{\sigma<\lambda} \mathrm{Re}[\mathbf{P}_{\lambda\sigma}(t)](\mu\lambda|\sigma\nu) + (\mu\sigma|\lambda\nu) \\ & + \sum_{\lambda} \mathrm{Re}[\mathbf{P}_{\lambda\lambda}(t)](\mu\lambda|\lambda\nu) \qquad (58)\end{aligned}$$

To check the symmetry, we switch the μ and ν indices

$$\mathrm{Re}[\mathbf{G}^K_{\nu\mu}(t)] = \sum_{\sigma<\lambda} \mathrm{Re}[\mathbf{P}_{\lambda\sigma}(t)](\nu\lambda|\sigma\mu) + (\nu\sigma|\lambda\mu) + \sum_{\lambda} \mathrm{Re}[\mathbf{P}_{\lambda\lambda}(t)](\nu\lambda|\lambda\mu) \qquad (59)$$

Permuting the two electron integrals (eq 45) gives

$$\mathrm{Re}[\mathbf{G}^K_{\nu\mu}(t)] = \sum_{\sigma<\lambda} \mathrm{Re}[\mathbf{P}_{\lambda\sigma}(t)](\mu\sigma|\lambda\nu) + (\mu\lambda|\sigma\nu) + \sum_{\lambda} \mathrm{Re}[\mathbf{P}_{\lambda\lambda}(t)](\mu\lambda|\lambda\nu) \qquad (60)$$

$$= \mathrm{Re}[\mathbf{G}^K_{\nu\mu}(t)] \qquad (61)$$

thus the real part of the exchange term is symmetric. The imaginary part is (c.f. eq 57)

$$\begin{aligned}\mathrm{Im}[\mathbf{G}^K_{\mu\nu}(t)] = & \sum_{\sigma<\lambda} \mathrm{Im}[\mathbf{P}_{\lambda\sigma}(t)](\mu\lambda|\sigma\nu) + \sum_{\lambda} \mathrm{Im}[\mathbf{P}_{\lambda\lambda}(t)](\mu\lambda|\lambda\nu) \\ & + \sum_{\sigma<\lambda} \mathrm{Im}[\mathbf{P}_{\sigma\lambda}(t)](\mu\sigma|\lambda\nu) \qquad (62)\end{aligned}$$

and since $\mathbf{P}(t)$ is Hermitian, the imaginary part is antisymmetric and has zeros on the on-diagonal. Therefore

$$\mathrm{Im}[\mathbf{G}^K_{\mu\nu}(t)] = \sum_{\sigma<\lambda} \mathrm{Im}[\mathbf{P}_{\lambda\sigma}(t)][(\mu\lambda|\sigma\nu) - (\mu\sigma|\lambda\nu)] \qquad (63)$$

As before, we examine the symmetry by swapping μ and ν

$$\mathrm{Im}[\mathbf{G}^K_{\nu\mu}(t)] = \sum_{\sigma<\lambda} \mathrm{Im}[\mathbf{P}_{\lambda\sigma}(t)][(\nu\lambda|\sigma\mu) - (\nu\sigma|\lambda\mu)] \qquad (64)$$

$$= \sum_{\sigma<\lambda} \mathrm{Im}[\mathbf{P}_{\lambda\sigma}(t)][(\mu\sigma|\lambda\nu) - (\mu\lambda|\sigma\nu)] \qquad (65)$$

$$= -\mathrm{Im}[\mathbf{G}^K_{\mu\nu}(t)] \qquad (66)$$

where again we used the permutation of the two-electron integrals. Therefore, *the exchange term is complex, Hermitian, and depends on the full complex density matrix.*

AUTHOR INFORMATION

Corresponding Author

*E-mail: kenneth.lopata@pnl.gov; niri.govind@pnl.gov.

ACKNOWLEDGMENT

The research was performed using EMSL, a national scientific user facility sponsored by the U.S. Department of Energy's Office of Biological and Environmental Research and located at Pacific Northwest National Laboratory (PNNL). PNNL is operated for the Department of Energy by the Battelle Memorial Institute under Contract DE-AC06-76RLO-1830. KL acknowledges the William Wiley Postdoctoral Fellowship from EMSL. Discussions with Daniel Neuhauser, Karol Kowalski, Hongfei Wang, Eric Bylaska, Bert de Jong, and Bill Shelton are gratefully acknowledged.

REFERENCES

(1) Runge, E.; Gross, E. K. U. *Phys. Rev. Lett.* **1984**, *52*, 997.

(2) Petersilka, M.; Gossmann, U. J.; Gross, E. K. U. *Phys. Rev. Lett.* **1996**, *76*, 1212–1215.

(3) Casida, M. E. In *Recent advances in density functional methods*; Chong, D. P., Ed.; World Scientific Publishing: River Edge, NJ, 1995; Vol. 1, Chapter 5, pp 155−192.

(4) Burke, K.; Werschnik, J.; Gross, E. K. U. *J. Chem. Phys.* **2005**, *123*, 062206.

(5) Harumiya, K.; Kawata, I.; Kono, H.; Fujimura, Y. *J. Chem. Phys.* **2000**, *113*, 8953–8960.

(6) Krause, P.; Klamroth, T.; Saalfrank, P. *J. Chem. Phys.* **2005**, *123*, 074105.

(7) Greenman, L.; Ho, P. J.; Pabst, S.; Kamarchik, E.; Mazziotti, D. A.; Santra, R. *Phys. Rev. A* **2010**, *82*, 023406.

(8) Deb, B.; Ghosh, S. *J. Chem. Phys.* **1982**, *77*, 342.

(9) Domps, A.; Reinhard, P.-G.; Suraud, E. *Phys. Rev. Lett.* **1998**, *80*, 5520–5523.

(10) Kulander, K. C. *Phys. Rev. A* **1987**, *36*, 2726–2738.

(11) Li, X.; Smith, S.; Markevitch, A.; Romanov, D.; Levis, R.; Schlegel, H. *Phys. Chem. Chem. Phys.* **2005**, *7*, 233–239.

(12) Bartell, L.; Wall, M.; Neuhauser, D. *J. Chem. Phys.* **2010**, *132*, 234106.

(13) Theilhaber, J. *Phys. Rev. B* **1992**, *46*, 12990–13003.

(14) Yabana, K.; Bertsch, G. F. *Phys. Rev. B* **1996**, *54*, 4484–4487.

(15) Baer, R.; Gould, R. *J. Chem. Phys.* **2001**, *114*, 3385–3392.

(16) Pi, M.; Ancilotto, F.; Lipparini, E.; Mayol, R. *Physica E* **2004**, *24*, 297–307.

(17) Castro, A.; Appel, H.; Oliveira, M.; Rozzi, C. A.; Andrade, X.; Lorenzen, F.; Marques, M. A. L.; Gross, E. K. U.; Rubio, A. *Phys. Status Solidi B* **2006**, *243*, 2465–2488.

(18) Sugino, O.; Miyamoto, Y. *Phys. Rev. B* **1999**, *59*, 2579–2586.

(19) Baer, R.; Neuhauser, D.; Weiss, S. *Nano Lett.* **2004**, *4*, 85–88.

(20) Qian, X.; Li, J.; Lin, X.; Yip, S. *Phys. Rev. B* **2006**, *73*, 035408.

(21) Soler, J.; Artacho, E.; Gale, J.; García, A.; Junquera, J.; Ordejón, P.; Sánchez-Portal, D. *J. Phys.: Condens. Mater.* **2002**, *14*, 2745.

(22) Tsolakidis, A.; Sánchez-Portal, D.; Martin, R. M. *Phys. Rev. B* **2002**, *66*, 235416.

(23) Takimoto, Y.; Vila, F. D.; Rehr, J. J. *J. Chem. Phys.* **2007**, *127*, 154114.

(24) Meng, S.; Kaxiras, E. *J. Chem. Phys.* **2008**, *129*, 054110.

(25) Niehaus, T. A.; Suhai, S.; Della Sala, F.; Lugli, P.; Elstner, M.; Seifert, G.; Frauenheim, T. *Phys. Rev. B* **2001**, *63*, 085108.

(26) Wang, F.; Yam, C. Y.; Chen, G.; Wang, X.; Fan, K.; Niehaus, T. A.; Frauenheim, T. *Phys. Rev. B* **2007**, *76*, 045114.

(27) Cheng, C.-L.; Evans, J. S.; Van Voorhis, T. *Phys. Rev. B* **2006**, *74*, 155112.

(28) Evans, J. S.; Voorhis, T. V. *Nano Lett.* **2009**, *9*, 2671–2675.

(29) Li, X.; Tully, J. C. *Chem. Phys. Lett.* **2007**, *439*, 199–203.

(30) Sun, J.; Song, J.; Zhao, Y.; Liang, W.-Z. *J. Chem. Phys.* **2007**, *127*, 234107.

(31) Isborn, C. M.; Li, X. *J. Chem. Phys.* **2008**, *129*, 204107.

(32) Isborn, C. M.; Li, X. *J. Chem. Theory Comput.* **2009**, *5*, 2415–2419.

(33) Micha, D. A.; Runge, K. *Phys. Rev. A* **1994**, *50*, 322–336.

(34) Li, X.; Tully, J. C.; Schlegel, H. B.; Frisch, M. J. *J. Chem. Phys.* **2005**, *123*, 084106.

(35) Livshits, E.; Baer, R. *J. Phys. Chem. A* **2006**, *110*, 8443–8450.

(36) Jakowski, J.; Morokuma, K. *J. Chem. Phys.* **2009**, *130*, 224106.

(37) Tully, J. C. *J. Chem. Phys.* **1990**, *93*, 1061–1071.

(38) Craig, C. F.; Duncan, W. R.; Prezhdo, O. V. *Phys. Rev. Lett.* **2005**, *95*, 163001.

(39) Duncan, W. R.; Stier, W. M.; Prezhdo, O. V. *J. Am. Chem. Soc.* **2005**, *127*, 7941–7951.

(40) Zhang, X.; Li, Z.; Lu, G. *Phys. Rev. B* **2010**, *82*, 205210.

(41) Stella, L.; Meister, M.; Fisher, A. J.; Horsfield, A. P. *J. Chem. Phys.* **2007**, *127*, 214104.

(42) Valiev, M.; Bylaska, E.; Govind, N.; Kowalski, K.; Straatsma, T.; Dam, H. V.; Wang, D.; Nieplocha, J.; Apra, E.; Windus, T.; de Jong, W. *Comput. Phys. Commun.* **2010**, *181*, 1477–1489.

(43) Baer, R.; Neuhauser, D. *Phys. Rev. Lett.* **2005**, *94*, 043002.

(44) Livshits, E.; Baer, R. *Phys. Chem. Chem. Phys.* **2007**, *9*, 2932–2941.

(45) Baer, R. *THEOCHEM* **2009**, *914*, 19–21.

(46) Baer, R.; Neuhauser, D. *J. Chem. Phys.* **2004**, *121*, 9803–9807.

(47) Yabana, K.; Nakatsukasa, T.; Iwata, J.; Bertsch, G. *Phys. Status Solidi B* **2006**, *243*, 1121–1138.

(48) Neuhauser, D.; Lopata, K. *J. Chem. Phys.* **2008**, *129*, 134106.

(49) Becke, A. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(50) Dreuw, A.; Weisman, J.; Head-Gordon, M. *J. Chem. Phys.* **2003**, *119*, 2943.

(51) Vydrov, O.; Heyd, J.; Krukau, A.; Scuseria, G. *J. Chem. Phys.* **2006**, *125*, 074106.

(52) Jacquemin, D.; Perpète, E. A.; Scuseria, G. E.; Ciofini, I.; Adamo, C. *Chem. Phys. Lett.* **2008**, *465*, 226–229.

(53) Yanai, T.; Tew, D. P.; Handy, N. C. *Chem. Phys. Lett.* **2004**, *393*, 51–57.

(54) Govind, N.; Valiev, M.; Jensen, L.; Kowalski, K. *J. Phys. Chem. A* **2009**, *113*, 6041–6043.

(55) (a) Jensen, L.; Govind, N. *J. Phys. Chem. A* **2009**, *113*, 9761–9765. (b) Jensen, L.; Govind, N. *J. Phys. Chem. A* **2009**, *113*, 11095–11095.

(56) Baer, R.; Livshits, E.; Salzner, U. *Annu. Rev. Phys. Chem.* **2010**, *61*, 85–109.

(57) Kosloff, R. *J. Phys. Chem.* **1988**, *92*, 2087–2100.

(58) Liu, J.; Guo, Z.; Sun, J.; Liang, W. *Front. Chem. China* **2010**, *5*, 11–28.

(59) Castro, A.; Marques, M. A. L.; Rubio, A. *J. Chem. Phys.* **2004**, *121*, 3425–3433.

(60) Magnus, W. *Commun. Pure Appl. Math.* **1954**, *7*, 649–673.

(61) de Jong, W. A.; Bylaska, E.; Govind, N.; Janssen, C. L.; Kowalski, K.; Müller, T.; Nielsen, I. M. B.; van Dam, H. J. J.; Veryazov, V.; Lindh, R. *Phys. Chem. Chem. Phys.* **2010**, *12*, 6896–6920.

(62) Eichkorn, K.; Treutler, O.; Ošhm, H.; Hašser, M.; Ahlrichs, R. *Chem. Phys. Lett.* **1995**, *240*, 283–289.

(63) EMSL Basis Set Exchange. https://bse.pnl.gov (accessed Dec 10, 2010).

(64) Tretiak, S.; Isborn, C. M.; Niklasson, A. M. N.; Challacombe, M. *J. Chem. Phys.* **2009**, *130*, 054111.

(65) Tsien, R. *Annu. Rev. Biochem.* **1998**, *67*, 509–544.

(66) Dong, J.; Solntsev, K. M.; Tolbert, L. M. *J. Am. Chem. Soc.* **2006**, *128*, 12038–12039.

(67) Kowalski, K.; Krishnamoorthy, S.; Villa, O.; Hammond, J.; Govind, N. *J. Chem. Phys.* **2010**, *132*, 154103.

(68) Sadlej, A. *Collect. Czech. Chem. Commun.* **1988**, *53*, 1995–2016.

(69) Mozer, A. J.; Griffith, M. J.; Tsekouras, G.; Wagner, P.; Wallace, G. G.; Mori, S.; Sunahara, K.; Miyashita, M.; Earles, J. C.; Gordon, K. C.; Du, L.; Katoh, R.; Furube, A.; Officer, D. L. *J. Am. Chem. Soc.* **2009**, *131*, 15621–15623.

(70) Maligaspe, E.; Sandanayaka, A. S. D.; Hasobe, T.; Ito, O.; D'Souza, F. *J. Am. Chem. Soc.* **2010**, *132*, 8158–8164.

(71) Palummo, M.; Hogan, C.; Sottile, F.; Bagalá, P.; Rubio, A. *J. Chem. Phys.* **2009**, *131*, 084102.

(72) Iikura, H.; Tsuneda, T.; Yanai, T.; Hirao, K. *J. Chem. Phys.* **2001**, *115*, 3540–3544.

(73) Samoylova, E.; Schultz, T.; Hertel, I.; Radloff, W. *Chem. Phys.* **2008**, *347*, 376–382.

(74) Szabo, A.; Ostlund, N. *Modern Quantum Chemistry*; Dover Publications: Mineola, NY, 1996; pp 142−145.

(75) *Blender*; The Blender Foundation: Amsterdam, The Netherlands, 2010.

JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Computational Mechanistic Studies Addressed to the Transimination Reaction Present in All Pyridoxal 5′-Phosphate-Requiring Enzymes

N. M. F. S. A. Cerqueira, P. A. Fernandes, and M. J. Ramos*

REQUIMTE, Faculdade de Ciências, Universidade do Porto, Rua Campo Alegre, s/n, 4169-007 Porto, Portugal

ABSTRACT: The pyridoxal-5′-phosphate-dependent enzymes (PLP enzymes) catalyze a myriad of biochemical reactions, being actively involved in the biosynthesis of amino acids and amino acid-derived metabolites as well as in the biosynthetic pathways of amino sugars and in the synthesis or catabolism of neurotransmitters. Although the scope of PLP-catalyzed reactions initially appears to be bewilderingly diverse, there is a simple unifying principle: In the resting state, the cofactor (PLP) is covalently bonded to the amino group of an active site lysine, forming an internal aldimine. Once the amino substrate interacts with the active site, a new Schiff base is generated, commonly referred to as the external aldimine. Only after this step, the mechanistic pathway for each PLP-catalyzed reaction diverges. In this paper, density functional methods have been applied to investigate this common step present in all PLP-dependent enzymes—the transimination reaction. The results indicate that the reaction involves three sequential steps: (i) formation of a tetrahedral intermediate with the active site lysine and the amino substrate bonded to the PLP cofactor; (ii) nondirect proton transfer between the amino substrate and the lysine residue; and (iii) formation of the external aldimine after the dissociation of the lysine residue. The overall reaction is exothermic (−12.0 kcal/mol), and the rate-limiting step is the second one with 12.6 kcal/mol for the activation energy.

1. INTRODUCTION

Pyridoxal 5′-phosphate (PLP) is a derivative of vitamin B6 and acts as a cofactor in a myriad of chemical reactions involving amino acids.[1] The enzyme commission (EC) has already catalogued more than 150 distinct enzymatic activities of this type of enzyme, which includes decarboxylations, racemisations, transiminations, retro−aldo cleavages, and β or γ eliminations.

The study of PLP enzymes is one of the most fascinating frontiers in enzymology, not only because of their unrivaled versatility as catalysts but also because they are involved in many cellular processes.[2] Their importance is further underscored by the number of receptors that have been identified as drug targets. For example, inhibitors of γ-aminobutyric acid aminotransferase (GABA ATase) are used in the treatment of epilepsy,[3] serine hydroxyl methyl transferase (SHMT) has been identified as a target for cancer therapy,[4] and inhibitors of ornithine decarboxylase (ODC) are employed in the treatment of African sleeping sickness.[5] Functional defects in PLP enzymes have also been implicated in several pathologies, such as homocystinuria.[6] Understanding the function of this important group of enzymes has thus become an important milestone in medicine and biorelated research areas to develop new molecules capable of impairing enzymatic activity and especially to design improved protein-based catalysts.

A comprehensive understanding of PLP-related enzymes or even their classification in different families is not a straightforward task, because beyond the wide variety of reactions that they can catalyze, these enzymes have diverse quaternary structures. Some enzymes can be found active as monomers, others as dimers, and some of them as tetramers or even hexamers.[7,8] Although the scope of the PLP-catalyzed reactions initially appears to be bewilderingly diverse, there is a simple unifying principle: In the resting state the cofactor is covalently attached to the ε-amino group of an active site lysine, forming an internal aldimine. Once the α-amino substrate interacts with the active site, the lysine residue dissociates from PLP, and the substrate becomes covalently bonded to it, generating a new Schiff base with PLP. This intermediate is commonly called the external aldimine. Only after this step, the mechanistic pathway for each PLP-catalyzed reaction diverges as it is depicted in Scheme 1.

The conversion between the internal (lysine PLP-imine) and external aldimine (substrate PLP-imine) has therefore a preponderant role in all PLP-related enzymes, and it is a crucial step in their activation. Moreover, as the inverse reaction is required for the enzymatic turnover, this reaction is critical toward the overall activity of these enzymes.

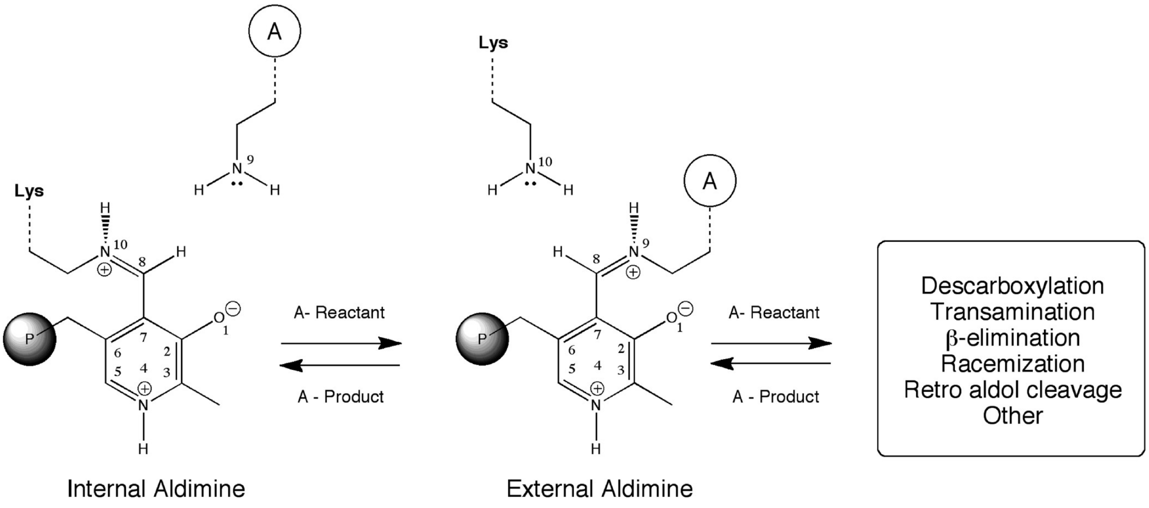
Experimental studies have provided some clues regarding this step. All PLP intermediates (internal and external aldimines) have distinct absorption bands and change only by a few nanometers from system to system, which makes their identification and characterization straightforward.[9,10]

The postulated mechanism proposes that at physiological conditions, the imine linkage of the internal aldimine is protonated in order to form a more electrophilic and reactive iminium ion (Scheme 1). Upon amino substrate binding, it is suggested that the nitrogen N10 of the substrate attacks carbon C8 of the internal aldimine. Subsequently, the lysine residue dissociates from the PLP endorsing the formation of the external aldimine. During this process, it is proposed that the dissociation of the lysine residues and the attachment of the amino acid substrate is not accomplished in a single step. Instead, it might involve the formation of a transient but stable geminal diamine intermediate.

Received: April 26, 2010
Published: March 29, 2011

dx.doi.org/10.1021/ct1002219 | *J. Chem. Theory Comput.* 2011, 7, 1356–1368

Scheme 1. Currently Accepted Mechanism for the Transimination Reaction: A Common Step in the Catalytic Mechanism of All PLP-Dependent Enzymes[a]



[a] P represents the phosphate group and A an α-amino substrate.

The presence of a geminal diamine intermediate is assumed to be common in most PLP-dependent enzymes. However, there is some controversy about the mechanism by which the transimination reaction occurs. Snell and Jenkins were the first ones to propose the formation of a geminal diamine intermediate.[11] This mechanism was supported by further experimental work.[12,13]

Other studies suggest the involvement of a two-fold addition−elimination type of mechanism instead[14,15] or even the possible involvement of the phosphate group of PLP in the proton shuttle mechanism.[14,7] However, the involvement of the 5′-phosphate of PLP in the transimination process was later on discarded, because this group is buried in the protein and cannot interact directly with the region where the transimination reaction occurs. Several studies proposed that the main function of this group is to act as an anchor to hold the PLP group in the active site.[16] This is rather important after the formation of the external aldimine, in which the PLP group becomes disconnected from the enzyme (it was earlier bonded through Lys69), and therefore it requires some sort of interaction that maintains it bonded to the active site and aligned in a proper orientation for effective catalysis. Yet, there are some PLP-dependent enzymes in which the 5′-phosphate group is indeed involved in the catalysis, as it is the case of the enzyme glucogen phosphorylase[17] and GDP-4-keto-deoxymannose-3-dehydratase (CoID).[18] However, in those cases the substrates are not amino acid-related molecules, and therefore the transimination reaction does not occur.

All the other mechanisms remain as open possibilities, and several studies were conducted aiming to elucidate the most favorable pathway for the transimination reaction present in the PLP-dependent enzymes.

The high rate complexity of the enzymatic transimination has made it very difficult to study the transimination reaction by experimental means. Even so several works have supported the presence of the geminal diamine intermediate in the transimination reaction.[19−22] However, the correct mechanism that could elucidate its formation and involvement in the transimination reaction is still poorly understood. During the past decade, several computational studies have emerged that have tried to explain the transimination reaction at the atomic level detail. The studies performed by Muñoz et al.[23] were pioneers in this field and revealed that the conversion between the internal aldimine into the external aldmine requires the direct participation of at least one water molecule. In this proposal, it is suggested that the phenolate oxygen of the PLP should receive a proton from the amino substrate (through the water molecule) to favor the formation of the geminal diamine intermediate.

Another study, performed by Zhao et al., proposes that the transimination reaction occurs through the direct proton transfer between both amino groups that are bonded to the PLP. However, prior to the formation of the geminal diamine intermediate, it is required the direct participation of the 5′-phosphate group of PLP.[24] Despite the enormous interest that this reaction has received, none of these studies have explored the participation of key active site amino acids in the transimination reaction.

In fact, analyzing several X-ray structures, we found that near the PLP cofactor there are two conserved residues that are capable of catalyzing this reaction, i.e., Cys360 and Tyr389 (considering the PDB code 2OO0). Both residues are pointing to the place in which the transimination reaction should occur. In addition, there is one water molecule that is pointing to the same reaction spot and can favor the transimination reaction without the direct involvement of the phenolic oxygen of the PLP, as it has been proposed by Muñoz et al.[25] In this paper, we explore all these possibilities using quantum mechanics calculations in order to enhance the knowledge about the transimination reaction and highlight the most favorable mechanism involved in this process.

2. METHODOLOGY

1. Model. The model system used in this work was based on the X-ray structure 2OO0 determined by Dufe et al.,[26] which contains the human ornithine decarboxylase. This structure has a good resolution (1.9 Å) but lacks the substrate inside the active site. To acquire this information, we superimposed the X-ray structure 1F3T, which contains the ornithine descarboxylase from *Trypanosoma brucei* complexed with the putrescine (the ODC's reaction product) in the binding pocket.[27] Near the active site region, the two structures can be almost superimposed,

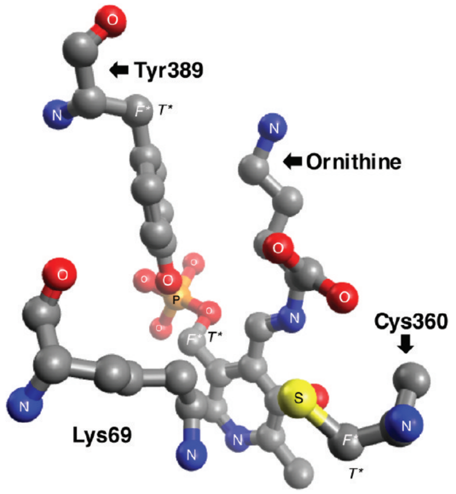


Figure 1. Active site model taken from the PDB structure 2OO0 used. All the residues that were used in this study are shown in ball and sticks. The substrate was built based on the X-ray structure 1F3T that contains the product of the reaction. The atoms marked with F* were kept frozen during the geometry optimizations, and the symbol T* highlights the place where the truncation of the residues took place.

allowing us to model the correct position of the substrate inside the active site of the X-ray structure 2OO0. The reactant of the reaction was subsequently modeled substituting the hydrogen atom, which is attached to the carbon atom that is closer to the PLP ring, by a carboxyl group.

This model was subsequently simplified, eliminating all the amino acids that do not interact directly with the PLP-imine. The final model contained the PLP, the substrate, Lys69, Cys360, and Tyr389 (Figure 1).

The selected amino acids were initially truncated at the α carbon. However, in order to maintain the net of hydrogen bonds within the residues of the model, we decided to keep the main chain of each residue and protonate the carboxylate and the amino groups. The calculations that we have performed have shown that this latter approach turned out to be more satisfactory than the reverse, because it improved the robustness of the model and, in some cases, resulted in the reduction of the activation energy by 3 kcal/mol and the reaction barrier by 1 kcal/mol.

In order to simplify the model, we decided to substitute the phosphate group of PLP by a methyl group. We have chosen the methyl group instead of a hydroxyl group because the first one ensures and maintains the stability of the active site, without requiring the inclusion of additional residues. Nevertheless, the differences between the barriers obtained for the first step of the studied mechanism were very similar in both cases [Ea = 5.3 and Er = 2.5 kcal/mol vs Ea = 4.5 and Er = 1.9 kcal/mol (values obtained with DFT6-31G(d))].

In spite of all simplifications the model had in total 108 atoms (Figure 1). This model was then subjected to geometry optimizations. To keep the optimized structures close to the X-ray structure, some atoms were kept frozen as depicted in Figure 1.

2. Methods. Density functional theory (DFT) calculations were performed with the Gaussian09 software package.[28] All structures were fully optimized and characterized both at the B3LYP level[29−32] and using the new hybrid exchange correlation functional proposed by Zhao and Truhlar, M06[33,34] together with the 6-31G(d) basis set. The 6-31+G(d,p) basis set was also used to optimize all geometries of the rate-limiting step and analyze the effect of polarization and diffusion functions in the hydrogen atoms. In all geometry optimizations, we first searched for the transition state starting from the reactants. This was obtained with a scan in which the reaction coordinate that we were interested in was shortened or stretched. The transition states were subsequently fully geometry optimized, starting from the structure of the higher energy point of the scans. The reactants and the products, associated with it, were determined through internal reaction coordinate (IRC) calculations. In all cases, the geometry optimizations and the stationary points were obtained with standard Gaussian convergence criteria. The transition-state structures were all verified by vibrational frequency calculations, having exactly one imaginary frequency with the correct transition vector, even using frozen atoms, which shows that the frozen atoms were almost free from steric strain.

The final electronic energies were calculated using the all-electron 6-311++G(3df,2pd) basis set, using the functionals B3LYP, M06, and M06-2X. These structures were the optimized geometries obtained with the M06/6-31G(d) level of theory. Zero-point corrections, thermal, and entropic effects ($T = 310.15$ K, $P = 1$ bar) were added to all calculated energies, with the 6-31G(d) basis set. To estimate the solvation effects of the rest of the enzyme, single point calculations on the optimized geometries were performed with IEF-PCM, as implemented in Gaussian 09,[28] with the 6-311++G(3df,2pd) basis set. This feature is of particular importance to the study of enzymatic catalysis because the use of a continuum model is normally taken as an approximation to the effect of the global enzyme environment in a reaction. A dielectric constant of $\varepsilon = 4$ was chosen to describe the protein environment of the active site in agreement with previous suggestions.[35−38] The atomic spin density distributions were calculated at the M06 level employing a Mulliken population analysis, using the 6-31G(d) basis set.

3. RESULTS AND DISCUSSION

This section will be divided in two main parts. In the first part, we will discuss the most favorable pathway for the transimination process at the M06-2X/6-311++G(3df,2pd)//M06/6-31G(d) level. In the second part, we will discuss a small benchmarking study that was performed to compare the results that were obtained with the very popular B3LYP and two hybrid meta exchange−correlation functionals of Zhao and Truhlar, M06 and M06-2X.

3.1. Transimination Mechanism. Our first task in this study was to review all the available experimental data concerning this subject. In this analysis, we found three interesting X-ray structures of the enzyme ornithine decarboxylase (enzyme that catalyzes the decarboxylation of ornithine to putrescine). Each PDB reveals the PLP in three different states that correspond to snapshots of the active site during the initial steps of the catalytic process (Figure 2). The first PDB structure (PDB code: 7ODC)[31] clearly shows a lysine residue bonded to the PLP cofactor. This should correspond to the initial state of the enzyme (internal aldimine) and is a common characteristic of all PLP-dependent enzymes. The PDB structure 1F3T[20] shows the lysine residue dissociated from the PLP cofactor. This structure resembles what should be found in the external aldimine but with the putrescine in place of the ornithine.

The active site region of the mutant PDB structure 1SZR[13] shows a structure where the PLP cofactor is bound simultaneously to the lysine residue and to the amino substrate. This

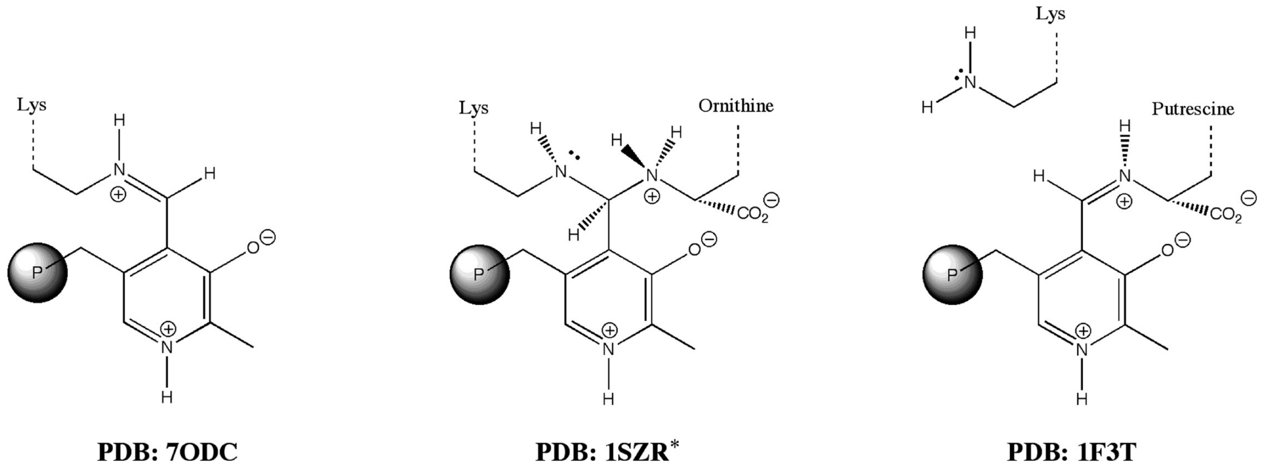
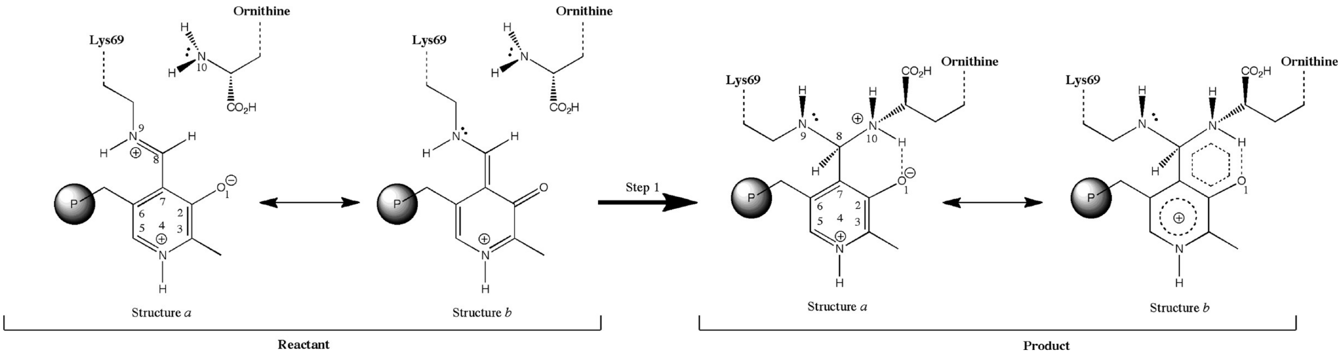


Figure 2. Active site topology of three PDB files of ornithine decarboxylase available in the protein databank (* mutant PDB structure of ODC).

Scheme 2. First Step of the Transimination Reaction of PLP-Dependent Enzymes[a]



[a] P stands for a phosphate group.

structure suggests that the transimination reaction might not occur in a single step but rather in three subsequent steps, i.e.,: (i) formation of the tetrahedral intermediate with the lysine and the ornithine bonded to the PLP cofactor; (ii) proton transfer between both amino substrates; and (iii) subsequent dissociation of the lysine residue with the concomitant formation of the external aldimine.

The enzyme ornithine descarboxylase (ODC) was used in the subsequent sections as a model of all PLP-dependent enzymes. Since the transimination reaction is a common feature in all PLP-dependent enzymes, the mechanism described in the following sections can be transferred unequivocally to any PLP-dependent enzyme.

Step 1: Formation of the Tetrahedral Geminal−Diamine Intermediate. Taking into account the mutant X-ray structure 1SZR we have tested in this step whether the formation of the geminal−diamine intermediate occurs, and if such an intermediate is a stable compound (Scheme 2).

The computational results have shown that once the ornithine enters the active site (the amino substrate), it interacts with the PLP nearby carbon C8 and oxygen O1 (2.60 Å).

The optimized structure of the reactant indicates that in the initial state, Lys69 remains tightly bonded to carbon C8 of the PLP through the NH group (1.33 Å). The bond length between atoms C8 and N9 is characteristic of secondary amines (1.31 Å), which means that the lysine residue is attached to PLP by a single bond (bond length$_{C=N}$ ∼ 1.28 Å). This is also emphasized by the covalent nature of the double bond between atoms carbon C2 (0.40 au) and oxygen O1 (−0.75 au), characteristic of the carbonyl group (1.25 vs 1.23 Å). All these results allow us to conclude that resonance structure b from scheme 2 is the one that describes better what must be found in the reactants of this reaction.

The positive charge of the system is mainly distributed along the extended π systems of the PLP ring (0.86 au). It is interesting to note that the charge of carbon C8 is slightly more positive (0.36 au) than the other atoms of the PLP ring (on average ∼ 0.28 au). This effect turns carbon C8 more prone to accept the nucleophilic attack of the amino substrate, therefore, favoring the reaction. This result is in agreement with previous suggestions[11] that point to the fact that atom C8 of PLP becomes bonded to the α-amino group of the substrate.

The transition-state structure (Figure 3) of this reaction is characterized by an imaginary frequency of 165i cm^{-1}. The optimized structure indicates that the ε-amino group of Lys69 remains tightly bonded to carbon C8 of the PLP (1.35 Å). The α-amino group of ornithine comes closer to the same center (1.92 Å), and carbon C8 changes its hybridization from sp^2 to sp^3, remaining slightly positively charged (0.39 au). The formation of such a tetrahedral intermediate is favored by the pyridine ring that acts as an electron sink, thus allowing to concentrate the excess of negative charge around oxygen O1 (−0.64 au) and to maintain the electropositive nature of carbon C8 (0.40 au).

In the optimized geometry of the products, both amino groups become covalently bound to carbon C8 of the PLP, generating a stable geminal−diamine intermediate, as it was observed in the mutant protein 1SZR[14]. The bond length between the nitrogen N9 of Lys69 and PLP slightly elongates to 1.41 Å (1.33 Å before), while the distance between the nitrogen N10 from the substrate and PLP gets shortened to 1.56 Å. Due to the formation of the tetrahedral intermediate, the bond length between carbons C7 and C8 becomes slightly elongated to 1.53 Å (1.44 Å before), and the same is also true for the bond length between carbon C2 and oxygen O1 (1.25 Å vs 1.23 Å in the reactants). Oxygen atom O1 starts to interact very closely with the proton of the α-amino group of ornithine through a hydrogen bond (1.59 Å), and this creates a pseudoring as depicted in Scheme 2. This rearrangement stabilizes this region, allowing the positive charge of the α-amino group of the substrate and the negative charge of oxygen O1 to spread along the pseudoring (0.14 au).

This reaction is characterized by an activation energy of 10.8 kcal/mol and is exothermic by −5.7 kcal/mol.

An overall evaluation of the reaction allows us to conclude that the PLP group has a central role in this reaction by delocalizing the negative charge through its π system. The charge becomes mainly concentrated at the carbonyl group that simultaneously allows guiding and aligning the amino substrate with carbon C8.

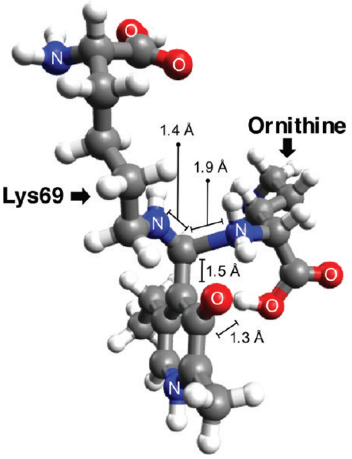


Figure 3. Optimized structure of the transition state of step 1.

These two effects have an important role in the reaction: (i) The first favors the ammoniumcation nature of the carbon located at position 8 (Scheme 2); and (ii) the second enhances the formation of a strong hydrogen bond between oxygen O1 and the amino group of ornithine that helps to align the substrate inside the active site. In the product of this reaction, this strong hydrogen bond is preserved, and it favors the creation of a pseudoring between PLP and the amino substrate that stabilizes the formation of the geminal−diamine intermediate.

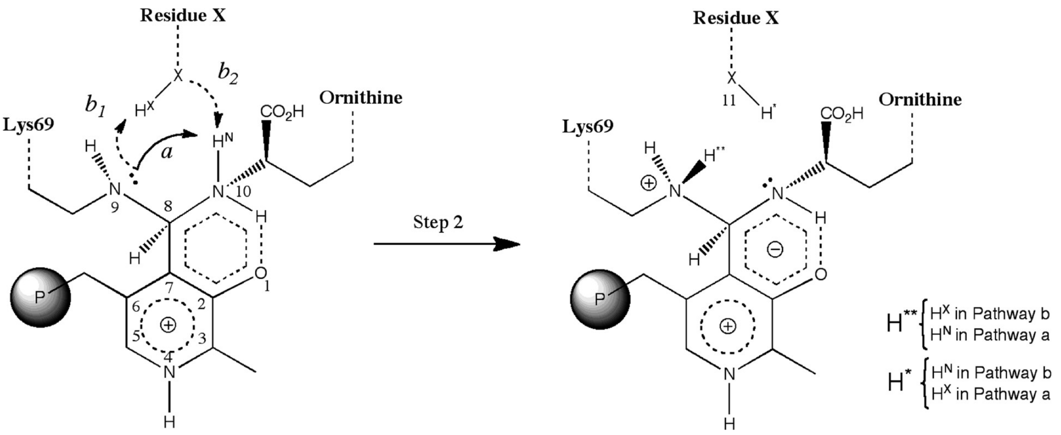
It may be interesting to note that the same type of interaction between the lysine residue and PLP was never observed in our calculations. This occurs because at the beginning of the reaction, the lysine is located in a perpendicular plane to that of the PLP and in the opposite direction of the carbonyl group. This behavior is in agreement with the available X-ray structures as depicted in Figure 4.

Comparing the optimized structure of the tetrahedral geminal diamine obtained by computational means and the one obtained experimentally (X-ray structure 1SZR), we can find many similarities (Figure 4). Both amino groups of Lys69 and ornithine become covalently attached to carbon C8 of the PLP at similar distances (1.41 and 1.56 Å vs 1.31 and 1.50 Å). The same is also true across all the bond lengths from the PLP ring, which emphasizes the equivalence between the theoretical model and the X-ray structure.

It must be said that there is a very good match between both structures presented in Figure 4. The only difference lies on the conformations adopted by Lys69 and ornithine for this study. The model contains all the residues that are directly involved in the reaction (in total they account to 108 atoms) but lacks all the others (much less important) that sometimes help to keep the conformation and orientation observed in the active site of the PDB structure 1SZR. However, this is not a crucial aspect for this reaction, because no active site residue is involved in it. Therefore taking into account the similarity observed between all the bond lengths between carbon C8 of PLP and Lys69 and ornithine, we can conclude that the structures presented in Figure 4 are very similar.

Step 2: Proton Transfer between the Amino Groups of the Tetrahedral Geminal−Diamine Intermediate. To complete the transimination process, Lys69 must dissociate from the PLP cofactor. Because of the formation of the tetrahedral intermediate, this reaction is not straightforward, and a proton must be first transferred from the α-amino group of the substrate to the ε-amino group of Lys69 (Scheme 3).

Scheme 3. Possible Proton-Transfer Pathways (a and b) between the Amino Acid Substrate and the Lysine Residue[a]



[a] Pathway a consists of one step. Pathway b consists of two steps (dashed line): b_1 and b_2 with X = Tyr, Cys, or H_2O.

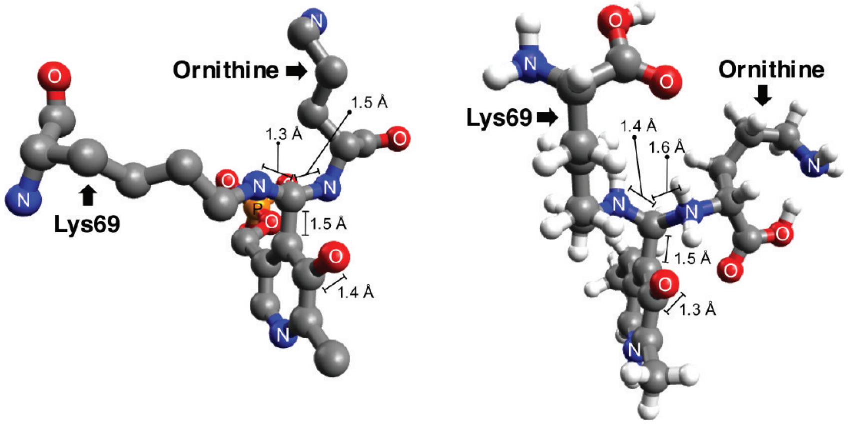


Figure 4. Left: X-ray structure of the PDB entry 1SZR. Right: Optimized structure of the products of the first step (the geminal−diamine intermediate).

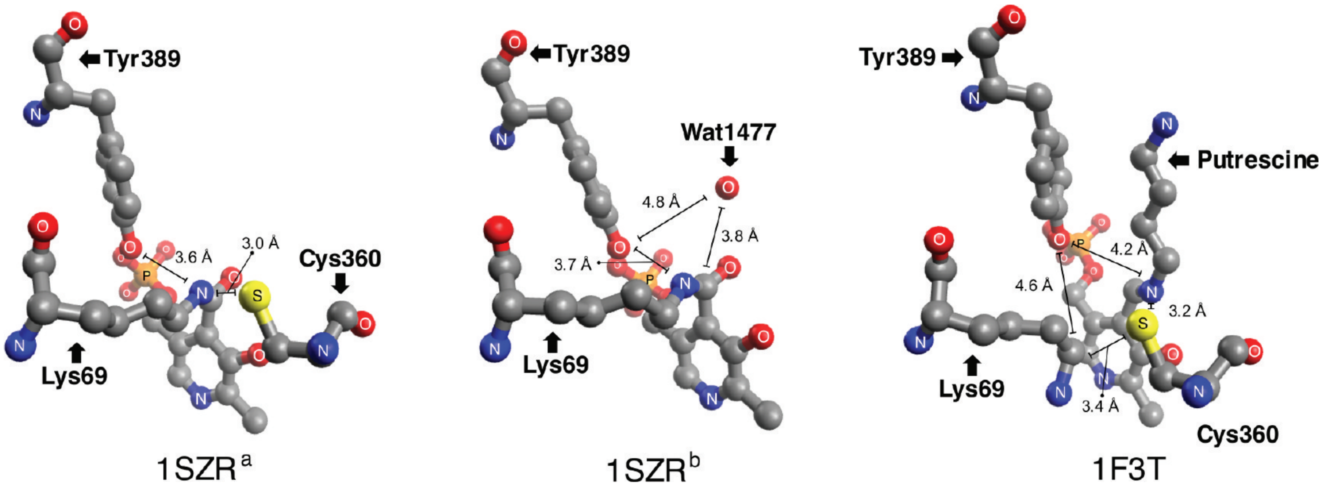


Figure 5. Active site topology of the X-ray structure 1SZR (1SZR^a:PLP from chain D; 1SZR^b:PLP from chain C) and 1F3T (PLP from chain A).

Multiple pathways can be drawn for this step, as is depicted in Scheme 3. The easiest one involves the direct proton transfer from the α-amino group of ornithine (N10) to the ε-amino group of Lys69 (N9). This reaction corresponds to pathway a of Scheme 3. Pathway b of the same scheme, involves the presence of a neighbor and proton donor/acceptor active site residue (residue X in Scheme 3) capable of catalyzing the proton transfer. Two residues can play such role in the mechanism, considering the PDB structures 1SZR and 1F3T of ornithine descarboxylase, and they are Cys360 and Tyr389 (Figure 5). The first residue is at 3.01 Å from carbon C8 of PLP. Tyr389 is not so close to this center (3.58 Å), but the flexibility of the side chain still allows it to interact with both amino groups (figure 5).

Another hypothesis is the involvement of one water molecule in the reaction (Figure 5), similar to what was proposed by Muñoz et al.[23] A closer inspection of the crystal structure 1SZR^b reveals the presence of the water molecule 1477 (1SZR numbering) that is very close to the site where the reaction happens (but different to what was modeled by Muñoz et al.). Moreover, this molecule is stabilized by several residues, such as Asp361, Asp332, and Tyr 389, suggesting that it might be conserved in the active site.

In order to model these reactions and understand which is, from a kinetic and thermodynamic point of view, the most favorable pathway, we have created four scenarios of the active site. Each scenario contains the residues/water molecule capable of catalyzing the proton transfer. The conformation of each residue was carefully chosen beforehand using a set of rotamer libraries and considering the protein surroundings.

Scenario 1 corresponds to pathway a of Scheme 3 and does not contain any residues between the two amino groups (direct pathway). Scenarios 2−4 contain a tyrosine (X = OH), a cysteine (X = SH), and a water molecule (X = H_2O), respectively. These three scenarios will be used to study pathway b of Scheme 3, and the acquired information will allow us to understand which residue is more likely to catalyze the proton transfer.

The optimized structures of the reactants are very similar between each scenario. Both amino groups remain covalently bonded to carbon C8 of the PLP, but while the ε-amino group of Lys69 (−0.25 au) is at 1.40 Å from carbon C8, the α-amino group of ornithine (0.15 au) is at 1.61 Å from the same atom. The discrepancies that are observed in the bond lengths are a consequence of the protonation state of each amino group. The α-amino group from ornithine has two protons, and when it is bonded to carbon C8, it becomes more electropositive (0.12 au), a situation that results in a weaker chemical bond. The ε-amino group of Lys69 has only one proton, and therefore the nitrogen is more electronegative (−0.54 au), a situation that results in a stronger chemical bond between the nitrogen atom and carbon C8 stronger.

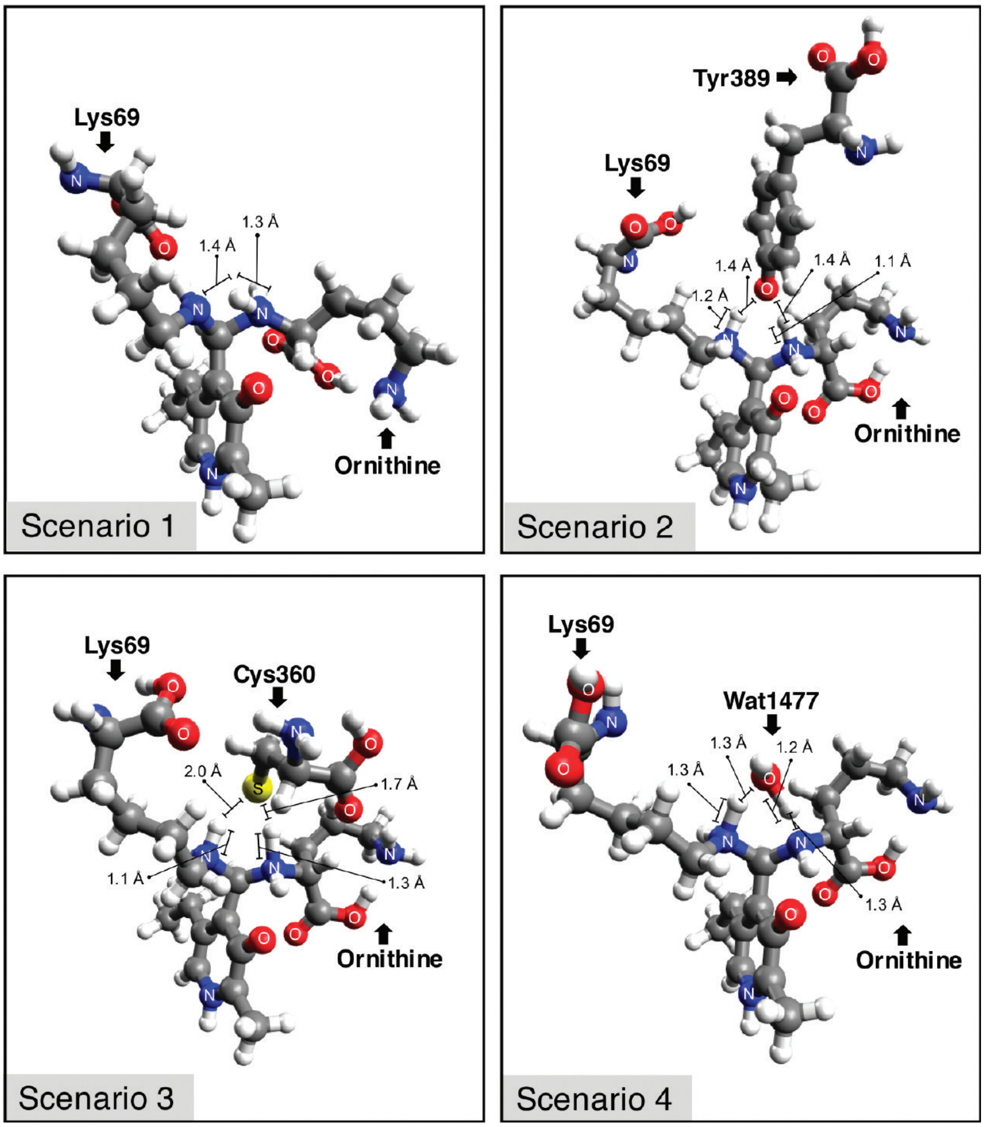


Figure 6. Optimized transition-state structures of step 2 of the transimination reaction of scenario 1 (direct pathway), 2 (catalyzed by Tyr389), 3 (catalyzed by Cys360), and 4 (catalyzed by water1477).

The distance between carbon C7 and carbon C8 of PLP remains stabilized around 1.52 Å, and the positive charge delocalized around the PLP ring (0.85 au), as before. The interaction between oxygen O1 and the α-amino group of ornithine is maintained similarly to what is observed in the products of the first step of the mechanism.

In all scenarios, hydrogen H^N (Scheme 3) of the α-amino group of ornithine is on average 2.25 Å away from the nitrogen atom of the ε-amino group of Lys69. In scenarios 2−4, this hydrogen atom makes an additional hydrogen bond with the sulfur atom of Cys360 (2.69 Å), with the oxygen atom of Tyr389 (2.15 Å), and with the oxygen atom of water molecule 1477 (1.88 Å), respectively.

Hydrogen H^X of the cysteine and tyrosine residues do not establish a hydrogen bond with the nitrogen atom from Lys69, as it would be expected. Such types of interaction only occur when the water molecule is placed between the two amino groups (3.3 Å). The transition state of scenario 1 (Figure 6) is characterized by an imaginary frequency of 1609i cm^{-1}. In this scenario, proton H^N is found halfway between both amino groups (~1.35 Å), and the charge distribution kept the same trend that was observed beforehand.

In scenarios 2 and 3 (Figure 6), the transition states are characterized by imaginary frequencies of −1350 and −1114 cm^{-1}, respectively. Proton H^X gets closer to the nitrogen atom of Lys69 (1.28 Å at scenario 2 and 1.08 Å at scenario 3), whereas hydrogen H^N is halfway between the nitrogen atom of ornithine and the oxygen atom of tyrosine (1.28 and 1.22 Å, respectively) in scenario 2 and the sulfur atom of cysteine (1.29 and 1.70 Å, respectively) in scenario 3. Accordingly, the sulfur atom of Cys360 (scenario 3) becomes more electronegative (−0.27 vs 0.04 au in the reactants), while the oxygen atom of Tyr389 (scenario 2) becomes more electropositive (0.16 vs −0.27 au in the reactants). The charge distribution around the PLP cofactor and the pseudoring remains unchanged in both cases.

The transition state of scenario 4 (Figure 6) is characterized by an imaginary frequency of 1476i cm^{-1}. Both protons, H^N (0.64 au) and H^X (0.60 au) are equally shared between both amino groups and the oxygen atom of the water molecule (1.25 Å on

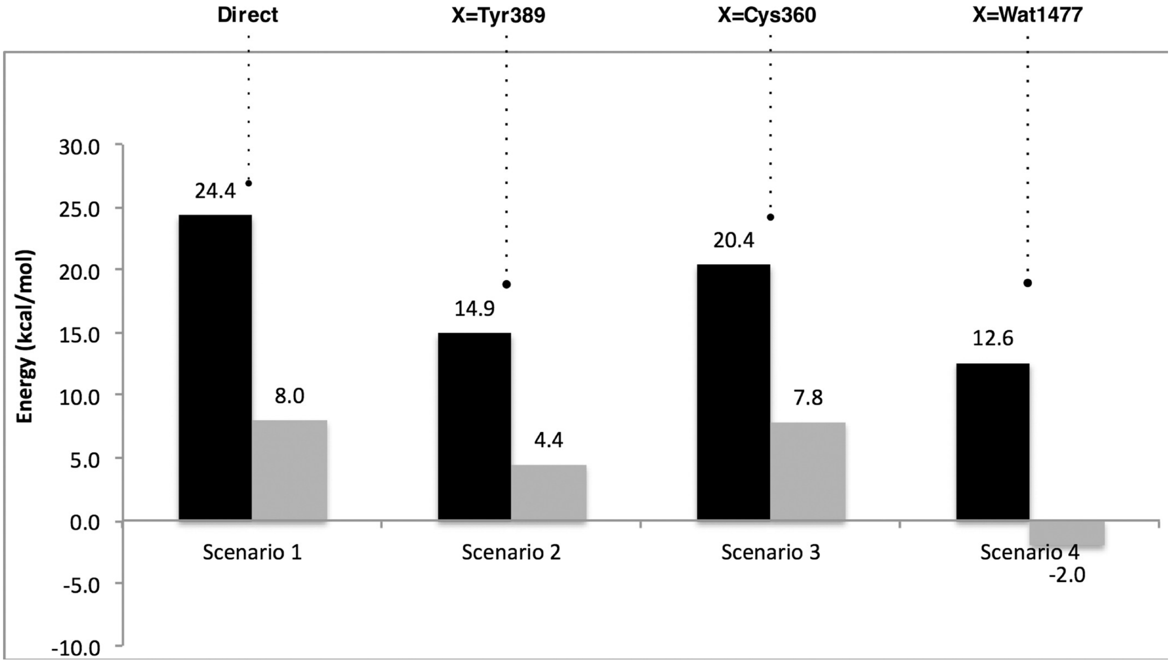


Figure 7. Energies involved in all of the four scenarios used to study the second step of the transimination reaction. Scenario 1: direct proton shuttle between Lys69 and the amino substrate. Scenarios 2−4: model reactions in which Tyr360, Cys360, or Water1477 are involved in the proton transfer between the amino substrate and Lys69, respectively (activation energies are colored in black and the reaction energies in light gray).

average), forming a tightly bound complex. This means that in the transition state of scenario 4, we have a H_3O^+ ion (0.55 au) within both amino groups (−0.70 au). Similarly to what happens in scenarios 2 and 3, the charge distribution around the PLP group and the pseudoring remains unchanged (+0.86 au),

In the product of each reaction, the proton that was previously bound to the α-amino group of ornithine (N10) transfers to the ε-amino group of Lys69 (N9) (H^X in scenarios 2−4 and H^N in scenario 1) (on average 1.03 Å bond length). Both amino groups remain bound to carbon C8 of PLP, but the bond length between the α-amino group of ornithine and carbon C8 of PLP is now shorter (1.42 vs 1.56 Å), whereas the bond length between the ε-amino group of Lys69 and carbon C8 of PLP elongated to 1.6 Å (1.41 Å beforehand). In scenarios 2−4, H^N becomes covalently bonded to the oxygen atom of Tyr389 (0.98 Å), to the sulfur atom of Cys360 (1.35 Å), and to the hydroxyl group of the water molecule (1.00 Å).

Similar to what happens in the reactants and in the transition state, the charge distribution around the PLP ring remains practically unchanged in all scenarios (0.85 au). The charge of the pseudoring in the products of the reaction becomes more electronegative (−0.22 au) than was observed in the reactant and the transition state of this reaction (−0.19 au).

The graphic shown in Figure 7 resumes the energetic profile obtained from all the studied reactions for step 2 of the transimination reaction. The results show that under physiological conditions, all of the studied pathways can occur and efficiently catalyze the proton transfer between both amino groups of the tetrahedral geminal−diamine intermediate.

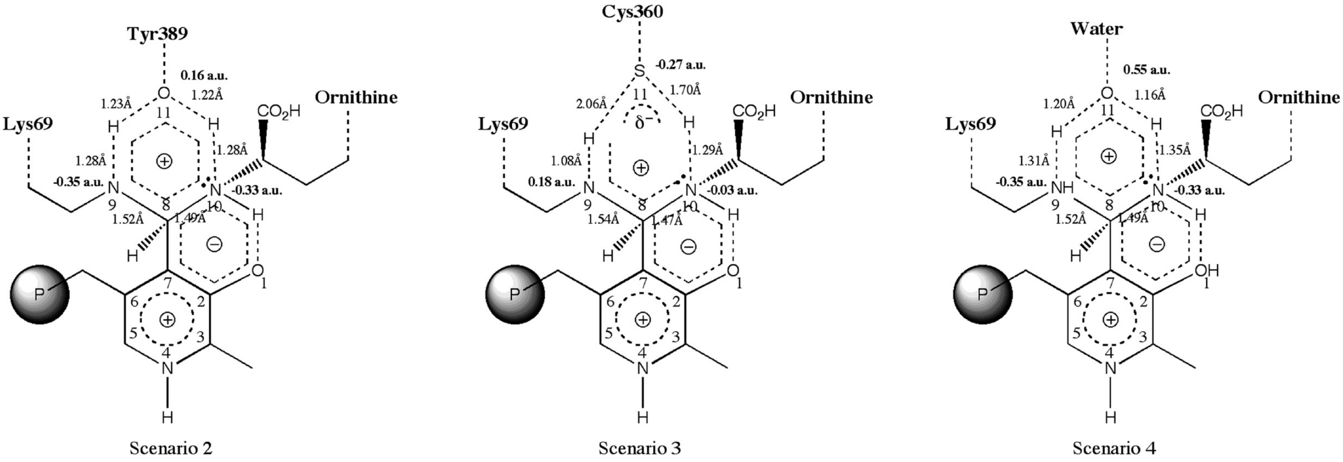
All the studied pathways have one characteristic in common: The reactions are almost thermoneutral, which indicates the feasibility of the reaction in both directions, as it is predicted experimentally.[12] However, some reactions are more favorable than others. Accordingly, the involvement of Tyr389 (scenario 2) or of the water molecule (scenario 4) in the proton transfer tends to decrease considerably the activation energy. The direct proton transfer of pathway a (scenario 1) or the involvement of Cys360 (scenario 3) in the reaction is less favorable, accounting to activations energies higher than 20 kcal/mol.

The energies involved in scenarios 2 and 4 are comparable ($Ea_{scenario2}$ = 14.9 vs $Ea_{scenario4}$ = 12.6 kcal/mol and $Er_{scenario2}$ = 4.4 vs $Er_{scenario4}$ = −2.0 kcal/mol), but it is evident that the reaction is more favored when the water molecule is directly involved in the reaction. This can be explained considering the volume and the flexibility of the water molecule that, when compared to Tyr398, promote a better and closer interaction of both amino groups (~1.2 Å). As a consequence, the difference between the energies of the reactants and the transition state tends to decrease, favoring the reaction. However, the differences observed in these energies are not sufficient to disclose which one should be preferred. Therefore, both of them should be equally capable of catalyzing the proton transfer. This means that if water is available in the active site, then it can catalyze the reaction. Otherwise, Tyr389 catalyzes the reaction without requiring a significant energetic cost.

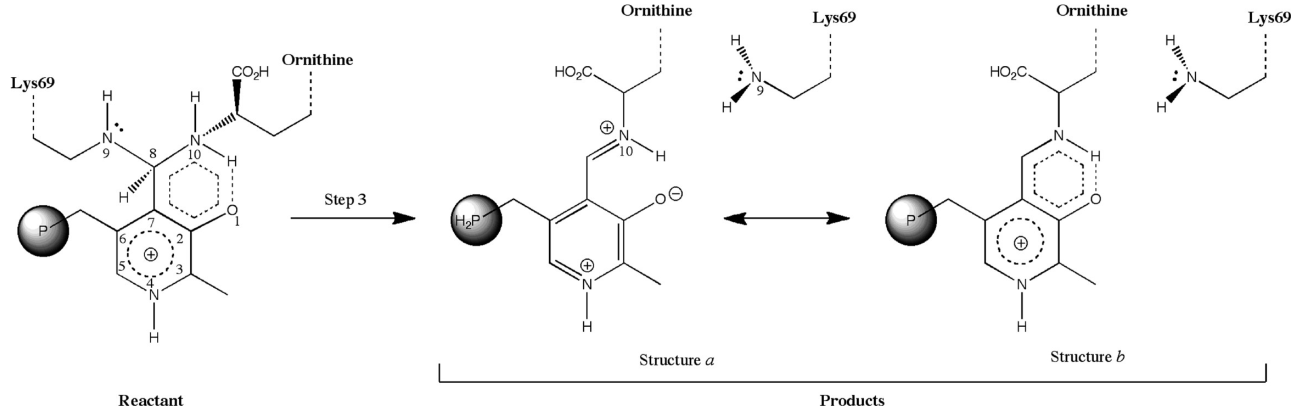
Looking at the pK_a values of the residues involved in the proton transfer, we can understand why Tyr389 and the water molecule are better proton shuttles than Cys360. Under physiological conditions, the pK_a values of a cysteine, a tyrosine, and the water molecule are respectively 8.09, 10.07, and 15.74. This means that Cys360 has a greater tendency to be found in the dissociated form, rather than Tyr389 or the water molecule. This behavior makes Cys360 a good proton donor but not a good proton acceptor. Accordingly, Cys360 is a worst proton shuttle than Tyrs389 and the water molecule.

Looking at the charge distribution and the bond lengths in the optimized models of the transition-state structures, we can conclude basically the same.

Scheme 4. Charge Distribution and Bond Lengths in the Transition State of Scenarios 4 and 3



Scheme 5. Third Step of the Transimination Reaction[a]



[a] P stands for phosphate group.

From Scheme 4, we can see that when the water molecule is used for the proton transfer, the bonds that are created/cleaved are placed within a pseudoring in which the charge is well delocalized (a similar behavior is observed with Typr389). The same does not occur with Cys360. The ring is largely deformed, and the charge is mainly located on the sulfur atom (−0.27 au), while the ε-amino of Lys69 remains slightly positively charged.

These results show that the proton shuttle seems to be only favored when the atom that is involved in the proton transfer can fit between both amino groups and that it can behave as a donor and acceptor of electrons. Only when these conditions are ensured, it is observed the formation of a pseudoring that enhances the delocalization of electrons in order to enhance the proton transfer.

Step 3: Formation of the External Aldimine. The third step of the transimination reaction involves the formation of the external aldimine. This reaction is the reverse of the first step and involves the dissociation of Lys69 from PLP (Scheme 5).

In the reactants, both amino groups remain covalently bonded to carbon C8 of PLP. However, due to the differences in the protonation state of both nitrogen atoms, the bond length between carbon C8 of PLP and the ε-amino group of Lys69 is more stretched (1.60 Å) than the bond length between carbon C8 and the amino group of ornithine (1.42 Å). The full complex retains the tetrahedral geometry around carbon C8 of PLP, and oxygen O1 continues to interact very closely with the proton of the α-amino group through a hydrogen bond (1.92 Å). This rearrangement continues to stabilize this region, allowing the positive charge of the amino group and the negative charge of oxygen O1 to spread along with the atoms of the ring, which in total accounts for ~0.03 au

The transition state of this reaction is characterized by an imaginary frequency of 1618i cm-1 (Figure 8). In this structure the bond length between the α-amino group of ornithine (N10) and carbon C8 of PLP decreases to 1.35 Å, whereas the bond length between the ε-amino group of Lys69 (N9) and carbon C8 elongates to 1.98 Å. The tetrahedral geometry around carbon C8 of PLP breaks up, and ornithine adopts a conformation parallel to the PLP ring and interacts with oxygen atom O1 (1.84 vs 1.91 Å in the reactants). Consequently, the nitrogen atom N9 becomes more electronegative (−0.01 au) than in the reactants (0.19 au). The charge around the pyridine ring and the pseudoring remains unchanged.

In the products of this step, the external aldimine is obtained. The tridimensional structure resembles what is observed in the PDB structure 1F3T that contains a similar intermediate but with putrescine instead of ornithine (Figure 5). Lys69 is now disconnected

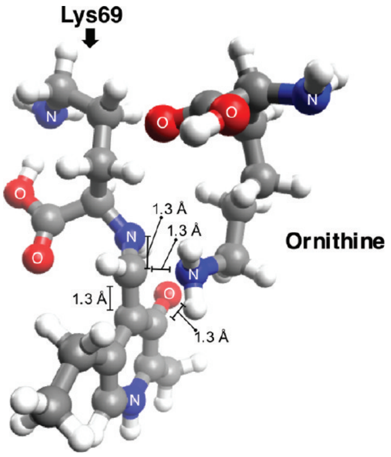


Figure 8. Optimized transition state of the third step of the transimination reaction.

from the PLP cofactor (4.40 Å), while ornithine becomes tightly bound to carbon C8 of PLP (1.31 Å). The hydrogen bond between oxygen atom O1 and the α-amino group is still present (1.74 Å). This rearrangement continues to stabilize this region, allowing the positive charge of the amino group and the negative charge of oxygen O1 to spread along the atoms that compose the pseudoring (~0.03 au). This means that the resonance structure b of Scheme 5 is the one that describes better the product of this reaction.

The formation of the external diamine requires a very small activation energy (2.2 kcal/mol), and the reaction is exothermic in −4.3 kcal/mol.

Comparing this step with the first step of the mechanism, this reaction is very similar but has a lower activation barrier of about 8.6 kcal/mol. Such behavior can be explained taking into account the type of bond that is cleaved/formed in each reaction. While in the first step, the bond that is formed/cleaved involves the ornithine residue (the substrate), in the last step the chemical bond that is cleaved/formed involves a lysine residue (which is part of the enzyme). Although these two residues are very similar (ornithine lacks only one CH_2 group in the side chain, when compared to lysine), the way they bind to the PLP group is quite different. Lysine binds to PLP through the NH_2 group located in the side chain (ε-amino group), whereas ornithine binds to PLP with the amino group located in the main chain (α-amino group). This means that the binding of ornithine to PLP is from a steric point of view less favorable, as the neighboring carboxylic group hinders the approach of the amino group to PLP. This explains why the activation energy of the first step is higher than that of the last step.

Looking at the reaction energies, we also see that both reactions are exothermic. In the first step, this means that in spite of the steric effect that results from the binding of ornithine to PLP, the formation of the geminal diamine intermediate is very favorable. In the last step, we have the opposite situation, i.e., the formation of the external aldimine is more stable than the geminal diamine intermediate. This happens because once ornithine binds to PLP, its amino group makes a strong hydrogen bond with the carbonyl group of PLP, and this interaction overcomes the steric penalty arising from the approach of ornithine. In the products, the same type of interaction exists, which stabilizes the formation of the external aldimine and favors the dissociation of the lysine residue. The stabilization of the ornithine−PLP complex arises from the formation of a pseudoring that seems to favor the delocalization of the charge around it.

It must be noted that the same type of interaction between the lysine residue and PLP was never observed. This occurs because at the beginning of the reaction, the lysine is located in a perpendicular plane to that of the PLP and in the opposite direction of the carbonyl group. This is in agreement with the available X-ray structures and the model used in this study kept the same orientation, which underscores its robustness.

3.2. Functional Benchmarking: B3LYP vs M06 Family. In order to understand if the energetic profile that was obtained in this study could be influenced by the functional that was used, we performed a small benchmarking exercise comprising B3LYP and the two hybrid meta exchange−correlation functionals M06 and M06-2X. The latter have shown to be very accurate for thermodynamics and kinetics. In addition, we have also compared the influence of the functional and the basis set in the geometries.

To evaluate the influence of the functional in the optimized geometries in this process, all the transition states were recalculated, and the products and reactants of each reaction were obtained through IRC calculations. The obtained results have shown that the differences between the optimized structures obtained with M06/6-31G(d) and with B3LYP/6-31G(d) amount to less than 0.7 Å.

In addition, a visual inspection of each minimum revealed that the atoms that are involved in the formation or cleavage of chemical bonds have a very small root-mean-square deviation (rmsd, below 0.12 Å). These results show, that in this type of system and when studying this type of chemistry, the optimized geometries that are obtained with B3LYP are similar to those that are obtained with the M06 functionals.

We have also tested the effect of the basis set in the geometry optimizations. For this purpose, we have reoptimized all the species of the second step of the mechanism (rate limiting step) with the functional M06 and the 6-31+G(d,p) basis set (a double-ζ basis set augmented with polarized functions for heavy atoms and hydrogens). The results revealed that there is not a significant difference in the optimized geometries. When compared with those obtained with the 6-31G(d) basis set, the rmsd of the reactants (rmsd = 0.15 Å), TS (rmsd = 0.08 Å) and products (rmsd = 0.07 Å) are indeed very small and on average below 0.1 Å.

In order to evaluate if the energetic profile of the transimination reaction did differ from the functional that was used, the very popular B3LYP functional was used instead of the hybrid meta exchange−correlation M06 functional to recalculate the final energies of the optimized models. The M06-2X functional was also used to check for the effect of doubling the HF exchange, which is known to affect the barriers. For this purpose we used the M06 geometries and recalculated the energy using B3LYP and M06-2X, with the 6-311++G(3df,2pd) basis set. The results are presented in Figure 9.

From all the employed functionals, the barriers obtained with M06-2X functional were always within the values obtained with B3LYP and M06. In general there are no significant differences between them. Exceptions are limited to the activation energy of the second step, scenario 2, and the reaction energy of the third step. Comparing the values obtained with M06-2X and M06, we observe that M06 tends to result in higher barriers. Comparing the values obtained with the M06-2X and the B3LYP functionals, we observe that B3LYP underestimates the barriers. As the experimental values for the barriers are unknown, we cannot pinpoint exactly which functional is giving us the most exact result.

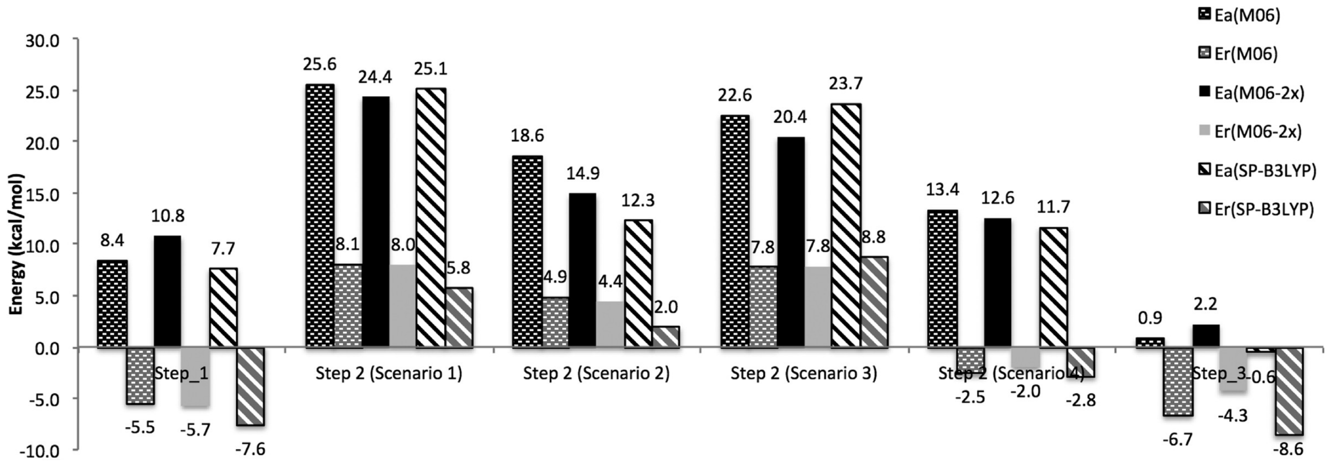


Figure 9. Energetic profile for the transimination reaction calculate with the B3LYP, M06, and M06-2X functionals with the 6-311++G(3df,2pd) basis set.

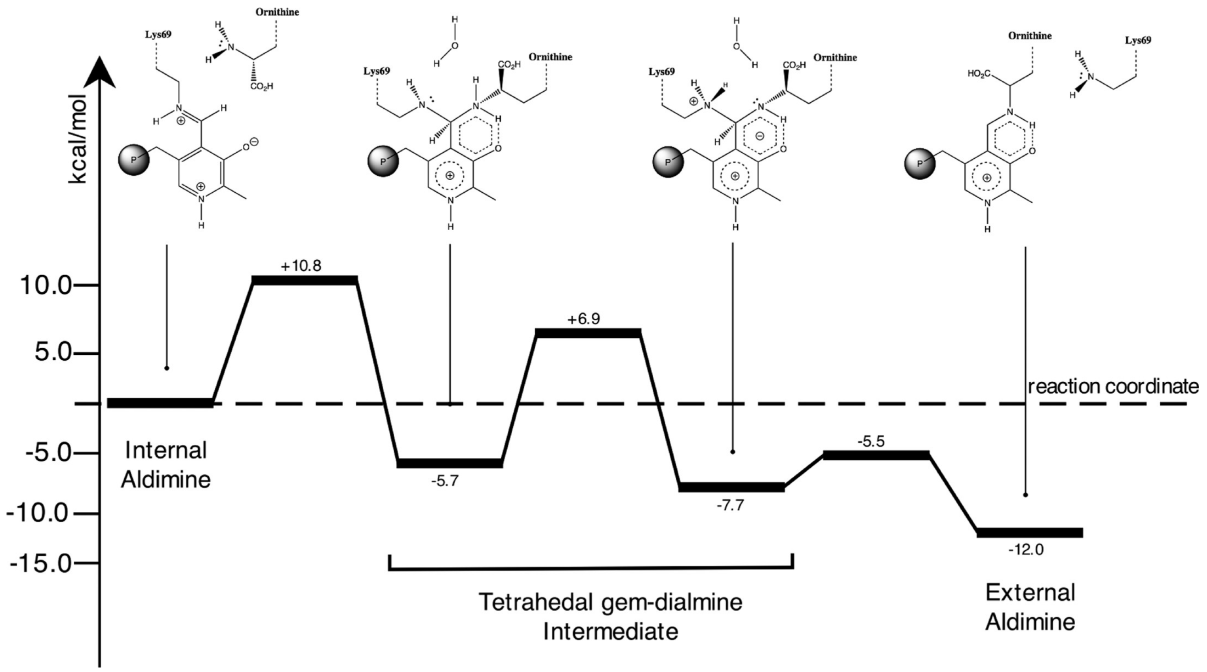


Figure 10. Most favorable pathway for the transimination reaction (P stands for phosphate group). The results were obtained withy the M06-2X/6-311++(3df,p2d) level of theory.

However, these results show that B3LYP tends to give lower values for the chemical barrier, when compared to the values obtained with the M06 family. This is in agreement with other previous studies, and therefore we may conclude that the barriers obtained with M06 and M06-2X should be more exact than those that are obtained with B3LYP. As the M06-2X functional includes twice the amount of exchange compared with M06 functional (which usually is favorable to describe activation energies) and resulted in energy values that were between the B3LYP and M06 extremes, we choose to use these values in the description of the reaction steps on the previous sections. The most important aspect of this study is that the chosen pathway (kinetically most favorable) is always the same, independently of the functional. The functionals affect the accuracy of the activation and the reaction energies but do not affect the discrimination between different hypotheses for the catalytic mechanism.

4. CONCLUSIONS

The external aldimine is the common central intermediate for all enzymatic and nonenzymatic reactions that are catalyzed by the PLP cofactor. Divergence in reaction specificity occurs from this point, which means that the formation of external aldimines from the internal PLP-aldimines represent the first level of catalysis in all PLP-dependent reactions.

The results obtained in this work show that the transimination reaction is very favorable, but it is not accomplished by a single step as it is generally accepted involving instead three subsequent

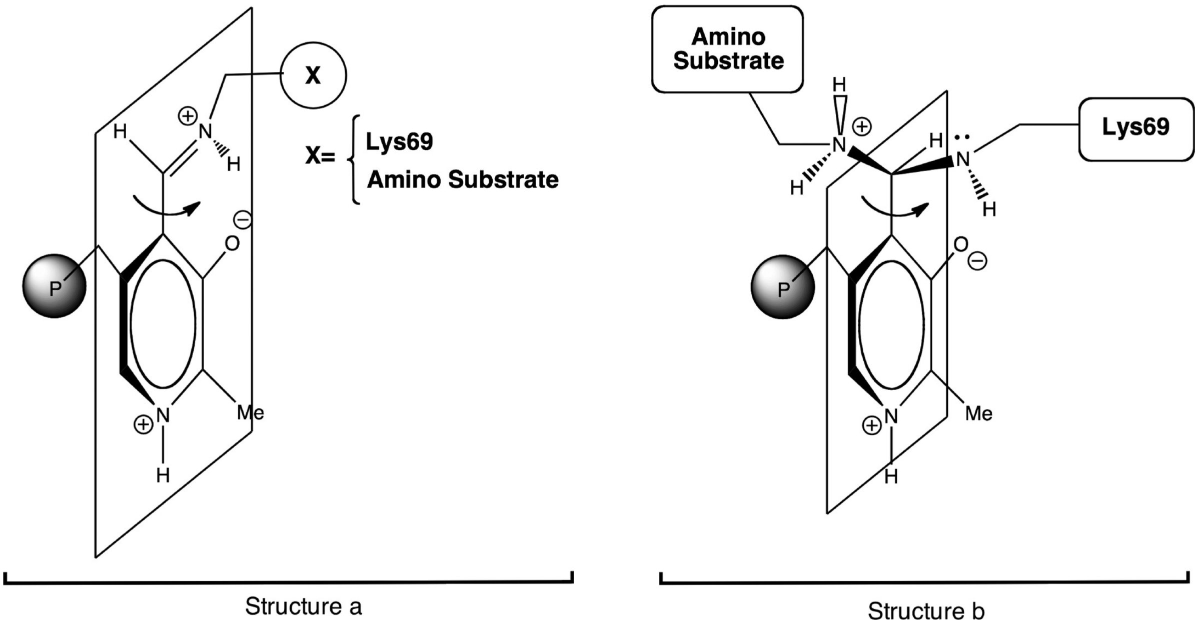


Figure 11. Conformational rearrangements adopted by the amino substrate (ornithine) and Lys69 around the PLP ring. Left: Conformational rearrangement observed in the external and internal aldimines. Right: Conformational rearrangement adopted in all the transition-state structures studied in this paper.

steps, as depicted in Figure 10. This is in agreement with previous NMR studies performed by Chan-Huot.[22]

Multiple pathways are possible for the conversion of the geminal—diamine intermediate into the external aldimine. In this article, we explored all reasonable pathways including the direct migration of the proton between both amino groups, the involvement of several active site residues, or even the participation of a water molecule. The results have shown that the rate of the reaction is lower if a water molecule or Tyr389 are involved in this process. The direct proton transfer or the involvement of the catalytic active site residue Cys360 were shown to be less favorable.

The most favorable pathway occurs when the water molecule is directly involved in the reaction. The overall reaction accounts for −10.8 kcal/mol and is exothermic by 12.0 kcal/mol. The second step of the reaction is the rate-limiting one amounting to 12.6 kcal/mol. In Figure 10, it is displayed the most favorable pathway obtained for the transimination reaction (only scenario 4 of step 2 is displayed).

These results also point out the importance of the PLP in the course of the reaction. It allows the interchange of carbon C8 between sp^2 and sp^3 hybridizations, without requiring a significant energetic cost. This key feature favors the formation of the tetrahedral intermediate that is the key driving force behind the conversion between the internal and external aldimines. During this process, the excess of charge in the system becomes lodged at oxygen O1 of the PLP ring. This effect is very important during the transimination reaction because it not only allows the attraction of the substrates to PLP but it also serves as a guide during the binding/dissociation of the amino substrate/lys69 residue to carbon C8 of the PLP cofactor. Moreover, this atom makes a strong hydrogen bond with the amino group of the amino substrates, favoring the alignment of both parts of the molecule on the same plane. This rearrangement, improves the stereoelectronic effect of the system, ensuring the maximum overlap of the extended π system and the stabilization of the full system. This is in agreement with Dunathan's hypothesis[39] postulated almost 50 years ago in which he predicted that the most active form of the external aldimine had a cisoide conformation underneath the same plane in order to favor the occurrence of the subsequent reactions (Figure 11, Structure a). Furthermore, Dunathan's proposed that the bonds that are formed/broken in the PLP system should adopt a perpendicular plane to that of the pyridoxal imine system. All the transition-state structures presented in this article show exactly the same sort of conclusion, which confirms once again the accuracy of the early proposals made by Dunathan (Figure 11, Structure b).

In the last 10 years, two different proposals for the transimination reaction have been suggested. One of those proposals was suggested by Salvà et al.,[23] in which he suggested that the transimination process occurs in seven steps, requiring the participation of one/two water molecules. That mechanism differs substantially from the one presented in this work since it required the direct participation of oxygen O1 as the key intermediate that shuttles the proton transfer between the Lys69 and the substrate. The mechanism presented here is accomplished in fewer steps (only three instead of seven) and is from an energetic point of view more favorable (Ea ∼ 10 vs Ea ∼ 20 kcal/mol). In addition, it shows that in the absence of water molecules, nearby the active site, Tyr389 can catalyze this reaction.

Another study was performed by Zhao et al.,[24] in which he proposed that the transimination reaction requires the direct participation of the phosphate group of PLP, adopting a similar role that is played by oxygen O1 in the mechanism proposed by Salvà. The energies involved in that process are comparable to those presented in this paper, but they do not involve the formation of the geminal—diamine intermediate that is observed experimentally.

From all the analyzed data, we can conclude that the mechanism presented here is, both from thermodynamic and kinetic points of view, more favorable than the previous suggestions and includes the formation of all intermediates that are observed experimentally. We believe therefore that this mechanism should be general for all PLP-dependent enzymes, corresponding to the first chemical transformation that is catalyzed by all PLP requiring enzymes that have amino acids as substrates.

AUTHOR INFORMATION

Corresponding Author

*E-mail: mjramos@fc.up.pt.

REFERENCES

(1) Toney, M. D. *Arch. Biochem. Biophys.* **2005**, *433*, 279–287.

(2) Eliot, A. C.; Kirsch, J. F. *Annu. Rev. Biochem.* **2004**, *73*, 383–415.

(3) Kleppner, S. R.; Tobin, A. J. *Expert Opin. Ther. Targets* **2001**, *5*, 219–239.

(4) Thorndike, J.; Pelliniemi, T.; Beck, W. *Cancer Res.* **1979**, *39*, 3435–3440.

(5) Wang, C. C. *Annu. Rev. Pharmacol. Toxicol.* **1995**, *35*, 93–127.

(6) Mudd, S. H.; Finkelstein, J. D.; Irreverre, F.; Laster, L. *Science* **1964**, *143*, 1443–1445.

(7) Jansonius, J. N. *Curr. Opin. Struc. Biol.* **1998**, *8*, 759–769.

(8) Grishin, N.; Phillips, M.; Goldsmith, E. *Protein Sci.* **1995**, *4*, 1291–1304.

(9) Korpela, T.; Mäkelä, M.; Lönnberg, H. *Arch. Biochem. Biophys.* **1981**, *212*, 581–8.

(10) Robitaille, P.; Scott, R.; Wang, J.; Metzler, D. *J. Am. Chem. Soc.* **1989**, *111*, 3034–3040.

(11) Lehtokari, M.; Puisto, J.; Raunio, R.; Korpela, T. *Arch. Biochem. Biophys.* **1980**, *202*, 533–539.

(12) Snell, E. E.; Jenkins, W. T. *J. Cell Comp. Physiol.* **1959**, *54*, 161–77.

(13) Tobias, P.; Kallen, R. *J. Am. Chem. Soc.* **1975**, *97*, 6530–6539.

(14) Vazquez, M.; Munoz, F.; Donoso, J.; Blanco, F. *J. Phys. Org. Chem.* **1992**, *5*, 142–154.

(15) Hershey, S.; Leussing, D. *J. Am. Chem. Soc.* **1977**, *99*, 1992–1993.

(16) Chang, Y. C.; McCalmont, T.; Graves, D. J. *Biochemistry* **1983**, *22*, 4987–4993.

(17) Fletterick, R. J.; Sprang, S. R. *Acc. Chem. Res.* **1982**, *15*, 361–369.

(18) Cook, P. D.; Holden, H. M. *J. Biol. Chem.* **2008**, *283*, 4295–4303.

(19) Counts, K. G.; Wong, I.; Oliveira, M. A. *Biochemistry* **2007**, *46*, 379–386.

(20) Phillips, R. S.; Miles, E. W.; Cohen, L. A. *Biochemistry* **1984**, *23*, 6228–6234.

(21) Schirch, L. *J. Biol. Chem.* **1975**, *250*, 1939–1945.

(22) Chan-Huot, M.; Sharif, S.; Tolstoy, P. M.; Toney, M. D.; Limbach, H. *Biochemistry* **2010**, *49* (51), 10818—10830.

(23) Salvà, A.; Donoso, J.; Frau, J.; Muñoz, F. *Int. J. Quantum Chem.* **2002**, *89*, 48–56.

(24) Zhao, Z.; Liu, H. *J. Phys. Chem. B* **2008**, *112*, 13091–13100.

(25) Ortega-Castro, J.; Adrover, M.; Frau, J.; Salvà, A.; Donoso, J.; Muñoz, F. *J. Phys. Chem. A* **2010**, *114*, 4634–4640.

(26) Dufe, V. T.; Ingner, D.; Heby, O.; Khomutov, A. R.; Persson, L.; Al-Karadagi, S. *Biochem. J.* **2007**, *405*, 261–268.

(27) Jackson, L. K.; Brooks, H. B.; Osterman, A. L.; GoldSmith, E.; Philips, M. *Biochemistry* **2000**, *39*, 11247–11257.

(28) Frisch, M. J.; Trucks, G. W.; Cheeseman, J. R.; Scalmani, G.; Caricato, M.; Hratchian, H. P.; Li, X.; Barone, V.; Bloino, J.; Zheng, G.; Vreven, T.; Montgomery, J. A.; Petersson, G. A.; Scuseria, G. E.; Schlegel, H. B.; Nakatsuji, H.; Izmaylov, A. F.; Martin, R. L.; Sonnenberg, J. L.; Peralta, J. E.; Heyd, J. J.; Brothers, E.; Ogliaro, F.; Bearpark, M.; Robb, M. A.; Mennucci, B.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Rendell, A.; Gomperts, R.; Zakrzewski, V. G.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H. *Gaussian 09*; Gaussian, Inc.: Wallingford, CT, 2009.

(29) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phy. Chem.* **1994**, *98*, 11623–11627.

(30) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200–1211.

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

(32) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.

(33) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *119*, 525–525.

(34) Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157–167.

(35) Chen, S.; Fang, W.; Himo, F. *Theor. Chem. Acc.* **2008**, *120*, 515–522.

(36) Cerqueira, N. M. F. S. A.; Fernandes, P. A.; Eriksson, L. A.; Ramos, M. J. *Biophys. J.* **2006**, *90*, 2109–2119.

(37) Amata, O.; Marino, T.; Russo, N.; Toscano, M. *J. Am. Chem. Soc.* **2009**, *131*, 14804–14811.

(38) Ramos, M. J.; Fernandes, P. A. Computational enzymatic catalysis. *Acc. Chem. Res.* **2008**, *41* (6), 689–698.

(39) Dunathan, H. *Proc. Natl. Acad. Sci. U.S.A.* **1966**, *55*, 712.

JCTC Journal of Chemical Theory and Computation ARTICLE

pubs.acs.org/JCTC

A Kirkwood-Buff Derived Force Field for Aqueous Alkali Halides

Moon Bae Gee,[†] Nicholas R. Cox,[‡] Yuanfang Jiao,[†] Nikolaos Bentenitis,[‡] Samantha Weerasinghe,[§] and Paul E. Smith*[,†]

[†]Department of Chemistry, Kansas State University, Manhattan, Kansas 66506, United States

[‡]Department of Chemistry and Biochemistry, Southwestern University, Georgetown, Texas 78626, United States

[§]Department of Chemistry, University of Colombo, Colombo 00300, Sri Lanka

S Supporting Information

ABSTRACT: A classical nonpolarizable force field is presented for the simulation of aqueous alkali halide solutions (MX), where M = Li^+, Na^+, K^+, Rb^+, and Cs^+ and X = F^-, Cl^-, Br^-, and I^-, and their interactions with biomolecules. The models are specifically designed to reproduce the experimental Kirkwood-Buff integrals, and thereby the solution salt activities, as a function of salt concentration. Additionally, we demonstrate that these models reasonably reproduce other experimental properties including ion diffusion constants, dielectric decrements, and the excess heats of mixing. The parameters are developed by considering the properties of aqueous NaX and MCl solutions using a previously established model for NaCl. Transferability of the parameters to other salts is then established by the successful simulation of additional aqueous salt solutions, KI and CsBr, not originally included in the parametrization procedure.

INTRODUCTION

Aqueous solutions of alkali metal halides are not only the simplest models for the aqueous electrolyte solutions but also play an important role in many biological systems. They can help to stabilize biomolecules, such as proteins, nucleic acids, and lipids, and are often involved in biological catalysis.[1−3] Because of their importance in biological phenomena, and the desire to study these more complicated ternary systems using computer simulation, many force fields for alkali metal cations and halide anions have been reported in the literature.[4−11] A recent comprehensive survey has also been provided by Joung and Cheatham.[4] The wide range of parameter sets available for salt systems is, in our opinion, a direct result of the fact that there is relatively little experimental data available that is both sensitive to changes in the ion parameters and also easily amenable to simulation. Furthermore, as our ability to access longer simulation time scales has improved, a number of problems with many of the existing ion force fields have recently come to light.[12,13] One approach to solving these problems is the use of models which explicitly include polarization effects.[14−16] However, as this significantly increases the computational demand, the vast majority of biomolecular simulations still do not include explicit polarization effects. Therefore, there remains a need for simple but reliable ion force fields, especially for systems displaying slow relaxation times.

Recently, there have been three major attempts to develop force fields for all alkali metals and halide ions. Jensen and Jorgensen have developed TIP4P water compatible alkali halide parameters using the ion hydration free energies and ion−water contact distances as target data.[11] Joung and Cheatham[4] have also used the free energy of hydration for individual ions, as well as the lattice energies and the lattice constants of alkali metal halides and gas phase ion−water interaction energies, in order to produce force fields for all of the alkali metal and halide ions which are compatible with three commonly used nonpolarizable water models, namely, SPC/E, TIP3P, and $TIP4P_{EW}$. Horinek et al.[17] have used both the free energy and the entropy of hydration of the individual ions in order to parametrize their force fields and focused on the nonpolarizable SPC/E water model. Horinek et al. argued that their force field would be more applicable for biomolecular simulations where the salt concentrations are low, while the Joung and Cheatham force fields would be more applicable when the salt concentrations are moderate. All three force fields attempt to reproduce a series of initial properties, including the free energies (and entropies) of hydration, the first peak of the ion−water radial distribution function (rdf), gas phase ion−water binding energies, and crystal lattice parameters. However, they were essentially developed using properties that that do not directly probe ion−ion interactions in solution. A subsequent study has since evaluated the solute activity for two salts using the Joung and Cheatham force fields obtained using thermodynamic integration.[18] This does probe ion−ion interactions. However, the study provided only moderate success—good results were obtained for KCl, but significant deviations from experimental results were observed for NaCl solutions above 0.5 m.[18] The comparison of simulated and experimental diffusion constants and solubilities also provided mixed results.

We have taken a very different approach in an attempt to develop accurate force fields for solution mixtures. Our approach is based on the thermodynamics of solution mixtures as described by Kirkwood-Buff (KB) theory.[19−26] Here, the central properties of interest are the Kirkwood-Buff integrals (KBIs)

Received: September 11, 2010
Published: April 26, 2011

defined by

$$G_{ij} = 4\pi \int_0^{\infty} [g_{ij}^{\mu VT}(r) - 1] r^2 \, \mathrm{d}r \qquad (1)$$

where G_{ij} is the KBI between species i and j, $g_{ij}^{\mu VT}(r)$ is the corresponding radial distribution function (rdf) in the grand canonical ensemble at the composition of interest, and r is the center of mass distance between the two species. An excess coordination number can be defined by $N_{ij} = \rho_j G_{ij}$, where $\rho_j = N_j/V$ is the number density of j particles. The physical meaning of the excess coordination number is the difference in the number of j species in the vicinity of a central i species on the addition of the i species from that found in an equivalent volume of bulk solution. Hence, a value of N_{ij} significantly greater than zero indicates an excess of species j in the vicinity of species i (over the random bulk distribution), while a significant negative value corresponds to a depletion of species j surrounding i. Combinations of KBIs provide expressions for a variety of thermodynamic properties of the solution of interest.[27,28]

Kirkwood-Buff theory can then be used to relate solution structure, in terms of the KBIs, to the thermodynamic behavior of the solution.[29−31] The expressions provided by KB theory are exact, and the theory involves no approximations beyond the usual statistical mechanical assumptions (larger number of molecules, thermodynamic limit, etc). The expressions can be applied to study any *stable* solution mixture involving any number of components of any type (small molecules through to proteins) at any composition and any temperature and pressure. The analysis of experimental data for solution mixtures using KB theory is well established and provides quantitative information concerning species distributions in solutions and how they vary with composition.[28,29,32] The resulting KBIs can also be obtained from computer simulations and thereby provide a rigorous test of the accuracy of current force fields.[31,33]

Our parameters were developed to reproduce the properties of solution mixtures and are therefore collectively known as Kirkwood-Buff derived force fields (KBFF).[19−25] The parameters for the KBFF models are determined using a combination of molecular dynamics simulation, the Kirkwood-Buff (KB) theory of solutions, and available experimental data concerning activity coefficients and solution densities. This approach has several advantages. First, KB theory is exact and includes no approximations. Second, KB theory can be applied to any stable solution mixture. Third, the KB integrals are easily obtained from the radial distribution functions (rdf) provided by MD simulations and are quite sensitive to the force field parameters. Fourth, the KB integrals help quantify the distributions arising from the relative strengths of the solute−solute and solute−solvent interactions.[25,34] Hence, the general philosophy of the Kirkwood-Buff derived force field approach is to use the KBIs obtained from an analysis of the experimental data as target values for the development of accurate force fields for a variety of solutes. The target data are composition-dependent, and this dependence is also used during the parametrization process. We have argued that reasonable agreement with experimental results is also obtained for other properties not included in the original parametrization.[19,20,22−25] In doing so, we favor the use of data for solution mixtures, primarily the KBIs, and are less influenced by gas phase data or infinite dilution data such as free energies of hydration. A model for aqueous NaCl solutions has already been developed using this type of approach,[25] and here we simply generalize this initial model to include other alkali halide salts.

Recently, two research groups also produced KB derived force fields for some of the alkali metal halides. Hess and van der Vegt used the SPC/E water model to develop KB-derived force fields for Li^+ and K^+ in order to explain the differential binding affinity of alkali metal ions to carboxylate ions.[35] Klasczyk and Knecht used the SPC water model and the KBFF force field for the chloride ion to develop force fields for Li^+, K^+, Rb^+, and Cs^+, but not for halide ions.[36] Therefore, the more extensive Klasczyk and Knecht force field is incomplete and, in principle, incompatible with our models because we use the SPC/E water model. In this paper, we present a KB derived force field for a wide variety of alkali metal and halide ions. The models are intended to be applicable over the whole concentration range and are consistent with our previous models for a variety of solutes in both aqueous and nonaqueous solutions.

METHODS

Kirkwood-Buff Analysis of Alkali Halide Solutions. The complete details concerning the extraction of the KBIs from the experimental data, the so-called Kirkwood-Buff inversion procedure, have been provided elsewhere.[27,28,37] For a binary solution consisting of water (w) and a salt cosolvent (c), a variety of thermodynamic quantities can be defined in terms of the KB integrals G_{ww}, G_{cc}, and $G_{cw} = G_{wc}$ and the number densities (or molar concentrations) ρ_w and ρ_c.[25] By use of the KB inversion procedure, one can also extract the composition-dependent KBIs from the corresponding experimental thermodynamic properties.[28] Specifically, the KB inversion approach uses composition-dependent experimental binary solution data for the isothermal compressibility, partial molar volumes, and cosolvent activity in order to extract the corresponding three KBIs using the expressions provided by KB theory. Hence, KB theory provides a link between measurable experimental data and the species distributions in solution, which are then quantified in terms of the KBIs. The relationships used for the present work are[27]

$$\begin{aligned} 1 + N_{cc} &= \rho_c RT\kappa_T + \rho_w^2 \frac{\overline{V_w}^2}{\mu_{cc}} \\ 1 + N_{ww} &= \rho_w RT\kappa_T + \rho_w \rho_c \frac{\overline{V_c}^2}{\mu_{cc}} \\ N_{wc} &= \rho_c RT\kappa_T - \rho_w \rho_c \frac{\overline{V_w V_c}}{\mu_{cc}} \end{aligned} \qquad (2)$$

where κ_T is the isothermal compressibility, $\overline{V_i}$ are partial molar volumes, and μ_{cc} represents a chemical potential (or activity) derivative given by

$$\mu_{cc} = \beta \left(\frac{\partial \mu_c}{\partial \ln m_c} \right)_{T,P} = 1 + \left(\frac{\partial \ln \gamma_c}{\partial \ln m_c} \right)_{T,P} \qquad (3)$$

where $\gamma_c = \gamma_{\pm}$ is the molal activity coefficient of the salt and m_i is the molality of i. Hence, the three KBIs can be obtained from a knowledge of the compressibility, partial molar volumes (or density), and activity as a function of the composition (three equations in three unknowns).

Experimental activity coefficient data at 298.15 K and 1 atm were taken from the literature,[38] and fitted to the following functional form,[38,39]

$$\ln \gamma_{\pm} = -\frac{1.18\sqrt{m_s}}{1 + a_1\sqrt{m_s}} - \ln(1 - a_2 m_s) + a_3 m_s + a_4 m_s^2 \quad (4)$$

where m_s is the salt molality and the a's represent fitting parameters with no particular physical meaning. The first term on the right-hand side of eq 4 is a Debye−Hueckel term for 1:1 salts which is required to fully capture the correct behavior of salts at low salt concentrations. Issues associated with the quality of fit for the experimental activity coefficient data provide the main source of error in the KB analysis. The final fitting parameters are provided in the Supporting Information. Previously established polynomial fitting expressions for the experimental density data of salts[40] were used to determine partial molar volumes using standard approaches.[41] The solution compressibility has a negligible effect on the resulting KBI values for solutions at moderate temperatures and pressure.[32] Hence, the compressibility was assumed to follow the simple relationship, $\kappa_T = \varphi_w \kappa^{\circ}_{Tw} + \varphi_c \kappa^{\circ}_{Tc}$, where φ_i is the volume fraction and κ°_{Ti} is the compressibility of the pure substance (water or salt). The compressibility of pure water was taken to be 4.6×10^{-10} m^2/N,[42] while the compressibilities of the salt crystals were taken to be zero. The experimental compressibility (approximated), partial molar volumes, and activity provided by eq 4 were then used with the expressions provided in eq 2 to isolate the experimental KBIs as a function of the composition. The results of the KB inversion analysis are presented in Figure 1.

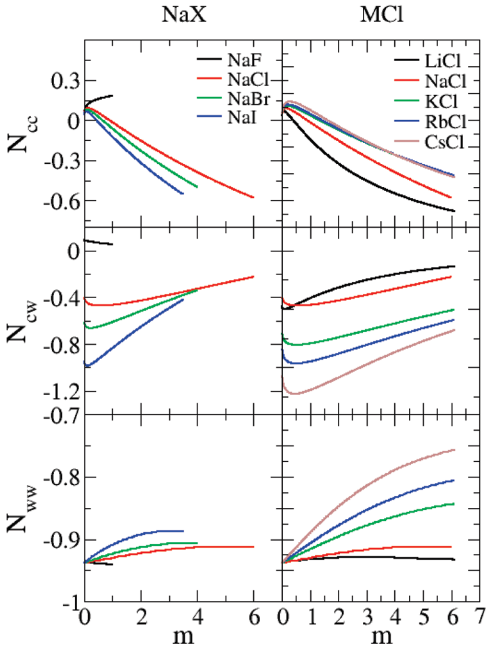


Figure 1. Experimentally derived excess coordination numbers for aqueous alkali halide solutions as a function of salt molality at 298.15 K and 1 atm.

Kirkwood-Buff Theory of Salt Solutions. Some complications arise when applying KB theory to salt solutions.[25,43] First, the salt can dissociate into free cations and anions (we will assume complete dissociation for the salts examined here). Second, electroneutrality constraints for regions of the solution surrounding each species provide additional relationships between the KBIs.[43] Let us consider a salt containing a total of n ions which will fully dissociate to provide n_+ cations and n_- anions. If one chooses the salt as the relevant thermodynamic species, then $d\mu_s = nRTd\ln(m_s\gamma_{\pm})$ and the activity derivatives provide a set of KBIs (G_{ss} and G_{sw}) involving the salt "molecules" when using the KB inversion approach. However, this choice is rather awkward from the simulation point of view as we typically observe free ions for strong electrolytes, and therefore the rdf's between salt "molecules" are difficult, if not impossible, to determine. Consequently, in this work, the salt solution is treated as a binary system of indistinguishable ions (c) and water (w), and we will distinguish between the cosolvent (total ion) concentration, m_c or ρ_c, and the classic salt concentration, m_s or ρ_s. Consequently, for a n_+:n_- salt, one has $nm_s = m_c$, $n\rho_s = \rho_c$, $\overline{V}_s = n\overline{V}_c$, and $\gamma_c = \gamma_{\pm}$. In addition, the following relationships are also obeyed: $d\mu_s = nd\mu_c$, $\rho_s d\mu_s = \rho_c d\mu_c$, $d \ln m_s = d \ln m_c$, $\rho_s\overline{V}_s + \rho_w\overline{V}_w = \rho_c\overline{V}_c + \rho_w\overline{V}_w = 1$, and $\rho_c d \ln a_c = \rho_w d \ln a_w = \rho_s d \ln a_s + \rho_w d \ln a_w = 0$, at constant p and T—the latter being the Gibbs−Duhem equation.

Hence, the experimental data can then be analyzed in terms of either salt molecules or a collection of indistinguishable ions. The resulting KBIs obtained from the two formalisms are related by

$$G_{ss} = \frac{1 - n}{\rho_c} + G_{cc} \qquad G_{sw} = G_{cw} \quad (5)$$

The KBIs obtained from the indistinguishable ion approach (G_{cc} and G_{cw}) involve rdf's between the ions (and water molecules), which ignore the ion identity (cation or anion). The relationships between the KBIs using the cosolvent label and those involving the anion/cation label are provided by

$$G_{cc} = \left(\frac{n_+}{n}\right)^2 G_{++} + \left(\frac{n_-}{n}\right)^2 G_{--} + \frac{n_+n_-}{n^2}(G_{+-} + G_{-+})$$
$$G_{cw} = G_{wc} = \frac{n_+}{n}G_{+w} + \frac{n_-}{n}G_{-w} \quad (6)$$

and were obtained in a similar manner as done previously.[25] Here, the KBI denoted as G_{++} refers to the integral over the cation−cation rdf in solution. We note that the above relationships merely reflect a change in indices and do not invoke the electroneutrality conditions.

If one then assumes that electroneutrality must be obeyed in the local regions surrounding each molecule or ion,[22,25,43] then one can show that the following relationships must also hold:

$$G_{cc} = -\frac{1}{\rho_c} + G_{+-} \qquad G_{cw} = G_{+w} = G_{-w}$$
$$G_{+-} = \frac{1}{\rho_+} + G_{++} \quad (7)$$
$$\frac{1}{\rho_+} + G_{++} = \frac{1}{\rho_-} + G_{--}$$

where ρ_+ is the number density of cations etc. Hence, all of the ion−ion KBIs are related, and there is only one independent KBI for a binary solution. We choose this to be G_{cc} for the present analysis.

Molecular Dynamics Simulations. All molecular dynamics simulations of alkali halide solutions were performed using the SPC/E water model[44] in the isothermal isobaric (NpT) ensemble at 300 K and 1 atm as implemented in the GROMACS program (v3.3.1).[45,46] A time step of 2 fs was used, and the geometry of the water molecules was constrained using SETTLE.[47] The weak coupling technique was used to modulate the temperature and pressure with relaxation times of 0.1 and 0.5 ps, respectively.[48]

The particle mesh Ewald technique (PME) was used to evaluate electrostatic interactions using a cubic interpolation and a grid spacing of 0.1 nm for the reciprocal space sum, coupled with tinfoil boundary conditions.[49] The initial cubic boxes for each solution at the required concentration were generated by randomly placing water molecules with ions starting from pure solvent boxes of length varying between 4 and 6 nm. During the simulations, configurations were saved every 0.1 ps for analysis. Diffusion constants were determined using the mean square fluctuation approach,[50,51] and relative permittivities were obtained from the dipole moment fluctuations.[52,53] The excess enthalpy of mixing (ΔH_{mix}) was determined using an established procedure which uses the average potential energies[54] and the configurational energies from the pure SPC/E water and the alkali halide lattice.

Kirkwood-Buff Analysis of the Simulation Data. Radial distribution functions were obtained for each system and composition. The pair rdf's thereby correspond to the ion–ion, ion–water, and water–water distributions after averaging over all other ions and water molecules at that particular composition. The indistinguishable ion treatment for salts involves the determination of ion–ion and ion–water rdf's, which ignore the identity of the ions involved. For example, in NaCl solutions, the ion–water rdf is determined after averaging over the ion–water distributions using both types of ion, sodium and chloride, at the origin. The Kirkwood-Buff integrals (KBIs) are defined for systems open to all the solution components. However, most simulations are performed in closed systems. Hence, one has to approximate the KBIs by truncating the integral after a certain distance

$$G_{ij} \approx 4\pi \int_0^R [g_{ij}^{NpT}(r) - 1]r^2 \, \mathrm{d}r \qquad (8)$$

where R represents a correlation distance within which the solution composition differs from the bulk composition. This approximation has been shown to be very reasonable as long as the systems are not too small ($L > 4$ nm) and sufficient sampling (>5 ns) is achieved.[26,29,55] The values of G_{ij} used here were determined by averaging the integral over a short-range of distances (1.2–1.5 nm), taken as approximately one water–water solvation shell. The final values were relatively insensitive to the exact distance and range used, but this approach helps to reduce statistical fluctuations associated with the integrals. Once the three simulated KBIs have been obtained from the trajectory at a particular bulk composition, one can then use these values in a series of expressions which provide thermodynamic properties of the solution mixture. The partial molar volumes of the components ($\overline{V_i}$) are given by[41]

$$\overline{V_{\mathrm{w}}} = \frac{1 + \rho_{\mathrm{c}}(G_{\mathrm{cc}} - G_{\mathrm{cw}})}{\eta}, \qquad \overline{V_{\mathrm{c}}} = \frac{1 + \rho_{\mathrm{c}}(G_{\mathrm{ww}} - G_{\mathrm{cw}})}{\eta}$$

$$\eta = \rho_{\mathrm{w}} + \rho_{\mathrm{c}} + \rho_{\mathrm{w}}\rho_{\mathrm{c}}(G_{\mathrm{ww}} + G_{\mathrm{cc}} - 2G_{\mathrm{cw}}) \qquad (9)$$

Using the simulated KBIs, one can determine a variety of derivatives of the chemical potential, depending on the concentration scale used. Here, we choose derivatives of the activity with respect to molarity.[25] Of primary interest is the following activity derivative:

$$a_{\mathrm{cc}} = \left(\frac{\partial \ln a_{\mathrm{c}}}{\partial \ln \rho_{\mathrm{c}}}\right)_{p,T} = 1 + \left(\frac{\partial \ln y_{\mathrm{c}}}{\partial \ln \rho_{\mathrm{c}}}\right)_{p,T} = \frac{1}{1 + \rho_{\mathrm{c}}(G_{\mathrm{cc}} - G_{\mathrm{cw}})} \qquad (10)$$

where a_{c} and y_{c} are the cosolvent (average ion) molar activity and molar activity coefficient, respectively. Hence, changes in the cosolvent activity can be determined directly from the simulations. Furthermore, accurate activity derivatives ensure reasonable activities are thereby obtained. The partial molar volumes and activities obtained in this manner have been shown to be in agreement with the results obtained using alternative computational approaches.[21,56]

Table 1. Experimental Data Used during the Initial Parameter Development[a]

	MCl					NaX			
	Li^+	Na^+	K^+	Rb^+	Cs^+	F^-	Cl^-	Br^-	I^-
r (nm)	0.115	0.101	0.138	0.149	0.170	0.133	0.181	0.196	0.220
a (nm)	0.257	0.282	0.319	0.332	0.412	0.239	0.282	0.299	0.324
d (nm)	0.213	0.240	0.280	0.289	0.314	0.263	0.319	0.338	0.365
ref	60, 62	25	60, 62	60, 62	60, 62	60, 62	25	60, 62	60, 62

[a] r, the ionic radii of alkali halide ions which are consistent with the crystal lattice dimensions; a, the crystal lattice unit cell dimension; and d, the ion to water oxygen contact distance.

Parameter Development. The KBFF models used in this study involve a simple classical nonpolarizable description for each molecule. The intermolecular interactions are described by the Coulomb and Lennard-Jones (LJ) 6–12 potentials, which contain just two adjustable parameters for ions, namely, the Lennard-Jones diameter (σ) and the interaction strength (ε). In this scheme, each pair of atoms i and j interact with an interaction energy given by

$$V_{ij} = \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} + 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \qquad (11)$$

Here, all of the symbols have their usual meaning.[1] This model was chosen so as to be computationally efficient, while maintaining compatibility with existing force fields and programs used for the simulation of biomolecules. The ion parameters are combined with the SPC/E model for water.[44] Geometric combination rules were used for both σ and ε. In order to obtain parameters for the LJ term, we have employed the same method described previously for NaCl.[25] This approach requires three pieces of experimental data: ionic radii of alkali and halide ions that are consistent with the crystal lattice dimensions, crystal lattice unit cell dimensions, and the ion to water oxygen contact distances (see Table 1). These data were chosen in an effort to be both consistent with our previous force fields and to help restrict the range of possible σ and ε values to be studied. However, satisfactory agreement with the experimental data was not possible for all ions using this simple approach (see below).

The first step was to parametrize the anions (F^-, Br^-, I^-) by studying the crystal structures and several aqueous solutions of NaF, NaBr, and NaI, using the same Na^+ parameters from our previous NaCl study.[25] The values of σ_{--} were determined by scaling the ionic radii of each ion with the same scaling factor as used previously (2.43).[25] The values of ε_{--} were then varied in an effort to reproduce the experimental lattice dimensions of the sodium halide crystals, and the anion–water contact distances, in the simulations. The final values determined for each ion were then used to provide the simulated KBIs for a variety of aqueous solutions. Unfortunately, in the case of the F^- anion, a reasonable

Table 2. Final Force Field Parameters Describing the KBFF Models for Alkali Halides[a]

model	atom	σ_{ii} (nm)	ε_{ii} (kJ/mol)	ε_{iO} (kJ/mol)	q (e)
KBFF	Li	0.1820	0.7000	0.2700	+1.0
	Na	0.2450	0.3200	0.3420	+1.0
	K	0.3340	0.1300	0.2327	+1.0
	Rb	0.3620	0.1500	0.2655	+1.0
	Cs	0.4130	0.0065	0.1954	+1.0
	F	0.3700	1.0000		−1.0
	Cl	0.4400	0.4700		−1.0
	Br	0.4760	0.3000		−1.0
	I	0.5350	0.2000		−1.0
SPC/E	O	0.3166	0.6506		−0.8476
	H	0.0000	0.0000		+0.4238

[a] The following combination rules used: $\sigma_{ij} = (\sigma_{ii} \times \sigma_{ij})^{1/2}$, $\varepsilon_{ij} = s(\varepsilon_{ii} \times \varepsilon_{ij})^{1/2}$. The value of s was set to unity for all interactions except for cation to water oxygen, where values of s = 0.4 (Li), 0.75 (Na), 0.8 (K), 0.85 (Rb), and 0.95 (Cs) were used. The NaCl ion and SPC/E water parameters were taken from previous studies.[25,44]

Table 3. Summary of the Alkali Halide Crystal Simulations Using the Final Parameters[a]

	E_{pot} (kJ/mol)	ρ_{sim} (g/cm^3)	ρ_{exp} (g/cm^3)	a_{sim} (nm)	a_{exp} (nm)
NaF	−1217.74	1.965	2.558	0.257	0.231
NaCl	−808.24	2.108	2.163	0.285	0.281
NaBr	−776.08	3.326	3.246	0.295	0.297
NaI	−750.94	3.878	3.665	0.303	0.323
LiCl	−1178.03	1.776	2.069	0.270	0.257
KCl	−725.29	1.980	1.990	0.315	0.314
RbCl	−692.73	2.800	2.859	0.325	0.327
CsCl	−650.12	3.990	3.973	0.419	0.412
KI	−663.23	3.406	3.125	0.343	0.353
CsBr	−628.80	4.582	4.453	0.433	0.429

[a] Symbols are E_{pot}, average total potential energy per molecule (N_s); ρ, mass density; and a, unit cell dimension. Subscripts sim and exp indicate simulation and experimental data,[70] respectively.

value for σ_{FF} which reproduced the crystal lattice dimensions could not be obtained by a simple scaling approach. Hence, we decided to develop specific values of σ_{FF} (and $\varepsilon_{\mathrm{FF}}$), which attempted to reproduce both the crystal lattice dimensions and solution KBIs.

Second, the initial cation parameters for Li$^+$, K$^+$, Rb$^+$, and Cs$^+$ were developed by reference to the crystal dimensions of LiCl, KCl, RbCl, and CsCl and the relevant cation−water contact distances. After the values of σ_{++} were determined by scaling the ionic radii of each ion, the values of ε_{++} were varied to reproduce the crystal unit cell dimensions and the cation−water contact distances. Unfortunately, and in agreement with our earlier study of NaCl,[25] we could not reproduce the experimental KBIs in aqueous solution by using standard combination rules for ε_{++} in aqueous solutions. Hence, modified ε parameters were developed specifically for the cation−water oxygen interactions. This interaction was subsequently modified by introducing a simple scale factor (s) for the interaction between metal ions and water oxygens such that $\varepsilon_{\mathrm{MO}} = s(\varepsilon_{\mathrm{MM}}\ \varepsilon_{\mathrm{OO}})^{0.5}$. This parameter scales the repulsive part of the LJ potential controlling the contact distance between an ion and first shell water molecules. The scale factor was set to unity for all other interactions. The final scaling factors for the metal ion−water interactions are provided in Table 2. Unfortunately, this simple approach did not work for LiCl. Hence, unique (not scaled) LJ values were determined for this salt by reference to the LiCl crystal dimensions and solution KBIs.

RESULTS

The main goal for the force fields developed here is to reproduce, as far as possible, the experimental KBIs for aqueous salt solutions as a function of salt concentration. Hence, we present this comparison first. This is followed by a comparison of a series of additional properties of solution mixtures, not included in the original parametrization, which is presented in an effort to both fully characterize the models and to establish the range of applicability of the models. As the solutions involve a variety of highly polarizing ions, the inherent many body interactions would be expected to vary substantially between different salts and also with concentration. Therefore, it should be obvious that it is essentially impossible to reproduce all the available experimental data using such a simple LJ 6−12 plus Coulomb model. Wherever possible, we have attempted to highlight any disagreement with experimental results and possible causes for these errors.

The experimental excess coordination numbers for sodium halides and alkali chlorides are displayed in Figure 1. The results presented in Figure 1 have been extracted from the experimental thermodynamic data on aqueous salt solutions and represent the primary target data for the current parametrization approach. The data display systematic trends between the different salts, which provide information concerning the underlying molecular distributions. At low concentrations (<0.1 m), the distributions are dominated by the Debye−Hueckel behavior leading to positive values for the ion−ion excess coordination numbers (N_{cc}). This behavior reverses at higher salt concentrations and indicates, with the exception of NaF, an increase in ion solvation by water. Similar results have been observed in other studies.[57,58]

Table 2 shows the final Lennard-Jones parameters used in our simulations. The LJ parameters for Na$^+$ and Cl$^-$ were taken from Weerasinghe and Smith.[25] As the size of the cation increased, the value of σ increased and that of ε essentially decreased. A similar trend is observed for the anions. Peng et al. have argued in favor of such trends in the LJ parameters, although the trend in ε parameters is the opposite of that expected (decreasing with atomic number, not increasing).[6] Their work used a LJ 9−12 potential, and hence the argument might not be so clear for the LJ 6−12 plus Coulomb models, or for systems with large polarization effects, where the ε parameter is linked to a scaling of the repulsive wall which resists the electrostatic attraction, rather than the usual relationship to dispersion interactions. The trend in the values of σ was also observed by both Joung and Cheatham[4] and Horinek et al.[17] However, any trend in the values of ε was absent from both these previous works.

Table 3 indicates the potential energy, density, and lattice constants obtained for the salt crystals studied in this work. The simulated crystal dimensions exhibit an average error of 3% with a maximum error of 10%. In the Supporting Information, the lattice energies of the Kirkwood-Buff models are compared to the experimental data and the force fields developed by Peng et al.[6] The KBFF models consistently overestimate the lattice energies. While reproducing the crystal lattice energies of salts was not a goal of the present parametrization, the results suggest that the

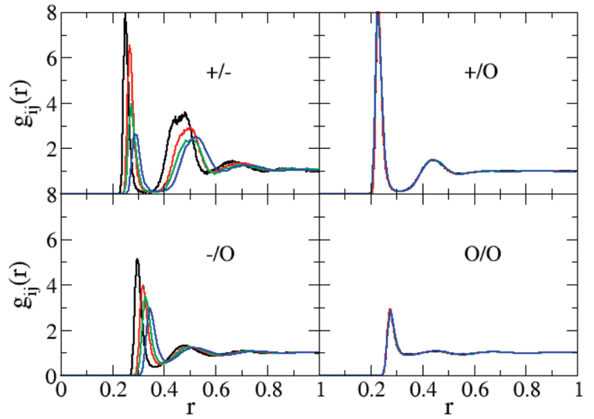


Figure 2. Radial distribution functions obtained from simulations of 1 M sodium salt solutions containing NaF (black lines), NaCl (red lines), NaBr (green lines), and NaI (blue lines). Cations, anions, and the water oxygen are denoted by the symbols +, −, and 0, respectively.

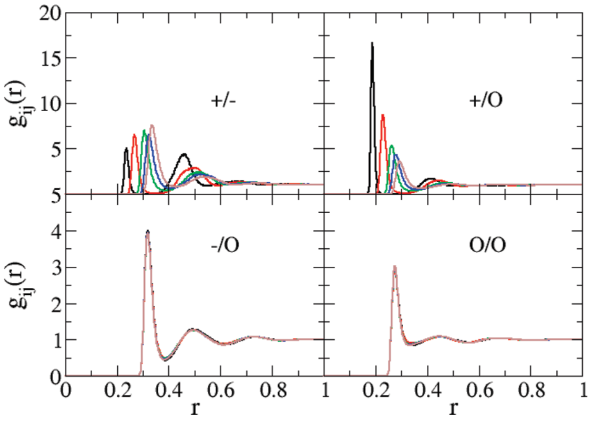


Figure 3. Radial distribution functions obtained from simulations of 1 M chloride salt solutions containing LiCl (black lines), NaCl (red lines), KCl (green lines), RbCl (blue lines), and CsCl (brown lines). Cations, anions, and the water oxygen are denoted by the symbols +, −, and 0, respectively.

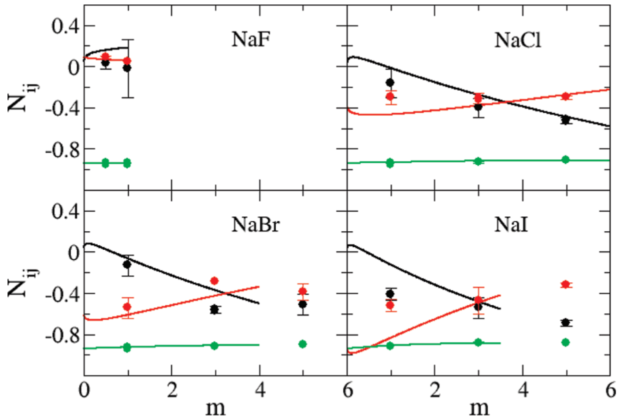


Figure 4. Excess coordination numbers as a function of salt molality. The N_{cc} (black lines), N_{cw} (red lines), and N_{ww} (green lines) are obtained from a KB analysis of the experimental data. The N_{cc} (black dots), N_{cw} (red dots), and N_{ww} (green dots) are obtained from simulations performed with the KBFF models.

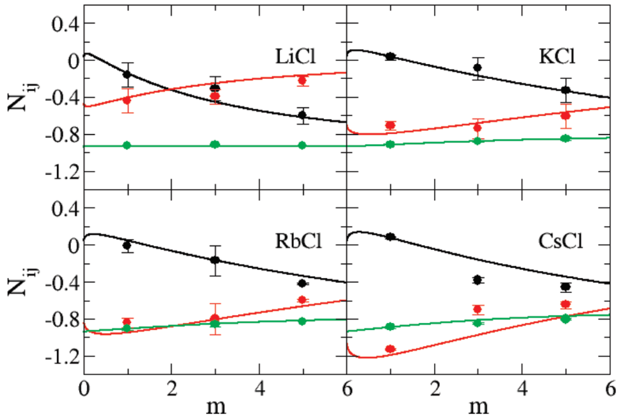


Figure 5. Excess coordination numbers as a function of salt molality. The N_{cc} (black lines), N_{cw} (red lines), and N_{ww} (green lines) are obtained from a KB analysis of the experimental data. The N_{cc} (black dots), N_{cw} (red dots), and N_{ww} (green dots) are obtained from simulations performed with the KBFF models.

current force fields may result in crystal lattices which are too stable with respect to the solution phase. This could be a concern for future simulations. However, a recent study of the KBFF model for NaCl indicates an approximate solubility of 7.9 m,[59] compared to the experimental value of 6.1 m.[60] The higher observed solubility suggests that, if anything, the opposite could be true. Some of these differences are probably related to the rather crude LJ 6−12 potential used in the current work which is known to fail for crystals.[6] Our main aim in studying the salt crystal lattice properties was to guide the systematic development of anion and cation LJ σ parameters. Furthermore, the enthalpies of mixing appear to be well reproduced (see later), indicating good compatibility with the SPC/E water model. Hence, we have not considered any further attempts to significantly improve the current data.

The radial distribution functions (rdf's) obtained from the 1 M salt simulations are displayed in Figure 2 for the sodium halides and in Figure 3 for the alkali metal chlorides. The sodium to halide anion−cation rdf's displayed a large first (ion pair) and a significant second (solvent separated ion pair) peak, which is in agreement with experimental results.[61] All rdf's approached unity beyond 1 nm. The first shell coordination numbers, n_{ij}, as well as the distances to the first rdf maximum (contact distance), R_{max}, and the first rdf minimum (first solvation shell), R_{min}, were calculated from the corresponding rdf's as a function of the solution molality and are presented in the Supporting Information. The final contact distances for Li^+, Na^+, K^+, Rb^+, Cs^+, F^-, Cl^-, Br^-, and I^- were 0.19, 0.23, 0.26, 0.28, 0.29, 0.27, 0.32, 0.33, and 0.35 nm, respectively. As expected, the radius of the first hydration shell increased as the size of the cation and anion increased. The simulated contact distances agree with the experimental values of 0.20, 0.24, 0.28, 0.29, 0.31, 0.26, 0.32, 0.34, and 0.36,[62] respectively, to within a 0.01 nm root-mean-square (rms) deviation—a similar deviation to that exhibited by the force field of Joung and Cheatham.[4] The first water shell coordination numbers of Na^+, K^+, Rb^+, and Cs^+ in ∼4 M aqueous solutions were determined to be 4.9, 5.9, 6.2, and 6.4, respectively. As expected, and similar to the trend in the radii of the first hydration shell, the hydration numbers increase as the

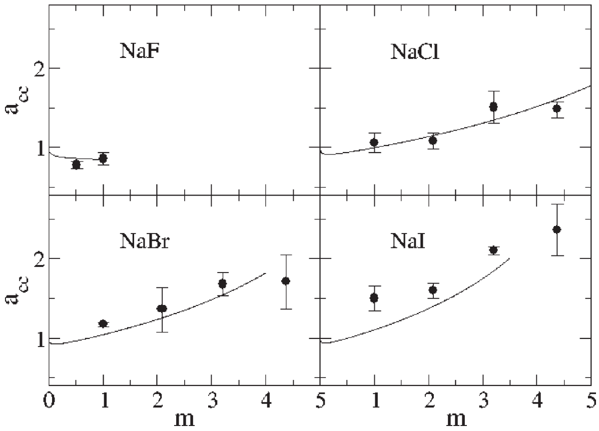


Figure 6. Activity derivatives for sodium salts as a function of salt molality. Lines are obtained from a KB analysis of the experimental data, while symbols correspond to the results obtained with the KBFF models.

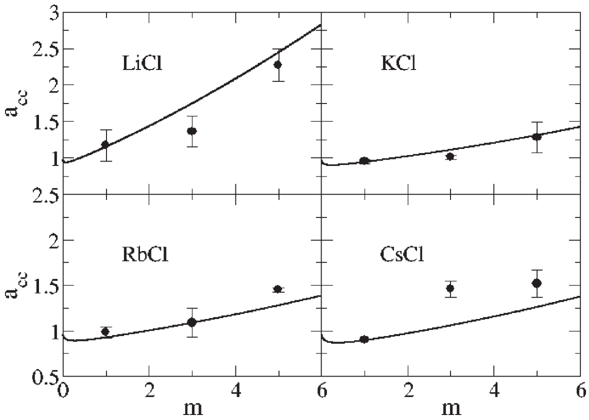


Figure 7. Activity derivatives for chloride salts as a function of salt molality. Lines are obtained from a KB analysis of the experimental data, while symbols correspond to the results obtained with the KBFF models.

size of the cation increases. The predicted hydration numbers agree with those determined from X-ray and neutron scattering data under the same conditions[61]—4.9, 5.3, 6.9, and 7.5, respectively—to within a 0.2 rms deviation. The Supporting Information also indicates that the coordination numbers are sensitive not only to the size of the alkali metal ion but also to changes in the salt concentration. The degree of ion pairing increases with increasing concentration. We note that no aggregation or crystallization was observed during any of the simulations.

The simulated and experimental excess coordination numbers, N_{ij}, are shown in Figure 4 for the sodium halides and in Figure 5 for the alkali metal chlorides, as a function of salt molality. The KBFF models quantitatively reproduce the experimental data, although the simulated values were somewhat less accurate for NaI and CsCl solutions. The correct trends (with salt concentration) are reproduced for all salts. The ion−ion excess coordination numbers (black lines) did not vary significantly from salt to salt when compared to the variation in the ion−water excess coordination numbers (red lines), which is in agreement with the experimental data (see Figure 1). This suggests that changes to the ion−water and water−water distributions determine the solution behavior to a large extent. However, it is very difficult to clearly relate these composition-dependent changes to the force field parameters used here. The relatively poor agreement for the NaI and CsCl solutions probably arises due to the high polarizability of the anion and cations, respectively, which would make the development of parameters suitable for both crystals and aqueous solutions quite challenging.

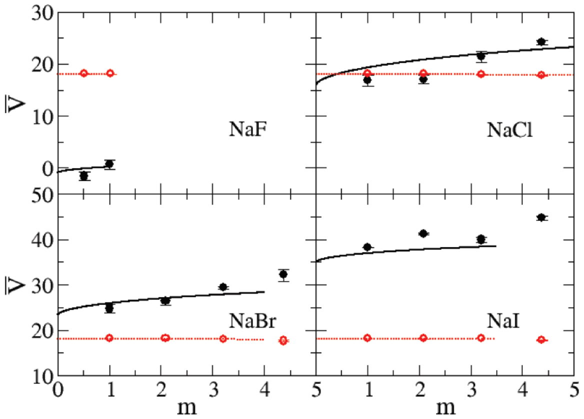


Figure 8. Partial molar volumes (cm^3/mol) for sodium salts as a function of salt molality. Lines are obtained from a KB analysis of the experimental data, while symbols correspond to the results obtained with the KBFF models. The partial molar volume of the salt is displayed in black with the partial molar volume of water displayed in red.

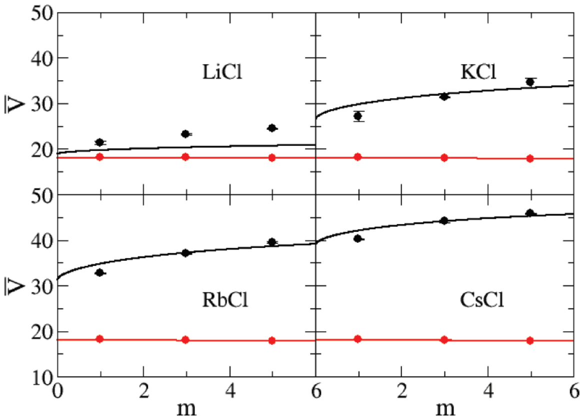


Figure 9. Partial molar volumes (cm^3/mol) for chloride salts as a function of salt molality. Lines are obtained from a KB analysis of the experimental data, while symbols correspond to the results obtained with the KBFF models. The partial molar volume of the salt is displayed in black with the partial molar volume of water displayed in red.

In Figures 6 and 7, the simulated activity derivatives (a_{cc}) as a function of molality are compared to the experimental values.[38] The KBFF model reproduced the correct increase in a_{cc} with concentrations at higher salt concentrations as indicated by the experimental data. We note that a_{cc} plays an important role for solutions as it characterizes the change in activity (chemical potential) of the salt with concentration.[31] Hence, accurate force fields are required to reproduce this data.[25] An expression for the molar activity coefficient ($y_c = y_\pm$) provided by the current force fields was obtained by taking appropriate derivatives of the fitting equations adopted for the experimental data (eq 4) and then obtaining parameters that best fit the simulated activity derivatives. The final fitting parameters are provided in the Supporting Information for most of the salt solutions studied here. It should

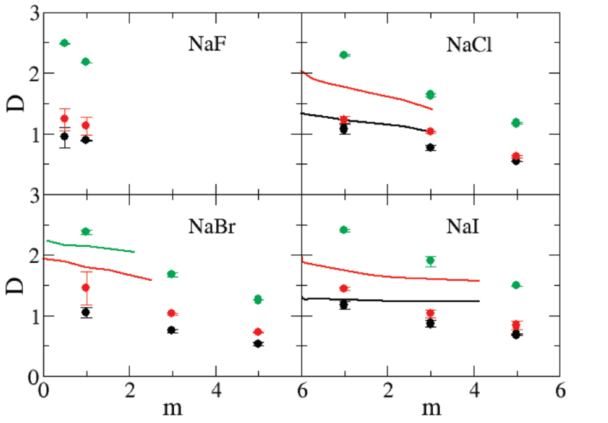


Figure 10. Diffusion constants ($\times 10^{-9}$ m^2/s) for sodium salts as a function of salt molality. The D_+ (black lines), D_- (red lines), and D_w (green lines) represent the experimental diffusion constant data,[71−74] while the D_+ (black dots), D_- (red dots), and D_w (green dots) were obtained from simulations using the KBFF models.

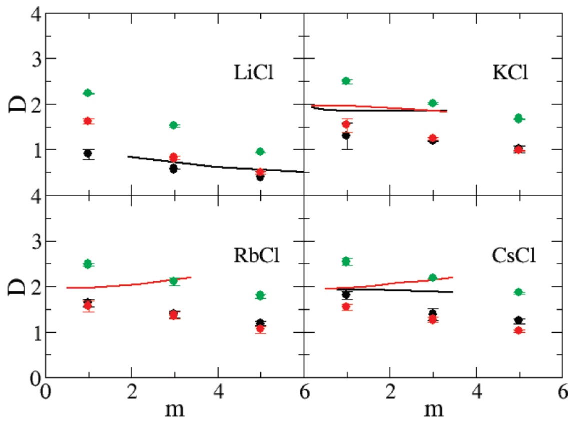


Figure 11. Diffusion constants ($\times 10^{-9}$ m^2/s) for chloride salts as a function of salt molality. The D_+ (black lines), D_- (red lines), and D_w (green lines) represent the experimental diffusion constant data,[75] while the D_+ (black dots), D_- (red dots), and D_w (green dots) were obtained from simulations using the KBFF models.

be noted that many common force fields do not correctly reproduce the experimental excess coordination numbers and activity derivatives.[20,22,23,25] For instance, in our previous work, we simulated 2 M NaCl solutions using a variety of salt force fields.[25] Many force fields provided values of a_{cc} < 0.5. Large deviations from experimental results are also observed for other solutes.[20,34,63] Hence, the data provided in Figures 6 and 7 for the present models, while not perfect, can be considered to be in good agreement with experimental results relative to typical results for similar force fields.

Figures 8 and 9 show the experimental and simulated partial molar volumes of both the water and salt as a function of the concentration. The experimental partial molar volumes of the salts generally increase monotonically, while that of water slightly decreases monotonically, as the salt concentration increases. The same trends were exhibited by the simulated values. Also, as expected, the partial molar volume of the salt increases as the size of the ions increases. The KBFF models reproduce the experimental data quantitatively except for LiCl, for which the salt partial molar volume is too large, presumably due to an overestimation of the cation size. This is also consistent with the low simulated crystal density. However, it was not possible to develop parameters using a smaller σ parameter for lithium and still reproduce the experimentally observed cation to water oxygen contact distance. Hence, we chose to correctly model this latter data.

The current models reproduce the excess coordination numbers, and therefore chemical potential derivatives and partial molar volumes, of a variety of salt solutions as a function of the concentration. This is the primarily goal for the KBFF models. However, it is important to test the models and their ability to reproduce other properties of salt solutions not included in the initial parametrization process, especially to see if they display significant deviations from experimental results, and to fully characterize the models in order to develop the exact range of properties for which the models will provide reliable results. The self-diffusion constants, calculated using the mean square fluctuation approach,[51] are displayed in Figures 10 and 11 as a function of alkali halide molality. The majority of the water, cation, and anion experimental diffusion constants all exhibit an essentially linear decrease with increasing salt molality. The notable exceptions are the diffusion constants for the chloride ion in RbCl and CsCl solutions. All the simulated diffusion constants decreased with salt concentration but typically displayed a stronger concentration dependence compared to experimental results. The self-diffusion constants of alkali metal cations increase with size even though the mass of the ions increases, confirming that the solvation of the cation is the most important factor for the diffusion constant.[64] In contrast, the self-diffusion constants of halide ions do not display any apparent correlation with the size of the ion. We note, however, that it is difficult to obtain quantitative agreement with the experimental data for most solutions, as even the diffusion constant of water varies considerably between water models and can be a factor of 2 too large.[65] The agreement with experimental results can be improved somewhat by correcting for finite size effects,[66] not included here, which typically result in larger (5−10%) diffusion coefficients. However, the simulated results would still appear to be more sensitive to changes in concentration compared to experimental results. It is unclear at present why this is the case. Comparison with diffusion data obtained for other models suggests the present models are reasonably competitive.[18]

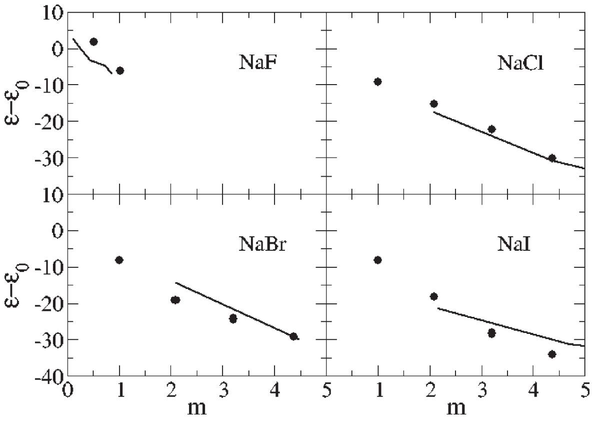


Figure 12. Dielectric decrements ($\varepsilon - \varepsilon_0$) for a series of sodium salts as a function of salt molality. Lines were obtained from the experimental dielectric constant data,[76−78] while the symbols correspond to data obtained from simulations using the KBFF models.

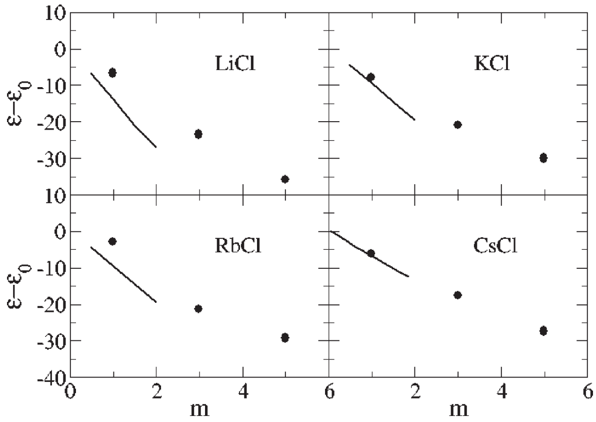


Figure 13. Dielectric decrements ($\varepsilon - \varepsilon_0$) for a series of chloride salts as a function of salt molality. Lines were obtained from the experimental dielectric constant data,[76−78] while the symbols correspond to data obtained from simulations using the KBFF models.

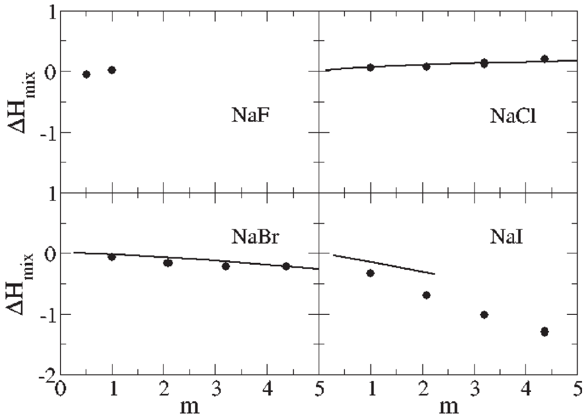


Figure 14. Excess enthalpy of mixing (kJ/mol) for sodium salts as a function of salt molality. Lines correspond to experimental data,[79] while symbols were obtained from simulations using the KBFF models.

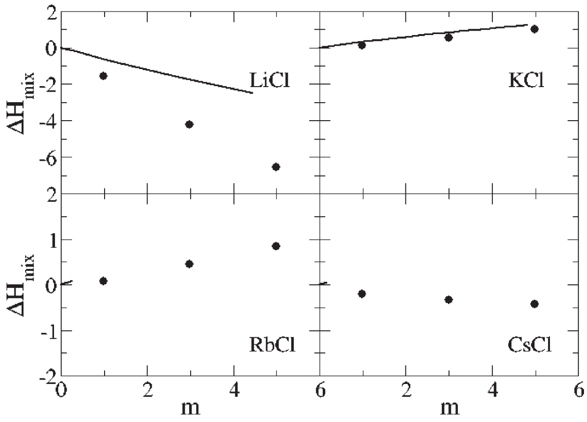


Figure 15. Excess enthalpy of mixing (kJ/mol) for chloride salts as a function of salt molality. Lines correspond to experimental data,[79] while symbols were obtained from simulations using the KBFF models.

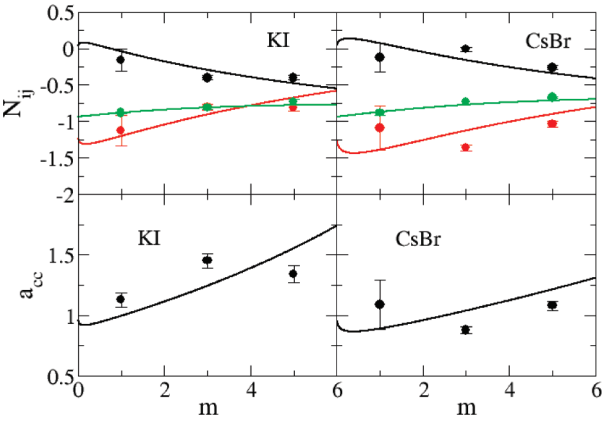


Figure 16. Excess coordination numbers as a function of salt molality (top). The N_{cc} (black lines), N_{cw} (red lines), and N_{ww} (green lines) are obtained from a KB analysis of the experimental data. The N_{cc} (black dots), N_{cw} (red dots), and N_{ww} (green dots) are obtained from simulations. Activity derivatives as a function of salt molality (bottom): Lines are obtained from a KB analysis of the experimental data, while symbols correspond to results obtained using the KBFF models.

The dielectric decrements ($\varepsilon - \varepsilon_0$) of alkali halide salts solutions, calculated from the dipole moment fluctuations,[53] are displayed in Figures 12 and 13. Here, ε is the relative permittivity of the solution, and ε_0 is the relative permittivity of pure water. The value of $\varepsilon_0 = 63$ obtained for pure water using the SPC/E model[67] is low compared to the experimental value of 78.[68] Hence, quantitative agreement for the absolute permittivities is not possible with this water model. The experimental relative permittivity for all salt solutions decreases as a function of molality, and this trend is clearly reproduced by the current models. The only exception appears to be NaF solutions at low concentrations where a small increase is observed. This increase was also reproduced in the present simulations. The KBFF models reproduce the experimental decrement data well, with the possible exception of LiCl solutions, compared to the simulated uncertainty of ±5.

The excess enthalpies of mixing for the sodium halides as a function of salt molality are displayed in Figures 14 and 15. The excess enthalpy of mixing for each sodium halide solution is calculated by the difference between the molar potential energy in the solution phase and in the crystal and pure water phases.[54] The data indicate that the models reproduce the experimental mixing enthalpies in a quantitative manner for NaCl, NaBr, and KCl, while the results for NaI and LiCl are somewhat too favorable. The simulated data for alkali chlorides become increasingly more unfavorable on moving from Li^+ to Rb^+ but then change sign for CsCl solutions. We presume this is due to a change in crystal structure from FCC to BCC for CsCl. It should be noted that reasonable agreement for both the free energy and enthalpy of mixing must therefore indicate good estimates for the entropy of mixing (data not shown).

In the previous sections, we have developed parameters for a series of sodium halides and alkali metal chlorides by using Kirkwood-Buff theory as a guide. In order to demonstrate the transferability of the parameters to a variety of alkali halides, we have used the same ion parameters to study two other systems, aqueous KI and aqueous CsBr, which were not included in the previous parametrization and for which there are no longer any free parameters. The results are presented in Figures 16−18 and clearly suggest that, to a high degree of accuracy, the parameters

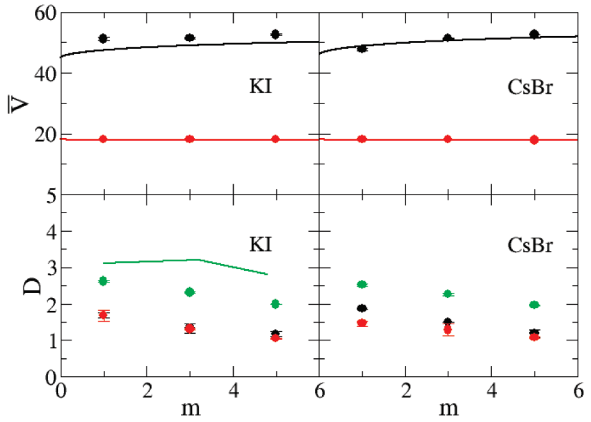


Figure 17. Partial molar volumes (cm^3/mol) as a function of salt molality (top). Lines are obtained from a KB analysis of the experimental data, while symbols correspond to results obtained using the KBFF models. The black lines and symbols represent the partial molar volume of the salt, while red lines and symbols indicate partial molar volume of water. Diffusion constants ($\times 10^{-9}$ m^2/s) as a function of salt molality (bottom): The D_+ (black lines), D_- (red lines), and D_w (green lines) are obtained from experimental diffusion constant data,[80] while the D_+ (black ●), D_- (red ○), and D_w (green ×) were obtained from simulations performed using the KBFF models.

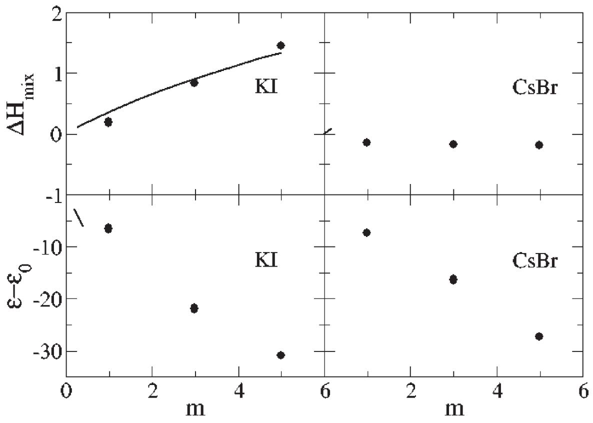


Figure 18. Excess enthalpy of mixing (kJ/mol) as a function of salt molality (top) and dielectric decrements as a function of salt molality (bottom). Lines correspond to the experimental data,[79] while symbols were obtained from simulations using the KBFF models.

developed here for the sodium and chloride salts are transferable to other alkali halide salts.

CONCLUSIONS

A series of models for aqueous alkali halide solutions have been developed by attempting to reproduce the experimentally derived Kirkwood-Buff integrals using molecular dynamics simulation. A major advantage of this type of approach is the ability to provide insight into salt activities in a computationally efficient manner and to ensure a reasonably accurate balance between solute—solute (N_{cc}) and solute—solvent (N_{cw}) distributions and, by inference, their interactions. Other physical and thermodynamic properties such as ion diffusion constants, relative permittivity, density, and heat of mixing have also been reasonably well reproduced. In addition, by examining the results obtained for aqueous KI and CsBr solutions, it has been clearly demonstrated that the parameters developed for sodium and chloride salts are transferable to other alkali halide salts. Unfortunately, not all the models provide good agreement for all the experimental data. To some degree, this is expected when using such simple models. The major issues involved the most highly polarizing ions (Li^+ an F^-), while the diffusion constant data also provided only modest agreement with experimental results. Hence, care should be taken when using the current models for these types of applications. The models are specifically designed to be used with the SPC/E water model, although, according to previous studies,[25,69] other simple point charge models should provide similar results. The recent models contribute to a consistent set of parameters that can eventually be used to study salt effects on peptides and proteins.

The solutions studied in this work include a variety of polarizable and polarizing anions and cations over a range of compositions. It is encouraging that one can reproduce much of the experimental data with the simple nonpolarizable models used here. However, to achieve this goal, it was necessary to break the standard combination rules when determining the cation—water interactions. The modified ε parameters actually lead to an increase in the cation—water interaction and can be thought of, to some degree, as a crude approach to incorporate polarization effects, which undoubtedly play a significant role in these solutions.

The present models provide an alternative to other recent ion force fields developed using more traditional approaches—such as the free energy of hydration. We have argued that the use of the experimental KBIs provides a rigorous test of force field accuracy and thereby provides ideal target data for the parametrization.[31] Furthermore, this can be achieved without a significant sacrifice in agreement with other solution properties. Whether the current models are substantially better than other, more traditional, models remains to be seen. This issue requires a more thorough and comprehensive study than is feasible here. The present models should be viewed as providing a reasonable balance between solute—solute, solute—solvent, and solvent—solvent interactions, as inferred by their resulting distributions, and are therefore suitable for studies of solute activities and cosolvent interactions with biomolecules.[30,63] Of course, one should always test that any potential model reasonably reproduces any specific properties of interest before use.

ASSOCIATED CONTENT

Supporting Information. Supporting Information is provided which includes tables containing a summary of all the simulations performed in this study, first shell coordination numbers, fitting constants for both the experimental and simulated activity data (eq 4), and a comparison of the present lattice energies with experimental and other simulation data. Additional figures are provided illustrating the RDFs as a function of composition. This information is available free of charge via the Internet at http://pubs.acs.org/.

AUTHOR INFORMATION

Corresponding Author

*Tel.: 785-532-5109. Fax: 785-532-6666. E-mail: pesmith@ksu.edu.

■ ACKNOWLEDGMENT

The project described was supported by Grant Number R01GM079277 from the National Institute of General Medical Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health.

■ REFERENCES

(1) Mclaughlin, S. *Annu. Rev. Biophys. Biol.* **1989**, *18*, 113–136.

(2) Anderson, C. F.; Record, M. T. *Annu. Rev. Phys. Chem.* **1995**, *46*, 657–700.

(3) Baldwin, R. L. *Biophys. J.* **1996**, *71*, 2056–2063.

(4) Joung, I. S.; Cheatham, T. E. *J. Phys. Chem. B* **2008**, *112*, 9020–9041.

(5) Aqvist, J. *J. Phys. Chem.* **1990**, *94*, 8021–8024.

(6) Peng, Z. W.; Ewig, C. S.; Hwang, M. J.; Waldman, M.; Hagler, A. T. *J. Phys. Chem. A* **1997**, *101*, 7243–7252.

(7) Rasaiah, J. C. *J. Chem. Phys.* **1970**, *52*, 704–715.

(8) Lee, S. H.; Rasaiah, J. C. *J. Chem. Phys.* **1994**, *101*, 6964–6974.

(9) Du, H.; Rasaiah, J. C.; Miller, J. D. *J. Phys. Chem. B* **2007**, *111*, 209–217.

(10) Lamoureux, G.; Roux, B. *J. Phys. Chem. B* **2006**, *110*, 3308–3322.

(11) Jensen, K. P.; Jorgensen, W. L. *J. Chem. Theory Comput.* **2006**, *2*, 1499–1509.

(12) Chen, A. A.; Pappu, R. V. *J. Phys. Chem. B* **2007**, *111*, 11884–11887.

(13) Auffinger, P.; Cheatham, T. E.; Vaiana, A. C. *J. Chem. Theory Comput.* **2007**, *3*, 1851–1859.

(14) Smith, D. E.; Dang, L. X. *J. Chem. Phys.* **1994**, *100*, 3757–3766.

(15) Dang, L. X. *J. Chem. Phys.* **1992**, *96*, 6970–6977.

(16) Dang, L. X.; Garrett, B. C. *J. Chem. Phys.* **1993**, *99*, 2972–2977.

(17) Horinek, D.; Mamatkulov, S. I.; Netz, R. R. *J. Chem. Phys.* **2009**, *130*, 124507.

(18) Joung, I. S.; Cheatham, T. E. *J. Phys. Chem. B* **2009**, *113*, 13279–13290.

(19) Bentenitis, N.; Cox, N. R.; Smith, P. E. *J. Phys. Chem. B* **2009**, *113*, 12306–12315.

(20) Kang, M.; Smith, P. E. *J. Comput. Chem.* **2006**, *27*, 1477–1485.

(21) Weerasinghe, S.; Smith, P. E. *J. Phys. Chem. B* **2005**, *109*, 15080–15086.

(22) Weerasinghe, S.; Smith, P. E. *J. Chem. Phys.* **2004**, *121*, 2180–2186.

(23) Weerasinghe, S.; Smith, P. E. *J. Chem. Phys.* **2003**, *118*, 10663–10670.

(24) Weerasinghe, S.; Smith, P. E. *J. Phys. Chem. B* **2003**, *107*, 3891–3898.

(25) Weerasinghe, S.; Smith, P. E. *J. Chem. Phys.* **2003**, *119*, 11342–11349.

(26) Chitra, R.; Smith, P. E. *J. Phys. Chem. B* **2002**, *106*, 1491–1500.

(27) Smith, P. E. *J. Chem. Phys.* **2008**, *129*, 124509.

(28) Ben-Naim, A. *J. Chem. Phys.* **1977**, *67*, 4884–4890.

(29) Ben-Naim, A. *Statistical Thermodynamics for Chemists and Biochemists*; Plenum Press: New York, 1992.

(30) Pierce, V.; Kang, M.; Aburi, M.; Weerasinghe, S.; Smith, P. E. *Cell Biochem. Biophys.* **2008**, *50*, 1–22.

(31) Weerasinghe, S.; Gee, M. B.; Kang, M.; Bentenitis, N.; Smith, P. E. In *Modeling Solvent Environments*; Feig, M., Ed.; Wiley-VCH: Weinheim, Germany, 2010.

(32) Matteoli, E.; Lepori, L. *J. Chem. Phys.* **1984**, *80*, 2856–2863.

(33) Ploetz, E. A.; Bentenitis, N.; Smith, P. E. *Fluid Phase Equilib.* **2010**, *290*, 43–47.

(34) Chitra, R.; Smith, P. E. *J. Chem. Phys.* **2001**, *115*, 5521–5530.

(35) Hess, B.; van der Vegt, N. F. A. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 13296–13300.

(36) Klasczyk, B.; Knecht, V. *J. Chem. Phys.* **2010**, *132*, 024109.

(37) Ben-Naim, A. *Molecular Theory of Solutions*; Oxford University Press: New York, 2006.

(38) Robinson, R. A.; Stokes, R. H. *Electrolyte Solutions*; 2nd ed.; Butterworths: London, 1959.

(39) Rosgen, J.; Pettitt, B. M.; Perkyns, J.; Bolen, D. W. *J. Phys. Chem. B* **2004**, *108*, 2048–2055.

(40) Sohnel, O.; Novotny, P. *Densities of Aqueous Solutions of Inorganic Substances*; Elsevier: Amsterdam, 1985.

(41) Kirkwood, J. G.; Buff, F. P. *J. Chem. Phys.* **1951**, *19*, 774–777.

(42) Fine, R. A.; Millero, F. J. *J. Chem. Phys.* **1973**, *59*, 5529–5536.

(43) Kusalik, P. G.; Patey, G. N. *J. Chem. Phys.* **1987**, *86*, 5110–5116.

(44) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269–6271.

(45) Van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *J. Comput. Chem.* **2005**, *26*, 1701–1718.

(46) Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Model.* **2001**, *7*, 306–317.

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

(48) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Dinola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

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

(50) Chandrasekhar, S. *Rev. Mod. Phys.* **1943**, *15*, 1–89.

(51) Chitra, R.; Smith, P. E. *J. Phys. Chem. B* **2000**, *104*, 5854–5864.

(52) Allen, M. P.; Tildesley, D. J.; *Computer Simulation of Liquids*; Oxford University Press: Oxford, U. K., 1987.

(53) Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1994**, *100*, 577–585.

(54) Walser, R.; Mark, A. E.; van Gunsteren, W. F.; Lauterbach, M.; Wipff, G. *J. Chem. Phys.* **2000**, *112*, 10450–10459.

(55) Chitra, R.; Smith, P. E. *J. Chem. Phys.* **2001**, *114*, 426–435.

(56) Kokubo, H.; Rosgen, J.; Bolen, D. W.; Pettitt, B. M. *Biophys. J.* **2007**, *93*, 3392–3407.

(57) Newman, K. E. *Chem. Soc. Rev.* **1994**, *23*, 31–40.

(58) Shimizu, S.; Boon, C. L. *J. Chem. Phys.* **2004**, *121*, 9147–9155.

(59) Smith, P. E. *Fluid Phase Equilib.* **2010**, *290*, 36–42.

(60) Weast, R. C. *CRC Handbook of Chemistry and Physics*; 66th ed.; CRC Press, Inc.: Boca Raton, FL, 1985.

(61) Ansell, S.; Barnes, A. C.; Mason, P. E.; Neilson, G. W.; Ramos, S. *Biophys. Chem.* **2006**, *124*, 171–179.

(62) Marcus, Y. *Chem. Rev.* **1988**, *88*, 1475–1498.

(63) Kang, M.; Smith, P. E. *Fluid Phase Equilib.* **2007**, *256*, 14–19.

(64) Atkins, P. W.; De Paula, J. *Physical Chemistry*; 7th ed.; W.H. Freeman: New York, 2002.

(65) Mark, P.; Nilsson, L. *J. Phys. Chem. A* **2001**, *105*, 9954–9960.

(66) Yeh, I. C.; Hummer, G. *J. Phys. Chem. B* **2004**, *108*, 15873–15879.

(67) Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1994**, *100*, 3169–3174.

(68) Heger, K.; Uematsu, M.; Franck, E. U. *Ber. Bunsen Phys. Chem.* **1980**, *84*, 758–762.

(69) Patra, M.; Karttunen, M. *J. Comput. Chem.* **2004**, *25*, 678–689.

(70) Davey, W. P. *Phys. Rev.* **1923**, *21*, 143–161.

(71) Wang, J. H.; Kennedy, J. W. *J. Am. Chem. Soc.* **1950**, *72*, 2080–2083.

(72) Nelson, F.; Marcinkowsky, A. E.; Kraus, K. A. In *Research and development progress report/office of saline water*; 302nd ed.; U.S. Dept. of the Interior, Office of Saline Water: Washington, DC, 1968.

(73) Tyrrell, H. J. V.; Harris, K. R. *Diffusion in Liquids*; Butterworths: London, 1984.

(74) Easteal, A. J.; Woolf, L. A. *J. Phys. Chem.* **1986**, *90*, 2441–2445.

(75) Kumamoto, E.; Kimizuka, H. *Bull. Chem. Soc. Jpn.* **1979**, *52*, 2145–2146.

(76) Haggis, G. H.; Hasted, J. B.; Buchanan, T. J. *J. Chem. Phys.* **1952**, *20*, 1452–1465.

(77) Harris, F. E.; Okonski, C. T. *J. Phys. Chem.* **1957**, *61*, 310–319.

(78) Buchner, R.; Hefter, G. T.; May, P. M. *J. Phys. Chem. A* **1999**, *103*, 1–9.

(79) Beggerow, G. In *Landolt-Boernstein*; Springer-Verlag: Berlin, 1976; Vol. 2.

(80) Matyash, I. V.; Toryanik, A. I.; Yashkichev, V. I. *Zh. Strukt. Khim.* **1964**, *5*, 777–778.

Keep It Flexible: Driving Macromolecular Rotary Motions in Atomistic Simulations with GROMACS

Carsten Kutzner,* Jacek Czub, and Helmut Grubmüller

Department of Theoretical and Computational Biophysics, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany

S Supporting Information

ABSTRACT: We describe a versatile method to enforce the rotation of subsets of atoms, e.g., a protein subunit, in molecular dynamics (MD) simulations. In particular, we introduce a "flexible axis" technique that allows realistic flexible adaptions of both the rotary subunit as well as the local rotation axis during the simulation. A variety of useful rotation potentials were implemented for the GROMACS 4.5 MD package. Application to the molecular motor F_1-ATP synthase demonstrates the advantages of the flexible axis approach over the established fixed axis rotation technique.

1. INTRODUCTION

Biomolecular function often rests on or is performed through motions of subunits. Rotary motions, in particular, are essential for the function of many motor proteins. These nanomotors use the free energy of chemical reactions or ion concentration gradients to generate mechanical torque. Rotary mechanisms were unequivocally demonstrated for three molecular engines, the F_o and F_1 motors in F-ATP synthase (F-ATPase)[1,2] and the bacterial flagellar motor.[3] Recently, rotary motion was also shown for the V_1 portion of the prokaryotic homologue of the vacuolar ATPase (V-ATPase).[4] Other motor proteins that are assumed to be rotary include DNA helicases[5] and proteins that translocate viral DNA into preformed capsids.[6−8]

The molecular mechanisms by which chemical reactions or transmembrane gradients drive protein rotary motions are in most cases not understood in full detail.[9] Also, these often quite complex motions are typically too slow or infrequent to be accessible to equilibrium molecular dynamics (MD) simulations. To overcome this limitation, techniques have been developed to exert external forces[10−12] or torques[13−15] to certain subunits to induce rotation and/or to increase its rate without severely perturbing the nature of the involved structural changes. This approach has also been used to simulate experiments in which biomolecules, such as proteins or DNA, are mechanically driven to rotate by externally applied torques by single molecule manipulation techniques.[16] In one impressive example, the F_1 portion of ATP synthase (F_1-ATPase) has been shown to produce ATP when the γ subunit is enforced to rotate using magnetic tweezers.[17]

With exceptions,[18] in most simulations involving external torque, a fixed, "stiff" rotation axis has been used so far[15,19,20] (dashed line in Figure 1A). As shown in the figure, this approach does not properly describe situations such as F_1-ATPase, where the rotating part flexibly adapts (dotted lines) to the steric restraints set by the bearing (gray). To more realistically describe biomolecular rotations, we have therefore developed a flexible axis rotation technique that (i) exerts torque with a curved axis that flexibly fits the shape of an arbitrarily shaped cavity (Figure 1A), (ii) avoids any impact or bias previously introduced by the necessary choice of the pivot for the axis, (iii) perturbs the internal dynamics and flexibility of the rotated structure as little as possible, and (iv) allows the curvature of the axis to adapt to structural changes of the bearing. In summary, a rotated fragment such as the γ subunit inside the ATPase $\alpha_3\beta_3$ stator should deform like a rotating pipe-cleaner.

To clarify notation and to explain the basic ingredients needed for the flexible technique, we start with a recapitulation of the established fixed axis rotation, as implemented, e.g., in NAMD[21] or EGO.[22] From these notions, several more complex potentials will be developed and characterized, and the resulting forces will be derived. We will then motivate and describe in detail the flexible axis approach, for which we present two different variants. After outlining details of our GROMACS[23,24] implementation, we will apply flexible axis rotation to the F_1-ATPase molecular motor and test if our approach is indeed capable of providing more accurate torque or free energy profiles.

2. FIXED AXIS ROTATION

Stationary Axis with an Isotropic Potential. In the established fixed axis approach[15,19−22] (Figure 1B), torque on a group of N atoms with positions $\mathbf{x}_i$ (denoted "rotation group") is applied by rotating a reference set of atomic positions—usually their initial positions $\mathbf{y}_i^0$—at a constant angular velocity ω around an axis defined by a direction vector $\hat{\mathbf{v}}$ and a pivot point $\mathbf{u}$. To that aim, each atom with position $\mathbf{x}_i$ is attracted by a "virtual spring" potential to its moving reference position $\mathbf{y}_i = \Omega(t)\ (\mathbf{y}_i^0 - \mathbf{u})$, where $\Omega(t)$ is a matrix that describes the rotation around the axis. In the simplest case, the "springs" are described by a harmonic potential

$$V^{\text{iso}} = \frac{k}{2}\sum_{i=1}^{N} w_i[\Omega(t)(\mathbf{y}_i^0 - \mathbf{u}) - (\mathbf{x}_i - \mathbf{u})]^2 \qquad (1)$$

Received: November 17, 2010
Published: March 31, 2011

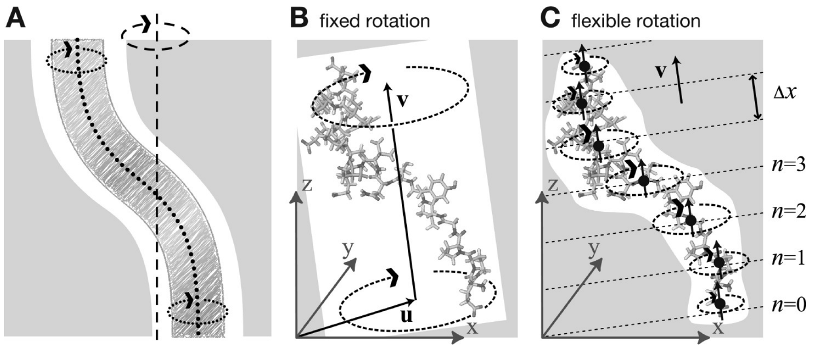


Figure 1. Comparison of fixed and flexible axis rotation. (A) Rotating the sketched shape inside the white tubular cavity creates severe artifacts when a conventional fixed rotation axis (dashed) is used. More realistically, the shape would revolve like a flexible pipe-cleaner (dotted) inside the bearing (gray). (B) Fixed rotation around an axis **v** with a pivot point specified by the vector **u**. (C) Subdividing the rotating fragment into slabs with separate rotation axes (↑) and pivot points (•) for each slab allows for the required flexibility. The distance between two slabs with indices n and $n + 1$ is Δx.

with optional mass-weighted prefactors $w_i = Nm_i/M$ with total mass $M = \sum_{i=1}^{N} m_i$. The rotation matrix $\Omega(t)$ is

$$\Omega(t) = \begin{pmatrix} \cos \omega t + v_x^2 \xi & v_x v_y \xi - v_z \sin \omega t & v_x v_z \xi + v_y \sin \omega t \\ v_x v_y \xi + v_z \sin \omega t & \cos \omega t + v_y^2 \xi & v_y v_z \xi - v_x \sin \omega t \\ v_x v_z \xi - v_y \sin \omega t & v_y v_z \xi + v_x \sin \omega t & \cos \omega t + v_x^2 \xi \end{pmatrix}$$

where v_x, v_y, and v_z are the components of the normalized rotation vector $\hat{\mathbf{v}}$ and $\xi := 1 - \cos(\omega t)$. As illustrated in Figure 2A for a single atom j, the rotation matrix $\Omega(t)$ operates on the initial reference positions $\mathbf{y}_j^0 = \mathbf{x}_j(t_0)$ of atom j at $t = t_0$. At a later time t, the reference position has rotated away from its initial place (along the blue dashed line), resulting in the force

$$\mathbf{F}_j^{\text{iso}} = -\nabla_j V^{\text{iso}} = k w_j [\Omega(t)(\mathbf{y}_j^0 - \mathbf{u}) - (\mathbf{x}_j - \mathbf{u})] \tag{2}$$

which is directed toward the reference position.

Pivot Free Isotropic Potential. We first address the bias introduced by an arbitrary choice of the pivot vector **u**. This arbitrariness is avoided by defining as the pivot the center of mass $\mathbf{x}_c$ of the rotation group

$$\mathbf{x}_c = \frac{1}{M}\sum_{i=1}^{N} m_i \mathbf{x}_i \text{ and } \mathbf{y}_c^0 = \frac{1}{M}\sum_{i=1}^{N} m_i \mathbf{y}_i^0 \tag{3}$$

which yields the "pivot-free" potential

$$V^{\text{iso-pf}} = \frac{k}{2}\sum_{i=1}^{N} w_i [\Omega(t)(\mathbf{y}_i^0 - \mathbf{y}_c^0) - (\mathbf{x}_i - \mathbf{x}_c)]^2 \tag{4}$$

with forces

$$\mathbf{F}_j^{\text{iso-pf}} = k w_j [\Omega(t)(\mathbf{y}_j^0 - \mathbf{y}_c^0) - (\mathbf{x}_j - \mathbf{x}_c)] \tag{5}$$

Without mass-weighting, the pivot $\mathbf{x}_c$ is the geometrical center of the group.

Parallel Motion Potential Variant. Obviously, the forces generated by the isotropic potentials (eqs 1 and 4) also contain components parallel to the rotation axis and thereby restrain motions along the axis of either the whole rotation group (in case of V^{iso}) or within the rotation group (in case of $V^{\text{iso-pf}}$). For cases where unrestrained motion along the axis is preferred, we have implemented a "parallel motion" variant by eliminating all components parallel to the rotation axis for the potential. This is achieved by projecting the distance vectors between reference and actual positions:

$$\mathbf{r}_i = \Omega(t)(\mathbf{y}_i^0 - \mathbf{u}) - (\mathbf{x}_i - \mathbf{u}) \tag{6}$$

onto the plane perpendicular to the rotation vector

$$\mathbf{r}_i^\perp := \mathbf{r}_i - (\mathbf{r}_i \cdot \hat{\mathbf{v}})\hat{\mathbf{v}} \tag{7}$$

yielding

$$\begin{aligned} V^{\text{pm}} &= \frac{k}{2}\sum_{i=1}^{N} w_i (\mathbf{r}_i^\perp)^2 = \frac{k}{2}\sum_{i=1}^{N} w_i \{\Omega(t)(\mathbf{y}_i^0 - \mathbf{u}) - (\mathbf{x}_i - \mathbf{u}) \\ &\quad - \{[\Omega(t)(\mathbf{y}_i^0 - \mathbf{u}) - (\mathbf{x}_i - \mathbf{u})]\cdot\hat{\mathbf{v}}\}\hat{\mathbf{v}}\}^2 \end{aligned} \tag{8}$$

and similarly

$$\mathbf{F}_j^{\text{pm}} = k w_j \mathbf{r}_j^\perp \tag{9}$$

Pivot-Free Parallel Motion Potential. Replacing in eq 8 the fixed pivot **u** with the center of mass $\mathbf{x}_c$ yields the pivot-free variant of the parallel motion potential. With

$$\mathbf{s}_i = \Omega(t)(\mathbf{y}_i^0 - \mathbf{y}_c^0) - (\mathbf{x}_i - \mathbf{x}_c) \tag{10}$$

the respective potential and forces are

$$V^{\text{pm-pf}} = \frac{k}{2}\sum_{i=1}^{N} w_i (\mathbf{s}_i^\perp)^2 \tag{11}$$

$$\mathbf{F}_j^{\text{pm-pf}} = k w_j \mathbf{s}_j^\perp \tag{12}$$

Radial Motion Potential. In the above variants, the minimum of the rotation potential is either a single point at the reference position $\mathbf{y}_i$ (for the isotropic potentials) or a single line through $\mathbf{y}_i$ parallel to the rotation axis (for the parallel motion potentials). As a result, radial forces restrict radial motions of the atoms. The two subsequent types of rotation potentials, V^{rm} and V^{rm2}, drastically reduce or even eliminate this effect. The first variant,

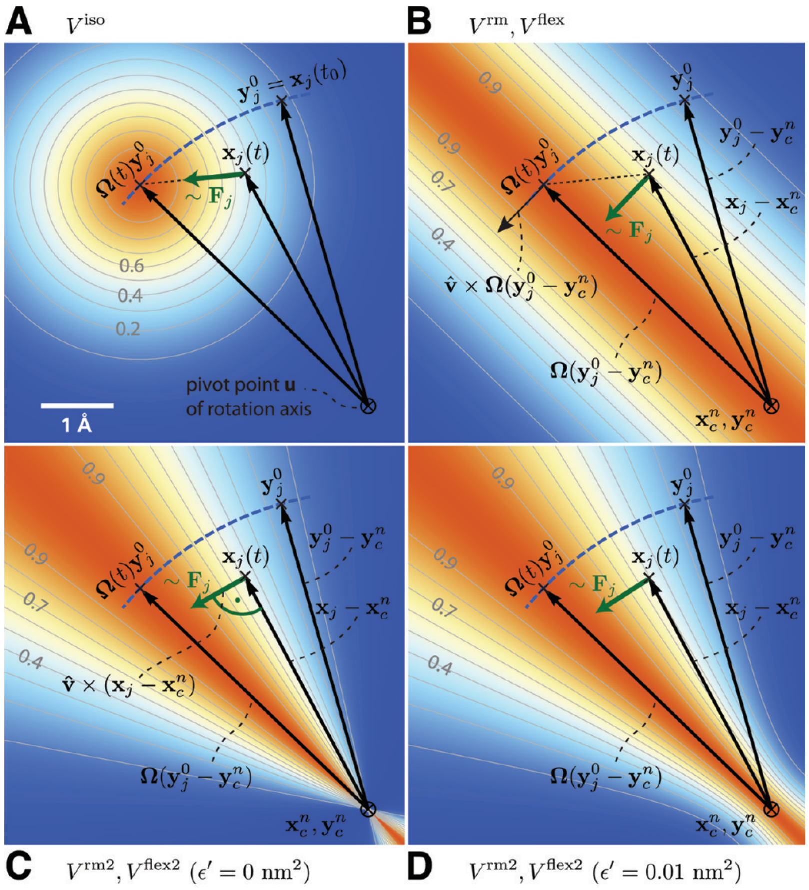

A
$V^{\rm iso}$
$\mathbf{y}_j^0 = \mathbf{x}_j(t_0)$
$\mathbf{x}_j(t)$
$\Omega(t)\mathbf{y}_j^0$
$\sim \mathbf{F}_j$
0.6
0.4
0.2
pivot point **u** of rotation axis
1 Å
B
$V^{\rm rm}, V^{\rm flex}$
0.9
0.9
0.7
0.4
$\mathbf{y}_j^0$
$\mathbf{y}_j^0 - \mathbf{y}_c^n$
$\mathbf{x}_j(t)$
$\Omega(t)\mathbf{y}_j^0$
$\mathbf{x}_j - \mathbf{x}_c^n$
$\sim \mathbf{F}_j$
$\hat{\mathbf{v}} \times \Omega(\mathbf{y}_j^0 - \mathbf{y}_c^n)$
$\Omega(\mathbf{y}_j^0 - \mathbf{y}_c^n)$
$\mathbf{x}_c^n, \mathbf{y}_c^n$
0.9
0.9
0.7
0.4
$\mathbf{y}_j^0$
$\mathbf{y}_j^0 - \mathbf{y}_c^n$
$\Omega(t)\mathbf{y}_j^0$
$\mathbf{x}_j(t)$
$\sim \mathbf{F}_j$
$\mathbf{x}_j - \mathbf{x}_c^n$
$\hat{\mathbf{v}} \times (\mathbf{x}_j - \mathbf{x}_c^n)$
$\Omega(\mathbf{y}_j^0 - \mathbf{y}_c^n)$
$\mathbf{x}_c^n, \mathbf{y}_c^n$
0.9
0.9
0.7
0.4
$\mathbf{y}_j^0$
$\mathbf{y}_j^0 - \mathbf{y}_c^n$
$\Omega(t)\mathbf{y}_j^0$
$\mathbf{x}_j(t)$
$\sim \mathbf{F}_j$
$\mathbf{x}_j - \mathbf{x}_c^n$
$\Omega(\mathbf{y}_j^0 - \mathbf{y}_c^n)$
$\mathbf{x}_c^n, \mathbf{y}_c^n$
C
$V^{\rm rm2}, V^{\rm flex2}$ $(\epsilon' = 0\ {\rm nm}^2)$
D
$V^{\rm rm2}, V^{\rm flex2}$ $(\epsilon' = 0.01\ {\rm nm}^2)$

Figure 2. Selection of different rotation potentials discussed in the text and definition of notation. All four potentials V (color coded) are shown for a single atom at position $\mathbf{x}_j(t)$. (A) Isotropic potential $V^{\rm iso}$, (B) radial motion potential $V^{\rm rm}$ and flexible potential $V^{\rm flex}$, (C,D) radial motion 2 potential $V^{\rm rm2}$ and flexible 2 potential $V^{\rm flex2}$ for $\varepsilon' = 0$ nm^2 (C) and $\varepsilon' = 0.01$ nm^2 (D). The rotation axis is perpendicular to the plane and marked by ⊗. The light gray contours indicate Boltzmann factors $e^{-V/(k_BT)}$ in the $\mathbf{x}_j$ plane for $T = 300$ K and $k = 200$ kJ/(mol·nm^2). The green arrow shows the direction of the force $\mathbf{F}_j$ acting on atom j; the blue dashed line indicates the motion of the reference position.

$V^{\rm rm}$ (Figure 2B), eliminates all force components parallel to the vector connecting the reference atom and the rotation axis

$$V^{\rm rm} = \frac{k}{2}\sum_{i=1}^{N} w_i[\mathbf{p}_i \cdot (\mathbf{x}_i - \mathbf{u})]^2 \tag{13}$$

with

$$\mathbf{p}_i := \frac{\hat{\mathbf{v}} \times \Omega(t)(\mathbf{y}_i^0 - \mathbf{u})}{||\ \hat{\mathbf{v}} \times \Omega(t)(\mathbf{y}_i^0 - \mathbf{u})\ \ ||} \tag{14}$$

This variant depends only on the distance $\mathbf{p}_i \cdot (\mathbf{x}_i - \mathbf{u})$ of atom i from the plane spanned by $\hat{\mathbf{v}}$ and $\Omega(t)(\mathbf{y}_i^0 - \mathbf{u})$. The resulting force is

$$\mathbf{F}_j^{\rm rm} = -kw_j[\mathbf{p}_j \cdot (\mathbf{x}_j - \mathbf{u})]\mathbf{p}_j \tag{15}$$

Pivot-Free Radial Motion Potential. Proceeding similar to the pivot-free isotropic potential yields a pivot-free version of the above potential. With

$$\mathbf{q}_i := \frac{\hat{\mathbf{v}} \times \Omega(t)(\mathbf{y}_i^0 - \mathbf{y}_c^0)}{||\hat{\mathbf{v}} \times \Omega(t)(\mathbf{y}_i^0 - \mathbf{y}_c^0)||} \tag{16}$$

the potential and force for the pivot free variant of the radial motion potential read

$$V^{\rm rm\text{-}pf} = \frac{k}{2}\sum_{i=1}^{N} w_i[\mathbf{q}_i \cdot (\mathbf{x}_i - \mathbf{x}_c)]^2 \tag{17}$$

$$\mathbf{F}_j^{\rm rm\text{-}pf} = -kw_j[\mathbf{q}_j \cdot (\mathbf{x}_j - \mathbf{x}_c)]\mathbf{q}_j + k\frac{m_j}{M}\sum_{i=1}^{N} w_i[\mathbf{q}_i \cdot (\mathbf{x}_i - \mathbf{x}_c)]\mathbf{q}_i \tag{18}$$

Radial Motion 2 Alternative Potential. As seen in Figure 2B, the force resulting from $V^{\rm rm}$ still contains a small, second-order radial component. In most cases, this perturbation is tolerable; if not, the following alternative, $V^{\rm rm2}$, fully eliminates the radial contribution to the force, as depicted in Figure 2C,

$$V^{\rm rm2} = \frac{k}{2}\sum_{i=1}^{N} w_i \frac{[(\hat{\mathbf{v}} \times (\mathbf{x}_i - \mathbf{u})) \cdot \Omega(t)(\mathbf{y}_i^0 - \mathbf{u})]^2}{||\ \hat{\mathbf{v}} \times (\mathbf{x}_i - \mathbf{u})\ \ ||^2 + \varepsilon'} \tag{19}$$

where a small parameter ε' has been introduced to avoid singularities. For $\varepsilon' = 0$ nm^2, the equipotential planes are spanned by $\mathbf{x}_i - \mathbf{u}$ and $\hat{\mathbf{v}}$, yielding a force perpendicular to $\mathbf{x}_i - \mathbf{u}$, thus not contracting or expanding structural parts that moved away from or toward the rotation axis.

We note that this variant is particularly suitable for free energy calculations via umbrella sampling techniques,[25] because the radial orientation of the equipotential planes shown in Figure 2C guarantees statistically consistent sampling of adjacent umbrella windows, as required for a consistent definition of the free energy profile via subspace projection. To see why this is actually the case, note that consistent umbrella sampling requires that for adjacent umbrella windows the "stack" of ($3N - 1$ dimensional) configurational subspaces defined by the values of the chosen reaction coordinate agrees, subspace by subspace, with the one defined by the values of the umbrella potential. This in turn requires that the equipotential planes shown in Figure 2 coincide with those of a rotated potential, which is obviously the case for Figure 2C, but not for Figure 2A or B.

Choosing a small positive ε' (e.g., $\varepsilon' = 0.01\ \mathrm{nm}^2$, Figure 2D) in the denominator of eq 19 yields a well-defined potential and continuous forces also close to the rotation axis, which is not the case for $\varepsilon' = 0\ \mathrm{nm}^2$ (Figure 2C). With

$$\mathbf{r}_i := \Omega(t)(\mathbf{y}_i^0 - \mathbf{u}) \tag{20}$$

$$\mathbf{s}_i := \frac{\hat{\mathbf{v}} \times (\mathbf{x}_i - \mathbf{u})}{\|\ \hat{\mathbf{v}} \times (\mathbf{x}_i - \mathbf{u})\ \|} \equiv \Psi_i \hat{\mathbf{v}} \times (\mathbf{x}_i - \mathbf{u}) \tag{21}$$

$$\Psi_i^* := \frac{1}{\|\ \hat{\mathbf{v}} \times (\mathbf{x}_i - \mathbf{u})\ \|^2 + \varepsilon'} \tag{22}$$

the force on atom j reads

$$\mathbf{F}_j^{\mathrm{rm2}} = -k \left\{ w_j (\mathbf{s}_j \cdot \mathbf{r}_j) \left[\frac{\Psi_j^*}{\Psi_j} \mathbf{r}_j - \frac{\Psi_j^{*2}}{\Psi_j^3} (\mathbf{s}_j \cdot \mathbf{r}_j) \mathbf{s}_j \right] \right\} \times \hat{\mathbf{v}} \tag{23}$$

Pivot-Free Radial Motion 2 Potential. The pivot free variant of the above potential is

$$V^{\mathrm{rm2\text{-}pf}} = \frac{k}{2} \sum_{i=1}^{N} w_i \frac{[(\hat{\mathbf{v}} \times (\mathbf{x}_i - \mathbf{x}_c)) \cdot \Omega(t)(\mathbf{y}_i^0 - \mathbf{y}_c)]^2}{\|\ \hat{\mathbf{v}} \times (\mathbf{x}_i - \mathbf{x}_c)\ \|^2 + \varepsilon'} \tag{24}$$

with

$$\mathbf{r}_i := \Omega(t)(\mathbf{y}_i^0 - \mathbf{y}_c) \tag{25}$$

$$\mathbf{s}_i := \frac{\hat{\mathbf{v}} \times (\mathbf{x}_i - \mathbf{x}_c)}{\|\ \hat{\mathbf{v}} \times (\mathbf{x}_i - \mathbf{x}_c)\ \|} \equiv \Psi_i \hat{\mathbf{v}} \times (\mathbf{x}_i - \mathbf{x}_c) \tag{26}$$

$$\Psi_i^* := \frac{1}{\|\hat{\mathbf{v}} \times (\mathbf{x}_i - \mathbf{x}_c)\|^2 + \varepsilon'} \tag{27}$$

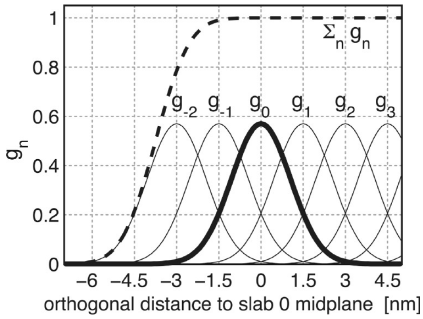


Figure 3. Gaussian functions g_n centered at $n\Delta x$ for a slab distance $\Delta x = 1.5$ nm and $n \geq -2$. Gaussian function g_0 is highlighted in bold; the dashed line depicts the sum of the shown Gaussian functions.

the force on atom j reads

$$\begin{aligned} \mathbf{F}_j^{\mathrm{rm2\text{-}pf}} = & -k \left\{ w_j (\mathbf{s}_j \cdot \mathbf{r}_j) \left[\frac{\Psi_j^*}{\Psi_j} \mathbf{r}_j - \frac{\Psi_j^{*2}}{\Psi_j^3} (\mathbf{s}_j \cdot \mathbf{r}_j) \mathbf{s}_j \right] \right\} \times \hat{\mathbf{v}} \\ & + k \frac{m_j}{M} \left\{ \sum_{i=1}^{N} w_i (\mathbf{s}_i \cdot \mathbf{r}_i) \left[\frac{\Psi_i^*}{\Psi_i} \mathbf{r}_i - \frac{\Psi_i^{*2}}{\Psi_i^3} (\mathbf{s}_i \cdot \mathbf{r}_i) \mathbf{s}_i \right] \right\} \times \hat{\mathbf{v}} \end{aligned} \tag{28}$$

3. FLEXIBLE AXIS ROTATION

As sketched in Figure 1A,B, the rigid body behavior of the fixed axis rotation scheme is a drawback for many applications. In particular, deformations of the rotation group are suppressed when the equilibrium atom positions directly depend on the reference positions. To avoid this limitation, eqs 18 and 24 will now be generalized toward a "flexible axis", as sketched in Figure 1C. This will be achieved by subdividing the rotation group into a set of equidistant slabs perpendicular to the rotation vector, and by applying a separate rotation potential to each of these slabs. Figure 1C shows the midplanes of the slabs as dotted straight lines and the centers as thick black dots.

To avoid discontinuities in the potential and in the forces, we define "soft slabs" by weighing the contributions of each slab n to the total potential function V^{flex} by a Gaussian function

$$g_n(\mathbf{x}_i) = \Gamma \exp\left(-\frac{\beta_n^2(\mathbf{x}_i)}{2\sigma^2} \right) \tag{29}$$

centered at the midplane of the nth slab. Here, σ is the width of the Gaussian function, Δx the distance between adjacent slabs, and

$$\beta_n(\mathbf{x}_i) := \mathbf{x}_i \cdot \hat{\mathbf{v}} - n\Delta x \tag{30}$$

A most convenient choice is $\sigma = 0.7\Delta x$ and

$$1/\Gamma = \sum_{n \in \mathbb{Z}} \exp\left(-\frac{\left(n - \frac{1}{4}\right)^2}{2 \times 0.7^2} \right) \approx 1.75464$$

which yields a nearly constant sum, essentially independent of $\mathbf{x}_i$ (dashed line in Figure 3), i.e.,

$$\sum_{n \in \mathbb{Z}} g_n(\mathbf{x}_i) = 1 + \varepsilon(\mathbf{x}_i) \tag{31}$$

with $|\varepsilon(\mathbf{x}_i)| < 1.3 \times 10^{-4}$. This choice also implies that the individual contributions to the force from the slabs add up to unity such that no further normalization is required.

To each slab center $\mathbf{x}_c^n$, all atoms contribute by their Gaussian-weighted (optionally also mass-weighted) position vectors $g_n(\mathbf{x}_i)\mathbf{x}_i$. The instantaneous slab centers $\mathbf{x}_c^n$ are calculated from the current positions $\mathbf{x}_i$

$$\mathbf{x}_c^n = \frac{\sum_{i=1}^{N} g_n(\mathbf{x}_i) m_i \mathbf{x}_i}{\sum_{i=1}^{N} g_n(\mathbf{x}_i) m_i} \tag{32}$$

while the reference centers $\mathbf{y}_c^n$ are calculated from the reference positions $\mathbf{y}_i^0$

$$\mathbf{y}_c^n = \frac{\sum_{i=1}^{N} g_n(\mathbf{y}_i^0) m_i \mathbf{y}_i^0}{\sum_{i=1}^{N} g_n(\mathbf{y}_i^0) m_i} \tag{33}$$

Due to the rapid decay of g_n, each slab will essentially involve contributions from atoms located within ∼3Δx from the slab center only.

Flexible Axis Potential. We consider two flexible axis variants. For the first variant, the slab segmentation procedure with Gaussian weighting is applied to the radial motion potential (eq 18/Figure 2B), yielding as the contribution of slab n

$$V^n = \frac{k}{2} \sum_{i=1}^{N} w_i g_n(\mathbf{x}_i) [\mathbf{q}_i^n \cdot (\mathbf{x}_i - \mathbf{x}_c^n)]^2$$

and a total potential function

$$V^{\text{flex}} = \sum_n V^n \tag{34}$$

Note that the global center of mass $\mathbf{x}_c$ used in eq 18 is now replaced by $\mathbf{x}_c^n$, the center of mass of the slab. With

$$\mathbf{q}_i^n := \frac{\hat{\mathbf{v}} \times \Omega(t)(\mathbf{y}_i^0 - \mathbf{y}_c^n)}{||\ \hat{\mathbf{v}} \times \Omega(t)(\mathbf{y}_i^0 - \mathbf{y}_c^n)\ ||} \tag{35}$$

$$b_i^n := \mathbf{q}_i^n \cdot (\mathbf{x}_i - \mathbf{x}_c^n) \tag{36}$$

the resulting force on atom j reads

$$\begin{aligned} \mathbf{F}_j^{\text{flex}} = & -k w_j \sum_n g_n(\mathbf{x}_j)\ b_j^n \left\{ \mathbf{q}_j^n - b_j^n \frac{\beta_n(\mathbf{x}_j)}{2\sigma^2} \hat{\mathbf{v}} \right\} \\ & + k m_j \sum_n \frac{g_n(\mathbf{x}_j)}{\sum_h g_n(\mathbf{x}_h)} \sum_{i=1}^{N} w_i g_n(\mathbf{x}_i)\ b_i^n \\ & \left\{ \mathbf{q}_i^n - \frac{\beta_n(\mathbf{x}_j)}{\sigma^2} [\mathbf{q}_i^n \cdot (\mathbf{x}_j - \mathbf{x}_c^n)] \hat{\mathbf{v}} \right\} \end{aligned} \tag{37}$$

Note that for V^{flex}, as defined, the slabs are fixed in space and so are the reference centers $\mathbf{y}_c^n$. If during the simulation the rotation group moves too far in the $\mathbf{v}$ direction, it may enter a region where—due to the lack of nearby reference positions—no reference slab centers are defined, rendering the potential evaluation impossible. We therefore have included a slightly modified version of this potential that avoids this problem by attaching the midplane of slab $n = 0$ to the center of mass of the rotation group, yielding slabs that move with the rotation group. This is achieved by subtracting the center of mass $\mathbf{x}_c$ of the group from the positions

$$\tilde{\mathbf{x}}_i = \mathbf{x}_i - \mathbf{x}_c, \text{ and } \tilde{\mathbf{y}}_i^0 = \mathbf{y}_i^0 - \mathbf{y}_c^0 \tag{38}$$

such that

$$V^{\text{flex-}t} = \frac{k}{2} \sum_n \sum_{i=1}^{N} w_i g_n(\tilde{\mathbf{x}}_i) \left[\frac{\hat{\mathbf{v}} \times \Omega(t)(\tilde{\mathbf{y}}_i^0 - \tilde{\mathbf{y}}_c^n)}{||\hat{\mathbf{v}} \times \Omega(t)(\tilde{\mathbf{y}}_i^0 - \tilde{\mathbf{y}}_c^n)\ ||} \cdot (\tilde{\mathbf{x}}_i - \tilde{\mathbf{x}}_c^n) \right]^2 \tag{39}$$

To simplify the force derivation, and for efficiency reasons, we here assume $\mathbf{x}_c$ to be constant, and thus $\partial \mathbf{x}_c/\partial x = \partial \mathbf{x}_c/\partial y = \partial \mathbf{x}_c/\partial z = 0$. The resulting force error is small (on the order of $O(1/N)$ or $O(m_j/M)$ if mass-weighting is applied) and can therefore be tolerated. With this assumption, the forces $\mathbf{F}^{\text{flex-t}}$ have the same form as eq 37.

Flexible Axis 2 Alternative Potential. In our second variant, slab segmentation is applied to V^{rm2} (eq 24), resulting in a flexible axis potential without radial force contributions (Figure 2C)

$$V^{\text{flex2}} = \frac{k}{2} \sum_{i=1}^{N} \sum_n w_i g_n(\mathbf{x}_i) \frac{[(\hat{\mathbf{v}} \times (\mathbf{x}_i - \mathbf{x}_c^n)) \cdot \Omega(t)(\mathbf{y}_i^0 - \mathbf{y}_c^n)]^2}{||\ \hat{\mathbf{v}} \times (\mathbf{x}_i - \mathbf{x}_c^n)\ ||^2 + \varepsilon'} \tag{40}$$

with

$$\mathbf{r}_i^n := \Omega(t)(\mathbf{y}_i^0 - \mathbf{y}_c^n) \tag{41}$$

$$\mathbf{s}_i^n := \frac{\hat{\mathbf{v}} \times (\mathbf{x}_i - \mathbf{x}_c^n)}{||\ \hat{\mathbf{v}} \times (\mathbf{x}_i - \mathbf{x}_c^n)\ ||} \equiv \psi_i \hat{\mathbf{v}} \times (\mathbf{x}_i - \mathbf{x}_c^n) \tag{42}$$

$$\psi_i^* := \frac{1}{||\ \hat{\mathbf{v}} \times (\mathbf{x}_i - \mathbf{x}_c^n)\ ||^2 + \varepsilon'} \tag{43}$$

$$W_j^n := \frac{g_n(\mathbf{x}_j) m_j}{\sum_h g_n(\mathbf{x}_h) m_h} \tag{44}$$

$$\mathbf{S}^n := \sum_{i=1}^{N} w_i g_n(\mathbf{x}_i)(\mathbf{s}_i^n \cdot \mathbf{r}_i^n) \left[\frac{\psi_i^*}{\psi_i} \mathbf{r}_i^n - \frac{\psi_i^{*2}}{\psi_i^3} (\mathbf{s}_i^n \cdot \mathbf{r}_i^n) \mathbf{s}_i^n \right] \tag{45}$$

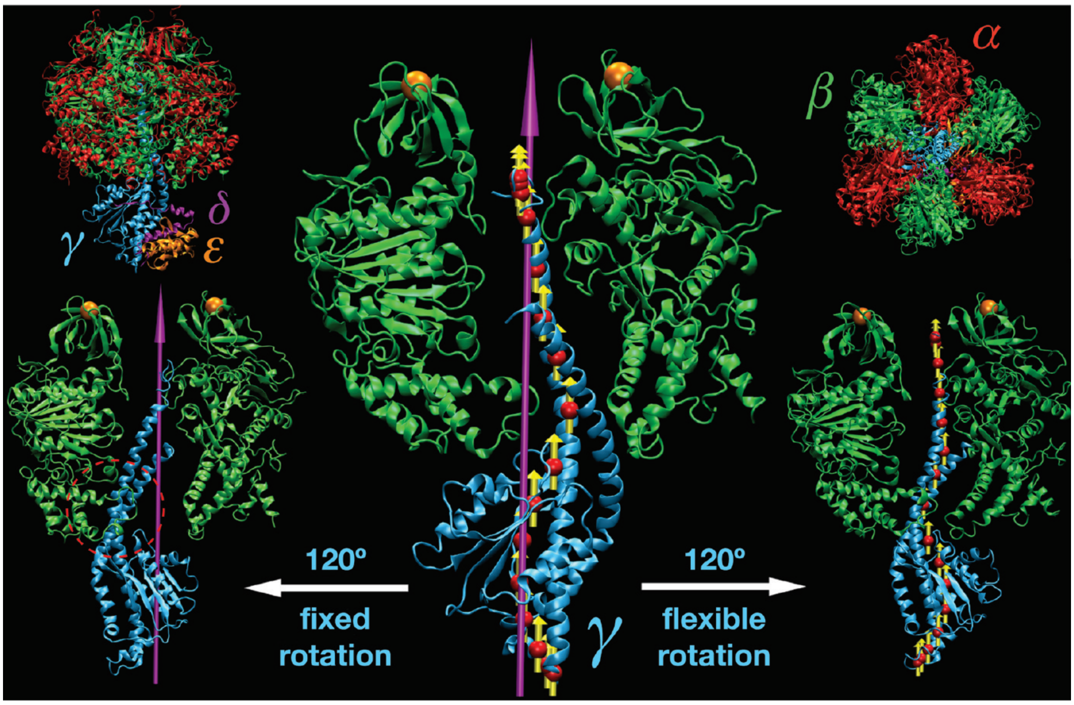


Figure 4. F_1-ATPase structure. In the upper left (right) corners, the full protein structure ($\alpha_3\beta_3\delta\varepsilon$) is shown in a side (top) view. Subunit color-coding is α, red; β, green; γ, cyan; δ, magenta; and ε, orange. The central panel illustrates the initial orientation of the rotor domain ($\gamma\delta\varepsilon$) with respect to the stator ($\alpha_3\beta_3$); for the sake of simplicity, only the γ and two β subunits are shown. The 3-fold symmetry axis of $\alpha_3\beta_3$ that was used as a rotation axis in V^{iso} is shown in magenta. The red spheres and yellow arrows depict slab centers and local rotation axes as used by the flexible potentials. The left and right side panels show the orientation of the rotor after 120° of enforced rotation using V^{iso} and V^{flex2}, respectively. The two orange spheres denote harmonic restraints applied to the N-terminal tags of the β subunits. This is to prevent co-rotation of the $\alpha_3\beta_3$ stator in close resemblance to single-molecule force probe experiments, in which the stator is immobilized by attaching the protein to the surface via His tags attached to one subunit type (usually the β chains). Figure prepared with VMD.[38]

the force on atom j reads

$$\begin{aligned}\mathbf{F}_j^{\text{flex2}} = & -k\left\{\sum_n w_j g_n(\mathbf{x}_j)(\mathbf{s}_j^n\cdot\mathbf{r}_j^n)\left[\frac{\psi_j^*}{\psi_j}\mathbf{r}_j^n - \frac{\psi_j^{*2}}{\psi_j^3}(\mathbf{s}_j^n\cdot\mathbf{r}_j^n)\mathbf{s}_j^n\right]\right\} \\ & \times\hat{\mathbf{v}} + k\left\{\sum_n W_j^n\mathbf{S}^n\right\}\times\hat{\mathbf{v}} - k\left\{\sum_n W_j^n\frac{\beta_n(\mathbf{x}_j)}{\sigma^2}\frac{1}{\psi_j}\mathbf{s}_j^n\cdot\mathbf{S}^n\right\}\hat{\mathbf{v}} \\ & + \frac{k}{2}\left\{\sum_n w_j g_n(\mathbf{x}_j)\frac{\beta_n(\mathbf{x}_j)}{\sigma^2}\frac{\psi_j^*}{\psi_j^2}(\mathbf{s}_j^n\cdot\mathbf{r}_j^n)^2\right\}\hat{\mathbf{v}} \qquad (46)\end{aligned}$$

Applying transformation 38 yields a translation-tolerant version of the flexible 2 potential, $V^{\text{flex2-}t}$. Again, assuming that $\partial\mathbf{x}_c/\partial x$, $\partial\mathbf{x}_c/\partial y$, and $\partial\mathbf{x}_c/\partial z$ are small, the resulting equations for $V^{\text{flex2-}t}$ and $\mathbf{F}^{\text{flex2-}t}$ are similar to those of V^{flex2} and $\mathbf{F}^{\text{flex2}}$.

4. GROMACS IMPLEMENTATION

For an efficient implementation, the following issues were taken into account. GROMACS 4 distributes the atoms among the parallel processors by domain-decomposing[24] the simulation box and assigning each domain to a processor. Depending on van der Waals and Coulomb cutoff settings, positions of atoms near the domain boundaries are communicated such that each processor can compute the forces assigned to its domain. However, the calculation of some of the proposed potentials and forces requires atom positions not present on the local processor. For instance, the pivot free potentials require the center of mass of the rotation group, while the flexible potentials require all N positions of the rotation group. The required coordinates are therefore distributed to all processors before the force calculations, which entails one extra communication step in the rotation module. Further, repeated expressions such as the last terms in eqs 18 and 28 are precalculated whenever possible. For the efficient computation of the forces $\mathbf{F}^{\text{flex}}$, the inner sum of the last term of eq 37

$$\sum_{i=1}^{N} w_i g_n(\mathbf{x}_i) b_i^n \left\{\mathbf{q}_i^n - \frac{\beta_n(\mathbf{x}_j)}{\sigma^2}[\mathbf{q}_i^n\cdot(\mathbf{x}_j - \mathbf{x}_c^n)]\hat{\mathbf{v}}\right\} \qquad (47)$$

is rewritten as

$$\mathbf{s}_n - \frac{\beta_n(\mathbf{x}_j)}{\sigma^2}[\mathbf{s}_n\cdot(\mathbf{x}_j - \mathbf{x}_c^n)]\cdot\hat{\mathbf{v}} \qquad (48)$$

such that the repeated terms

$$\mathbf{s}_n = \sum_{i=1}^{N} w_i g_n(\mathbf{x}_i) b_i^n \mathbf{q}_i^n \qquad (49)$$

are also precomputed for each relevant slab n and then used for the calculation of each $\mathbf{F}_j$ term. Likewise, for $\mathbf{F}^{\text{flex2}}$, the terms $\mathbf{S}^n$ (eq 45) of eq 46 are precalculated.

Moreover, for the flexible potentials, only significant contributions to V and $\mathbf{F}$ are computed, defined by a cutoff value of $g_n(\mathbf{x}) \geq g_n^{\min}$ with a default value $g_n^{\min} = 0.001$, which is checked according to a simple distance criterion. Also, the atoms of the rotation group are sorted according to their position along the rotation vector such that for each slab n, a first and a last index i between $g_n(\mathbf{x}_i) \geq g_n^{\min}\ \forall i \in [i_{\text{first}}...i_{\text{last}}]$ is stored, and all contributions outside that range can safely be ignored.

Special care has been taken for periodic boundary conditions. Here, the appropriate periodic image for each of the particles of the rotation group has to be chosen such that groups are not split. For fixed axis rotation, each atom is put closest to its current reference position. For the flexible and pivot-free radial motion potentials, each atom is put next to its position at the previous time step, thereby ensuring the integrity of all rotation fragments.

5. APPLICATION TO F_1-ATP SYNTHASE

As a sample application of our flexible axis approach, and to compare results obtained by fixed and flexible axis rotation, a series of all-atom MD simulations was performed in which the γ subunit of F_1-ATPase was enforced to rotate with respect to its stator part, $\alpha_3\beta_3$ (Figure 4).

F_1-ATPase is the soluble domain of the F_oF_1-ATP synthase, a rotary motor protein that synthesizes ATP from ADP using the electrochemical proton gradient across the membrane as its energy source.[26] The mitochondrial F_1-ATPase is an oligomeric protein consisting of nine polypeptide chains, $\alpha_3\beta_3\gamma\delta\varepsilon$.[27] In synthesis direction, F_1-ATPase is driven by the membrane-embedded proton-translocating F_o motor while the F_1 mobile subunit, $\gamma\delta\varepsilon$, rotates clockwise (seen from the membrane) within the bearing formed by the hexagonally arranged α and β chains.[28,26] The energy transmitted mechanically via the rotating subunit is subsequently used at the catalytic sites of $\alpha_3\beta_3$ for ATP synthesis. To prevent co-rotation, the $\alpha_3\beta_3$ hexamer is connected to the membrane-embedded F_o motor by a peripheral linker stalk. Despite numerous theoretical[29−31] and simulation studies,[15,32−36] the molecular mechanism of energy transmission between the rotor subunit and the ligand binding sites in the stator is still not fully understood.[37]

Simulation Setup. The initial configuration of the F_1 motor was based on the X-ray structure of bovine F_1-ATPase determined at 2.4 Å resolution[39] (Protein Data Bank entry 1E79). The covalently bound inhibitor as well as the glycerol and sulfate molecules were removed, leaving only Mg·ATP and Mg·ADP ligands in their respective binding sites. All crystal water molecules were retained. Two five-residue-long loops missing from the γ subunit were modeled with tCONCOORD.[40] Protonation states of ionizable groups were set according to the pK_a shifts calculated with the DelPhi[41] interface of WhatIf.[42] The protein structure was solvated with 87 321 water molecules in a 16.7 × 13.8 × 13.8 nm rectangular unit cell. To neutralize the system and to obtain physiological ionic strength, 261 Na^+ and 216 Cl^- ions were added. The system was energy-minimized using the steepest descent method in two stages. First, all heavy atoms of the protein and the protein's ligands were kept fixed; subsequently, all atoms in the system were allowed to relax.

All simulations were performed with GROMACS 4.0[24] in which the potentials V^{iso}, V^{flex}, and V^{flex2} were implemented. For convenience, we here also describe the newer 4.5 version, which produces the same results for the V^{iso}, V^{flex}, and V^{flex2} potentials but includes nine additional rotation potentials.

For the protein as well as its ligands and ions, the OPLS/AA force field[43,44] was used, and TIP4P[45] was used for the water. All production runs were carried out in the NPT ensemble at 300 K and 1 bar. Temperature and pressure were controlled by Nosé−Hoover[46,47] (coupling constant τ_t = 0.5 ps) and Parrinello−Rahman[48,49] (τ_p = 2.0 ps) schemes, respectively. To avoid severe density oscillations, the first 5 ns of the NPT equilibration run were performed with Berendsen weak coupling[50] for temperature and pressure. Periodic boundary conditions were applied in 3D, and electrostatic forces were calculated with the particle mesh Ewald (PME) method[51,52] using a real-space cutoff of 1 nm and an FFT grid density of 10 nm^{-1}. Lennard-Jones interactions were truncated at 1 nm. Covalent bond lengths in the protein and ligand were constrained to their reference values with P-LINCS.[53] SETTLE was used to constrain the water geometry.[54] Equations of motion were integrated using the leapfrog scheme with a time step of 2 fs. Prior to enforcing the rotor movement, the system was equilibrated for 10 ns at the target temperature and pressure. During the first 1 ns of this run, all protein heavy atoms were harmonically restrained to their initial positions.

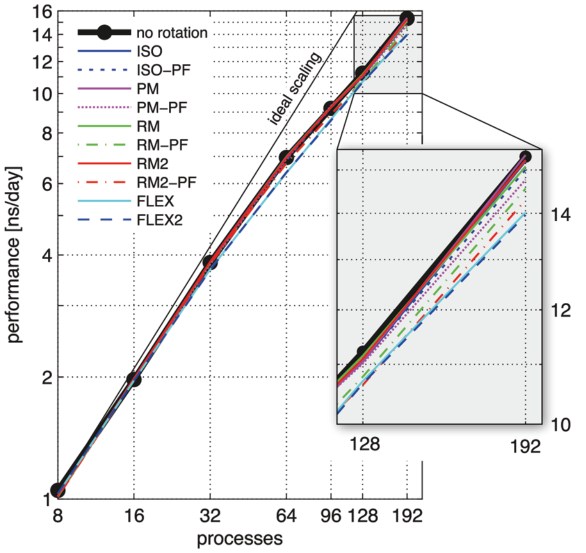


Figure 5. GROMACS 4.5 performance for various rotation potentials (colors) compared to a simulation without rotation. The system comprises 401 152 atoms in total, of which 2116 are subjected to the rotation potential. For the flexible potentials, slab distance Δx = 1 nm and $g_n^{\min}$ = 0.001 have been chosen. The thin black line denotes ideal scaling.

To mimic the effect exerted on the F_1 subunit by the rotation of the F_o motor, a potential of the form V^{iso} (eq 1), V^{flex} (34), or V^{flex2} (40) was applied during the production runs. All 272 C_α atoms of the γ subunit were chosen as a rotation group. The longest principal axis of the $\alpha_3\beta_3$ stator, i.e., the eigenvector of the inertia tensor of $\alpha_3\beta_3$ corresponding to the largest eigenvalue, was used as a rotation vector $\mathbf{v}$. For the fixed variant, the pivot vector $\mathbf{u}$ of the axis was placed at the center of mass of the $\alpha_3\beta_3$ units, thus defining the 3-fold pseudosymmetry axis of the stator subunit (Figure 4). For the flexible axis runs, a slab distance of Δx = 1 nm, a Gaussian function cutoff of $g_n^{\min}$ = 0.001, and ε' = 0 nm^2 were chosen. The γ reference positions were rotated counter-clockwise around $\mathbf{v}$ at an angular rate of ω = 0.021°/ps over 6 ns of the simulation time, yielding a 120° rotation of the $\gamma\delta\varepsilon$ domain. Due to its symmetry, this covers a complete synthesis cycle, as also seen from the observed stepped motion of the $\gamma\delta\varepsilon$ domain.[28] To examine the effect of the chosen spring constant k, for each of the three potentials, five runs were performed with k values ranging from 100 to 800 kJ/(mol·nm^2). In each case, all heavy atoms of the N-terminal six-residue sequences of each β subunit were harmonically restrained to their initial positions using a force constant of 1500 kJ/(mol·nm^2).

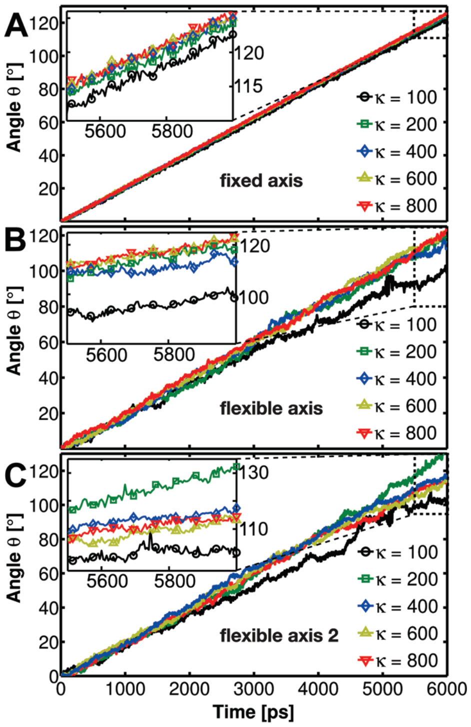


Figure 6. Time evolution of the γ rotor angle with respect to the $\alpha_3\beta_3$ symmetry axis for the F_1-ATPase motor enforced to rotate in the synthesis direction using the potentials V^{iso} (A), V^{flex} (B), and V^{flex2} (C) with spring constants k of 100−800 kJ/(mol·nm^2).

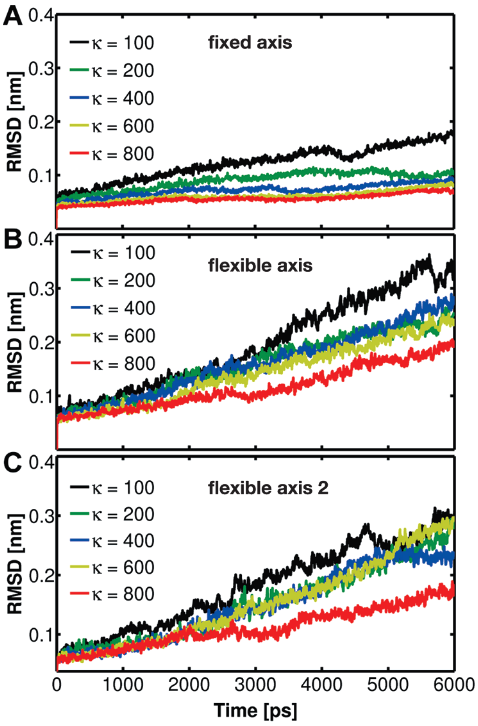


Figure 7. RMSD of the γ subunit backbone atoms with respect to the X-ray structure as a function of time for the F_1 motor driven to rotate in the synthesis direction using the potentials V^{iso} (A), V^{flex} (B), and V^{flex2} (C) with spring constants k of 100−800 kJ/(mol·nm^2).

Additionally, for k = 600 kJ/(mol·nm^2), a complete 360° rotation of the $\gamma\delta\varepsilon$ domain was simulated.

Performance. We demonstrate that, due to the optimizations described in the implementation section, simulation performance is nearly unaffected for typical setups such as the F_1-ATPse system, where only a small fraction of all atoms are subjected to a rotation potential. The described ATPase example with the implemented rotation types was benchmarked (Figure 5) on a cluster of Intel Xeon L5430 nodes connected by a DDR Infiniband network. Each node comprised eight processor cores running at 2.66 GHz. An Intel MPI 3.2.1 was used with the Intel 11.1 compiler and the FFTW 3.2 library. For the benchmarks, Coulomb and van der Waals cutoffs were set to 0.9 nm and the Fourier grid to 144 × 120 × 120 points, yielding a grid spacing of less than 0.12 nm in each dimension. Separation into long-range (PME-only) and short-range (particle−particle) processes was allowed. The optimal number of PME-only processes was derived with the g_tune_pme[55] tool using 2000 equilibration steps for the dynamic load balancing, with run times taken from 2000 subsequent steps.

With the $N = 272$ C_α atoms of the γ subunit as the rotation group, none of the potentials significantly reduced the MD performance. To be able to analyze the scaling behavior (Figure 5), the rotation group was therefore enlarged to contain all $N = 2116$ atoms of the γ subunit. As seen, the overall performance decreases only slightly compared to the case without rotation. For the most computationally demanding flexible potentials, on eight processors, a 2% decrease is seen and a 9% decrease on 192 processors.

6. RESULTS

Evolution of the Rotor Angle. To verify that the proposed methods properly control the motion of the rotary subunit, we first determined the time evolution of the rotor angle θ. The actual rotation angle $\theta(t)$ of the γ subunit was determined by a mass-weighted root-mean-square deviation (RMSD) fit to the initial ($\theta = 0°$) configuration of the γ backbone. Figure 6 shows $\theta(t)$ with respect to the $\alpha_3\beta_3$ symmetry axis.

The results show that in all 6-ns-long enforced rotation runs the rotor changes its orientation with respect to the stator by the expected 120°. The angle increases nearly linearly with time, with the slope reflecting the constant angular velocity of 0.021°/ps, at which the reference is rotated. For fixed axis rotation, the subunit closely follows the reference for all tested force constants k = 100−800 kJ/(mol·nm^2). In contrast, for both flexible variants, a less regular evolution is observed, as indicated by the large fluctuations of θ for k = 100 kJ/(mol·nm^2). These result from conformational changes of the rotor that occur because the flexible method allows for structural relaxations and adaptations to the bearing. Additionally, at high rotation velocities, frictional forces occur, which cause further conformational changes.

Movies illustrating the effect of the fixed and flexible axis methods have been included within the Supporting Information. In the movies, a V^{iso} and a V^{flex2} rotation potential with k = 600 kJ/(mol·nm^2) is applied to all C_α atoms of the γ subunit.

For a quantitative comparison of the γ subunit internal deformation, Figure 7 shows the time evolution of the RMSD of the γ backbone atoms from their initial configuration. Relatively small RMSD variations are observed for the fixed method, confirming nearly rigid-body like rotation. In contrast, both flexible axis methods allow for structural rearrangements particularly for small k values. A secondary structure analysis shows that for the F_1-ATPase flexibly rotating at 0.021°/ps the force constant k should be 200 kJ/(mol·nm^2) or larger to preserve the rotor

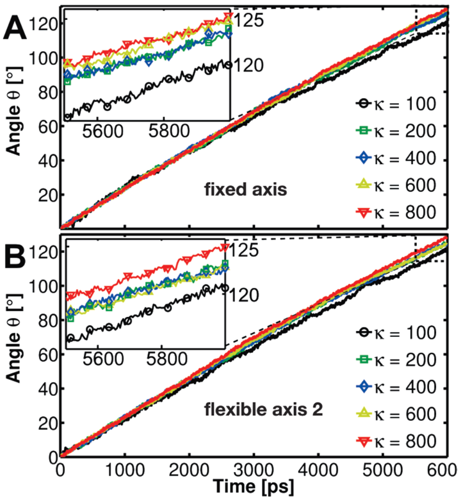


Figure 8. Time evolution of the angular position of the γ rotor computed as the best-fit angle with respect to the γ longest principal axis for the F_1 motor enforced to rotate in the synthesis direction using V^{iso} (A) and V^{flex2} (B).

coiled-coil conformation of the crystal structure (Figure 4). For any of the rotation potentials, the force constant will depend on the studied system and on the rotation rate. Generally, higher rotation rates will require larger force constants that stabilize the rotation group with the help of a stronger coupling to its reference. Yet, the decrease in conformational freedom with increasing k (Figure 7) shows that when using the flexible axis approach one can optimize the tradeoff between structural flexibility and mechanical resistance of the rotary subunit. Note that the 120° rotations, in principle, cannot perfectly reproduce the starting configuration of the F_1-ATPase, as in our simulations the rotor motion is not accompanied by occupancy changes of the active sites.

Because for the flexible potentials the local rotation axis adapts dynamically, it is interesting to monitor the evolution of the F_1 rotor angle θ also with respect to a variable axis. Figure 8 shows the time dependence of θ computed in the same manner as previously but now with the instantaneous (longest) principal axis of the γ subunit used as the reference axis. Significantly smoother variation of θ with time is seen in Figure 8 compared to using a fixed symmetry axis (Figure 6B,C). This result illustrates the ability of the flexible methods to adapt the rotation geometry to the structure and conformational changes of the stator.

Torque Profiles. We will now characterize the different rotation methods in terms of torque and energetics. Because the efficiency of the chemomechanical energy transmission in the F_1-ATPase, when studied in single molecule measurements, is close to 100%,[28] and due to the implied tight coupling between the mechanical reaction coordinate (e. g., the θ angle) and conformational changes in the catalytic subunit, the work necessary to enforce a 120° rotation of F_1-ATPase in the synthesis direction should approach the free energy of ~50−70 kJ/mol required for ATP synthesis.[26,37]

Figure 9 therefore compares the torque profiles along the mechanical reaction coordinate θ (eq 53) for the three methods considered above. As can be seen, for both flexible axis potentials, the average torque along θ is about 5 times smaller than that for the fixed axis potential. Assuming that for infinitely slow rotation the (equilibrium) torque curve is smaller than the observed torques, and, further, that this difference is due to dissipation or other nonequilibrium effects, this result implies that the flexible axis approach reduces the dissipated energy by at least a factor of 5 with respect to the fixed axis potential.

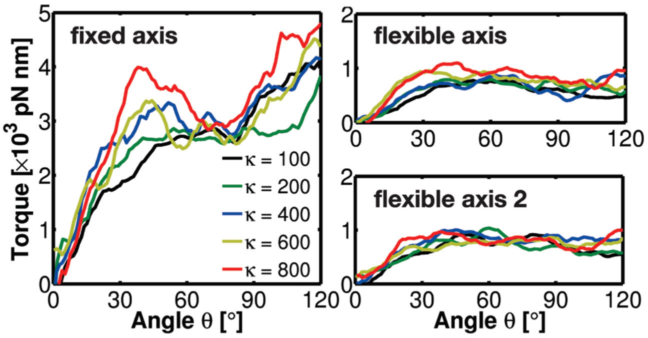


Figure 9. The angular dependence of the driving torque for the γ subunit enforced to rotate in the synthesis direction using V^{iso} (fixed axis), V^{flex2} (flexible axis), and V^{flex2} (flexible axis 2), using five different spring constants k = 100−800 kJ/(mol·nm^2). All torque profiles were smoothed using a running average window of 8°.

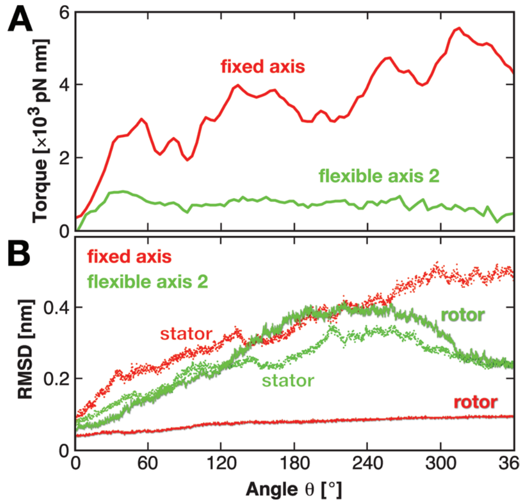


Figure 10. Evolution of the driving torque for the γ subunit enforced to rotate in the synthesis direction using V^{iso} (red) and V^{flex2} (green) with k = 600 kJ/(mol·nm^2) (A). RMSD of the γ rotor backbone atoms (solid) and of the $\alpha_3\beta_3$ stator backbone atoms (dotted) with respect to their respective X-ray structure (B).

Integrated over 120°, the corresponding work is 5900 ± 300, 1400 ± 100, and 1490 ± 80 kJ/mol, for the fixed, flexible, and flexible 2 potential, respectively. Due to the large angular velocity applied as well as the resulting nonequilibrium nature of this process, this work is still much larger than the free energy of ATP synthesis but clearly shows the dramatic reduction by the flexible axis method. For much lower velocities of 0.00042°/ps, the integrated work reduces further to 350 ± 50 kJ/mol (data not shown).

Already for the short simulations, the dependence of the torque on the angular position of the rotor reveals details of the free energy landscape governing the F_1 rotation. In all simulations with a flexible axis potential (Figures 9 and 10), only small variations of the average torque with respect to the rotor angle are observed. This suggests that the underlying energy landscape is smooth and nearly linear, which is in agreement with recent experiments.[26,56] However, due

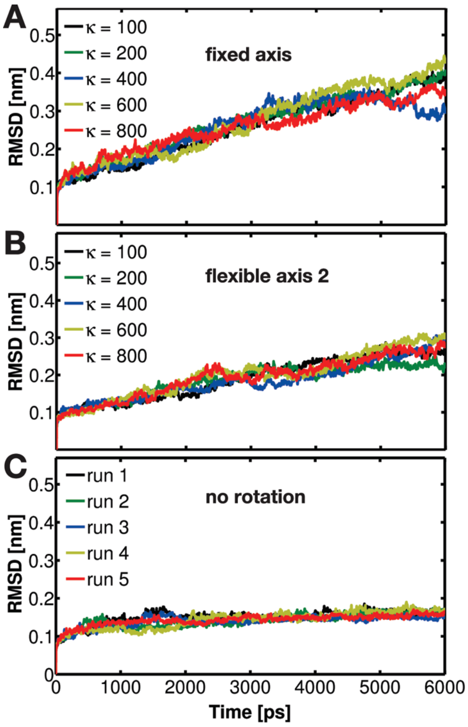


Figure 11. RMSD of the α_3 subunit backbone atoms with respect to the X-ray structure when driving γ subunit rotation using V^{iso} (A) and V^{flex2} (B). For comparison, the corresponding RMSD evolution for five independent free MD runs is also shown (C).

to the much larger angular velocity employed here, the torque values calculated from our simulations are at least one order of magnitude larger than the F_1-generated torque measured under a viscous load (40−50 pN·nm).[28,56] In addition, the calculated torque profiles indicate that the free energy landscape is steepest at $\theta \approx 40°$. The increase of the torque for $\theta > 90°$ that shows up in the fixed type points to shortcomings of this particular method, which will be discussed in the next section.

Since in the simulations the motion along the mechanical reaction coordinate is not synchronized with changes in chemical occupancy of the stator binding sites, we do not expect the torque to drop to zero after 120° of rotation. To examine how far our nonequilibrium simulations are from the reversible limit, Figure 10A shows torque profiles of full 360° rotations applying V^{iso} and V^{flex2} with $k = 600$ kJ/(mol·nm^2). As can be seen, the torque determined for the fixed axis case increases strongly, whereas for the flexible case, after a small increase up to $\theta \approx 40°$, the torques decrease toward considerably smaller values. This result underscores that the flexible potential perturbs the system to a much lesser extent, such that it remains much closer to equilibrium than for a fixed axis rotation.

The fixed axis potential induces structural changes almost exclusively in the bearing (stator in Figure 10B) while in the flexible axis case, the structural changes are distributed rather equally among the rotating subunit and its bearing. Moreover, in the flexible axis case, both structures nearly approach the starting structure ($\theta = 0°$) at the end of a whole 360° turn with an RMSD below 2.5 Å, which is not seen for the fixed axis simulation.

Origin of Differences in Energetics of Fixed and Flexible Axis Rotation. When using a simple fixed axis rotation potential, the rotating part behaves like a rigid body. In combination with the fixed axis, this behavior can cause unphysical close contacts and strong torques between the rotor and the bearing, which may cause extensive artificial structural changes of the bearing. The flexible axis approach, in contrast, keeps the system closer to the equilibrium for two reasons. First, the self-adjusting local rotation axis ensures an overall optimal position of the pivot; second, the built-in flexibility allows for structural relaxation of the rotating part and thus locally minimizes sterical hindrances. As the F_1-ATPase motor components are strongly coupled and leave only little room for the rotating subunit inside the bearing, both reasons allow for the necessary tight adaption of the γ rotor to the $\alpha_3\beta_3$ bearing.

To quantify this effect, Figure 11 displays the enforced conformational changes of the bearing, in terms of stator RMSD with respect to its X-ray structure as a function of time for the α_2 subunit. This subunit was chosen because it interacts most closely with the γ rotor throughout the whole runs. It is evident that the structural changes induced in α_2 are considerably larger for fixed axis rotation than for the flexible potentials. Secondary structure analysis reveals that in the former case the structural motifs exposed to the center of the $\alpha_3\beta_3$ hexamer are distorted by the rotating γ subunit. Also in Figure 4 one can notice partial disruption of the helices in the C-terminal part of the β_3 subunit (the bottom part of β on the left side of γ, red dashed circle) when it is pressed upon by the rotor driven to rotate around the fixed axis. The torque increase for angles $\theta > 90°$ (Figure 9, left, and Figure 10) reflects this effect, which is mainly due to wrapping of the β_3 C-terminal domain around the γ subunit.

7. CONCLUSIONS

We have developed, implemented, and tested a new method to enforce the rotation of protein subunits that allows for (i) a flexible rotation axis and (ii) structural adaptions of the rotated subunit to its environment. For γ subunit rotation in F_1-ATPase, we have shown that our flexible axis method reduced the frictional dissipation of the γ subunit within the $\alpha_3\beta_3$ bearing by more than a factor of 5. As a result, also the induced torque was 5-fold smaller compared to the one using a fixed axis.

Concerning the use of the flexible axis potentials developed here, we should like to point out two possible caveats. The first caveat is due to the fact that, while the pivot vector is free to adapt flexibly, the orientation of the direction vector is fixed. For systems where the subsystem subjected to the rotation potential is embedded within a curved "bearing", the flexible adaptation will work properly only as long as the angle between the orientations of the bearing axis and the direction vector is not too large, i.e., for not too strong bending of the bearing. In extreme cases such as a complete U-shaped bearing, artificial structural changes of the bearing similar to those induced by fixed axes may occur. This problem can be addressed by subdividing the system into several parts and using a separate flexible rotation axis for each of these parts, with orientation vectors locally adapted to the respective part of the bearing.

The second caveat regards the proper choice of the slab thickness. If chosen too small, only a few atoms will be assigned to each slab, thus compromising the averaging that defines the pivot vector of each slab. In contrast, if chosen too large, the slabs might stretch over regions that would require changing pivot vectors, in which case the enforced rotation would induce, albeit to a lesser extent, the artifacts caused by fixed axis approaches.

Obviously, in the limit of just a single slab covering the complete rotating subsystem, the fixed axis potential is recovered.

With these limitations and caveats in mind, our flexible axis potentials are applicable to a broad range of quite diverse biomolecular systems, processes, and functions. Apart from mimicking molecular rotary motors, it can also serve to restrain the orientation of a protein or ligand, or, in combination with umbrella sampling, to calculate the preferred orientation of transmembrane proteins or membrane-active agents within a lipid bilayer. Further, the method is expected to yield more accurate free energy profiles along circular reaction coordinates via umbrella sampling. In the long run, our flexible axis approach might prove useful for the study and design of synthetic nanodevices with rotating elements, such as those considered for molecular nanotechnology.[57]

■ APPENDIX: USING GROMACS FOR ENFORCED ROTATION SIMULATIONS

All methods and potentials described in this paper have been implemented into GROMACS and will be part of the next major release. For immediate use, the rotation repository branch should be checked out from the GROMACS git repository. See www.gromacs.org for how to access the repository.

To use one of these potentials, the particles i that are to be subjected to rotation potentials are defined via index groups rot_group0, rot_group1, etc., in the grompp preprocessor mdp input file. The reference positions $\mathbf{y}_i^0$ are read from a file provided to grompp. If no such file is found, $\mathbf{x}_i(t = 0)$ are used as reference positions and written to file such that they can be used for subsequent setups. All parameters of the potentials such as k, ε', etc. (Table 1) are provided via input file parameters; rot_type selects the type of the potential. The option rot_massw allows one to choose whether or not to use mass-weighted averaging. Table 2 summarizes observables that are written to additional output files, which are described below.

Angle of Rotation Groups: Fixed Axis. For fixed axis rotation, the average angle $\theta_{\text{av}}(t)$ of the group relative to the reference group is determined via the distance-weighted angular deviation of all rotation group atoms from their reference positions

$$\theta_{\text{av}} = \frac{\sum_{i=1}^{N} r_i \theta_i}{\sum_{i=1}^{N} r_i} \tag{50}$$

Here, r_i is the distance of the reference position to the rotation axis, and the difference angles θ_i are determined from the atomic positions, projected onto a plane perpendicular to the rotation axis through pivot point $\mathbf{u}$ (see eq 7 for the definition of $\perp$)

$$\cos\theta_i = \frac{(\mathbf{y}_i - \mathbf{u})^{\perp} \cdot (\mathbf{x}_i - \mathbf{u})^{\perp}}{||(\mathbf{y}_i - \mathbf{u})^{\perp} \cdot (\mathbf{x}_i - \mathbf{u})^{\perp}||} \tag{51}$$

The sign of θ_{av} is chosen such that $\theta_{\text{av}} > 0$ if the actual structure rotates ahead of the reference.

Angle of Rotation Groups: Flexible Axis. For flexible axis rotation, two outputs are provided, the angle of the entire rotation group and separate angles for the segments in the slabs. The angle of the entire rotation group is determined by an RMSD fit of $\mathbf{x}_i$ to the reference positions $\mathbf{y}_i^0$ at $t = 0$, yielding θ_{fit} as the angle by which the reference has to be rotated around $\hat{\mathbf{v}}$ for the optimal fit

$$\text{RMSD}(\mathbf{x}_i, \Omega(\theta_{\text{fit}})\mathbf{y}_i^0) \overset{!}{=} \min \tag{52}$$

To determine the local angle for each slab n, both reference and actual positions are weighted with the Gaussian function of slab n, and $\theta_{\text{fit}}(t,n)$ is calculated as in eq 52 from the Gaussian-weighted positions.

For all angles, the input option rot_fit_method controls whether a normal RMSD fit is performed or whether for the fit each position $\mathbf{x}_i$ is put at the same distance to the rotation axis as its reference counterpart $\mathbf{y}_i^0$. In the latter case, the RMSD measures only angular differences, not radial ones.

Table 1. Parameters Used by the Various Rotation Potentials Defined Above[a]

parameter			k	$\hat{\mathbf{v}}$	$\mathbf{u}$	ω	ε'	Δx	g_n^{min}
grompp input			k	vec	pivot	rate	eps	slab_dist	min_gauss
unit	variable name	eq	[(kJ)/(mol·nm^2)]	[-]	[nm]	[deg/ps]	[nm^2]	[nm]	[-]
				fixed axis:					
isotropic	V^{iso}	1	X	X	X	X	-	-	-
−pivot-free	$V^{\text{iso-pf}}$	4	X	X	-	X	-	-	-
parallel motion	V^{pm}	8	X	X	X	X	-	-	-
−pivot-free	$V^{\text{pm-pf}}$	12	X	X	-	X	-	-	-
radial motion	V^{rm}	13	X	X	X	X	-	-	-
−pivot-free	$V^{\text{rm-pf}}$	18	X	X	-	X	-	-	-
radial motion2	V^{rm2}	19	X	X	X	X	X	-	-
−pivot-free	$V^{\text{rm2-pf}}$	24	X	X	-	X	X	-	-
				flexible axis:					
flexible	V^{flex}	34	X	X	-	X	-	X	X
−transl. tol.	$V^{\text{flex-t}}$	39	X	X	-	X	-	X	X
flexible2	V^{flex2}	40	X	X	-	X	X	X	X
−transl. tol.	$V^{\text{flex2-t}}$		X	X	-	X	X	X	X

[a] X's indicate which parameter is actually used.

Table 2. Quantities Recorded in Output Files during Enforced Rotation

quantity	unit	equation	output file	fixed	flexible	controlled by
$V(t)$	kJ/mol	see Table *1*	rotation	X	X	nstrout
$\theta_{\text{ref}}(t)$	deg	$\theta_{\text{ref}}(t) = \omega t$	rotation	X	X	nstrout
$\theta_{\text{av}}(t)$	deg	50	rotation	X	-	nstrout
$\theta_{\text{fit}}(t)$, $\theta_{\text{fit}}(t,n)$	deg	52	rotangles	-	X	nstsout
$\mathbf{y}_0(n)$, $\mathbf{x}_0(t,n)$	nm	32, 33	rotslabs	-	X	nstsout
$\tau(t)$	kJ/mol	53	rotation	X	-	nstrout
$\tau(t,n)$	kJ/mol	53	rottorque	-	X	nstsout

Angle Determination by Searching the Energy Minimum. Alternatively, for rot_fit_method=potential, the angle of the rotation group is determined as the angle for which the rotation potential energy is minimal. Therefore, the used rotation potential is additionally evaluated for a set of angles around the current reference angle. In this case, the rotangles.log output file contains the values of the rotation potential at the chosen set of angles, while rotation.xvg lists the angle with minimal potential energy.

Torque. The torque $\tau(t)$ exerted by the rotation potential is calculated for fixed axis rotation via

$$\tau(t) = \sum_{i=1}^{N} \mathbf{r}_i(t) \times \mathbf{f}_i^{\perp}(t) \qquad (53)$$

where $\mathbf{r}_i(t)$ is the distance vector from the rotation axis to $\mathbf{x}_i(t)$ and $\mathbf{f}_i^{\perp}(t)$ is the force component perpendicular to $\mathbf{r}_i(t)$ and $\hat{\mathbf{v}}$. For flexible axis rotation, torques τ_n are calculated for each slab using the local rotation axis of the slab and the Gaussian-weighted positions.

■ ASSOCIATED CONTENT

Supporting Information. Movies illustrating the effect of the fixed and flexible axis methods. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: ckutzne@gwdg.de.

■ ACKNOWLEDGMENT

This work has been supported by the EU (NANOMOT/NEST 029084) and the Deutsche Forschungsgemeinschaft (DFG GR 1590/2-1). We thank Anna Neumann for the derivation of eqs 23 and 46, and Marcus Kubitzki and Berk Hess for helpful discussions.

■ REFERENCES

(1) Yoshida, M.; Muneyuki, E.; Hisabori, T. *Nat. Rev. Mol. Cell Biol.* **2001**, *2*, 669–677.

(2) Stock, D.; Gibbons, C.; Arechaga, I.; Leslie, A.; Walker, J. *Curr. Opin. Struct. Biol.* **2000**, *10*, 672–679.

(3) Sowa, Y.; Berry, R. *Q. Rev. Biophys.* **2008**, *41*, 103–132.

(4) Nakano, M.; Imamura, H.; Toei, M.; Tamakoshi, M.; Yoshida, M.; Yokoyama, K. *J. Biol. Chem.* **2008**, *283*, 20789–20796.

(5) Laskey, R.; Madine, M. *EMBO Rep.* **2003**, *4*, 26–30.

(6) Smith, D.; Tans, S.; Smith, S.; Grimes, S.; Anderson, D.; Bustamante, C. *Nature* **2001**, *413*, 748–752.

(7) Simpson, A.; Tao, Y.; Leiman, P.; Badasso, M.; He, Y.; Jardine, P.; Olson, N.; Morals, M.; Grimes, S.; Anderson, D.; Baker, T.; Rossmann, M. *Nature* **2000**, *408*, 745–750.

(8) Moore, S. *Curr. Biol.* **2002**, *12*, R96–R98.

(9) Weber, J.; Senior, A. *Biochim. Biophys. Acta* **2000**, *1458*, 300–309.

(10) Grubmüller, H.; Heymann, B.; Tavan, P. *Science* **1996**, *271*, 997–999.

(11) Izrailev, S.; Stepaniants, S.; Balsera, M.; Oono, Y.; Schulten, K. *Biophys. J.* **1997**, *72*, 1568–1581.

(12) Lu, H.; Isralewitz, B.; Krammer, A.; Vogel, V.; Schulten, K. *Biophys. J.* **1998**, *75*, 662–671.

(13) Izrailev, S.; Stepaniants, S.; Isralewitz, B.; Kosztin, D.; Lu, H.; Molnar, F.; Wriggers, W.; Schulten, K. *Comput. Mol. Dynamics: Challenges, Methods, Ideas* **1998**, *4*, 39–65.

(14) Kale, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K. *J. Comput. Phys.* **1999**, 283–312.

(15) Böckmann, R.; Grubmüller, H. *Nat. Struct. Biol.* **2002**, *9*, 198–202.

(16) Meglio, A.; Praly, E.; Ding, F.; Allemand, J.-F.; Bensimon, D.; Croquette, V. *Curr. Opin. Struct. Biol.* **2009**, *19*, 615–622.

(17) Itoh, H.; Takahashi, A.; Adachi, K.; Noji, H.; Yasuda, R.; Yoshida, M.; Kinosita, K., Jr. *Nature* **2004**, *427*, 465–468.

(18) Mazur, A. K. *J. Chem. Theory Comput.* **2009**, *5*, 2149–2157.

(19) Aksimentiev, A.; Balabin, I.; Fillingame, R.; Schulten, K. *Biophys. J.* **2004**, 1332–1344.

(20) Saam, J.; Tajkhorshid, E.; Hayashi, S.; Schulten, K. *Biophys. J.* **2002**, *83*, 3097–3112.

(21) Phillips, J.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.; Kalé, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781–1802.

(22) Eichinger, M.; Heller, H.; Grubmüller, H. *Molecular Dynamics on Parallel Computers*; World Scientific: River Edge, NJ, 2000; pp 154—174.

(23) van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A.; Berendsen, H. *J. Comput. Chem.* **2005**, *26*, 1701–1718.

(24) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, 435–447.

(25) Torrie, G.; Valleau, J. *J. Comput. Phys.* **1977**, *23*, 187–199.

(26) Kinosita, K., Jr.; Adachi, K.; Itoh, H. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 245–268.

(27) Abrahams, J.; Leslie, A.; Lutter, R.; Walker, J. *Nature* **1994**, *370*, 621–628.

(28) Yasuda, R.; Noji, H.; Kinosita, K.; Yoshida, M. *Cell* **1998**, *93*, 1117–1124.

(29) Wang, H.; Oster, G. *Nature* **1998**, *396*, 279–282.

(30) Oster, G.; Wang, H. *Biochim. Biophys. Acta* **2000**, 482–510.

(31) Gao, Y.; Yang, W.; Karplus, M. *Cell* **2005**, *123*, 195–205.

(32) Dittrich, M.; Hayashi, S.; Schulten, K. *Biophys. J.* **2003**, *85*, 2253–2266.

(33) Böckmann, R.; Grubmüller, H. *Biophys. J.* **2003**, *85*, 1482–1491.

(34) Gao, Y.; Yang, W.; Marcus, R.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 11339–11344.

(35) Pu, J.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 1192.

(36) Czub, J.; Grubmüller, H. *Biophys. J.* **2010**, 168a.

(37) Junge, W.; Sielaff, H.; Engelbrecht, S. *Nature* **2009**, *459*, 364–370.

(38) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33–38.

(39) Gibbons, C.; Montgomery, M.; Leslie, A.; Walker, J. *Nat. Struct. Mol. Biol.* **2000**, *7*, 1055–1061.

(40) Seeliger, D.; de Groot, B. *J. Comput. Chem.* **2009**, *30*, 1160–1166.

(41) Nicholls, A.; Honig, B. *J. Comput. Chem.* **1991**, *12*, 435–445.

(42) Vriend, G. *J. Mol. Graph.* **1990**, *8*, 52–56.

(43) Jorgensen, W.; Maxwell, D.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.

(44) Kaminski, G.; Friesner, R.; Tirado-Rives, J.; Jorgensen, W. *J. Phys. Chem. B.* **2001**, *105*, 6474–6487.

(45) Jorgensen, W.; Chandrasekhar, J.; Madura, J.; Impey, R.; Klein, M. *J. Chem. Phys.* **1983**, *79*, 926–935.

(46) Nosé, S. *Mol. Phys.* **1984**, 255–268.

(47) Hoover, W. *Phys. Rev. A* **1985**, *31*, 1695–1697.

(48) Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981**, *52*, 7182–7190.

(49) Nosé, S.; Klein, M. *Mol. Phys.* **1983**, *50*, 1055–1076.

(50) Berendsen, H.; Postma, J.; van Gunsteren, W.; DiNola, A.; Haak, J. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

(51) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, 10089–100092.

(52) Essmann, U.; Perera, L.; Berkowitz, M.; Darden, T.; Lee, H.; Pedersen, L. *J. Chem. Phys.* **1995**, *103*, 8577–8593.

(53) Hess, B. *J. Chem. Theory Comput.* **2008**, *4*, 116–122.

(54) Miyamoto, S.; Kollman, P. *J. Comput. Chem.* **1992**, *13*, 952–962.

(55) van der Spoel, D.; Lindahl, E.; Hess, B.; van Buuren, A.; Apol, E.; Meulenhoff, P.; Tieleman, D.; Sijbers, A.; Feenstra, K.; van Drunen, R.; Berendsen, H. *Gromacs User Manual version 4.5.* www.gromacs.org (accessed March 2011).

(56) Pänke, O.; Cherepanov, D.; Gumbiowski, K.; Engelbrecht, S.; Junge, W. *Biophys. J.* **2001**, *81*, 1220–1233.

(57) Browne, W.; Feringa, B. *Nat. Nanotechnol.* **2006**, *1*, 25–35.

JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Efficient Solvation Free Energy Calculations of Amino Acid Analogs by Expanded Ensemble Molecular Simulation

Andrew S. Paluch,[†] Jindal K. Shah,[†,‡] and Edward J. Maginn*[,†]

[†]Department of Chemical and Biomolecular Engineering and [‡]Center for Research Computing, University of Notre Dame, Notre Dame, Indiana 46556, United States

S Supporting Information

ABSTRACT: We present an efficient, automated expanded ensemble method to calculate the residual chemical potential or solvation free energy by molecular dynamics simulation. The methodology is validated by computing the residual chemical potential of 13 amino acid analogs in water at 300 K and 1 bar and comparing to reference simulation data. Overall agreement is good, with the methodology of the present study reaching limiting precisions of less than 0.1 k_BT in half of the total simulation time of the reference simulation study which utilized Bennett's acceptance ratio method. The apparent difference in the efficiencies is a result of the inherent advantages of the expanded ensemble method, which creates an improved decorrelation of simulation data and improves the sampling of the important regions of the configurational phase space of each subensemble. The present adaptation utilizes histograms of proposed transition energies collected throughout the entire simulation, to make extremely precise calculations of the relative free energy between neighboring subensembles.

1. INTRODUCTION AND MOTIVATION

The thermodynamic behavior of all chemical and biological systems at equilibrium may be fundamentally understood in terms of the underlying free energy or chemical potential. Knowledge of the free energy is crucial to understanding the phase equilibria between solids, liquids, and gases, which in turn is key for the design of separation processes and the selection of solvents for synthesis reactions. Likewise, the transport of drug molecules between cell membranes and partitioning between multiple environments may be explained in terms of the relative chemical potentials, whose knowledge is crucial for drug design.[1] The importance of such information in drug design is emphasized by the fact that entire monographs have been devoted to the topic.[2,3]

Proteins are required by the body for the growth, repair, and maintenance of cells. They are vital for virtually every process within the human body, such as metabolism and digestion, and are necessary for the production of antibodies to fight off infections and diseases. Proteins may be regarded as large biopolymers of amino acids, creating an expansive range of possible chemical compositions.[4] Insight into the native structure and the folding mechanism of proteins in solution may be obtained by examining the solvation free energy of individual constituent amino acid analogs.[5,6]

To make the link between the solvation free energy of a given amino acid analog and a particular solvent, one must account for the molecular-level details that occur during solvation. One way to do this is via molecular simulation. Recently, several studies have looked at the ability of molecular simulation to predict hydration free energies of amino acid analogs.[7−11] All of the studies employed either thermodynamic integration (TI)[12] or Bennett's acceptance ratio (BAR) method,[13] and results were presented with an unprecedented level of precision.[7] Despite the success of the studies, all of the employed methods were computationally intensive. To facilitate the calculations, a stratification or staging strategy was used in which intermediate states were constructed in between the target and the reference state, so as to increase the phase space overlap between neighboring states.[14,15] That is, many simulations at varying coupling strengths of the amino acid analogs were conducted, followed by postsimulation data analysis.

An additional method that may help facilitate the exploration of phase space for computing solvation free energies is the expanded ensemble (EE) method originally developed by Lyubartsev and co-workers.[16−19] While the free energy between states is calculated with an appropriate technique, a single simulation is performed in which a random walk is constructed over the reference, intermediate, and target states. The phase space of each state is sampled according to a unique Hamiltonian, and the propagation of configurations between neighboring states helps prevent quasi-nonergodicity, improving the rate of exploration of phase space. The EE method has been combined previously with various techniques to compute free energy changes in an efficient, automated fashion[20−23] with a high level of precision.[24,25] In addition, the EE method has been applied to study the solvation free energies of drug molecules.[26] Together, these previous studies suggest that automated EE calculations may be applied to obtain precise solvation free energies of biological systems in an extremely efficient manner.

In the present study, we evaluate the use of EE to calculate the solvation free energy of amino acid analogs in an efficient, automated fashion, reaching levels of precision comparable to previous studies.[7,27] We have accomplished this by combining the strengths of the flat histogram method of Wang and Landau

Received: November 19, 2010
Published: April 13, 2011

(WL)[28−30] with the BAR technique. The approach extends our previous work[25] in two ways: First, the current implementation of EE is in a molecular dynamics (MD) framework rather than a Monte Carlo (MC) framework.[17] The choice of sampling configurational phase space with MD resulted from the overwhelming preference of the biological modeling community for MD versus MC as a means of generating configurations. Furthermore, the numerous, highly efficient, freely available MD codes reinforced this motive. Second, the current study employs BAR rather than transition-matrix Monte Carlo (TMMC).[31−35] This was done because BAR is straightforward to implement within MD. By using BAR, the method may be employed using either MC or MD. We note that TMMC and BAR have been shown to be intimately related,[36,37] and the MD implementation of EE[17] has similarities with the independently formulated λ-dynamics method.[38] In Section 2 of the paper we will present an overview of the employed methodology, followed by the relevant computational details in Section 3. Results and discussion are given in Section 4, followed by a summary of our findings in Section 5.

2. METHODS

EE. The main idea behind the EE method is to construct an augmented ensemble as a sum of subensembles.[16−19] This series of subensembles connects two systems of interest by gradually performing transitions between the two systems. In the current study, the systems of interest are pure solvent (water) and solvent with the addition of a single solute (amino acid analog) molecule at the same temperature and pressure. These systems are connected through a series of subensembles that begin with a noninteracting solute molecule in a pure solvent (i.e., an ideal gas reference state at the same density of the pure solvent) and end with a fully interacting solute molecule in solution. The intermediate subensembles serve to scale the intermolecular interaction potential of the solute. A specific subensemble is designated by index m. Intermolecular Lennard-Jones (LJ) and electrostatic (elec) interactions are regulated by the subensemble dependent coupling parameters λ_m^{LJ} and $\lambda_m^{\mathrm{elec}}$, respectively, which vary from $0 \le \lambda_m^{\mathrm{LJ}} \le 1$ and $0 \le \lambda_m^{\mathrm{elec}} \le 1$.

While within a given subensemble, configurational phase space is sampled by MD. Periodically, a MC random walk is performed in which moves consist of transitions to neighboring subensembles. In this way, a probability distribution over subensembles is generated. In the isothermal−isobaric expanded ensemble (EE-NpT), a specific microstate (or configuration within a subensemble m) is observed with probability:

$$\pi_m(\mathbf{r}) = \frac{1}{Z_{\mathrm{NpT}}} \exp\{-\beta[U_m(\mathbf{r}) + pV(\mathbf{r})]\} \tag{1}$$

where Z_{NpT} is the EE-NpT configurational partition function (or configurational integral), U_m is the subensemble dependent potential energy, p and V are the pressure and the volume, respectively, $\beta = 1/k_{\mathrm{B}}T$, where k_{B} is Boltzmann's constant and T is the temperature, and $\mathbf{r}$ is a $3(N_{\mathrm{solv}} + N_{\mathrm{solute}})$ dimensional vector representing the positions of the solvent and solute molecules, where N_{solv} and N_{solute} are the number of solvent and solute molecules, respectively. Note that for all of the cases examined here $N_{\mathrm{solute}} = 1$. A transition from subensemble m to subensemble n is accepted with probability:[39]

$$a_{m \to n} = \min\left\{1, \frac{\pi_n(\mathbf{r})}{\pi_m(\mathbf{r})}\right\} \tag{2}$$

Transitions attempting to take the system outside the range of subensembles are rejected. Further, the probability Π_m of finding the system in a given macrostate (or subensemble m) is the sum over all microstates in the subensemble:

$$\Pi_m = \sum_{\mathbf{r}} \pi_m(\mathbf{r}) \tag{3}$$

The probability of visiting a macrostate is characterized by the configuration of the system and the subensemble. Thus, each microstate maps to a single macrostate. The relative Gibbs free energy between any two subensembles m and n is related to the relative macrostate probabilities as:[16−18]

$$\beta G_n(N_{\mathrm{solv}}, N_{\mathrm{solute}}, T, p) - \beta G_m(N_{\mathrm{solv}}, N_{\mathrm{solute}}, T, p) = -\ln\left(\frac{\Pi_n}{\Pi_m}\right) \tag{4}$$

It follows from the definition of the chemical potential and finite difference arguments that[18]

$$\mu_{\mathrm{solute}}^{res}(N_{\mathrm{solv}}, N_{\mathrm{solute}}, T, p) = -\ln\left(\frac{\Pi_{M_{\mathrm{Total}}}}{\Pi_0}\right) \tag{5}$$

where $\mu_{\mathrm{solute}}^{\mathrm{res}}$ is the residual chemical potential of the solute (i.e., chemical potential of the solute relative to an ideal gas reference state), and the subscripts M_{Total} and 0 are the subensemble indices corresponding to the fully interacting solute in solution and the noninteracting solute in solution, respectively. The residual chemical potential is equivalent to the Gibbs free energy of transfer reported in many studies of biological systems.[7−9,11]

As a result of eq 4, we find that as the free energy difference between subensembles increases, the frequency of transitions between subensembles decreases exponentially. To ensure that the system sufficiently samples the entire range of subensembles, a subensemble dependent weighting function η_m is employed to bias the acceptance probability.[40] Trial moves between subensembles are accepted according to a biased acceptance probability:

$$a_{\eta, m \to n} = \min\left\{1, \frac{\pi_n(\mathbf{r})}{\pi_m(\mathbf{r})} \exp(\eta_n - \eta_m)\right\} \tag{6}$$

A uniform sampling of subensembles is obtained if the weighting functions are set according to

$$\eta_n - \eta_m = -\ln\left(\frac{\Pi_n}{\Pi_m}\right) \tag{7}$$

Unfortunately, Π_n and Π_m are the unknown macrostate probabilities that one seeks to calculate and, in general, are not known a priori. In the original implementation of the EE method,[16−19] a multicanonical algorithm[41] was adopted to estimate the weighting functions in an iterative manner through a series of short simulations until a relatively flat visited states histogram was achieved. Advances have been made with the use of histogram based methods that aim to calculate directly the macrostate probabilities with a high level of precision and obtain the relevant weights by use of eq 7. These histogram methods include WL and TMMC and the intimately related BAR.[36,37,42] Recent studies have also successfully combined multiple methods, namely WL and TMMC.[25,43] In the present study, we will employ a combined WL-BAR approach, as described in the next subsection.

WL-BAR Scheme. To obtain an initial estimate of the weighting functions in eq 6, WL is used to estimate the macrostate probability and hence the weighting functions via eq 7. The EE-NpT simulation is started with the phase space of each subensemble being sampled by MD, and periodic attempts are made to transition between subensembles. After each attempted transition, the current estimate of the macrostate probability is updated as

$$\ln \Pi_m^{\text{new}} = \ln \Pi_m^{\text{old}} + \upsilon_{\text{WL}} \qquad (8)$$

where υ_{WL} is a convergence factor greater than 0. After a specified period of time, the convergence factor is reduced according to the following expression:

$$\upsilon_{\text{WL}}^{\text{new}} = \kappa \cdot \upsilon_{\text{WL}}^{\text{old}} \qquad (9)$$

where κ is an update factor less than 1. The entire process is then repeated. The implementation of WL has been studied extensively in the past,[28−30,44,45] including a detailed discussion with regards to combining WL with TMMC.[43] Given the close resemblance of TMMC and BAR,[36,37,42] we have followed the recommendations of Shell et al.[43] Namely, the convergence factor should initially be large enough to sample a broad range of subensembles, allowing for the collection of an expansive amount of transition energies. However, the convergence factor should not be excessively large and should then be quickly reduced, minimizing the time and the extent at which our random walk violates detailed balance. The heuristics of this update scheme are provided in the next section.

The WL procedure quickly samples a broad range of subensembles but converges to a limiting, nonprecise estimate of the macrostate probabilities that are not improved with additional steps.[29] On the other hand, TMMC and BAR methods may be slower to sample a broad range of subensembles[35] but converge upon an extremely precise estimate of the macrostate probabilities.[34,36] Therefore, in an effort to utilize the strengths of both WL and BAR, after sufficient sampling has been achieved with WL, the WL calculated weights are refreshed with weights calculated with BAR. It is important to emphasize that the role of WL is only to quickly sample a broad range of macrostates; the free energy is ultimately calculated using BAR.

BAR is an optimal method to calculate free energy differences between neighboring states and has been derived previously using several different criteria: by minimizing the variance of the acceptance ratio between neighboring states,[13,46] as an optimal overlap-sampling method,[47] and by using maximum likelihood arguments.[48] In addition, Ferrenburg and Swendsen[49] showed that the optimal combination of histogram data reduces to BAR in the limit that only two states are sampled. Also, Escobedo and co-workers[36,37] have shown that TMMC is a limiting case of BAR. As a result of the many derivations in previous studies, only the relevant working equations will be presented here. To calculate the free energy difference between subensembles n and m, BAR considers perturbations from both m to n and n to m. The difference in free energy is then calculated as

$$\beta G_n(N_{\text{solv}}, N_{\text{solute}}, T, p) - \beta G_m(N_{\text{solv}}, N_{\text{solute}}, T, p)$$
$$= -\ln\left[\frac{\langle f(U_n - U_m - C)\rangle_m}{\langle f(U_m - U_n + C)\rangle_n}\right] + \beta C \qquad (10)$$

where $f(x)$ is the Fermi−Dirac function $[1 + \exp(\beta x)]^{-1}$, C is an adjustable parameter, and the brackets correspond to an ensemble average taken with respect to the probability distribution of subensemble n or m, as indicated by the subscript. While in principle any value of C may be used,[13] the optimal value is found by the following relationship:

$$\sum_{N_{n\to m}} f(U_m - U_n + C) = \sum_{N_{m\to n}} f(U_n - U_m - C) \qquad (11)$$

where $N_{n\to m}$ and $N_{m\to n}$ are the number of perturbations from n to m and m to n, respectively. During the course of the simulation, histograms are collected to track the transition energies (i.e., $U_n - U_m$) between each subensemble. While configurations (and hence energies) are correlated for a finite time, when a transition is accepted, the system begins sampling from a different Hamiltonian. This transitioning reduces the configurational correlation time of the system relative to a simulation in a fixed ensemble.

Periodically, an estimate of the free energy difference between neighboring subensembles is calculated by self-consistently solving eq 11 and then using eq 10 to obtain the free energy difference. From eqs 4 and 7, the free energy difference may be used as a new estimate of the weighting functions. While the weighting functions are continuously changing, the transition energies required for BAR are continuously collected from the unbiased target and reference subensembles. In this way, the transition energy histograms do not need to be rezeroed.[31,33,36]

Physical insight into the success of BAR is presented nicely in the work of Kofke and co-workers,[47,50,51] who present BAR as an optimized overlap sampling method. Rather than viewing BAR in the context of a perturbation from both m to n and n to m, it is better viewed as a two stage perturbation in which perturbations in each direction are performed to a mutual intermediate state. The optimized BAR will select an intermediate state between the two neighboring states that is inside the overlapping region of their respective phase spaces; the free energy prediction from BAR is then the sum of the free energy difference between each neighbor and the intermediate state.[47,50,51] Therefore, so long as there exists a phase space overlap, with sufficient sampling, BAR will be successful.

3. COMPUTATIONAL DETAILS

Molecular Models. To model the intermolecular and intramolecular interactions of the systems, a proper force field is required. While many models exist, it was not the objective of the present work to evaluate which models reproduce experimental data the best. Rather, the objective was to test the simulation method itself, and therefore models were selected to allow for comparison with previous work of Dill and co-workers.[27] Consistent with that work, water was modeled with the rigid three-point transferable intermolecular potential function (TIP3P) of Jorgensen and co-workers.[52] Parameters for the amino acid analogs were taken from the general AMBER force field (GAFF)[53,54] with AM1-BCC partial charges.[55] A complete listing of parameters may be found in the Supporting Information of the paper

Table 1. Studied Amino Acid Analogs and the Corresponding Amino Acid

amino acid	analog
$NH_2(R)CHCOOH$	RH
Ala	methane
Val	propane
Ile	*n*-butane
Leu	isobutane
Ser	methanol
Thr	ethanol
Phe	toluene
Tyr	*p*-cresol
Cys	methanethiol
Met	methylethylsulfide
Asn	acetamide
Trp	3-methylindole
His	4-methylimidazole

by Dill and co-workers.[27] The studied amino acid analogs and the corresponding amino acids are summarized in Table 1. In all of these cases, nonbonded intermolecular interactions were treated using a combined LJ and fixed point charge model of the form:

$$U_{\mathrm{nb}}(r_{ij}) = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \frac{1}{4\pi\varepsilon_0}\frac{q_i q_j}{r_{ij}} \qquad (12)$$

where r_{ij}, ε_{ij}, σ_{ij}, q_i, and q_j are the site separation distance between atoms i and j, well-depth of the LJ interaction, distance at which the LJ interaction is zero, and partial charge values, respectively. For interactions between unlike LJ sites, Lorentz–Berthelot combining rules[39] were employed. To prevent instabilities in the trajectory when the solute is nearly decoupled from the system (i.e., when $\lambda_m^{\mathrm{LJ}} \approx 0$), solute–solvent intermolecular nonbonded LJ interactions were modeled with a modified, "soft-core" potential of the form:[8,56,57]

$$U_{\mathrm{LJ}}^{\mathrm{sc}}(r_{ij}; m) = 4\lambda_m^{\mathrm{LJ}}\varepsilon_{ij}\left\{\frac{\sigma_{ij}^{12}}{[(1-\lambda_m^{\mathrm{LJ}})\alpha_{\mathrm{LJ}}\sigma_{ij}^{6} + r_{ij}^{6}]^{2}} - \frac{\sigma_{ij}^{6}}{[(1-\lambda_m^{\mathrm{LJ}})\alpha_{\mathrm{LJ}}\sigma_{ij}^{6} + r_{ij}^{6}]}\right\} \qquad (13)$$

where r_{ij}, ε_{ij}, and σ_{ij} are the same LJ parameters as in eq 12, λ_m^{LJ} is the subensemble dependent coupling strength of the LJ potential, and α_{LJ} is a constant, taken in this study to be 1/2. Note that when the solute molecule is fully coupled to the system, $\lambda_m^{\mathrm{LJ}} = 1$, and eq 13 reduces to the normal LJ potential given by eq 12. When the solute is nearly decoupled, λ_m^{LJ} approaches 0, and eq 13 represents a smooth interaction function that allows solvent molecules to overlap the solute with finite energy. When the solute is decoupled from the system, $\lambda_m^{\mathrm{LJ}} = 0$, and the potential is 0. Thus, the potential form in eq 13 correctly represents the limiting behavior of the solute–solvent interactions, while eliminating instabilities when λ_m^{LJ} approaches 0. Solute–solvent intermolecular electrostatic interactions are decoupled linearly as

$$U_{\mathrm{elect}}(r_{ij}; m) = \lambda_m^{\mathrm{elect}}\frac{1}{4\pi\varepsilon_0}\frac{q_i q_j}{r_{ij}} \qquad (14)$$

where r_{ij}, q_i, and q_j are the same as in eq 12, and $\lambda_m^{\mathrm{elec}}$ is the subensemble dependent coupling strength of the electrostatic interactions.

The same standard LJ and electrostatic interaction potentials (eq 12) and combining rules are used for all intramolecular nonbonded interactions by all pairs of atoms separated by three or more bonds. For the case in which the intramolecular sites are separated by exactly three bonds, the LJ and electrostatic interactions are scaled by a factor of 1/2 and 5/6, respectively.

While the TIP3P water model is completely rigid, the amino acid analogs were modeled with fixed bond lengths but flexible bond and dihedral angles. The bond angle bending intramolecular interaction between sites separated by two bonds was modeled by a simple harmonic potential of the form:

$$U_{\mathrm{angle}}(\theta_{ijk}) = k_{ijk}(\theta_{ijk} - \theta_{ijk}^{0})^{2} \qquad (15)$$

where k_{ijk}, θ_{ijk}, and θ_{ijk}^0 are the force constant, angle between sites i, j, and k, and corresponding nominal bond angle, respectively. The torsional potential describing the intramolecular interaction between sites separated by three bonds was modeled by a potential of the form:

$$U_{\mathrm{tors}}(\varphi_{ijkl}) = \sum_{n=0}^{5} K_n \cos^n(\varphi_{ijkl} - 180^{o}) \qquad (16)$$

where ϕ_{ijkl} is the dihedral angle between sites i, j, k, and l, and the K_n coefficients are constants. The same torsional potential form was used to describe improper dihedral angles, meant to keep planar groups planar. All of the amino acid analog force field files used in the present study are provided in the Supporting Information.

Simulation Details. All simulations were performed with a modified version of the MD simulation package M.DynaMix 5.2.[58,59] For all systems studied in this work, LJ interactions were truncated at a distance of r_{cut} = 12 Å, and standard uniform fluid tail corrections were applied to both the energy and the pressure, assuming $g(r)$ = 1 beyond the cutoff.[39,46] Electrostatic interactions were evaluated with an Ewald summation with tin foil boundary conditions,[39,46] with real space interactions truncated at r_{cut}. A damping parameter of αr_{cut} = 3.72 was used, and the maximum number of reciprocal space lattice vectors was set by K_{max} = 11.0. Integration of the equations of motion was performed with the Verlet leap-frog algorithm in Cartesian coordinates[39,46] with a time step of 2 fs. All bond lengths and the H–O–H angle of water were constrained with the SHAKE algorithm[60] with a tolerance of 10^{-6}. An Andersen thermostat,[61] as implemented by Andrea et al.,[62] and Andersen–Hoover barostat[61,63] were used to sample the phase space of an isothermal–isobaric (NpT) ensemble at 300 K and 1 bar. The collision time for the thermostat was set at 0.4 ps, and the time constant for the barostat was 1.5 ps. Modifications to M.DynaMix include implementation of the Andersen thermostat, the "soft-core" potential (eq 13), separate decoupling of LJ and electrostatic interactions for EE calculations, WL-BAR, modification of the Ewald summation with EE solute molecules as described in the Appendix, and other minor additions.

The systems were set up by randomly placing a single gas phase minimized solute molecule in each of the five independently equilibrated cubic boxes of 900 TIP3P water molecules. Production runs were carried out in an EE-NpT ensemble at 300 K and 1 bar for a total of 10 ns. Each of the five independent systems were initialized with a unique random number seed for the thermostat and for the MC random walk, with all velocities

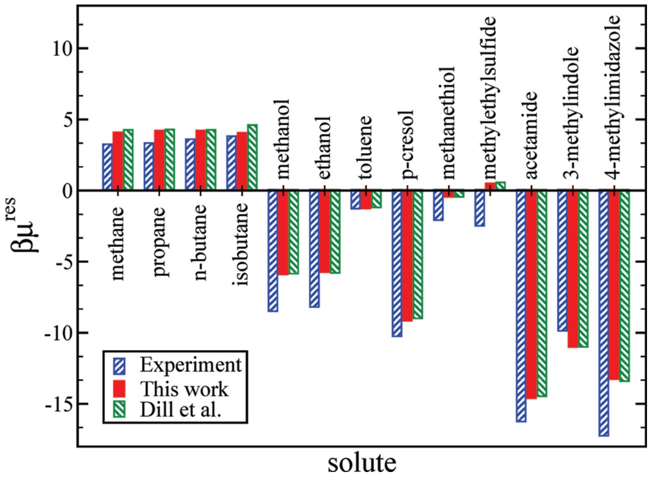


Figure 1. Comparison of the residual chemical potential ($\beta\mu^{\text{res}}$) of the 13 studied amino acid analogs (solutes) from experiment,[67] computed in this study and reference simulation results of Dill and co-workers.[27]

initialized from a Maxwell—Boltzmann distribution at 300 K. The system began in the subensemble with a noninteracting solute molecule and attempts to change subensembles were made every 10 fs. Over the first 0.5 ns, the random walk was carried out with WL biasing, in which the WL weight factor was initially taken to be $\upsilon_{\text{WL}} = 0.25$ and reduced as $\upsilon_{\text{WL}}^{\text{new}} = 0.25\upsilon_{\text{WL}}^{\text{old}}$ every 0.1 ns. The initial WL weight factor was chosen to be an order of magnitude smaller than $\mathcal{O}$ (1) values typically used,[28−30,44,45] yet large enough to sample a broad range of subensembles. The update scheme quickly reduced the weight factor to a value of 0.004 after 0.4 ns for the last WL biasing cycle. During the entire course of the simulation, transition energies (in both directions) were computed each time a transition between subensembles was attempted/proposed, and new subensemble weights were computed from BAR every 0.5 ns. The solute was taken from noninteracting ($m = 0$) to fully interacting ($m = 20$) by first bringing the intermolecular LJ interaction to full strength over 15 subensembles ($1 \leq m \leq 15$) and then adding in intermolecular electrostatic interactions in the final five subensembles ($16 \leq m \leq 20$), for a total of 20 subensembles. For the first 15 subensembles, the intermolecular electrostatic interactions were turned off, and the intermolecular LJ interactions were strengthened as $\lambda_m^{\text{LJ}} = \{0.05, 0.10, 0.20, 0.30, 0.40, 0.50, 0.60, 0.65, 0.70, 0.75, 0.80, 0.85, 0.90, 0.95, 1.0\}$. Next, while the LJ intermolecular interactions were fully restored, the intermolecular electrostatic interactions were strengthened as $\lambda_m^{\text{elec}} = \{0.2, 0.4, 0.6, 0.8, 1.0\}$. Care must be taken to properly decouple the solute—solvent intermolecular interactions with Ewald summation; a detailed description of how this was done is given in the Appendix.

The reported residual chemical potentials of the present work are the mean value of the five independent productions runs for each solute, and the uncertainty is taken as the bootstrap standard error.[64−66] To compute the bootstrap standard error for each solute—solvent combination, the estimate of the chemical potential from each of the five production runs was taken to be an independent data point. Next, 1000 sets containing 5 data points each were created by randomly selecting 5 of our independent data points, with replacement. The mean of each set was computed, creating a bootstrap sample of 1000 estimates of the residual chemical potential. The bootstrap standard error was then found as the standard error of the bootstrap sample relative to the mean of the five independent production runs for each solute—solvent combination.

4. RESULTS AND DISCUSSION

A summary of the computed residual chemical potentials and a comparison to experiment[67] and the simulation work of Dill and co-workers[27] are provided in Figure 1 and in Table 2. The results of the present study are in good overall agreement with the previous simulation results, with an average absolute difference between the computed residual chemical potentials of the present study and Dill and co-workers[27] of 0.13 k_BT. If we exclude isobutane from this calculation, which will be discussed in further detail, the average difference decreases to 0.09 k_BT, which is the same order of magnitude of the reported uncertainties. The excellent agreement of the results suggests that the proposed method yields correct residual chemical potentials. In all cases, an estimate of the residual chemical potential may be obtained within a few ns of simulation time. Further time beyond this serves only to decrease the uncertainty of the calculation.

The largest discrepancy between the present study and that of Dill and co-workers[27] is found for isobutane, corresponding to a discrepancy of 0.53 k_BT, which warrants further investigation. While results of EE calculations are not known to be biased, a potential source of error in the calculations may result from a lack of configurational phase space overlap between neighboring states;[50,51,68,69] the degree of phase space overlap between neighboring states is related to the relative entropy and the energy histograms between neighboring states,[50,51] and for the case of EE, it is related to the observed visited states[16,17] and transition[20] probabilities. As mentioned previously, BAR is limited to applications in which neighboring states have at least some phase space overlap. So long as the states have some phase space overlap, with sufficient sampling, BAR will select an intermediate state between the two neighboring states that is inside the overlapping region of their respective phase spaces; the free energy prediction from BAR is then the sum of the free energy difference between each neighbor and the intermediate state.[47,50,51] Therefore, if there is little or no phase space overlap, BAR will be unsuccessful, and we would expect there to be little or no transitions between neighboring subensembles. The result would be a relatively large uncertainty in the predicted residual chemical potential between independent simulations stemming from inadequate sampling.

To this end, Figure 2 shows the observed visited state and transition probabilities for our EE simulation of isobutane. In the bottom pane it is observed that in all cases, the forward transition probability from subensemble m to subensemble $m + 1$ is nearly indistinguishable from the reverse transition probability to subensemble m from subensemble $m + 1$. Therefore, as a result of MC detailed balance,[39] there is nearly a uniform probability of visiting each subensemble, as shown in the top panel. In addition, the observed transition probabilities are fairly high, between 20 and 50%, ensuring an adequate sampling of each subensemble. As shown in Table 2 and Figure 3, the predicted free energy for isobutane converges to a limiting value with a precision of 0.09 k_BT, dismissing concerns with regards to phase space overlap.

Furthermore, in Figure 4 we can compare to the results of Pande and co-workers[7] who employed the same TIP3P water model but used a previous version of the AMBER force field and a different treatment of long-range electrostatic interactions.

Table 2. Summary of the Residual Chemical Potential ($\beta\mu^{res}$) of the 13 Studied Amino Acid Analogs (Solutes) from Experiment,[67] Computed in This Study, and Reference Simulation Results of Dill and Co-Workers[27] [a]

		$\beta\mu^{res}$			
solute	amino acid	experiment[67]	this work	Dill et al.[27]	absolute difference
methane	Ala	3.25	4.09 ± 0.03	4.26 ± 0.02	0.17 ± 0.04
propane	Val	3.34	4.22 ± 0.05	4.29 ± 0.03	0.07 ± 0.06
n-butane	Ile	3.61	4.22 ± 0.08	4.26 ± 0.03	0.04 ± 0.08
isobutane	Leu	3.82	4.07 ± 0.09	4.60 ± 0.03	0.53 ± 0.09
methanol	Ser	−8.49	−5.89 ± 0.04	−5.84 ± 0.02	0.05 ± 0.04
ethanol	Thr	−8.19	−5.73 ± 0.04	−5.79 ± 0.05	0.06 ± 0.06
toluene	Phe	−1.27	−1.26 ± 0.05	−1.19 ± 0.03	0.07 ± 0.06
p-cresol	Tyr	−10.25	−9.16 ± 0.08	−8.99 ± 0.03	0.17 ± 0.08
methanethiol	Cys	−2.08	−0.33 ± 0.01	−0.44 ± 0.02	0.11 ± 0.02
methylethylsulfide	Met	−2.48	0.51 ± 0.04	0.57 ± 0.03	0.06 ± 0.05
acetamide	Asn	−16.24	−14.61 ± 0.05	−14.46 ± 0.05	0.15 ± 0.07
3-methylindole	Trp	−9.86	−11.01 ± 0.07	−10.99 ± 0.05	0.02 ± 0.09
4-methylimidazole	His	−17.24	−13.26 ± 0.04	−13.40 ± 0.05	0.14 ± 0.06

[a] The last column is the absolute difference of the present study relative to Dill and co-workers,[27] with the uncertainty computed from propagation of errors.

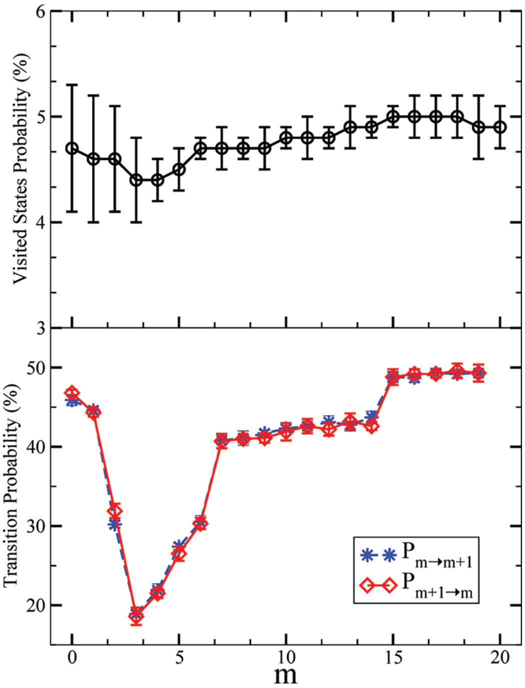


Figure 2. Summary of the observed visited states (top) and forward and reverse transition probabilities (bottom) for the EE simulations of isobutane as a function of subensemble *m*. For forward transition probabilities, the *x*-axis refers to the subensemble the transition is attempted from. For reverse transition probabilities, the *x*-axis refers to the subensemble the transition is attempted to. The error bars are the standard deviation of five independent simulations.

While we would not expect quantitative agreement, for alkanes in which the electrostatics interactions have a small contribution to the overall free energy, we would expect qualitative agreement. The observed trends for the alkanes in both the present study and that of Pande and co-workers[7] are consistent. Both studies observe that the solvation free energy of methane is less than

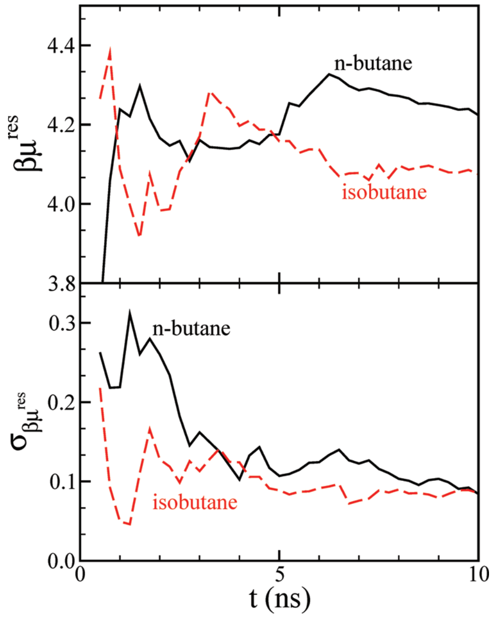


Figure 3. Performance of the EE method for predicting the residual chemical potential ($\beta\mu^{res}$) of *n*-butane and isobutane. The top panel is the estimate of the residual chemical potential as a function of simulation time, and the bottom pane is the boot strap standard error[64−66] of the five independent simulations as a function of simulation time.

propane and that of isobutane is less than *n*-butane. This observation is contrary to the results of Dill and co-workers.[27] This suggests that perhaps the uncertainty of the present results and for those of Dill and co-workers[27] may be larger than reported.

Overall, given the good agreement between the present study and that of Dill and co-workers,[27] we next draw our attention to evaluating the efficiency of the methodology employed in the current study. Figures 3 and 5−7 show the convergence of eight representative compounds from the current study. At the end of

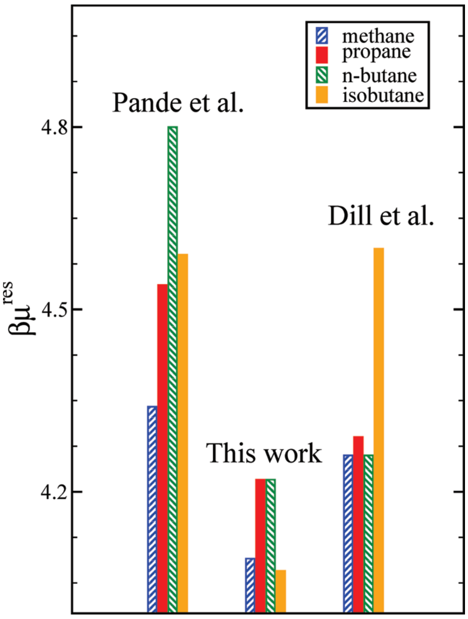


Figure 4. Comparison of the computed residual chemical potential ($\beta\mu^{res}$) of the four alkane molecules from Pande and co-workers,[7] this study, and Dill and co-workers.[27] Pande and co-workers[7] employed the same TIP3P water model but used a previous version of the AMBER force field and a different treatment of long-range electrostatic interactions.

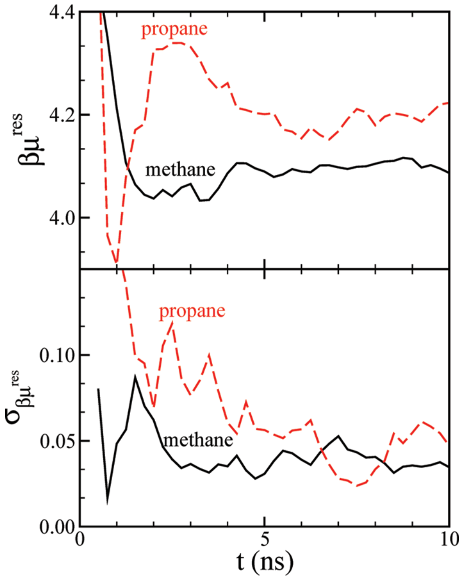


Figure 5. Performance of the EE method for predicting the residual chemical potential ($\beta\mu^{res}$) of methane and propane. The top panel is the estimate of the residual chemical potential as a function of simulation time, and the bottom panel is the boot strap standard error[64−66] of the five independent simulations as a function of simulation time.

the 10 ns simulation, all of the systems have converged to final residual chemical potentials (or hydration free energies) in agreement with Dill and co-workers[27] and with comparable precision. In the BAR study of Dill and co-workers,[27] independent simulations needed to be performed in each subensemble. Throughout the course of each simulation, transitions energies were periodically collected, and the free energy difference was computed postsimulation using all of the collected data. While the present study utilized 5 independent, 10 ns simulations, the work of Dill and co-workers[27] required a 5 ns simulation at each of the 20 coupling strengths used. Therefore, the current study required half of the total simulation time, reducing the computational time by a factor of 2. Both studies used the same LJ scaling

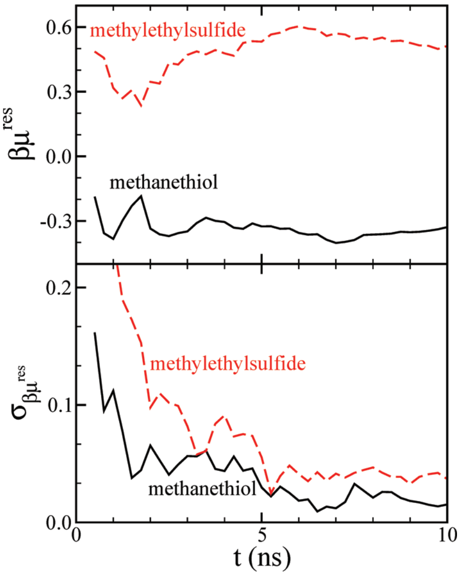


Figure 6. Performance of the EE method for predicting the residual chemical potential ($\beta\mu^{res}$) of methanethiol and methylethylsulfide. The top panel is the estimate of the residual chemical potential as a function of simulation time, and the bottom panel is the boot strap standard error[64−66] of the five independent simulations as a function of simulation time.

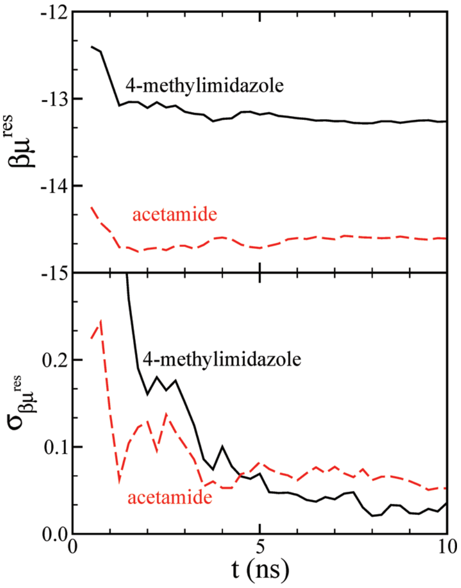


Figure 7. Performance of the EE method for predicting the residual chemical potential ($\beta\mu^{res}$) of 4-methylimidazole and acetamide. The top panel is the estimate of the residual chemical potential as a function of simulation time, and the bottom panel is the boot strap standard error[64−66] of the five independent simulations as a function of simulation time.

and number of LJ subensembles, and both used the same linear, evenly spaced electrostatic scaling scheme, but the present study used an extra electrostatic subensemble. With the close agreement of the scaling and staging schemes, we will focus our attention on the employed methods themselves. Given the similarities of the two methods, namely both sampled configurational space with MD and employed similar BAR algorithms to calculate the free energy, a natural question arises: What is the source of the apparent difference in efficiencies of the two studies? The answer to this question gives insight into the clear advantages of employing EE in a MD framework.

First, equilibrium MD studies using BAR are limited by the configurational correlation time of the system. On the other hand, in the present study, subensemble transitions are periodically attempted. When a transition is accepted, the system begins sampling in a different subensemble. This transitioning reduces the configurational correlation time of the system and increases the rate of phase space sampling, akin to parallel tempering (or replica exchange) simulations.[70−72]

Second, in order for BAR to be successful, all of the important phase space of the target and reference subensemble need to be sampled.[15,47,50,51,68,69] In addition to stratification, the use of an importance weighted[14] MC sampling procedure[39] allows the system to bridge free energy barriers separating neighboring subensembles. This is accomplished by weighting the subensemble transition acceptance probabilities so as to artificially enhance the occurrence of important configurations of intrinsically low probability. Put differently, within a given subensemble, there are likely important regions of phase space of low probability of being observed or separated from other important regions by large energetic barriers. If we again draw analogy to parallel tempering (or replica exchange),[70−72] in subensembles in which we sample from a different Hamiltonian, these states may be sampled with a much greater probability. By importance weighting the transition acceptance probabilities, the likelihood of propagating these configurations is enhanced.

In addition to the theoretical advantages mentioned, although not utilized in the present study, the EE method with BAR is readily amenable to parallel processing. The present study utilized 5 independent 10 ns simulations for a total of 50 ns of simulation time. The MD BAR method of Dill and co-workers[27] required 20 independent 5 ns simulations or 100 ns of total simulation time. As all of the MD BAR simulations are independent from each other, with sufficient computational resources, they could be performed faster in real time using multiple processors. However, for the EE method the range of subensembles studied may be broken into various overlapping windows, in which an independent simulation may be conducted. Within each window the relevant free energy change may be calculated, and then all of the results may be stitched together by enforcing that the free energy be a continuous function of subensemble.[73,74] With an intelligent choice of windowing, the wall clock time necessary to calculate the residual chemical potential with EE may be readily decreased.[25]

Another advantage of the EE method is that it requires a single simulation whose convergence can be monitored during the simulation. Once the desired level of uncertainty is reached, the simulations may be terminated. For example, many of the simulations of the present study could have been stopped after 5 ns, and the precision would not have suffered (see Figures 3 and 5−7). The fact that no postprocessing is required to obtain a solvation free energy makes the method particularly easy to use.

5. CONCLUSION

Results have been presented for a refined expanded ensemble[16−19] algorithm that combines the flat histogram method of Wang and Landau[28−30] with the Bennett's acceptance ratio methodology[36,37,42] in a MD framework. The methodology was inspired by our previous success[25] with a combined transition-matrix Monte Carlo[31−35] and Wang−Landau approach. Use of the Bennett's acceptance ratio methodology is advantageous as a result of the ability to implement the methodology in either a MC or MD framework. The method was validated by computing the residual chemical potential of 13 amino acid analogs and by comparing to reference simulation data.[27] Overall agreement is good, with the methodology of the present study reaching a comparable precision in half of the total simulation time of the previous study.[27] The apparent difference in the efficiencies is a result of the inherent advantages of the expanded ensemble method. The proposed method creates an improved decorrelation of simulation data and enhances the sampling of the important regions of the configurational phase space of each subensembles. Furthermore, the present method enables the solvation free energy to be computed from a single simulation with no postprocessing. The encouraging results of the present study suggest consideration of the employed methodology in future studies requiring free energy calculations. Moreover, the proposed methodology is highly adaptable and may be used in any type of multicanonical framework.

■ APPENDIX

Ewald Summation with Expanded Ensemble. While the use of the Ewald sum for electrostatic interactions in molecular simulations has been described extensively in the literature,[39,46,75,76] we will briefly overview the necessary implementation when decoupling electrostatic interactions of a solute molecule using the EE method. With the EE method, modifications to the standard Ewald sum are necessary to ensure that electrostatic interactions are properly decoupled such that solute−solvent interactions are treated with scaled charges of the solute (corresponding to a given coupling strength), while intramolecular interactions are treated with unaltered charge interactions.

For charge neutral systems that are periodic in three dimensions, the electrostatic potential energy, U_{elect}, may be divided into four parts with the Ewald summation:

$$U_{\text{elect}} = \frac{1}{4\pi\varepsilon_0}[U_{\text{real}} + U_{\text{recip}} + U_{\text{intra - self}} + U_{\text{point - self}}] \quad (17)$$

where U_{real} and U_{recip} are the real- and reciprocal-space terms, respectively, and $U_{\text{intra-self}}$ and $U_{\text{point-self}}$ are the intramolecular- and point-self energies, respectively.

The real-space term is given by

$$U_{\text{real}} = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}{}^{\dagger} q_i q_j \frac{\text{erfc}(\sqrt{\alpha} r_{ij})}{r_{ij}} \quad (18)$$

where α is the damping parameter, and N is the total number of charge sites. The "dagger" (†) summation indicates the exclusion of all pairs $i = j$, and intramolecular interactions separated by one, two, and three bonds. When evaluating the real-space term, all intramolecular electrostatic interactions separated by more than three bonds and all intermolecular electrostatic interactions involving only solvent molecules are treated with full charges, and all intermolecular electrostatic interactions between the solute and solvent are treated with scaled charges.

The general form of the reciprocal-space term is given by

$$U_{\text{recip}}^{\text{general}} = \frac{2\pi}{V}\sum_{\mathbf{k}\neq 0}\frac{1}{k^2}\exp\left(-\frac{k^2}{4\alpha}\right)\left[\left|\sum_{i=1}^{N} q_i\cos(\mathbf{k}\cdot\mathbf{r}_i)\right|^2 + \left|\sum_{i=1}^{N} q_i\sin(\mathbf{k}\cdot\mathbf{r}_i)\right|^2\right] \qquad (19)$$

where $\mathbf{r}_i$ is the position vector of site i, and $\mathbf{k}$ is the reciprocal lattice vector of the periodic cell images. Regardless if scaled or full charges are used for the solute molecule, eq 19 will be inconsistent with the real-space term in which different charges are used for intramolecular and intermolecular interactions. As a result, eq 19 is extended to involve three contributions:

$$U_{\text{recip}} = U_{\text{recip}}^{\text{scaled}} + U_{\text{recip}}^{\text{solute, full}} - U_{\text{recip}}^{\text{solute, scaled}} \qquad (20)$$

where the first term $U_{\text{recip}}^{\text{scaled}}$ is the evaluation of eq 19 using the actual intermolecular charges of the system. That is, the charges are full for the solvent and are scaled for the solute. To remain consistent with the formulation of the Ewald sum, the last two additional terms correct for different charges used for the intramolecular and intermolecular electrostatic interactions of the solute. $U_{\text{recip}}^{\text{solute,full}}$ is the evaluation of eq 19 but using full charges for the solute and performing the sine and cosine sum over all of the charge sites of the solute molecule (not the entire system). Similarly, $U_{\text{recip}}^{\text{solute,scaled}}$ is the evaluation of eq 19 but using scaled charges for the solute and performing the sine and cosine sum over all of the charge sites of the solute molecule. These additional two terms may be thought of as computing the reciprocal-space term for a system containing only the solute molecule at the same position, with both full and scaled charges. In the absence of interactions with the solvent, the difference of these two terms, $U_{\text{recip}}^{\text{solute,full}} - U_{\text{recip}}^{\text{solute,scaled}}$, gives the desired net effect of using different intermolecular and intramolecular charges. The necessary corrections to eq 19 may readily and efficiently be implemented into existing reciprocal-space Ewald subroutines with minor modifications. In addition, the extension to applications of particle-mesh Ewald[46] is straightforward.

The intramolecular-self energy is given by

$$U_{\text{intra-self}} = -\frac{1}{2}\sum_{j=1}^{M}\sum_{k=1}^{N_j}\sum_{l=1}^{N_j}{}^{\dagger^{-1}} q_k q_l \frac{\text{erf}(\sqrt{\alpha}\, r_{kl})}{r_{kl}} \qquad (21)$$

where M is the total number of molecules in the system, and N_j is the number of charge sites on molecule j. The "inverse dagger" ($\dagger^{-1}$) summation indicates that the sum is only over intramolecular sites excluded in the real-space term of eq 18 (i.e. intramolecular interactions separated by one, two, and three bonds). While it is straightforward to exclude these terms in the real-space term in the central simulation cell, they are implicitly included in the reciprocal-space term. Equation 21 corrects the reciprocal-space term by removing these interactions.[76] Since all intramolecular interactions use full charges, full charges are used in eq 21.

Lastly, the point-self energy is given by

$$U_{\text{point-self}} = -\sqrt{\frac{\alpha}{\pi}}\sum_{i=1}^{N} q_i^2 \qquad (22)$$

Similar to the intramolecular self energy, the point self energy corrects the reciprocal space term for self interactions. That is, interactions of a charge site with itself are straightforward to exclude in the real-space term in the central simulation cell, but they are implicitly included in the reciprocal-space term. Therefore, eq 22 corrects the reciprocal-space term by removing these interactions.

For completeness, many force fields, including those used in the present study for the amino acid analogs, use scaled electrostatic interactions between intramolecular sites separated by exactly three bonds to complement the dihedral potential. These "1-4" electrostatic interactions are computed using a direct Coulombic interaction and are hence excluded from the Ewald sum.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information. M.DynaMix 5.2 force field files for the studied amino acid analogs. This material is available free of charge via the Internet at http://pubs.acs.org/.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: ed@nd.edu; telephone: (574) 631-5687.

■ ACKNOWLEDGMENT

A.S.P. would like to thank Dr. David Mobley for informative discussions regarding his experience performing free energy calculations and gratefully acknowledges a fellowship from the Arthur J. Schmitt Foundation and additional funding from the Notre Dame Sustainable Energy Initiative. J.K.S. acknowledges funding from NSF-CBET-096758 and the Notre Dame Center for Research Computing. Computing support was provided by the Notre Dame Center for Research Computing.

■ REFERENCES

(1) Prausnitz, J. M.; Lichtenthaler, R. N.; de Azevedo, E. G. *Molecular Thermodynamics of Fluid-phase Equilibria*, 3rd ed.; Prentice-Hall PTR: Upper Saddle River, NJ, 1999.

(2) Connors, K. A.; Mecozzi, S. *Thermodynamics of Pharmaceutical Systems: An Introduction to Theory and Applications*, 2nd ed.; John Wiley and Sons, Inc.: Hoboken, NJ, 2010.

(3) *Water-Insoluble Drug Formulation*, 2nd ed.; Liu, R., Ed.; CRC Press: Boca Raton, FL, 2008.

(4) Wade, L. G. *Organic Chemistry*, 6th ed.; Pearson Education, Inc.: Upper Saddle River, NJ, 2006.

(5) Anfinsen, C. B. *Science* **1973**, *181*, 223–230.

(6) Dill, K. A.; Ozkan, S. B.; Shell, M. S.; Weikl, T. R. *Annu. Rev. Biophys.* **2008**, *37*, 289–316.

(7) Shirts, M. R.; Pitera, J. W.; Swope, W. C.; Pande, V. S. *J. Chem. Phys.* **2003**, *119*, 5740–5761.

(8) Shirts, M. R.; Pande, V. S. *J. Chem. Phys.* **2005**, *122*, 134508.

(9) Hess, B.; van der Vegt, N. F. A. *J. Phys. Chem. B* **2006**, *110*, 17616–17626.

(10) Chang, J.; Lenhoff, A. M.; Sandler, S. I. *J. Phys. Chem. B* **2007**, *111*, 2098–2106.

(11) Mobley, D. L.; Dumont, E.; Chodera, J. D.; Dill, K. A. *J. Phys. Chem. B* **2007**, *111*, 2242–2254.

(12) Kirkwood, J. G. *J. Chem. Phys.* **1935**, *3*, 300–313.

(13) Bennett, C. H. *J. Comput. Phys.* **1976**, *22*, 245–268.

(14) Valleau, J. P.; Card, D. N. *J. Chem. Phys.* **1972**, *57*, 5457–5462.

(15) Lu, N.; Kofke, D. A. *J. Chem. Phys.* **1999**, *111*, 4414–4423.

(16) Lyubartsev, A. P.; Martsinovski, A. A.; Shevkunov, S. V.; Vorontsov-Velyaminov, P. N. *J. Chem. Phys.* **1992**, *96*, 1776–1783.

(17) Lyubartsev, A. P.; Laaksonen, A.; Vorontsov-Velyaminov, P. N. *Mol. Phys.* **1994**, *82*, 455–471.

(18) Lyubartsev, A. P.; Laaksonen, A.; Vorontsov-Velyaminov, P. N. *Mol. Sim.* **1996**, *18*, 43–58.

(19) Lyubartsev, A. P.; Forrisdahl, O. K.; Laaksonen, A. *J. Chem. Phys.* **1998**, *108*, 227–233.

(20) Aberg, K. M.; Lyubartsev, A. P.; Jacobsson, S. P.; Laaksonen, A. *J. Chem. Phys.* **2004**, *120*, 3770–3776.

(21) Shah, J. K.; Maginn, E. J. *J. Phys. Chem. B* **2005**, *109*, 10395–10405.

(22) Martinez-Veracoechea, F. J.; Escobedo, F. A. *J. Phys. Chem. B* **2008**, *112*, 8120–8128.

(23) Chang, J. *J. Chem. Phys.* **2009**, *131*, 074103.

(24) Cichowski, E. C.; Schmidt, T. R.; Errington, J. R. *Fluid Phase Equilib.* **2005**, *236*, 58–65.

(25) Paluch, A. S.; Jayaraman, S.; Shah, J. K.; Maginn, E. J. *J. Chem. Phys.* **2010**, *133*, 124504.

(26) Lyubartsev, A. P.; Jacobsson, S. P.; Sundholm, G.; Laaksonen, A. *J. Phys. Chem. B* **2001**, *105*, 7775–7782.

(27) Mobley, D. L.; Bayly, C. I.; Cooper, M. D.; Shirts, M. R.; Dill, K. A. *J. Chem. Theory Comput.* **2009**, *5*, 350–358.

(28) Wang, F.; Landau, D. P. *Phys. Rev. Lett.* **2001**, *86*, 2050–2053.

(29) Yan, Q.; Faller, R.; de Pablo, J. J. *J. Chem. Phys.* **2002**, *116*, 8745–8749.

(30) Shell, M. S.; Debenedetti, P. G.; Panagiotopoulos, A. Z. *Phys. Rev. E* **2002**, *66*, 056703.

(31) Fitzgerald, M.; Picard, R. R.; Silver, R. N. *Europhys. Lett.* **1999**, *46*, 282–287.

(32) Fitzgerald, M.; Picard, R. R.; Silver, R. N. *J. Stat. Phys.* **2000**, *98*, 321–345.

(33) Errington, J. R. *Phys. Rev. E* **2003**, *67*, 012102.

(34) Errington, J. R. *J. Chem. Phys.* **2003**, *118*, 9915–9925.

(35) Paluch, A. S.; Shen, V. K.; Errington, J. R. *Ind. Eng. Chem. Res.* **2008**, *47*, 4533–4541.

(36) Fenwick, M. K.; Escobedo, F. A. *J. Chem. Phys.* **2004**, *120*, 3066–3074.

(37) Escobedo, F. A.; Abreu, C. R. A. *J. Chem. Phys.* **2006**, *124*, 104110.

(38) Kong, X., III; C., L. B. *J. Chem. Phys.* **1996**, *105*, 2414–2423.

(39) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press Inc.: New York, 1987.

(40) Torrie, G. M.; Valleau, J. P. *Chem. Phys. Lett.* **1974**, *28*, 578–581.

(41) Berg, B. A.; Neuhaus, T. *Phys. Rev. Lett.* **1992**, *68*, 9–12.

(42) Fenwick, M. K.; Escobedo, F. A. *J. Chem. Phys.* **2003**, *119*, 11998–12010.

(43) Shell, M. S.; Debenedetti, P. G.; Panagiotopoulos, A. Z. *J. Chem. Phys.* **2003**, *118*, 9406–9411.

(44) Wang, F.; Landau, D. P. *Phys. Rev. E* **2001**, *64*, 056101.

(45) Yan, Q.; de Pablo, J. J. *Phys. Rev. Lett.* **2003**, *90*, 035701.

(46) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*, 2nd ed.; Academic Press: San Diego, CA, 2002.

(47) Lu, N.; Singh, J. K.; Kofke, D. A. *J. Chem. Phys.* **2003**, *118*, 2977–2984.

(48) Shirts, M. R.; Bair, E.; Hooker, G.; Pande, V. S. *Phys. Rev. Lett.* **2003**, *91*, 140601.

(49) Ferrenberg, A. M.; Swendsen, R. H. *Phys. Rev. Lett.* **1989**, *63*, 1195–1198.

(50) Wu, D.; Kofke, D. A. *J. Chem. Phys.* **2005**, *123*, 054103.

(51) Wu, D.; Kofke, D. A. *J. Chem. Phys.* **2005**, *123*, 084109.

(52) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D. *J. Chem. Phys.* **1983**, *79*, 926–935.

(53) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157–1174.

(54) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. *J. Mol. Graphics Modell.* **2006**, *25*, 247–260.

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

(56) Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. *Chem. Phys. Lett.* **1994**, *222*, 529–539.

(57) Steinbrecher, T.; Mobley, D. L.; Case, D. A. *J. Chem. Phys.* **2007**, *127*, 214108.

(58) Lyubartsev, A. P.; Laaksonen, A. *Comput. Phys. Commun.* **2000**, *128*, 565–589.

(59) Lyubartsev, A. P.; Laaksonen, A. *MDynaMix: a Molecular Dynamics Program*; Universitet Stockholms: Stockholm, Sweden; http://www.mmk.su.se/~sasha/mdynamix/. Accessed February 1, 2010).

(60) Ryckaert, J.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327–341.

(61) Andersen, H. C. *J. Chem. Phys.* **1980**, *72*, 2384–2393.

(62) Andrea, T. A.; Swope, W. C.; Andersen, H. C. *J. Chem. Phys.* **1983**, *79*, 4576–4584.

(63) Martyna, G. J.; Tobias, D. J.; Klein, M. L. *J. Chem. Phys.* **1994**, *101*, 4177–4189.

(64) Efron, B. *SIAM Review* **1979**, *21*, 460–480.

(65) Efron, B. *Biometrika* **1981**, *68*, 589–599.

(66) Moore, D. S.; McCabe, G. P.; Craig, B. *Introduction to the Practice of Statistics*, 6th ed.; W.H. Freeman and Company: New York, 2009.

(67) Wolfenden, R.; Andersson, L.; Cullis, P. M.; Southgate, C. C. B. *Biochemistry* **1981**, *20*, 849–855.

(68) Kofke, D. A.; Cummings, P. T. *Mol. Phys.* **1997**, *92*, 973–996.

(69) Kofke, D. A.; Cummings, P. T. *Fluid Phase Equilib.* **1998**, *150−151*, 41–49.

(70) Swendsen, R. H.; Weng, J.-S. *Phys. Rev. Lett.* **1986**, *57*, 2607–2609.

(71) Hansmann, U. H. E. *Chem. Phys. Lett.* **1997**, *281*, 140–150.

(72) Earl, D. J.; Deem, M. W. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3910–3916.

(73) Errington, J. R. *Langmuir* **2004**, *20*, 3798–3804.

(74) Shen, V. K.; Errington, J. R. *J. Phys. Chem. B* **2004**, *108*, 19595–19606.

(75) de Leeuw, S. W.; Perram, J. W.; Smith, E. R. *Proc. R. Soc. Lond. A* **1980**, *373*, 27–56.

(76) Heyes, D. M. *CCP5 Quarterly Newsletter* **1983**, *8*, 29–36.

JCTC Journal of Chemical Theory and Computation
ARTICLE
pubs.acs.org/JCTC

Conformational Dependence of Isotropic Polarizabilities

Pär Söderhjelm,[†] Jacob Kongsted,[‡] and Ulf Ryde*,[§]

[†]Department of Chemistry and Applied Biosciences, Computational Science, ETH Zürich, Via Giuseppe Buffi 13, CH-6900 Lugano, Switzerland

[‡]Department of Physics and Chemistry, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark

[§]Department of Theoretical Chemistry, Lund University, Chemical Centre, P.O. Box 124, SE-221 00 Lund, Sweden

S *Supporting Information*

ABSTRACT: We perform a statistical and energetic analysis of atomic polarizabilities obtained with the LoProp approach for all atoms in the avidin tetramer for 70 snapshots from molecular dynamics simulations with seven different biotin analogues, and from the crystal structure of the photosynthetic reaction center (in total 560 698 individual polarizabilities). Dynamic effects give a variation of the polarizabilities of 0.09 Å^3 on average. Atoms at different positions in the sequence show a variation of 0.14 Å^3 on average, caused by the conformational dependence of the polarizabilities. This variation gives errors of 2 and 1 kJ/mol for relative conformational and ligand-binding induction energies. Averaged elementwise or atom-type polarizabilities give larger errors, e.g., 9 and 7 kJ/mol, respectively, for the relative conformational energies. Therefore, we recommend that polarizabilities should be assigned atomwise (i.e., individual polarizabilities for each atom in all residues), in the same way as for charges. We provide such a set of extensively averaged polarizabilities (xAvPol) for all atoms in avidin and the photosynthetic reaction center, applicable at the B3LYP/aug-cc-pVTZ level, which is converged with respect to the basis-set limit.

INTRODUCTION

During the latest decades, molecular simulations have become a powerful alternative and complement to experiments to obtain information about the structure and function of macromolecules. Such simulations are mainly based on the molecular mechanics (MM) approach, employing empirical force fields.[1] One of the most crucial issues in these force fields is the treatment of electrostatics. The great majority of such MM force fields for macromolecules employ a simple Coulomb interaction between atom-centered fixed partial charges. The atomic charges are typically obtained from quantum mechanics (QM) calculations, by fitting them to reproduce either the QM electrostatic potential or intermolecular interaction energies.[2−5]

It has long been recognized that this provides a quite crude description of the electrostatics. In particular, induction effects are completely ignored or treated in an implicit average sense, although it is well-known that polarization typically constitutes 6−30% of the electrostatic interaction energy.[6−11] Consequently, there has been great interest in incorporating induction effects in the MM force field,[11−15] e.g., by using fluctuating charges,[16,17] induced dipoles,[18−20] or Drude oscillators.[21−23] The first polarizable force field appeared as early as in the mid-1970s,[19] and specialized and accurate force fields such as SIBFA, EFP, and NEMO also early employed polarizabilities (and higher-order multipoles).[13,24,25] During the past decade, polarized variants of the more widely used macromolecular force fields have started to appear, e.g., Amber02, PFF, and Amoeba,[11,26−28] all three of which are based on atomic isotropic dipole polarizabilities.

Naturally, the accuracy of polarizable force fields depends on the accuracy of the atomic polarizabilities employed. As for atomic partial charges,[2] atomic polarizabilities are not observables, meaning that there are no reference values that could be obtained from experiments or QM calculations.[11] Instead, atomic polarizabilities have to be determined by some (arbitrary) method that is optimized in a specific way. Several methods to obtain distributed polarizability from QM calculations have been suggested.[11] For example, the atomic polarizabilities can be obtained by partitioning molecular polarizabilities, either in real space (e.g., the atoms-in-molecules approach[29]) or in terms of the basis set.[30,31] Moreover, there are also several ways to apply the perturbing field.[32−34] Alternatively, the polarizabilities can be determined by fitting to a property calculated by QM methods, e.g., the molecular polarizabilities or induction energy.[18,35−42]

There are several sets of atomic polarizabilities available. Some of them are listed in Table 1.[18−20,26−28,35,43−45] Apparently, there is little agreement in the values used or how the polarizabilities should be assigned. Thole and van Duijnen have argued that good reproduction of molecular polarizabilities can be obtained by a single isotropic polarizability for each element,[20,46] and Warshel simply uses 0.5 Å^3 for hydrogen atoms and 1 Å^3 for all other atoms.[19] Other force fields use 8−15 atom types, with one to four different polarizabilities for each element for the normal amino acids. This is in sharp contrast to atomic charges, for which most general-purpose macromolecular force fields today employ individual charges on each distinct (by symmetry) atom in each amino acid. In fact, Woods and co-workers have shown that improved accuracy is obtained using specific atomic polarizabilities, rather than polarizabilities determined by the atom type.[42] They also tested the conformational dependence of the fitted polarizabilities and showed that it was quite small, ∼1%.

Received: December 13, 2010
Published: April 11, 2011

Table 1. Comparison of 10 Different Sets of Atomic Polarizabilities (Å^3)

atom	Vogel[43]	Applequist[18]	Thole[20]	Dykstra[35]	Enzymix[19]	Charmm[45 a]	Amber02[26 b]	Amoeba[28 c]	PFF[27 d]	Amber09[11]
HC alkyl			0.514	0.00	0.5	0.044	0.135	0.496	0.25	0.443
HC aromatic	0.407	0.135	0.514	0.00	0.5	0.10	0.167	0.800	0.39	0.443
HO alcohol	0.405	0.135	0.514	0.00	0.5	0.044	0.135	0.496	0.22	0.443
HN amides		0.161	0.514	0.00	0.5	0.044	0.161	0.496	0.24	0.443
HN amines			0.514	0.00	0.5	0.044	0.135	0.496	0.24	0.443
HN in RNH_3^+			0.514	0.00	0.5	0.044	0.135	0.496	0.24	0.443
C alkyl	1.027	0.878	1.405	1.87	1.0	0.98	0.878	1.334	1.22	0.920
C aromatic			1.405	1.61	1.0	2.07	0.360	1.334	1.49	1.298
C amide	1.027	0.616	1.405	1.88	1.0	1.65	0.616	1.334	0.83	1.298
C in COO^-			1.405	1.88	1.0	1.65	0.616	1.334	0.82	1.298
N amine			1.105	1.64	1.0	1.10	0.530	1.073	1.33	0.934
N aromatic			1.105	1.29	1.0	1.10	0.530	1.073	1.42	0.934
N amide		0.530	1.105	1.29	1.0	1.10	0.530	1.073	1.15	0.934
OH aliphatic alcohol	0.604	0.465	0.862	0.75	1.0	0.84	0.465	0.834	0.77	0.606
OH aromatic alcohol			0.862	0.75	1.0	0.84	0.465	0.873	0.77	0.593
O backbone amide	0.841	0.434	0.862	0.25	1.0	0.84	0.434	0.837	0.91	0.593
O side-chain amide			0.862	0.25	1.0	0.84	0.434	0.834	0.91	0.593
O in COO^-			0.862	0.25	1.0	2.14	0.434	0.837	0.97	0.593
S					1.0	0.34	2.900	3.300	2.872	3.183

[a] Listed data for CHARMM are from an old but complete listing.[13] Newer developments for alcohols, alkanes, and amides[7−9] have used either slightly modified Applequist parameters[18] or the Thole parameters.[20] [b] Data from the parm99.dat file in the Amber10 distribution. [c] Data from the amoebapro.prm files in the Amber10 distribution. [d] Data from Table 8 in ref 27.

In this paper, we address these issues in a more systematic way. In previous investigations of the influence of the protein electrostatics on excitation and ligand-binding energies, we have calculated polarizabilities for all atoms in several proteins with QM calculations,[47,48] using the LoProp approach.[34] Here, we analyze those data, collecting statistics over the polarizabilities of each atom in the sequence. Thereby, we can address questions such as the following: How large is the conformational dependence of atomic polarizabilities? How are polarizabilities best assigned: by element, by atom type, or by atom? Can transferable polarizabilities be obtained by simply averaging over all calculated values?

METHODS

In this paper, we analyze polarizabilities calculated in two studies, viz., a study of the binding affinity of seven biotin analogues to the protein avidin[48] and new calculations for the photosynthetic reaction center (PRC) from *Rhodobacter sphaeroides*. Both these studies employed a multicenter−multipole expansion up to quadrupoles and anisotropic polarizabilities, obtained with the LoProp approach[34] using the Molcas software.[49] The LoProp method has been shown to be better than other related methods to calculate polarizabilities.[50] The calculations were performed at the density functional B3LYP[51] level, using either the 6-31G*,[52] aug-cc-pVDZ, aug-cc-pVTZ, or aug-cc-pVQZ basis sets.[53] These basis sets are of sizes smaller than, similar to, larger than, and much larger than, respectively, the popular Sadlej basis set designed for the calculation of polarizabilities.[54] Each basis set was turned into the atomic natural orbital form (as required by the LoProp procedure) by a linear transformation that does not affect the orbital optimization.

The properties were calculated for the whole protein by dividing it into the individual amino acid residues, which were capped with CH_3CO- and $-NHCH_3$ groups (dipeptides). The effects of the capping groups were removed by calculating the properties also of the overlapping $CH_3CONHCH_3$ fragments and subtracting them from the properties of the corresponding dipeptides—the molecular fractionation with conjugate caps approach,[55] which has been shown to give errors of 1 kJ/mol or less.[10] A separate calculation was performed on every residue in the structure, with the actual geometry obtained either from the crystal structure (PRC) or from 10 snapshots from a molecular dynamics (MD) simulation with the Amber02 force field (avidin[56]).

In the standard LoProp approach, anisotropic polarizabilities are obtained both for atoms and for bond isocenters. To facilitate the present comparison, we restricted this study to isotropic polarizabilities, because this is the form used in the Amber02, PFF, and Amoeba force fields. The isotropic polarizabilities were obtained as the average of the three diagonal elements of the anisotropic tensor. Moreover, only atomic polarizabilities were considered by partitioning the bond polarizabilities equally on the two bonded atoms.

Interaction energies were calculated with the Amber10 software,[57] using Amber exclusion rules, i.e., that polarization between atoms separated by one or two bonds is ignored, whereas for atoms separated by three bonds, the electric field was scaled by a factor of 1.2.[26] The induction energy was calculated iteratively until successive estimates of the induced dipoles agreed within 0.0001 D, using a second-order extrapolation scheme (indmeth=1).

The exclusion rules are important, because they influence the molecular polarizability resulting from a given set of atomic polarizabilities. Therefore, polarizabilities derived with a specific set of rules are in principle not comparable to those derived with other rules, and they cannot be directly transferred. Nevertheless, such transferability has sometimes been assumed, as in the development of the Amber 2002 force field,[26] in which Applequist polarizabilities, derived using coupling between all atoms, were adopted into the much more restricted coupling scheme of Amber. One can therefore expect that these polarizabilities are too small.

The same problem also occurs in this investigation, because the LoProp polarizabilities add up to the molecular polarizability

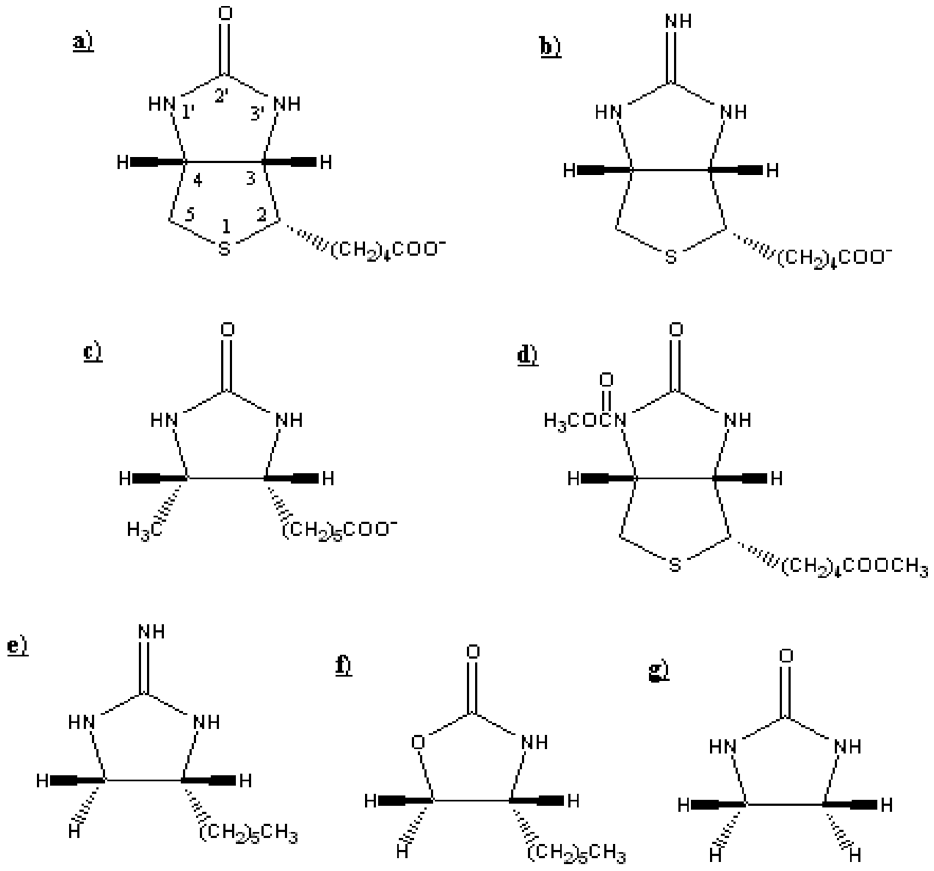


Figure 1. The seven biotin analogues used in this study. (a) Btn1 (biotin); (b−g) Btn2−Btn7.

and thus should not be coupled within the molecule used to calculate them, in our case a protein residue. Thus, when they are used with the Amber exclusion rules or numerically compared to Amber polarizabilities, they should in principle be scaled down to reproduce the (isotropic) molecular polarizability. To investigate the magnitude of this effect, we assumed a uniform scaling over all atoms in a molecule and calculated the required scale factor for each of the 991 molecules used to compute the LoProp polarizabilities for an avidin snapshot. On average, this factor was 0.987, with a standard deviation of 0.007. Because the influence of such scaling on the results would be negligible, we did not modify the polarizabilities. It should also be noted that the choice of exclusion rules also has an effect on the polarization caused by the static charges. However, in the Amber polarizable force field, the charges are derived by taking the statically induced dipoles into account so that the major part of this effect is canceled. Because of this connection, we did not specifically study this issue.

We studied the binding of the seven biotin analogues (Btn1−Btn7) in Figure 1 to avidin. The setup of the molecular dynamics simulations has been described before.[10,56] We used 10 snapshots (sampled every 20 ps) for each analogue taken from this investigation, performed by the polarizable Amber 2002 force field[11,26] (the 02ohp simulation in ref 56).

■ RESULT AND DISCUSSION

Polarizabilities. First, we studied the conformational dependence of the polarizabilities calculated with the LoProp approach[34] for all atoms in 10 snapshots from MD simulations of avidin bound to the seven different biotin analogues in Figure 1 using the B3LYP/6-31G* method. The LoProp polarizabilities range from 0.05 to 2.45 Å^3 (H in Phe-70 to SG in Cyx-452; Cyx denotes Cys in cystine linkages). For individual atoms, the range of the polarizability (i.e., the maximum minus the minimum value of the polarizability of the same atom) among the 70 snapshots varies from 0.008 to 0.35 Å^3 (for HH2 in Trp-219 and CD2 in Trp-68; average 0.09 Å^3). This illustrates the expected variation of the polarizabilities caused by dynamic effects. There is little similarity between the calculated polarizabilities and those in the Amber02 force field: In fact, for 6796 of the 7708 protein atoms

Table 2. Polarizabilities Calculated for Each Element in the 70 Snapshots of Avidin (Only Protein Atoms) (Å^3)[a]

	LoProp						Amber02	
element	no.	Aver	Stdev	Min	Max	Range	Min	Max
H	267 820	0.22	0.04	0.05	0.33	0.27	0.14	0.17
C	169 400	1.13	0.13	0.82	1.59	0.77	0.36	0.88
N	48 160	0.91	0.13	0.49	1.24	0.75	0.53	0.53
O	53 060	0.54	0.03	0.41	0.68	0.27	0.43	0.47
S	1 120	2.16	0.13	1.88	2.45	0.57	2.90	2.90

[a] no. is the number of individual polarizabilities obtained for each element. Aver, Stdev, Min, Max, and Range are the average, standard deviation, minimum, and maximum values for each element. Range is Max − Min. For comparison, the Min and Max values of the Amber02 polarizabilities are also included.

(88%), the Amber value is outside the range of the calculated polarizabilities in the various snapshots.

An interesting question is how polarizabilities are best assigned to atoms in a protein. Are they the same for each element, for each atom type, or should they be assigned atomwise, like point charges? Statistics for elemental polarizabilities are given in Table 2. It can be seen that the LoProp polarizabilities of all elements show a quite large variation, ranging from 0.27 Å^3 for H and O to ∼0.75 Å^3 for N and C. Thus, it does not seem to be a good idea to assign polarizabilities only on the basis of the element. For all elements, except sulfur, the averaged LoProp polarizabilities are higher than the corresponding Amber values. For H, N, and O, the Amber values are within the calculated range, but for C and S, at least some of the Amber values are outside the range of the LoProp values. The same applies to all the other sets of polarizabilities in Table 1, although with different elements.

The corresponding statistics for the Amber02 atom types are shown in Tables 3 and 4. Amber02 employs 27 atom types for a normal protein, which are all included and described in Table 4. However, most of the Amber02 atom types of the same element use the same polarizabilities. In fact, there are only 10 distinct polarizabilities in Amber (taken from Applequist;[18] three for C and H, two for O, and one for N and S). These are shown in Table 3. It can be seen that the LoProp polarizabilities still show large ranges, e.g., up to 0.77 Å^3 for carbon, and 0.57 and 0.75 Å^3 for S and N. Hydrogen has the lowest ranges (0.10−0.23 Å^3), followed by oxygen (0.23−0.27 Å^3). There is a fair correlation between the average calculated values and the Amber values (r^2 = 0.78).

The corresponding statistics for all the 27 Amber02 atom types are given in Table 4. It can be seen that the range is still large for most atom types, up to 0.77 Å^3 for CT (sp^3 carbon). In fact, the range is below 0.1 Å^3 only for three of the Amber atom types, H4, H5, and HP (explained in Table 3). For 20 of the 27 atom types, the Amber polarizabilities are outside the range of the calculated ones. In many cases, it is obvious that the Amber

Table 3. Statistics for LoProp Polarizabilities over the Amber02 Atom Types That Have Distinct Polarizabilities (Å^3)[a]

atom type	no.	Aver	Stdev	Min	Max	Range	Amber
C	41 020	1.15	0.06	0.94	1.39	0.45	0.62
CT	103 040	1.12	0.15	0.82	1.59	0.77	0.88
C other	25 340	1.17	0.09	0.88	1.52	0.64	0.36
H	59 920	0.17	0.02	0.05	0.23	0.18	0.16
HA, H4, H5	17 080	0.28	0.02	0.23	0.33	0.10	0.17
H other	190 820	0.24	0.03	0.09	0.32	0.23	0.14
N	48 160	0.91	0.13	0.49	1.24	0.75	0.53
O, O2	44 380	0.54	0.03	0.41	0.68	0.27	0.43
OH	8 680	0.53	0.03	0.45	0.68	0.23	0.47
S	1 120	2.16	0.13	1.88	2.45	0.57	2.90

[a] The columns have the same meaning as in Table 2. The atom types are explained in Table 4.

Table 4. Statistics for the LoProp Polarizabilities over All the Amber02 Atom Types for Proteins (Å^3)[a]

atom type	no.	Aver	Stdev	Min	Max	Range	Amber	description
C	41 020	1.15	0.06	0.94	1.39	0.45	0.62	sp^2 C in carbonyl groups
CA	20 020	1.15	0.08	0.98	1.51	0.53	0.36	aromatic C
CB	1 120	1.25	0.06	1.01	1.47	0.46	0.36	CD2 in Trp
CC	280	1.19	0.03	1.11	1.26	0.15	0.36	CG in His
CN	1 120	1.27	0.05	1.10	1.41	0.31	0.36	CE2 in Trp
CR	280	0.96	0.03	0.89	1.03	0.14	0.36	CE1 in His
CT	103 040	1.12	0.15	0.82	1.59	0.77	0.88	sp^3 aliphatic C
CV	280	0.97	0.04	0.88	1.07	0.19	0.36	CD2 in Hid
CW	1 120	1.15	0.04	1.02	1.25	0.23	0.36	CD2 in Hie and Hip, CD1 in Trp
C*	1 120	1.34	0.05	1.19	1.52	0.34	0.36	CG in Trp
H	59 920	0.17	0.02	0.05	0.23	0.18	0.16	H bound to N
H1	57 540	0.23	0.03	0.15	0.30	0.15	0.14	aliphatic H bound to C with one electron-withdrawing group
H4	1 400	0.28	0.02	0.23	0.30	0.07	0.17	HD1 in Trp, HD2 in Hid
H5	280	0.29	0.01	0.28	0.30	0.03	0.17	HE1 in Hid
HA	15 400	0.28	0.02	0.23	0.33	0.10	0.17	aromatic H
HC	119 840	0.25	0.02	0.18	0.32	0.13	0.14	aliphatic H bound to C without electron-withdrawing groups
HO	8 680	0.16	0.02	0.09	0.21	0.12	0.14	H in hydroxyl groups
HP	4 760	0.22	0.01	0.17	0.27	0.09	0.14	HE in Lys
N	37 660	0.96	0.09	0.64	1.24	0.61	0.53	sp^2 N in amide groups
N2	6 300	0.74	0.12	0.57	1.02	0.44	0.53	NE and NH in Arg
N3	2 520	0.64	0.02	0.49	0.70	0.21	0.53	NZ in Lys
NA	1 400	0.94	0.07	0.76	1.17	0.41	0.53	protonated N in aromatic rings
NB	280	0.87	0.04	0.79	0.97	0.18	0.53	nonprotonated N in aromatic rings
O	37 660	0.54	0.03	0.41	0.64	0.23	0.43	O in carbonyl groups
O2	6 720	0.58	0.04	0.42	0.68	0.26	0.43	O in carboxyl groups
OH	8 680	0.53	0.03	0.45	0.68	0.23	0.47	O in hydroxyl group
S	1 120	2.16	0.13	1.88	2.45	0.57	2.90	S

[a] The columns have the same meaning as in Table 2.

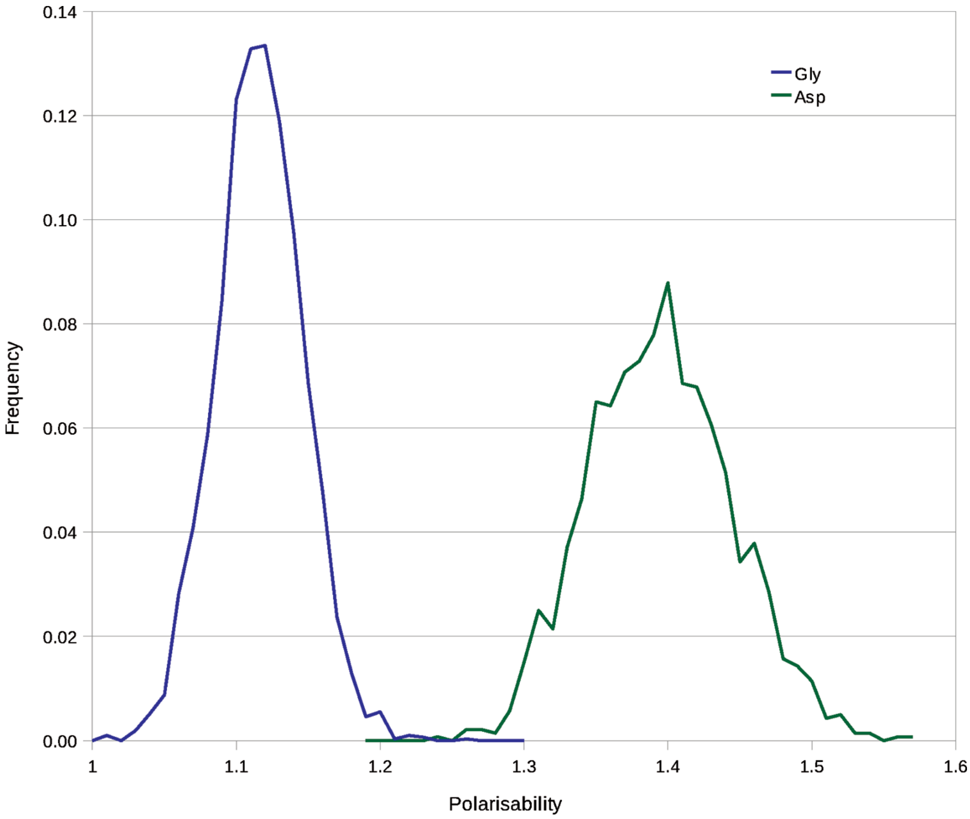


Figure 2. Frequency plot for the LoProp polarizabilities (Å^3) of the CA atom in Gly and Asp in avidin (3080 and 1400 individual polarizabilities, respectively).

atom types still are too crude to give accurate and transferable polarizabilities. This is clearly illustrated for CA atoms of Gly and Asp (which share the same Amber atom type), shown in Figure 2, where the frequencies of the LoProp polarizabilities are shown for the 70 snapshots and the 44 and 22 atoms of each type, respectively. It is obvious that the two distributions are distinct and essentially nonoverlapping, so that different polarizabilities are appropriate for the CA atom in these two amino acids.

Finally, we calculated the average of the polarizabilities for the same atom in the same residue anywhere in the sequence and over the 70 snapshots. This suppressed some of the variation. Now, the average range was 0.14 Å^3. 229 of the 388 distinct atoms (59%) showed a range of less than 0.15 Å^3 and only 28 atoms showed a range over 0.3 Å^3, with CD2 of Trp showing the largest range (0.46 Å^3). Other atoms with large ranges are always carbon and nitrogen atoms, as well as the two sulfur atoms. These are also the atoms with the highest polarizabilities. In fact, there is a good correlation between the size of the polarizabilities and the range ($r^2 = 0.76$), as is shown in Figure 3. This shows that there is a significant conformational dependence of the polarizabilities (23% on average), much larger than for small model compounds (1%).[42] In fact, 70% of the polarizabilities of all possible pairs of atoms from the same residue at different places in the sequence were statistically different at the 95% level according to a simple *t* test.

The largest polarizabilities are those of the two S atoms in Cys and Met (2.27 and 2.04 Å^3). Next largest are those of some carbon atoms, typically CA atoms in various residues, but also some CB and CG atoms (up to 1.39 Å^3 for CA in Asp). The smallest C polarizability is that of the CG atoms of Val (0.90 Å^3). The largest nitrogen polarizability is that of the backbone amide in Pro (1.14 Å^3), and the smallest one is that of the side-chain NZ of Lys (0.62 Å^3). The largest oxygen polarizability is that of the OH group in Tyr (0.64 Å^3). The smallest one is that of the amide backbone O of Cyx (0.43 Å^3). The hydrogen polarizabilities are well separated from those of the other elements. The largest one is that of HH2 in Trp (0.32 Å^3), and the smallest is that of the amide backbone H of Phe (0.16 Å^3).

There are several obvious groups of the calculated polarizabilities. For O, they are distinct and not overlapping: hydroxyl and backbone carbonyl groups (0.50−0.55 Å^3), side-chain carbonyl groups and all carboxyl groups (0.56−0.60 Å^3), and the hydroxyl group of Tyr (0.64 Å^3). The same applies to N atoms, although the ranges are larger: N in Lys side chains and in NH of Arg (0.62−0.69 Å^3), N in side-chain amides (0.71 Å^3), N in His and NE in Arg (0.86−0.91 Å^3), N in the backbone amides and NE in Trp (0.84−1.06 Å^3), and N in Pro (1.14 Å^3). However, for the hydrogen atoms, the ranges are large and overlapping: H in amide and NH_3^+ groups (0.14−0.18 Å^3), H in hydroxyl groups (0.16−0.17 Å^3), H in side-chain amide groups (0.18−0.22 Å^3), HC with electron-withdrawing neighbors (0.19−0.26 Å^3), H in aromatic groups (0.26−0.32 Å^3), and other HC (0.22−0.28 Å^3).

Finally, for carbon atoms, it becomes even harder to find natural groups: methyl groups, as well as CB and CD in Pro and CE1 in His have 0.90−0.96 Å^3, CD2 in His, CD in Arg, and CG and CD in Lys have 0.97−1.06 Å^3, C in side-chain carbonyl groups and all carboxyl groups have 1.04−1.13 Å^3, C in backbone carbonyl groups, as well as CD2 in Hie and Hip, and CD1 in Trp give 1.10−1.20 Å^3. However, the remaining aliphatic and aromatic C

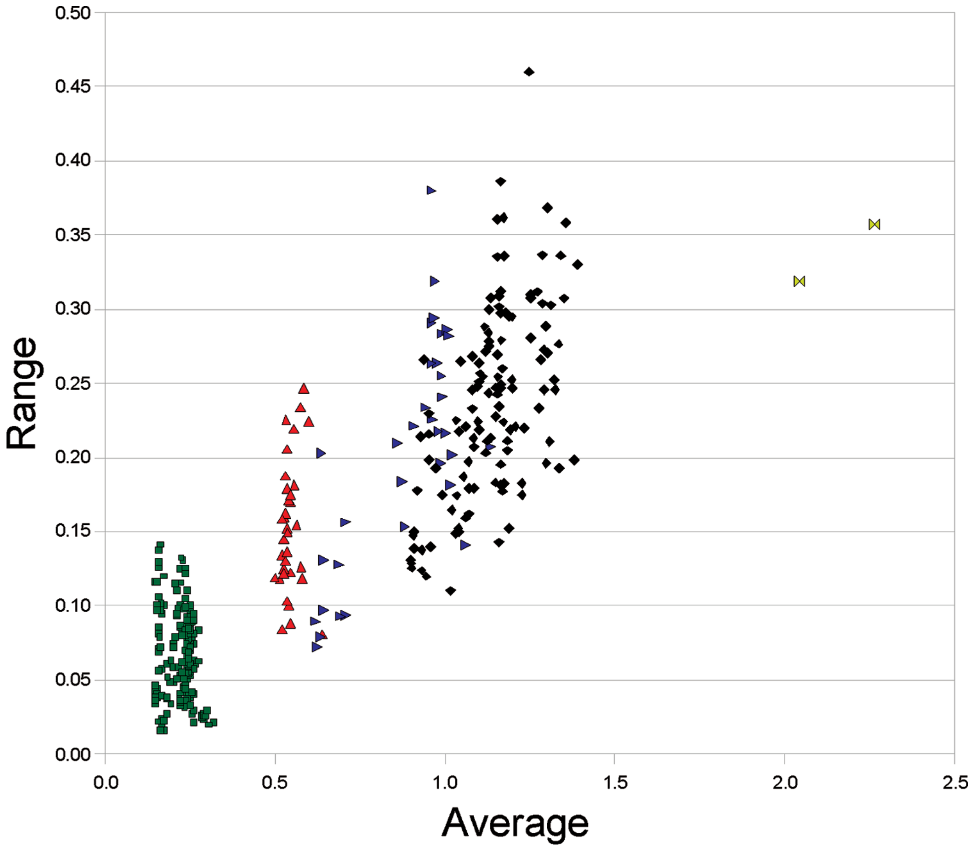


Figure 3. Correlation between the average size of the LoProp atomic polarizabilities and their range (both in units of Å^3). The points are coded according to the element: H, green squares; C, black diamonds; N, blue right-pointing triangles; O, red up triangles; S, yellow double triangles.

atoms still give large and overlapping ranges (1.03–1.39 and 1.07–1.34 Å^3, respectively), without any obvious grouping.

Figure 4 shows the correlation between the atomic polarizabilities and the Amber polarizabilities. It can be seen that there is some correlation ($r^2 = 0.72$), but there is room for significant improvement, in particular for the carbon, nitrogen, and sulfur atoms. Apparently, the polarizabilities of the atoms are very sensitive to their neighboring atoms in a way that is hard to describe without introducing very many atom types. Therefore, we suggest that, for accurate results, it is better to assign separate polarizabilities to each atom in every amino acid, rather than using atom types, in exactly the same way as done for the charges in most force fields, including Amber. In analogy with the extensively averaged electrostatic potential (xAvESP) charges obtained in a similar way,[58,59] we call these averaged LoProp atomic polarizabilities from avidin xAvPol1 in the following and they are provided in the Supporting Information, Table S1.

Basis-Set Dependence. It is well-known that calculated polarizabilities are sensitive to the specific electronic-structure method and the one-electron basis sets.[60] Owing to the presence of the electric-dipole operator in the second-order perturbation theory expression for the dipole–dipole polarizability, use of diffuse basis functions in accurate calculations of polarizabilities is usually of great importance. In the avidin calculations, we have used the B3LYP density functional combined with the middle-sized 6-31G* basis set. In order to check the reproducibility of these results, we need to ensure that polarizabilities calculated with other methods are not widely different. Fortunately, we have also polarizabilities calculated at the B3LYP/aug-cc-pVTZ level for one snapshot of two of the biotin analogues (Btn1 and Btn7; the results for the two ligands are very similar). Therefore, we can make a direct comparison of the polarizabilities obtained with this more accurate but much more expensive method. The polarizabilities calculated with the two methods differ by 0.12 Å^3 on average, with the larger basis set giving larger polarizabilities (only for ~5% of the atoms does the calculation with the smaller basis set give larger polarizabilities, and only by up to 0.04 Å^3). As expected, the largest differences are obtained for the negatively charged carboxylate groups and for the sulfur atoms: The difference is 0.61 Å^3 for SD in Met, 0.44 Å^3 for SG in Cyx, 0.48–0.55 Å^3 for the carboxylate O atoms, and 0.42–0.51 Å^3 for the carboxylate C atoms (with slight differences between Asp, Glu, and the carboxy terminals). Other atoms with large differences are OE1 of Gln (0.31 Å^3), CE1 and NE2 of Hid (0.29 Å^3), OD1 of Asn, and CH2 and CZ3 of Trp (0.28 Å^3).

Again, there is a significant variation between the various atoms, which is impossible to describe elementwise and also hard to describe by atom types. Instead, it is best described by atomic polarizabilities. Then, the differences are highly reproducible: Only three atomic polarizabilities give differences over 0.01 Å^3 between the Btn1 and Btn7 simulations (SD in Met, OD1 in Asp, and C in the carboxy terminal, with differences of 0.04, 0.02, and 0.02 Å^3, respectively). Thus, the effect of the basis set is quite small and highly consistent and therefore the polarizabilities can quite easily be extrapolated to the larger basis set. This will increase the polarizabilities for all except five atoms (CA in Lys and Arg, CB in Ile and Val, and CG in Leu). Therefore, the difference toward the Amber polarizabilities will increase, except for the two S atoms,

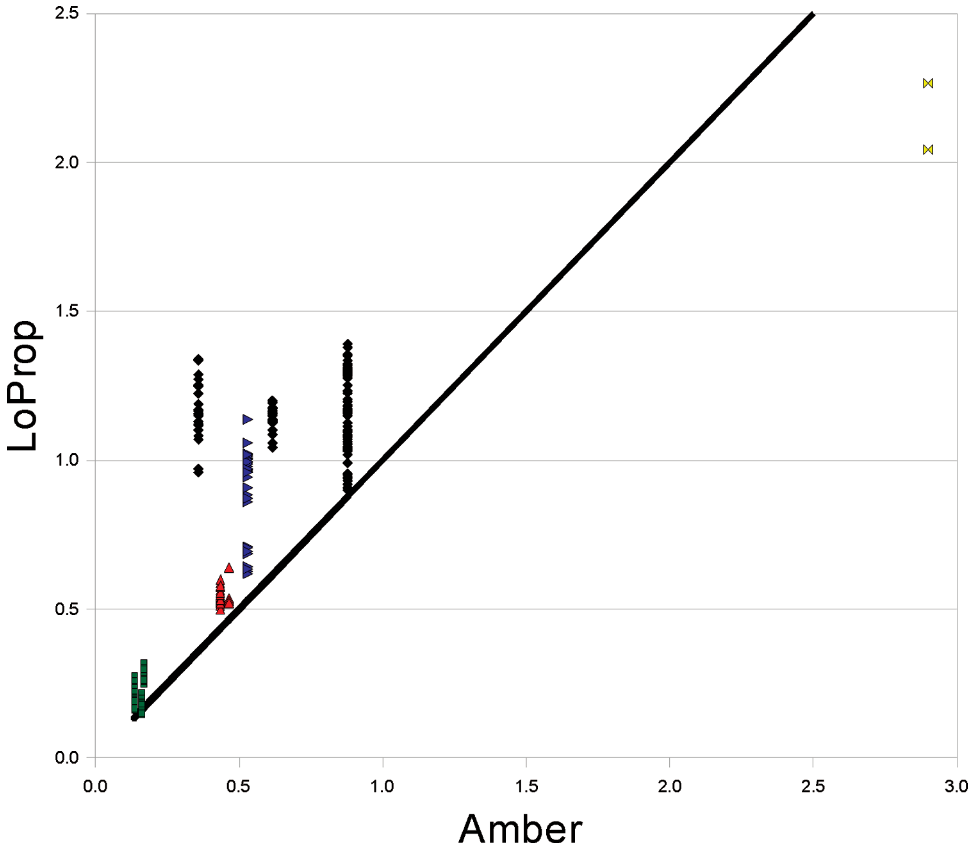


Figure 4. Comparison between the atomic LoProp and the Amber polarizabilities (both in units of Å^3). The points are coded according to the element: H, green squares; C, black diamonds; N, blue right-pointing triangles; O, red up triangles; S, yellow double triangles. The line is $x = y$.

which become similar with the larger basis set, 2.65 and 2.67 Å^3, for SD in Met and SG in Cyx, respectively (2.9 Å^3 in Amber).

To further study the basis-set dependence of the polarizabilities, we performed some additional calculations with the aug-cc-pVDZ, aug-cc-pVTZ, and aug-cc-pVQZ basis sets (still with the B3LYP method) for the groups that showed the largest dependence with respect to the basis set: Cys, Cyx, Met, Asp, and a carboxy terminal. The results show that the polarizabilities are reasonably converged at the aug-cc-pVTZ level: The polarizabilities calculated at the aug-cc-pVTZ and aug-cc-pVQZ differ by only 0.02 Å^3 on average, with a maximum difference of 0.09 Å^3 for SD in Met (the polarizability decreases when the basis set is increased). The SG atoms in Cys and Cyx also show rather large differences, 0.04−0.08 Å^3, whereas the polarizabilities of the carboxylate O atom change by only 0.03 Å^3 (but those of the carboxylate C atom change by 0.05 Å^3). Besides these atoms, the largest change is 0.04 Å^3 for some carbonyl O atoms. In fact, the polarizabilities are fairly converged already at the aug-cc-pVDZ level, with average and maximum differences of 0.03 and 0.15 Å^3 (again SD of Met gives the largest change) toward the aug-cc-pVQZ data. This shows that it is probably better to calculate the polarizabilities with the aug-cc-pVDZ or Sadlej basis set than with 6-31G*.

On the other hand, it is normally assumed that polarizabilities in the condensed phase are lower than those calculated in a vacuum,[27,61,62] e.g., by 7−9% for water. Therefore, the Friesner group uses a basis set without diffuse functions (cc-pVTZ-f[63]) for the calculation of polarizabilities, whereas MacKerell and coworkers scale down polarizabilities by a factor of 0.724.[62] However, the primary aim of this paper is not to establish a proper level to calculate polarizabilities, but rather to quantify the extent and effect of conformational dependence of polarizabilities in proteins.

Different Proteins. Next, we performed the same analysis for another protein, viz., the photosynthetic reaction center from *Rhodobacter sphaeroides*. We calculated the LoProp isotropic atom-centered polarizabilities for each atom (in total 12 818), but only for a single structure (crystal structure with added hydrogen atoms). From these, we calculated atomic polarizabilities by averaging over all residues of each type in the protein (xAvPol2; also included in Table S1 in the Supporting Information). For the 325 atoms that are common to avidin, the average difference between the two sets is only 0.02 Å^3, indicating that the LoProp atomic polarizabilities are remarkably transferable between different proteins. In particular, the largest differences (up to 0.13 Å^3) were observed for C and N atoms in Hid and Tyr residues, for which there is only one occurrence in the avidin monomer, showing that the deviation is mainly statistical in the nature (but it also indicates that there is a significant conformational dependence of the polarizabilities).

Finally, we constructed a set of atomic polarizabilities by averaging over the two proteins, weighting the average after the number of residues of each type in the monomer of each protein. For example, there are 79 Ala residues in PRC and four in the avidin monomer, so we summed the polarizability from PRC multiplied by 79 and that of avidin multiplied by 4 and divided the sum by 83. Note, however, that this weighting of the average has a maximum effect of 0.06 Å^3, so it is of little importance. This averaged set of atomic polarizabilities will be

Table 5. Description of the Various Sets of Polarizabilities Considered in the Work

			polarizabilities different for		
charge set	no. distinct polarizabilities	description	snapshots	same residue	based on protein
LoProp	547 880	LoProp atomic polarizabilities	yes	yes	avidin
Aver	7916	LoProp average over snapshots	no	yes	avidin
xAvPol1	459	Aver, averaged over residues	no	no	avidin
xAvPol2	309	like xAvPol1 but from PRC	no	no	PRC
xAvPol3	521	weighted average over xAvPol1 and xAvPol2	no	no	avidin, PRC
xAvPol4	395	xAvPol3 corrected to aug-cc-pVTZ basis	no	no	avidin, PRC
Element	5	LoProp averaged over elements (Table 2)	no	no	
Type	27	LoProp averaged over atom types (Table 4)	no	no	
Amber02	10	Amber FF02 polarizabilities[26]	no	no	
Amber09	7	new Amber polarizabilities[11]	no	no	
Charmm	9	CHARMM polarizabilities[45]	no	no	
Amoeba	8	Amoeba polarizabilities[28]	no	no	
Enzymix	2	Enzymix polarizabilities[19]	no	no	

called "xAvPol3" in the following. We also constructed a fourth set of polarizabilities by extrapolating the xAvPol3 polarizabilities to the aug-cc-pVTZ basis set with the atomic correction factors obtained in the previous section. The resulting set, xAvPol4, is also included in Table S1 in the Supporting Information.

Induction Energies. Up to now, we have only discussed the actual values of the polarizabilities. To put these into a more interesting perspective, we studied how these differences in the polarizabilities affect electrostatic interaction energies. Therefore, we have calculated three types of energies for avidin and its complexes with the seven biotin analogues in Figure 1. We tested 13 different sets of polarizabilities, viz., the original LoProp polarizabilities for avidin (LoProp), polarizabilities averaged over the 10 snapshots (Aver), xAvPol1, xAvPol2, xAvPol3, xAvPol4, the average elemental polarizabilities in Table 2 (Element), the averaged polarizabilities for the 27 Amber atom types in Table 4 (Type), and the Amber02, Amber09, Charmm, Amoeba, and Enzymix polarizabilities listed in Table 1. The polarizabilities are briefly described in Table 5. All the other MM parameters, including the atomic charges, were identical in the calculations. The calculations were performed with the Amber software[57] and the Amber02 charges.[26]

First, we studied the total induction energy within the whole avidin tetramer without any ligand and water molecules in the 70 snapshots. The absolute energies are not comparable, because different polarizabilities are used, but the fluctuations around the average value should be similar if the different force fields are to sample the same configurational space. Interestingly, all polarizabilities give fluctuations with a range (maximum minus minimum value among the 70 snapshots) of 1515−1651 kJ/mol. The force fields based on the LoProp B3LYP/6-31G* polarizabilities give a smaller range (1515−1533 kJ/mol) than the other polarizabilities (1553−1595 kJ/mol), and Enzymix gives the largest range (1651 kJ/mol).

Second, we compared these relative interaction energies for each snapshot, using the Aver polarizabilities as a reference (we cannot use the LoProp polarizabilities as a reference, because they change for each snapshot). Several conclusions can be drawn from the results presented in Table 6. First, the various force fields give mean absolute differences (MADs) of 2−65 kJ/mol in the order xAvPol1, xAvPol3, Type, Element, xAvPol2, Amber02, xAvPol4, Amber09, Charmm, Amoeba, and Enzymix. Thus, the polarizabilities are much less sensitive to the conformation than charges: The MAD between the Aver and xAvPol1, xAvPol2, or xAvPol3 sets is only 2 kJ/mol, and both the Type and Element polarizabilities give MADs less than 10 kJ/mol, which may be acceptable in many applications.

Third, the B3LYP/6-31G* polarizabilities are clearly not converged, because the B3LYP/aug-cc-pVTZ polarizabilities (xAvPol4) give induction energies that differ by 27 kJ/mol on the average. This shows that larger basis sets should be used for the calculation of the polarizabilities or they should be corrected in the same way as for xAvPol4.

Fourth, different standard force fields give widely differing results, differing from Aver by 26−65 kJ/mol, or up to 4% of the total variation. In most cases, the crude Enzymix polarizabilities give the largest difference. Of course, some of this difference may be caused by the fact that the Aver polarizabilities are based on calculations with a too small basis set. Therefore, we have added an extra row in Table 6 (MAD′) where we instead use the xAvPol4 results (which are close to the basis-set limit) as the reference. It can be seen that the MAD for Amber09, Amoeba, and Enzymix are reduced to 15, 35, and 42 kJ/mol, whereas the MAD for Charmm is not changed and that of Amber02 actually increases. This shows that there still are extensive differences between the polarizabilities of the various force fields, far beyond what is caused by the conformational dependence.

Finally, we note that the variation in the relative induction energies is appreciably smaller than the corresponding variation in relative electrostatic energies when the atomic charges were varied in a similar manner (up to 150 kJ/mol).[58] This is in accordance with the observation that induction energies typically are 6−30% of the electrostatic energies.[6−11] Still, differences of over 10 kJ/mol in relative energies may have a strong influence on the phase space visited during a MD simulation.

Ligand Binding Energies. Next, we studied the induction contribution to the binding energies of the seven biotin analogues in Figure 1 with 10 snapshots for each ligand and the same 13 sets of polarizabilities (and still with the same Amber02 charges). The energy was calculated as the difference between the interaction energies in the complex, the protein, and the ligand:

$$E(\mathrm{PL}) - E(\mathrm{P}) - E(\mathrm{L})$$

Table 6. Differences in Relative Polarization Energies Relative to Aver (kJ/mol)

	xAvPol1	xAvPol2	xAvPol3	xAvPol4	Element	Type	Amber02	Amber09	Charmm	Amoeba	Enzymix
MAD	2	2	2	27	9	7	26	35	36	58	65
Min	−5	−10	−9	−72	−36	−12	−69	−97	−133	−142	−151
Max	4	6	5	61	22	17	53	91	99	166	181
Range	9	15	14	134	57	29	122	188	232	308	333
MAD′ [a]							42	15	34	35	42

[a] Mean absolute deviation from the xAvPol4 results.

Table 7. Differences in Ligand-Interaction Polarization Energies, Compared to LoProp (kJ/mol)[a]

	Aver	xAvPol1	xAvPol2	xAvPol3	xAvPol4	Element	Type	Amber02	Amber09	Charmm	Amoeba	Enzymix
MAD	0.6	0.8	1.1	1.1	12.6	4.1	1.5	4.8	17.2	12.4	24.9	25.7
Max	2.6	3.6	4.8	4.6	30.5	10.8	5.4	11.9	39.3	36.3	54.2	61.1
MAD1−3	1.0	1.4	1.8	1.7	19.4	7.8	2.1	4.2	28.7	24.6	36.9	40.6
Max1−3	2.6	3.6	4.8	4.6	30.5	10.8	5.4	10.6	39.3	36.3	54.2	61.1
MAD4−7	0.3	0.4	0.6	0.6	5.8	1.3	1.0	5.2	8.6	3.4	15.9	14.5
Max4−7	0.9	1.4	2.2	2.0	12.9	4.6	3.2	11.9	23.1	8.1	37.1	33.2
MAD′ [b]								12.7	6.9	5.2	14.5	15.7
Max′ [b]								32.7	14.9	12.7	24.7	30.6

[a] Mean absolute (MAD) and maximum differences (Max) compared to those obtained with the LoProp polarizabilities are listed, calculated either over all seven ligands or over the charged (1−3) or neutral ligands (4−7). [b] Deviations from the xAvPol4 results (only for Btn1 and Btn7).

without any solvation. Only one of the biotin ligands in the tetramer (the fourth) was considered, whereas the other three were considered as a part of the protein. The results in Table 7 show that the Aver polarizabilities give induction contributions to the binding energies that are most similar to those obtained with the LoProp polarizabilities, with a MAD of 1 kJ/mol and a maximum error of 3 kJ/mol for the three charged ligands (Btn1−Btn3) and a MAD of 0.3 kJ/mol and a maximum error of 0.9 kJ/mol for the neutral ligands, respectively. The xAvPol1 polarizabilities also give excellent results with only slightly higher deviations. If the xAvPol2, xAvPol3, or even the atom-type polarizabilities are instead used, the MADs increase to 2 and 1 kJ/mol, respectively, and the maximum errors increase to 5 and 2−3 kJ/mol. On the other hand, the elemental polarizabilities give much worse results, with a MAD of up to 8 kJ/mol for the charged ligands (but only 1 kJ/mol for the neutral ligands). Recalculating the polarizabilities with a larger basis set (xAvPol4) has a major effect on the interaction energies, with MADs of 19 and 6 kJ/mol, respectively, again indicating that 6-31G* is a too small basis set for polarizabilities.

Among the various standard force fields, Amber02 polarizabilities give results that are closest to the LoProp results, with MADs of 4−5 kJ/mol and maximum errors of 11−12 kJ/mol. The other force fields give larger differences, e.g., MADs of 25−41 kJ/mol for the charged ligands and 3−16 kJ/mol for the neutral ligands. If we instead compare to the xAvPol4 results (available only for Btn1 and Btn7), the results for all force fields are improved (to 2−8 kJ/mol average deviation for Btn7 and 8−17 kJ/mol for Btn1), except for Amber02. This indicates that the Amber02 polarizabilities are not compatible with high-level QM calculations, presumably because the force field employs artificially restrictive exclusion rules, as discussed in the Methods section.

Previously, we have observed that effects of variations of the charges are strongly screened by solvation.[58] Therefore, we studied the effect of solvation also for the polarizabilities. Unfortunately, neither of the continuum-solvation models available in Amber is compatible with a polarizable force field. Therefore, we instead simply included all explicit solvent molecules in the calculation of the energy terms for the complex and the free protein. Of course, this is not a fully consistent method, but it at least gives an indication of how much solvation may screen the effect of differences in the polarizabilities. The results in Table 8 show that solvation has a small effect on the induction-energy part of the ligand-binding energies. In particular, no clear screening by solvation is seen. In fact, if different solvation models are used in the calculations (i.e., polarizabilities for the explicit water molecules that are consistent with the respective force field), the differences are typically increased, whereas if the same (LoProp) water polarizabilities are used in all calculations, the results are similar to those obtained without solvation.

■ CONCLUSIONS

In this paper, we have made a statistical and energetic analysis of isotropic atom-centered polarizabilities calculated individually for all atoms in two different proteins and for 70 snapshots from molecular dynamics simulations (in total 560 698 individual polarizabilities). As mentioned in the Introduction, atomic polarizabilities are not observables, so there are no true reference values of these. It is also well-known that polarizabilities strongly depend on the method and basis sets used for their calculation and that polarizabilities in the condensed phase are different from those in the gas phase.[27,60−62] Moreover, the polarizabilities are closely connected to the model used for the permanent electrostatics and exclusion rules used in the force field.[11] Therefore, it is not meaningful to discuss whether one set of polarizabilities is better than another without defining all the other components of the force field. Instead, this article is concerned with more general aspects of the polarizabilities, viz., their variation with conformation and chemical environment, and how polarizabilities are best assigned (by element, by atom type, or by individual atoms).

First, we show that dynamic effects induce a variation in the polarizabilities of individual atoms of 0.01−0.35 Å^3, with an

Table 8. Differences in Ligand-Interaction Polarization Energies, Compared to LoProp, with Explicit Solvent (kJ/mol)[a]

	Aver	xAvPol1	xAvPol2	xAvPol3	xAvPol4	Element	Type	Amber02	Amber09	Charmm	Amoeba	Enzymix
					Force Field Specific Water Polarizabilities							
MAD	0.6	0.9	1.0	0.9	23.9	4.2	1.5	5.3	38.5	17.8	23.1	47.9
Max	3.2	4.2	4.9	4.8	54.4	19.6	4.8	13.8	171.4	47.9	94.3	245.5
MAD1−3	1.0	1.5	1.5	1.5	41.2	7.0	1.9	7.4	59.2	36.6	35.7	77.6
Max1−3	3.2	4.2	4.9	4.8	54.4	19.6	4.8	13.8	171.4	47.9	94.3	245.5
MAD4−7	0.3	0.4	0.6	0.5	6.7	2.0	1.1	3.8	23.0	3.3	13.7	25.6
Max4−7	0.9	1.3	1.8	1.7	13.5	5.3	3.3	11.1	47.3	9.0	27.7	63.0
					LoProp Water Polarizabilities							
MAD	0.6	0.9	1.0	0.9	15.7	3.6	1.5	4.8	14.5	27.6	22.0	25.7
Max	3.2	4.2	4.9	4.8	32.8	8.3	4.8	13.5	36.6	82.5	55.3	60.3
MAD1−3	1.0	1.5	1.5	1.5	25.4	6.2	1.9	5.6	22.2	58.7	31.5	40.6
Max1−3	3.2	4.2	4.9	4.8	32.8	8.3	4.8	13.5	36.6	82.5	55.3	60.3
MAD4−7	0.3	0.4	0.6	0.5	6.0	1.6	1.1	4.2	8.7	3.6	14.9	14.6
Max4−7	0.9	1.3	1.8	1.7	12.8	4.8	3.3	11.4	23.2	9.0	36.2	33.1

[a] Mean absolute (MAD) and maximum differences (Max) compared to those obtained with the LoProp polarizabilities are listed, calculated either over all seven ligands or over the charged (1−3) or neutral ligands (4−7).

average of 0.09 Å^3 for the 7827 atoms in the avidin tetramer. The standard deviation ranges from 0.002 to 0.07 Å^3 (average 0.02 Å^3), indicating that up to 50 snapshots are needed to obtain a standard error of less than 0.01 Å^3 for all polarizabilities. This clearly shows that it is not enough to calculate polarizabilities for a single structure.

Second, we show that it is very hard to assign transferable polarizabilities by element or atom types. Elementwise polarizabilities would have an uncertainty of up to 0.77 Å^3, i.e., 50% of the magnitude of the polarizabilities themselves. This would induce errors of up to 36 kJ/mol in relative conformational induction energies and of up to 11 kJ/mol in ligand-binding energies. Likewise, polarizabilities assigned by the 27 Amber protein atom types would still have an uncertainty of up to 0.77 Å^3, and it would induce errors of up to 17 kJ/mol in relative energies and of up to 5 kJ/mol for ligand-binding energies (7 and 2 kJ/mol on average). We have also tried to design better groups of atom types, but this is very hard, in particular for aliphatic and aromatic carbon atoms, for which the range is up to 0.36 Å^3.

Therefore, we suggest that polarizabilities should be assigned the same way as for charges, i.e., atomwise. This suppressed the variation of the polarizabilities to 0.14 Å^3 on average, with a maximum of 0.46 Å^3. The average and maximum standard deviations are 0.01 and 0.07 Å^3. This remaining variation reflects the conformational dependence of the polarizabilities, and it cannot be further suppressed unless the conformational dependence is explicitly modeled. The variation is related to the size of the polarizabilities, with an average of 23%. The conformational dependence induces average and maximum errors of 2 and 5 kJ/mol for relative conformational energies, and of 1 and 4 kJ/mol for ligand-binding energies. Polarizabilities calculated in the same way for a different protein (the photosynthetic reaction center) give similar results: 2 and 9 kJ/mol average and maximum error for relative conformational energies and 1 and 5 kJ/mol for ligand-binding energies.

On the other hand, the polarizabilities strongly depend on the basis sets used in the QM calculations. Clearly, the 6-31G* basis set is too small to give converged polarizabilities. Instead, at least the aug-cc-pVDZ (and preferably, the aug-cc-pVTZ) basis set should be used in the calculations. Fortunately, the atomic correction factors between the 6-31G* and aug-cc-pVTZ basis sets are transferable, so the results can be easily extrapolated from bulk calculations with the 6-31G* basis set. In the Supporting Information, we present a set of such polarizabilities (xAvPol4), averaged over 70 molecular dynamics snapshots for avidin and over two different proteins, and finally extrapolated to the aug-cc-pVTZ basis set. These are the best atomic polarizabilities obtained in this paper.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information. Table S1 showing the four sets of xAvPol polarizabilities (Å^3) for the various atoms. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: Ulf.Ryde@teokem.lu.se. Telephone: +46-46 2224502. Fax: +46-46 2228648.

■ ACKNOWLEDGMENT

This investigation has been supported by grants from the Swedish Research Council (Project 2010-5025) and by computer resources of Lunarc at Lund University. J.K. thanks the Danish Natural Science Research Council/The Danish Councils for Independent Research and the Lundbeck Foundation for financial support.

■ REFERENCES

(1) Ponder, J. W.; Case, D. A. *Adv. Protein Chem.* **2003**, *66*, 27–85.
(2) Bachrach, S. M. *Rev. Comput. Chem.* **1994**, *5*, 171–227.
(3) Sigfridsson, E.; Ryde, U. *J. Comput. Chem.* **1998**, *19*, 377–395.
(4) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269–10280.
(5) Brooks, B. R.; Brooks, C. L.; MacKerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archonitis, G.; Bartels, C.; Boersch, S.;

Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazardis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Veneable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. *J. Comput. Chem.* **2009**, *30*, 1545–1614.

(6) Gao, J.; Xia, X. *Science* **1992**, *258*, 631–635.

(7) Gao, J.; Habibollazadeh, D.; Shao, L. *J. Phys. Chem.* **1995**, *99*, 16460–16467.

(8) Gao, J.; Pavelites, J. J.; Habibollazadeh, D. *J. Phys. Chem.* **1996**, *100*, 2689–2697.

(9) Xie, W.; Pu, J.; MacKerell, A. D.; Gao, J. *J. Chem. Theory Comput.* **2007**, *3*, 1878–1889.

(10) Söderhjelm., P.; Ryde, U. *J. Phys. Chem. A* **2009**, *113*, 617–627.

(11) Cieplak, P.; Dupradeau, F.-Y.; Duan, Y.; Wang, J. *J. Phys.: Condens. Matter* **2009**, *21*, 333102.

(12) Halgren, R. A.; Damm, W. *Curr. Opin. Struct. Biol.* **2001**, *11*, 236–242.

(13) Gresh, N.; Cisneros, G. A.; Darden, T. D.; Piquemal, J.-P. *J. Chem. Theory Comput.* **2007**, *3*, 1960–1986.

(14) Warshel, A.; Kato, M.; Pisliakov, A. V. *J. Chem. Theory Comput.* **2007**, *3*, 2034–2045.

(15) Lopes, P. E. M.; Roux, B.; MacKerell, A. D. *Theor. Chem. Acc.* **2009**, *124*, 11–28.

(16) Rappé, A. K.; Goddard, W. A., III. *J. Phys. Chem.* **1991**, *95*, 3358–3363.

(17) Rick, S. W.; Stuart, S. J.; Berne, B. J. *J. Chem. Phys.* **1994**, *101*, 6141–6156.

(18) Applequist, J.; Carl, J. R.; Fung, K.-K. *J. Am. Chem. Soc.* **1972**, *94*, 2952–2960.

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

(20) Thole, B. T. *Chem. Phys.* **1981**, *59*, 341–350.

(21) Dick, B. G.; Overhauser, A. W. *Phys. Rev.* **1958**, *112*, 90–103.

(22) Cao, J.; Berne, B. J. *J. Chem. Phys.* **1993**, *99*, 6998–7011.

(23) Lamoureux, G.; Roux, B. *J. Chem. Phys.* **2003**, *119*, 3025–3039.

(24) Day, P. N.; Jensen, J. H.; Gordon, M. S.; Webb, S. P.; Stevens, W. J.; Kraus, M.; Garmer, D.; Basch, H.; Cohen, D. *J. Chem. Phys.* **1996**, *105*, 1968–1986.

(25) Engkvist, O.; Åstrand, P.-O.; Karlström, G. *Chem. Rev.* **2000**, *100*, 4087–4108.

(26) Cieplak, P.; Caldwell, J.; Kollman, P. *J. Comput. Chem.* **2001**, *22*, 1048–1057.

(27) Maple, J. R.; Cao, Y.; Damm, W.; Halgren, T. A.; Kaminski, G. A.; Zhang, L. Y.; Friesner, R. A. *J. Chem. Theory Comput.* **2005**, *1*, 694–715.

(28) Ren, P.; Ponder, J. W. *J. Comput. Chem.* **2002**, *23*, 1497–1506.

(29) Ángyán, J. G.; Jansen, G.; Loos, M.; Hättig, C.; Hess, B. A. *Chem. Phys. Lett.* **1994**, *219*, 267–273.

(30) Stone, A. J. *Mol. Phys.* **1985**, *56*, 1065–1082.

(31) Le Sueur, C. R.; Stone, A. J. *Mol. Phys.* **1994**, *83*, 293–307.

(32) Dehez, F.; Chipot, C.; Millot, C.; Ángyán, J. G. *Chem. Phys. Lett.* **2001**, *338*, 180–188.

(33) Williams, G. J.; Stone, A. J. *J. Chem. Phys.* **2003**, *119*, 4620–4628.

(34) Gagliardi, L.; Lindh, R.; Karlström, G. *J. Chem. Phys.* **2004**, *121*, 4494–4500.

(35) Stout, J. M.; Dykstra, C. E. *J. Am. Chem. Soc.* **1995**, *117*, 5127–5132.

(36) Zhou, T.; Dykstra, C. E. *J. Phys. Chem. A* **2000**, *104*, 2204–2210.

(37) Nakagawa, S.; Kosugi, N. P. *Chem. Phys. Lett.* **1993**, *210*, 180–186.

(38) Alkorta, I.; Bachs, M.; Perez, J. J. *Chem. Phys. Lett.* **1994**, *224*, 160–165.

(39) Celebi, N.; Ángyán, J. G.; Dehez, F.; Millot, C.; Chipot, C. *J. Chem. Phys.* **2000**, *112*, 2709–2717.

(40) Dehez, F.; Soetens, J. C.; Chipot, C.; Ángyán, J. G.; Millot, C. *J. Phys. Chem. A* **2000**, *104*, 1293–1303.

(41) Soteras, I.; Curutchet, C.; Bidon-Chanal, A.; Dehez, F.; Ángyán, J. G.; Orozco, M.; Chipot, C.; Luque, F. J. *J. Chem. Theory Comput.* **2007**, 3, 1901–1913.

(42) Elking, D.; Darden, T.; Woods, R. J. *J. Comput. Chem.* **2007**, *28*, 1261–1274.

(43) Vogel, A. I. *J. Chem. Soc.* **1948**, 1833–1847.

(44) Miller, K. J. *J. Am. Chem. Soc.* **1990**, *112*, 8533–8542.

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

(46) van Duijnen, P. T.; Swart, M. *J. Phys. Chem. A* **1998**, *102*, 2399–2407.

(47) Söderhjelm, P; Husberg, C.; Strambi, A.; Olivucci, M.; Ryde, U. *J. Chem. Theory Comput.* **2009**, *5*, 649–658.

(48) Söderhjelm, P.; Kongsted, J.; Ryde, U. *J. Chem. Theory Comput.* **2010**, *6*, 1726–1737.

(49) Karlström, G.; Lindh, R.; Malmqvist, P.-Å.; Roos, B. O.; Ryde, U.; Veryazov, V.; Widmark, P.-O.; Cossi, M.; Schimmelpfennig, B.; Neogrady, P.; Seijo, L. *Comput. Mater. Sci.* **2003**, *28*, 222–239.

(50) Söderhjelm, P.; Krogh, J. W.; Karlström, G.; Ryde, U.; Lindh, R. *J. Comput. Chem.* **2007**, *28*, 1083–1090.

(51) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372–1377.

(52) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213–222.

(53) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(54) Sadlej, A. J. *Collect. Czech. Chem. Commun.* **1988**, *53*, 1995–2016.

(55) Zhang, D. W.; Zhang, J. Z. H. *J. Chem. Phys.* **2003**, *119*, 3599–3605.

(56) Weis, A.; Katebzadeh, K.; Söderhjelm, P.; Nilsson, I.; Ryde, U. *J. Med. Chem.* **2006**, *49*, 6596–6606.

(57) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, M.; Walker, R. C.; Zhang, W.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossvary, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman, P. A. *Amber10*; University of California: San Francisco, 2008.

(58) Söderhjelm, P.; Ryde, U. *J. Comput. Chem.* **2009**, *30*, 750–760.

(59) Genheden, S.; Söderhjelm, P.; Ryde, U. *Int. J. Quantum Chem.* 2011, DOI: 10.1002/qua.22967.

(60) Giese, T. J.; York, D. M. *J. Chem. Phys.* **2004**, *120*, 9903–9906.

(61) Morita, A. *J. Comput. Chem.* **2002**, *23*, 1466–1471.

(62) Anisimov, V. M.; Lamoureux, G.; Vorobyov, I. V.; Huang, N.; Roux, B.; MacKerell, A. D. *J. Chem. Theory Comput.* **2005**, *1*, 153–168.

(63) Kaminski, G. A.; Stern, H. A.; Berne, B. J.; Friesner, R. A. *J. Phys. Chem. A* **2004**, *108*, 621–627.

JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Polarizable Simulations with Second-Order Interaction Model (POSSIM) Force Field: Developing Parameters for Alanine Peptides and Protein Backbone

Sergei Y. Ponomarev and George A. Kaminski*

Department of Chemistry and Biochemistry, Worcester Polytechnic Institute, Worcester, Massachusetts 01609, United States

S Supporting Information

ABSTRACT: Polarizable simulations with second-order interaction model (POSSIM) force field has been extended to include parameters for alanine peptides and protein backbones. New features were introduced into the fitting protocol, as compared to the previous generation of the polarizable force field for proteins. A reduced amount of quantum mechanical data was employed in fitting the electrostatic parameters. Transferability of the electrostatics between our recently developed *N*-methylacetamide model and the protein backbone was confirmed. Binding energy and geometry for complexes of alanine dipeptide with a water molecule were estimated and found in a good agreement with high-level quantum mechanical results (for example, the intermolecular distances agreeing within ca. 0.06 Å). Following the previously devised procedure, we calculated average errors in alanine di- and tetrapeptide conformational energies and backbone angles and found the agreement to be adequate (for example, the alanine tetrapeptide extended globular conformational energy gap was calculated to be 3.09 kcal/mol quantum mechanically and 3.14 kcal/mol with the POSSIM force field). However, we have now also included simulation of a simple α helix in both gas phase and water as the ultimate test of the backbone conformational behavior. The resulting alanine and protein backbone force field parameters are currently being employed in further development of the POSSIM fast polarizable force field for proteins.

I. INTRODUCTION

While quantum mechanical calculations offer valuable data in a variety of biological and biomedical calculations, applications of empirical force fields remain the only way of approaching the majority of problems of interest. On one hand, they require less computer resources. On the other hand, the issue of choosing the best level of quantum theory is still a nontrivial one, and the level of quantum mechanical accuracy in a specific application is far from being guaranteed.

When empirical force fields are employed, accurate assessment of energy often requires explicit treatment of the electrostatic polarization.[1] The properties which depend on it include dimerization energies and acidity constants of small molecules, energies of protein−ligand interactions, protein pK_a values, or even the very thermodynamic stability of complexes in solutions. For example, we have demonstrated that that pK_a values for acidic and basic residues of the turkey ovomucoid third domain (OMTKY3) can be reproduced within 0.6 and 0.7 pH units of the experimental data with a polarizable force field. The corresponding errors with the nonpolarizable orthogonal partial least-squares (OPLS) were 3.3 and 2.2 pH units.[2] Formation of sugar−protein complexes represents yet another example when polarization is critical for predicting a thermodynamically stable structure.[3] It is generally acknowledged that polarization is an important component in many computational studies of proteins and protein−ligand complexes, although it is sometimes included in surrogate forms, such as, for example, conformation-specific protein charges.[4]

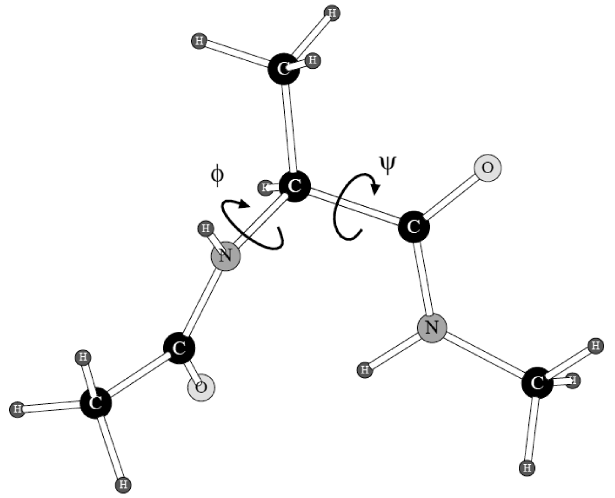


Figure 1. Protein backbone angles ϕ and ψ shown in the alanine dipeptide molecule.

There are two main issues related to the empirical polarizable force field development. The first one is in the functional form of the electrostatic polarization. Using fluctuating charges saves time and is computationally efficient in simulating uniform systems, such as pure liquid water.[5] However, it causes problems when out-of-plane polarization response is required or when a bifurcated hydrogen bond is formed. Therefore, the inducible dipoles approach is more adequate

Received: December 14, 2010
Published: April 14, 2011

when arbitrary systems have to be simulated with a high degree of accuracy. On the other hand, it is known that the inducible dipole technique slows down polariable calculations significantly. In order to reduce the severity of this problem, we are applying the second-order approximation in treatment of the electrostatic polarization. It has been demonstrated to increase the speed by ca. an order of magnitude without sacrificing the accuracy.[6] Moreover, this approximation makes the so-called polarization catastrophe (the resonance-like infinite growth of the induced dipole moment values) impossible. Our previous paper described development of the polarizable simulations with second order interaction model (POSSIM) software and the force field parameters for a series of small molecules, including water and *N*-methylacetamide (NMA). In this work, we describe creation of alanine and protein backbone parameter sets in the POSSIM framework.

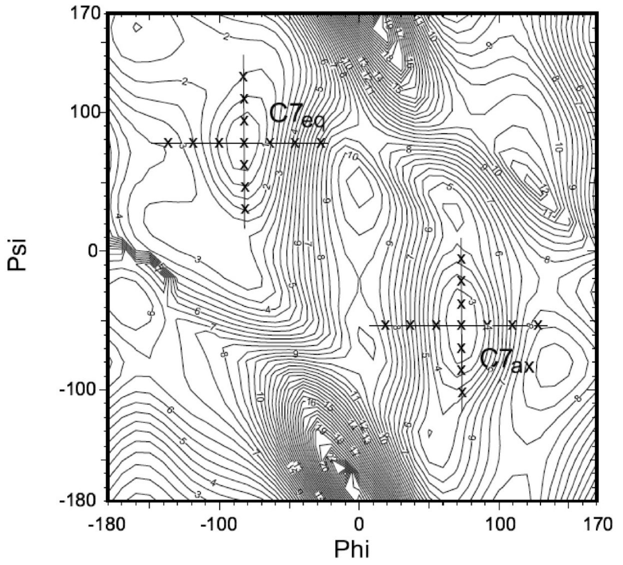


Figure 2. Torsional fitting subspace, for the alanine dipeptide ϕ/ψ potential energy surface. Such crosses were centered at each of the six minima, and each arm contained four fitting points (here some crosses and points are omitted for the sake of clarity).

Table 1. Backbone Torsional Parameters, Set tors.1[a]

parameter	V_1	V_2	V_3
C−N−αC−C, ϕ	0.667	−0.012	−4.003
N−αC−C−N, ψ	−2.011	2.528	−4.829
C−N−αC−βC, ϕ'	−2.165	0.024	4.221
βC−αC−C−N, ψ'	0.594	−0.386	4.378

[a] The coefficients are given in kcal/mol.

The second issue is choosing the source of fitting data for a polarizable force field. High-level quantum mechanical results are very attractive in this respect,[7,8] but experimental data can be more robust. We follow the middle-of-the-road path by relying on experimental data whenever possible and by making heavy use of quantum mechanical calculations when needed. One important issue is the standard procedure of producing torsional parameters for peptides by fitting to conformational energies of di- and tetrapeptides.[8] We include it in our work and are describing an improved procedure for creation of the torsional parameters in the Methods Section below. At the same time, the quantum mechanical conformers employed in such calculations are created by gas-phase quantum mechanical optimizations, and they often belong to parts of the conformational space which are not found in experimental protein structures. Therefore, we have included an additional conformational test in the alanine and the backbone parameter fitting. It is known that the tridecaalanine peptide (ala-13) forms a stable α-helix.[9] Therefore, we studied the stability of our POSSIM ala-13 α-helix and compared it to that of the OPLS-AA[8] for benchmarking. We have also discovered that the quality of the force field in reproducing the quantum mechanical di- and tetrapeptide conformational energies has a relatively weak effect on the stability of the tridecaalanine peptide in water.

Overall, the following has been derived, developed, or otherwise calculated in this work: (i) the torsional parameters for the alanine residues and the protein backbones have been produced; (ii) the binding energies of a water molecule with the alanine dipeptide as calculated with the POSSIM and OPLS-AA force fields have been compared with the quantum mechanical data to confirm transferability of the nonbonded parameters and to justify using the latter from the POSSIM NMA model in protein studies; (iii) the resulting parameters were employed in gas-phase and aqueous solution simulations of an α-helix to validate the resulting POSSIM parameters as

Table 2. Conformational Energies and Angles for Alanine Dipeptide[a]

	energy			ϕ			ψ		
conformer	QM	OPLS	POSSIM	QM	OPLS	POSSIM	QM	OPLS	POSSIM
$C7_{eq}$	0.00	0.00	0.00	−81.4	−79.5	−83.8	85.6	61.8	53.2
C5	1.00	0.91	0.78	−160.5	−149.8	−151.3	165.9	159.9	150.9
$C7_{az}$	2.71	2.40	2.85	70.3	77.5	76.5	−76.8	−46.6	−50.3
β_2	2.56	2.82	2.57	−105.1	−105.1	−105.1	10.6	10.6	10.6
α_L	4.21	5.96	5.41	68.3	68.3	68.3	22.4	22.4	22.4
α′	5.47	5.96	5.53	−162.0	−156.5	−149.5	−33.2	−48.5	−100.3
PII	2.78	2.18	3.96	−85.0	−85.0	−85.0	160.0	160	160.0
α_R	2.71	2.39	1.95	−83.7	−83.7	−83.7	−3.9	−-3.9	−3.9
error	–	0.73	0.67	–	3.2	3.8	–	9.4	17.6

[a] Energies are in kcal/mol, and angles are in degrees. POSSIM refers to the polarizable force field with the tors.1 version of the torsional parameters. Angles ϕ and ψ for conformers β_2, α_L, PII, and α_R were fixed at their quantum mechanical values. Quantum mechanical energy minimizations were unconstrained except for PII.

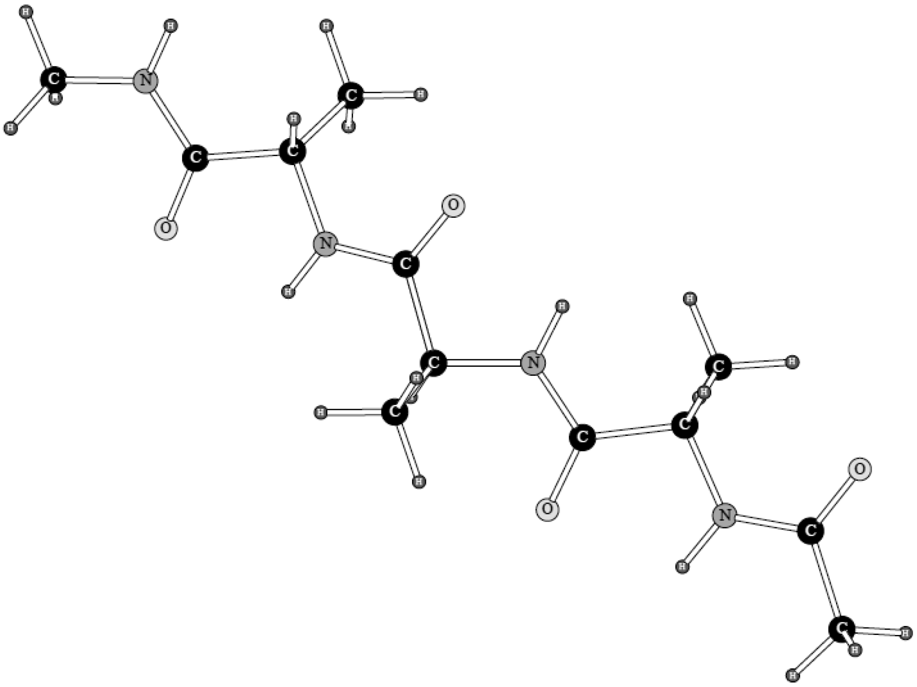


Figure 3. LMP2/cc-pVTZ(-f) geometry of the extended alanine tetrapeptide conformation.

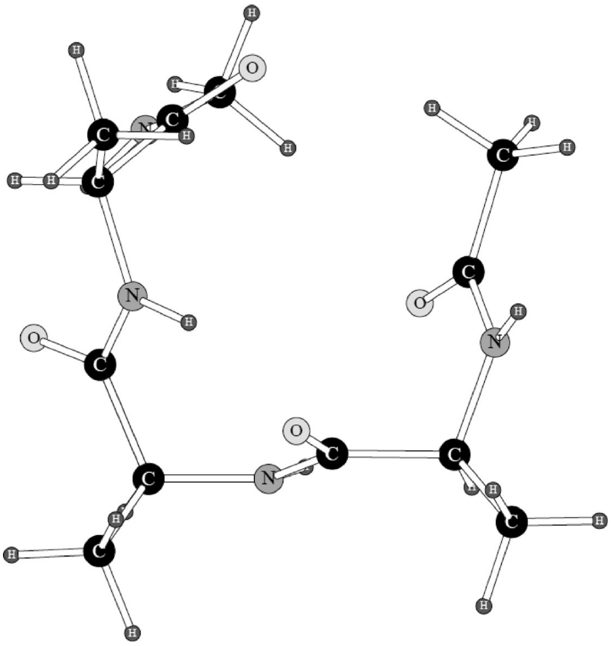


Figure 4. LMP2/cc-pVTZ(-f) geometry of the globular alanine tetrapeptide conformation.

acceptable in protein and peptide simulations. Moreover, the additional optimizations in (i) above and (ii) and (iii) altogether represent a novel development in our methodology of protein force field production, as compared to that used to create the previous version of the force field.

The rest of the paper is organized as follows: Section II is a description of the methodology involved. Section III contains results and discussion. Finally, Section IV presents the conclusions.

Table 3. Backbone Torsional Parameters, Set tors.final[a]

parameter	V_1	V_2	V_3
C−N−αC−C, ϕ	2.000	−0.500	−3.772
N−αC−C−N, ψ	−2.837	3.942	−3.328
C−N−αC−βC, ϕ'	−2.718	1.757	5.202
βC−αC−C−N, ψ'	0.372	−0.915	3.321

[a] The coefficients are given in kcal/mol.

II. METHODS

A. Force Field. The total energy E_{tot} is a sum of the electrostatic interactions $E_{\text{electrostatic}}$, van der Waals energy E_{vdW}, harmonic bond stretching and angle bending E_{stretch} and E_{bend}, and the torsional term E_{torsion}:

$$E_{\text{tot}} = E_{\text{electrostatic}} + E_{\text{vdW}} + E_{\text{stretch}} + E_{\text{bend}} + E_{\text{torsion}} \quad (1)$$

Electrostatic Energy. The electrostatic polarization energy as calculated with inducible point dipoles $\boldsymbol{\mu}$ is

$$E_{\text{pol}} = -\frac{1}{2}\sum_i \mu_i \mathbf{E}_i^0 \quad (2)$$

where $\mathbf{E}^0$ is the electrostatic field in the absence of the dipoles.

$$\boldsymbol{\mu}_i = \alpha_i \mathbf{E}_i^0 + \alpha_i \sum_{j \neq i} \mathbf{T}_{ij}\boldsymbol{\mu}_j \quad (3)$$

where α are scalar polarizabilities, and $\mathbf{T}_{ij}$ is the dipole−dipole interaction tensor. The self-consistent eq 3 is usually solved iteratively. Let us explicitly write down the first two iterations:

$$\boldsymbol{\mu}_i^{\text{I}} = \alpha_i \mathbf{E}_i^0 \quad (4a)$$

Table 4. Conformational Energies and Angles for Alanine Dipeptide[a]

	energy			ϕ			ψ		
conformer	QM	OPLS	POSSIM	QM	OPLS	POSSIM	QM	OPLS	POSSIM
C7$_{eq}$	0.00	0.00	0.00	−81.4	−79.5	−77.2	85.6	61.8	34.4
C5	1.00	0.91	1.37	−160.5	−149.8	−160.3	165.9	159.9	159.2
C7$_{az}$	2.71	2.40	2.17	70.3	77.5	78.1	−76.8	−46.6	−36.2
β_2	2.56	2.82	2.77	−105.1	−105.1	−105.1	10.6	10.6	10.6
α$_L$	4.21	5.96	5.79	68.3	68.3	68.3	22.4	22.4	22.4
α′	5.47	5.96	5.98	−162.0	−156.5	−162.9	−33.2	−48.5	−38.0
PII	2.78	2.18	3.52	−85.0	−85.0	−85.0	160.0	160	160.0
α$_R$	2.71	2.39	0.99	−83.7	−83.7	−83.7	−3.9	−-3.9	−3.9
error	–	0.73	0.97	–	3.2	1.6	–	9.4	12.9

[a] Energies are in kcal/mol, and angles are in degrees. POSSIM refers to the polarizable force field with the tors.final version of the torsional parameters. Angles ϕ and ψ for conformers β_2, α$_L$, PII, α$_R$ were fixed at their QM values. QM energy minimizations were unconstrained except for PII.

$$\boldsymbol{\mu}_i^{\mathrm{II}} = \alpha_i \mathbf{E}_i^0 + \alpha_i \sum_{j \neq i} \mathbf{T}_{ij} \boldsymbol{\mu}_j^{\mathrm{I}} = \alpha_i \mathbf{E}_i^0 + \alpha_i \sum_{j \neq i} \mathbf{T}_{ij} \alpha_j \mathbf{E}_j^0 \qquad (4b)$$

We are using the second-order expression in eq 4b. It has been previously shown to yield ca. an order of magnitude increase of the computational speed with no loss of accuracy.[6] The electrostatic energy also includes the pairwise additive contribution from interactions of permanent charges:

$$E_{\mathrm{additive}} = \sum_{i \neq j} \frac{q_i q_j}{R_{ij}} f_{ij} \qquad (5)$$

The factor f_{ij} is set to 0 for 1,2- and 1,3-pairs (atoms which belong to the same valence bond or angle), to 0.5 for 1,4-interactions (atoms in the same dihedral angle), and to 1.0 otherwise.

To avoid unphysical increase of the electrostatic interactions at short distances, each atom type has a cutoff parameter R_{cut}. When the overall distance R_{ij} is smaller than the sum of these parameters $R_{\mathrm{min}}^{ij} = R_{\mathrm{cut}}^{i} + R_{\mathrm{cut}}^{j}$ for the atoms i and j, R_{ij} is replaced by an effective smooth function:

$$R_{ij}^{\mathrm{eff}} = \left(1 - \left(\frac{R_{ij}}{R_{\mathrm{min}}^{ij}}\right)^2 + \left(\frac{R_{ij}}{R_{\mathrm{min}}^{ij}}\right)^3\right) \cdot R_{\mathrm{min}}^{ij} \qquad (6)$$

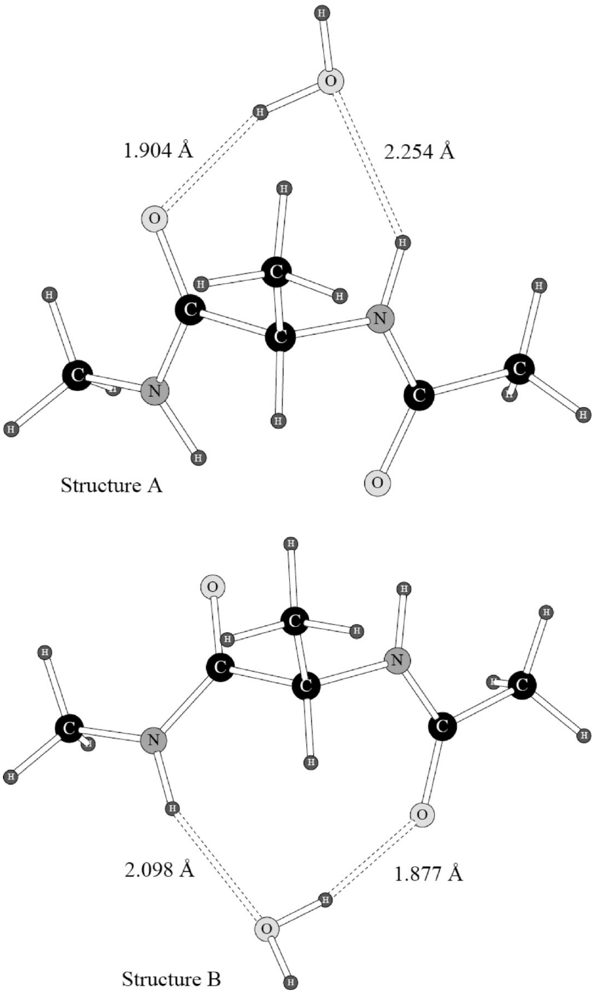


Figure 5. Two Alanine dipeptide hydrogen bonded complexes with a water molecule.

The following important points about the second-order approximation in eq 4b should be made: First of all, we do not fit parameters using the full-scale polarization solution to eq 3 to later employ eq 4b as an approximate technique during the simulations. For our practical purposes, eq 4b is, in fact, the representation of the many-body interactions. It does differ from the true physical point−dipole approximation, and thus we always carefully monitor whether any errors are introduced by not computing inducible dipoles with the complete iterative procedure. So far, simulations of gas-phase dimers, quantum mechanical electrostatic three-body energies, pure liquids, solutions and peptides have given us no indication that the second-order approximation leads to any deficient physical results, and we have always been able to produce fitting to quantum mechanical and experimental data which was as good as for the full-scale polarization.[6,10] Moreover, application of the second-order approximation given by eq 4b turns the expression for the inducible dipoles into an analytical one, thus eliminating the possibility of the polarization catastrophe. This can also become a very useful feature in future developments, e.g., in creating a continuum dielectric model, as convergence issues are known to be of importance for continuum solvation techniques.

Table 5. Results of Simulating Alanine Dipeptide Complexes with Water [a]

	structure A			structure B		
property	QM	OPLS	POSSIM	QM	OPLS	POSSIM
binding energy	−9.80	−9.48	−7.95	−9.73	−11.19	−9.34
R(O···O)	2.83	2.75	2.84	2.83	2.72	2.73
R(O···N)	3.10	2.89	3.09	3.05	2.86	2.95
ϕ, dimer	−83.4	−87.3	−80.2	−84.3	−89.4	−86.9
ψ, dimer	90.3	114.6	83.7	131.2	113.3	122.2
ϕ, monomer	−79.7	−79.5	−77.2	−79.7	−79.5	−77.2
ψ, monomer	88.1	61.8	34.4	88.1	61.8	34.4

[a] Energies are in kcal/mol, distances in Å, angles in degrees.

Table 6. Results of Simulating Alanine Dipeptide Complexes with Water[a]

	structure A			structure B		
property	QM	OPLS	POSSIM	QM	OPLS	POSSIM
binding energy	−10.71	−10.04	−9.75	−11.68	−11.79	−12.24
R(O···O)	2.83	2.81	2.82	2.83	2.75	2.74
R(O···N)	3.10	2.94	3.06	3.05	2.94	2.98
ϕ, dimer	−83.4	−83.4	−83.4	−84.3	−84.3	−84.3
ψ, dimer	90.3	90.3	90.3	131.2	131.2	131.2
ϕ, monomer	−83.4	−83.4	−83.4	−84.3	−84.3	−84.3
ψ, monomer	90.3	90.3	90.3	131.2	131.2	131.2

[a] ϕ and ψ of both dimers and monomers are fixed in the quantum mechanical dimer positions. Energies are in kcal/mol, distances in Å, angles in degrees.

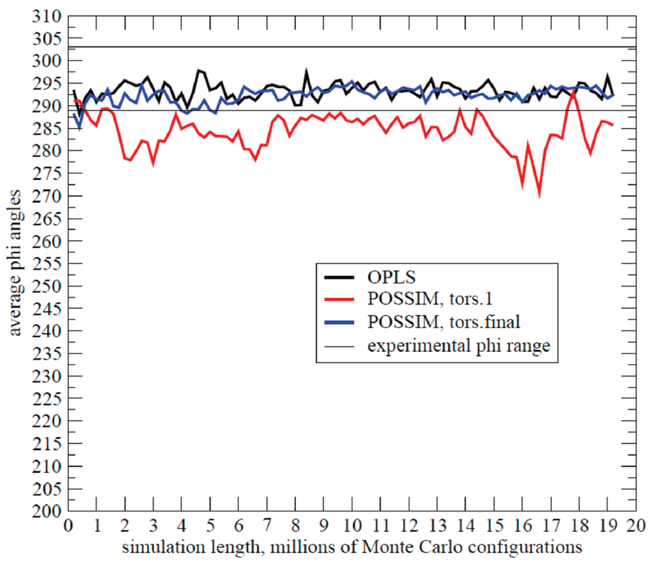


Figure 6. Average ϕ angles in the α-helix gas-phase simulations vs the simulation length.

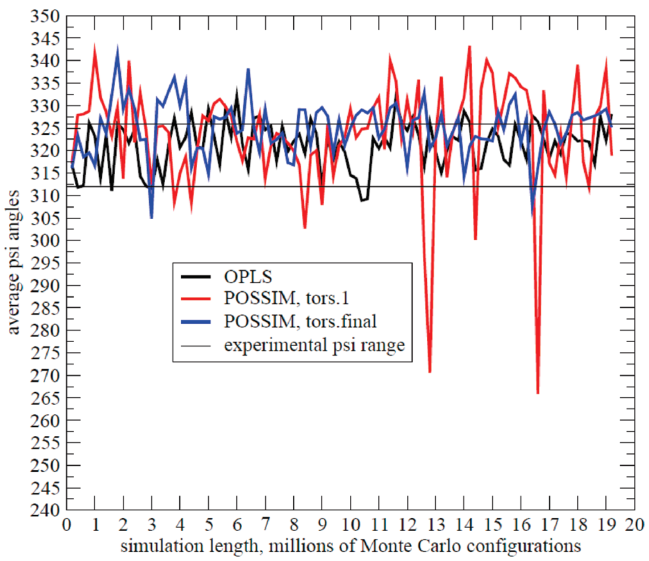


Figure 7. Average ψ angles in α-helix gas-phase simulations vs the simulation length.

The Rest of the Force Field. We are using the standard Lennard-Jones formalism for the van der Waals energy:

$$E_{\mathrm{vdW}} = \sum_{i \neq j} 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{R_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{R_{ij}}\right)^{6}\right] f_{ij} \qquad (7)$$

Geometric combining rules are applied ($\varepsilon_{ij} = (\varepsilon_i \cdot \varepsilon_j)^{1/2}$, $\sigma_{ij} = (\sigma_i \cdot \sigma_j)^{1/2}$). Bond stretching and angle bending are computed with the usual harmonic formalism, and the torsional term is calculated as

$$E_{\mathrm{torsion}} = \sum_i \frac{V_1^i}{2}[1 + \cos(\varphi_i)] + \frac{V_2^i}{2}[1 - \cos(2\varphi_i)] + \frac{V_3^i}{2}[1 + \cos(3\varphi_i)] \qquad (8)$$

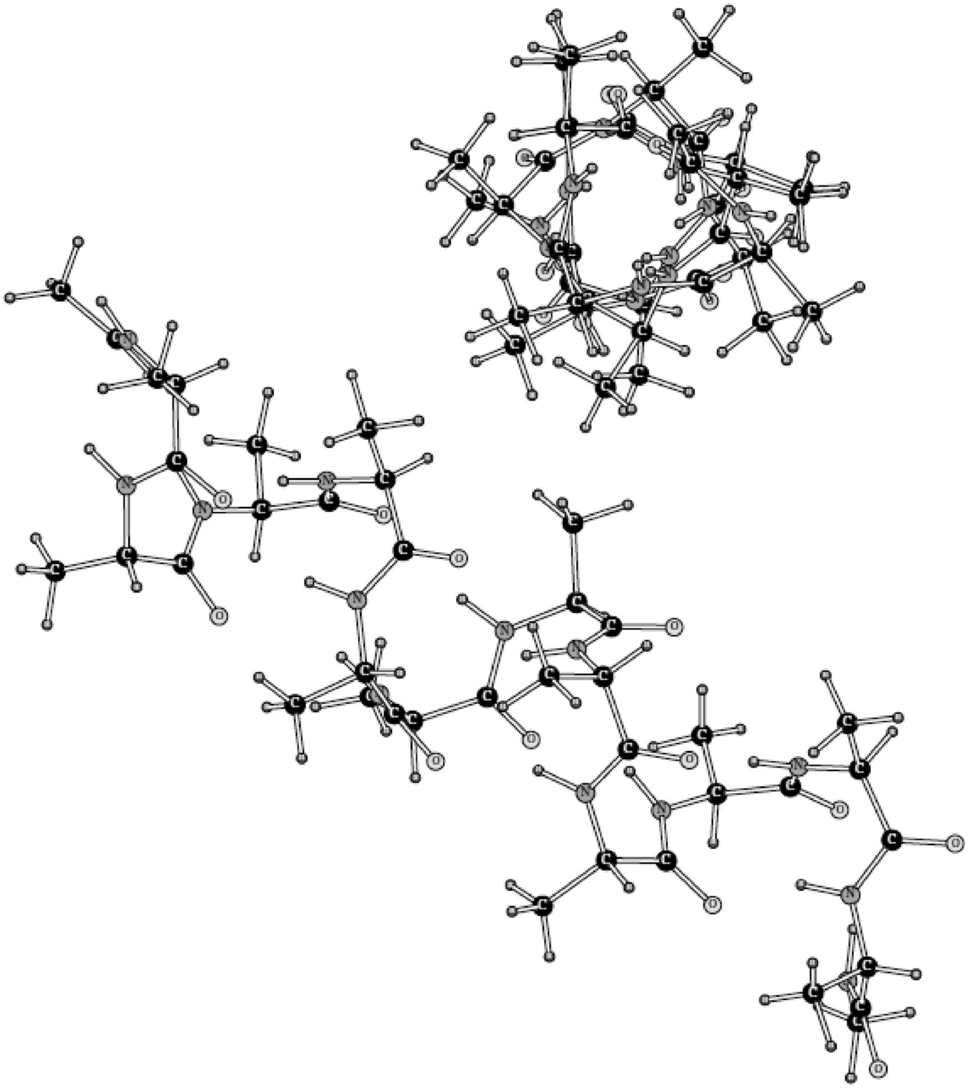


Figure 8. Structure of the ala-13 α-helix simulated with OPLS in gas phase, after 19 × 10^6 Monte Carlo configurations.

The fixed-charges OPLS-AA force field used for benchmarking is functionally the same, except that it lacks the polarization part of the electrostatic energy.

B. Parameterization of the Force Field. Whenever possible, the force field parameters for the alanine peptides were adopted directly from the previously created NMA parameter values.[10] The only completely new parameters were those for the backbone torsions. This is different from the previous version of the polarizable force field (PFF) for proteins in which electrostatic parameters for the alanine (and thus for the backbone) were also refitted.[7] Therefore, we believe that the present work demonstrates a greater degree of utilizing parameter transferability.

Fitting of torsional parameters for the protein backbone ϕ and ψ angles (Figure 1) cannot be done separately from each other, as the torsions are coupled.

The initial part of our torsional fitting was the same as used before.[7,8] (i) The fitting was done to an ab initio data obtained previously[8] at the LMP2/cc-pVTZ(-f)//HF-6-31G** level with Jaguar software suite.[11] (ii) The choice of the fitting subspace is illustrated in Figure 2. Out of the six alanine dipeptide local minima previously used,[7,8] only two are shown for the sake of clarity. (iii) We used the following non-Boltzmann weighting scheme for the error at the fitting points:

$$W_i = A \cdot \exp(-b \cdot G_i) \qquad (9)$$

Here G_i is the magnitude of the torsional surface gradient at the point i, and W_i is the weight. This way more importance was given to the points with low gradients (near the minima).

In the presented work, we used the procedure described above only to produce the initial guess for the torsional parameters to be employed in eq 8. After that, the following approach was taken. The errors in the conformational energies were combined with the errors in the conformational angles ϕ and ψ to produce the error function as shown in eq 10:

$$\text{erf} = \sum_i (E_i^0 - E_i)^2 + \sum_j (\varphi_j^0 - \varphi_j)^2 + (\psi_j^0 - \psi_j)^2 \qquad (10)$$

Here E_i^0 and E_i are the quantum mechanical and empirical conformational energies for all the conformers i, and the second sum contains the values of the backbone angles ϕ and ψ. The

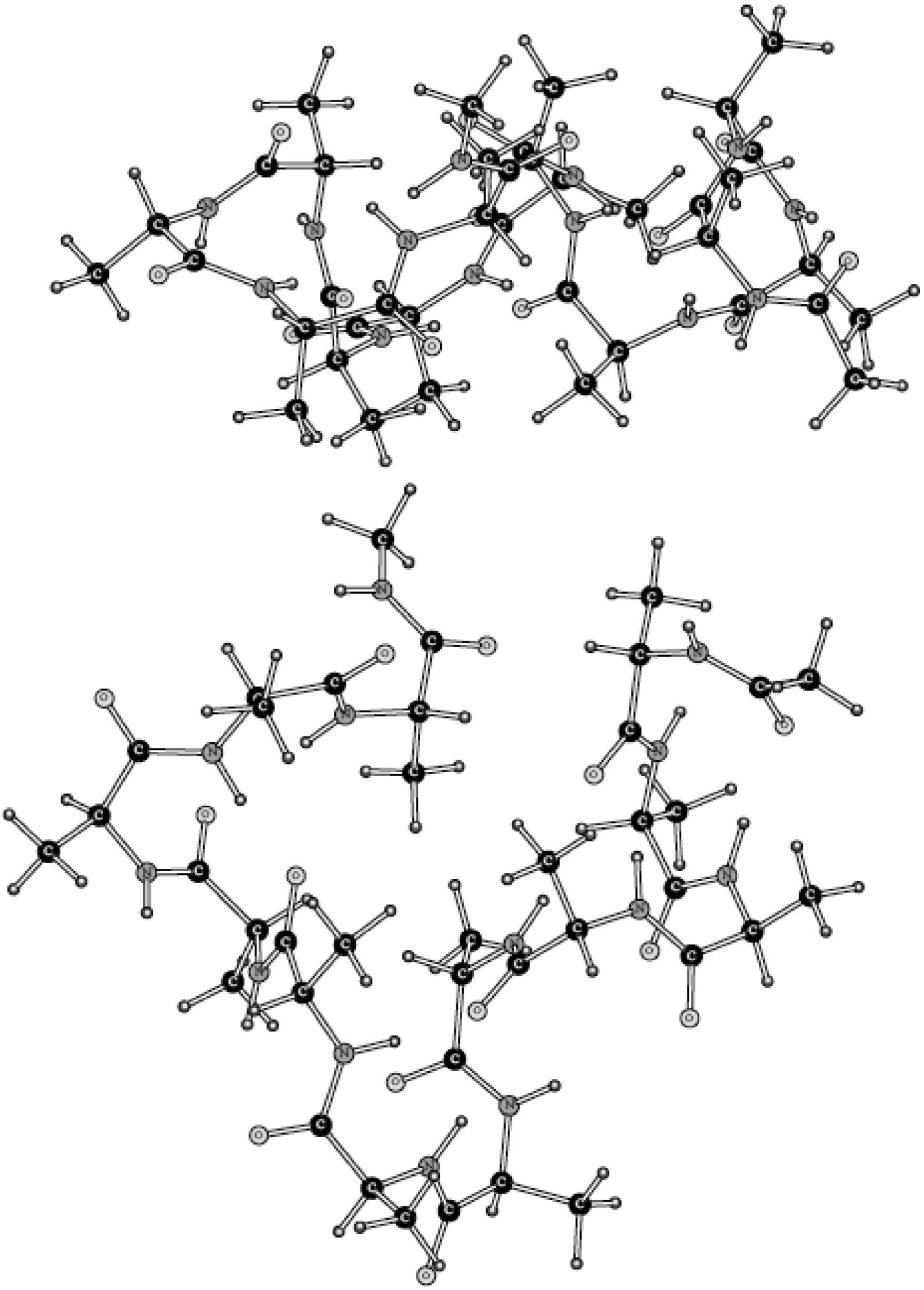


Figure 9. Structure of the ala-13 α-helix simulated with POSSIM, version tors.1, in gas phase, after 19 × 10^6 Monte Carlo configurations.

error function was minimized as a function of the torsional parameters in eq 8.

C. Calculating Dimerization Energies for the Alanine Dipeptide Complexes with Water. In order to test the transferability of the NMA nonbonded parameters employed for our alanine and protein backbone model, we calculated energies of interaction of the alanine–dipeptide with a water molecule. Structures and energies obtained for these systems with the POSSIM program were compared to the quantum mechanical results obtained with Jaguar.[11] For hydrogen bonds, a good level of accuracy can be achieved via MP2 calculations extrapolated to the basis set limit, where the contribution of higher level excitations (e.g., CCSD(T)) has been shown to be negligible (except for some cases, such as π stacking of aromatic rings, where the MP2 level has been shown to not be sufficient).

Briefly, dimer geometries were obtained by LMP2 optimizations with a cc-pVTZ(-f) basis set. The final quantum mechanical dimer binding energy E_{bind}, as used in this work, is a linear combination of the LMP2 binding energy for a smaller cc-pVTZ(-f) basis set (E_{ccpvtz}) and the LMP2 binding energy with a larger cc-pVQZ(-g) basis set (E_{ccpvq}).[15] This method has been previously demonstrated to produce a high-quality fitting and benchmarking data for force field development.[7,8]

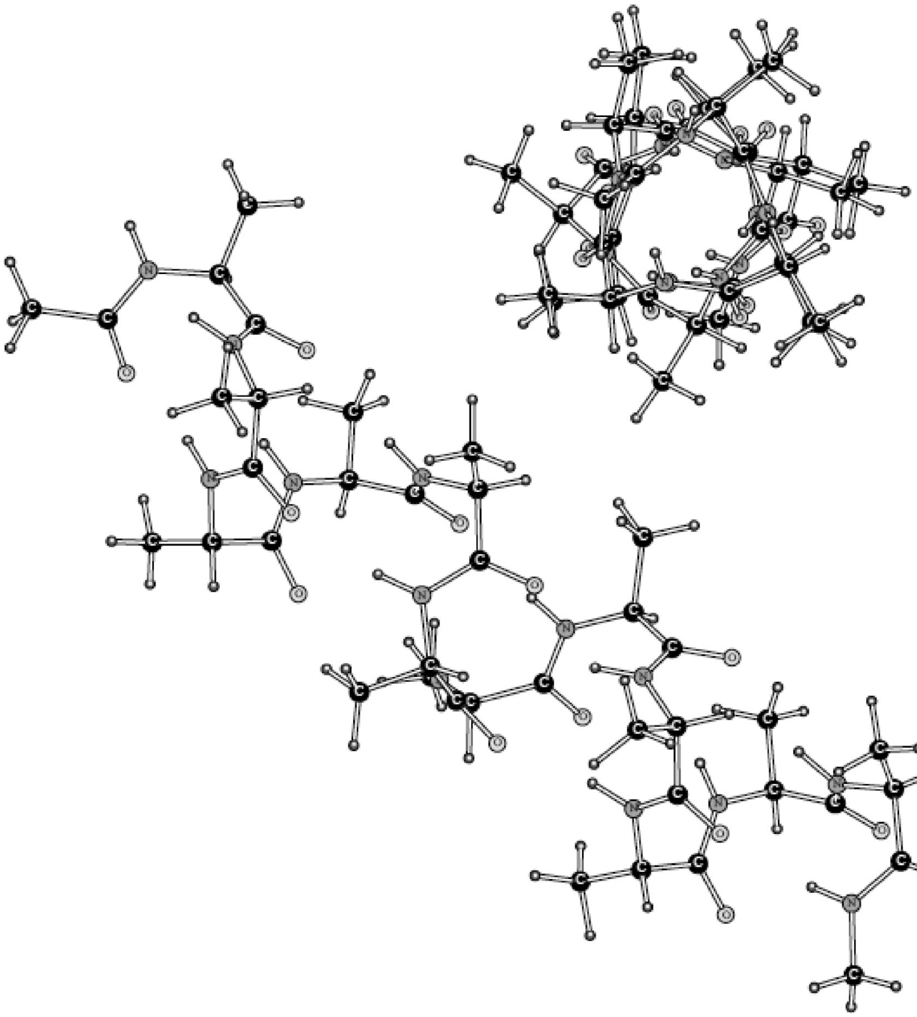


Figure 10. Structure of the ala-13 α-helix simulated with POSSIM, version tors.final in gas phase, after 19 × 10^6 Monte Carlo configurations.

D. Gas-Phase and Liquid-State Simulations of the Tridecaalanine Peptide. In order to give our alanine and backbone model a final test, we carried out simulations of a tridecaalanine (ala-13) peptide both in gas-phase and in aqueous solution at 25 °C and 1 atm. The initial structure was set at the α-helix conformation, with $\phi = 296°$ and $\psi = 319°$, and the simulations proceeded with all the degrees of freedom completely unconstrained. It is known experimentally that an α-helix represents a stable conformation of alanine peptides, including ala-13, both in gas-phase and in aqueous solution.[9] We intended to show that our POSSIM force field for the alanine and backbone protein systems performs reasonably well under these conditions and is thus sufficiently robust to be successfully employed in protein and protein−ligand studies.

Gas-phase and hydrated simulations consisted of at least 18 × 10^6 and 25 × 10^6 Monte Carlo configurations, respectfully, to ensure convergence. A 7 Å dipole−dipole cutoff was used. An 8 Å cutoff was employed for the intermolecular interactions in solution (including both the solute−solvent and solvent−solvent interactions). The standard correction procedure to account for the Lennard-Jones interactions beyond the cutoff was used. The electrostatic interactions were quadratically feathered over the last 0.5 Å before the cutoff distance. A rectangular box with periodic boundary conditions was used. The box contained 948 water molecules. The initial box setup was done to have 10 Å of water on each side of the hydrated ala-13 molecule. After that, the isobaric−isothermal (NPT) ensemble was used, with Metropolis Monte Carlo technique. In the case of the OPLS simulations, a three-site model was used with TIP3P[13] nonbonded parameters and flexible bond lengths and bond angles. A flexible three-site POSSIM water model[10] was employed in the polarizable runs.

All the calculations which did not involve quantum mechanics (i.e., geometry optimizations and Monte Carlo runs) were performed with our previously introduced POSSIM software suite.[10] Whenever possible, comparison with the fixed charges OPLS-AA force field was done, and the OPLS-AA results were also calculated with the POSSIM program.

III. RESULTS AND DISCUSSION

A. Alanine Dipeptide and Tetrapeptide Conformational Energies and Angles. We have followed the previously established procedure of calculating the alanine di- and tetrapeptide conformational energies and ϕ and ψ values as the initial assessment of the quality of the parameters for the alanine and protein backbones. The same set of the

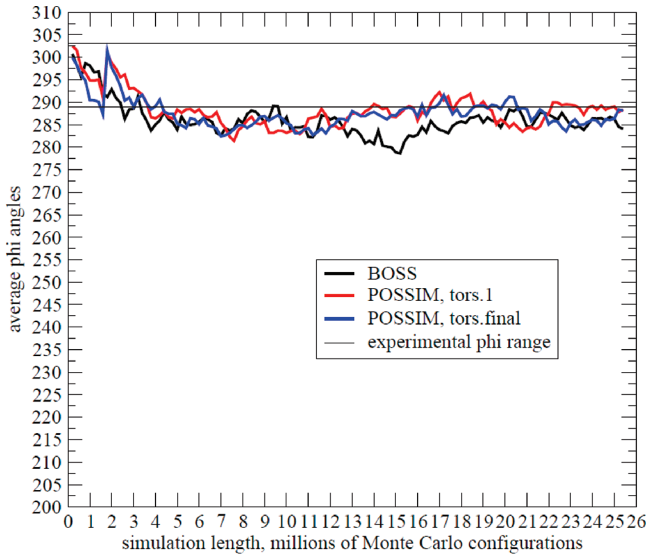


Figure 11. Average ϕ angles in α-helix simulations in aqueous solution vs the simulation length, in millions of Monte Carlo configurations.

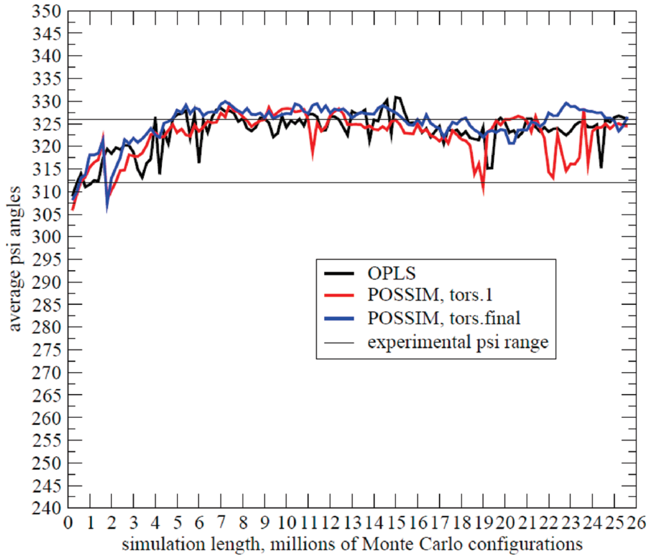


Figure 12. Average ψ angles in α-helix simulations is aqueous solution vs the simulation length, in millions of Monte Carlo configurations.

conformers that was employed in the previous studies was used.[7,8] The production of the torsional parameters proceeded as described in the Methods Section. First, weighted fitting to rotamer energies was carried out. The resulting parameters are shown in Table 1 (torsional parameters which are not listed were the same as in the NMA model).[10] We denote this set of parameters as tors.1, as opposed to the final set tors.final. Given in Table 2 are conformational energies and ϕ and ψ values, as computed with the quantum mechanics, POSSIM and OPLS. In addition to the six conformers used in previous studies, we have also added PII and α_R which are more relevant in aqueous solution.[14] Quantum mechanical optimizations were done at the LMP2/cc-pVTZ(-f) level. In both OPLS and POSSIM calculations, conformers β_2, α_L, PII, and α_R had the backbone dihedral angles fixed at the quantum mechanical values. It is known that molecular mechanics usually does not reproduce these conformers well. Overall, the performance of both POSSIM and OPLS is satisfactory. The POSSIM results have a slightly lower error in the conformational energies, while the OPLS results are closer to the quantum mechanics in terms of the geometries.

We have also calculated relative energies of the extended and globular conformations of the alanine tetrapeptide (shown on Figures 3 and 4, respectively). We determined the quantum mechanical energy difference for these conformers to be 3.09 kcal/mol, the globular form being the global energy minimum. At the same time, this quantity is known to have a relatively large range of calculated quantum mechanical energies. For example, ref 15 lists the globular−extended energy gap for the alanine tetrapeptide to be between 2.88 and 4.99 kcal/mol. The POSSIM result with the tors.1 torsional parameters set was 2.53 kcal/mol, and the OPLS result was 3.51 kcal/mol.[7,8]

We then further refined the backbone torsional parameters as described in the Methods section. The resulting values of the torsional Fourier coefficients and the conformational energies and angles are given in Tables 3 and 4, respectively. This set of the torsional parameters is termed tors.final, and this is the final set for the POSSIM protein backbone ϕ and ψ. The average dipeptide conformational energy error is now slightly higher at 0.97 kcal/mol, but the average errors in the backbone angles ϕ and ψ are reduced to 1.6° and 12.9°, respectively. Moreover, the globular−extended energy gap in the tetrapeptide is 3.14 kcal/mol, in a better agreement with the quantum mechanical results (3.09 kcal/mol with our calculations and 2.88−4.99 kcal/mol from the data ref 15). The value of the ψ for the C7$_{eq}$ conformer is lower now, but this part of the conformational space is not relevant in practical protein applications. The overall average error in both backbone angles was reduced.

B. Alanine Dipeptide−Water Dimerization Energies and Distances. There are four possible water hydrogen bonding sites in the alanine dipeptide—two NH hydrogens and two carbonyl oxygen atoms. However, our quantum mechanical energy minimizations have demonstrated that water molecules prefer to make two hydrogen bonds at the same time, one with the H and one with the O atoms. Therefore, there are only two water−alanine dipeptide heterodimer structures, as shown on Figure 5.

The quantum mechanical structures were used as the initial guesses for the POSSIM optimizations. Both POSSIM and OPLS-AA were utilized. We compared the binding energies, as well as the geometries of the complexes. Both hydrogen bonding distances (O···H−N) and H···O=C) and the ϕ and ψ angles of the alanine dipeptide backbone were used for the comparison. The results of these calculations are presented in Table 5. The quantum mechanical energy of the dimerization is reproduced slightly better with the OPLS, the average error being 0.89 kcal/mol vs 1.12 kcal/mol with POSSIM. The latter tends to underestimate the magnitude of the binding energy. This is not unexpected. The nonbonded parameters for the alanine dipeptide have been adopted from NMA fitting.

And the same tendency was also present in the NMA case, with the POSSIM underestimating the NMA−water binding energy by an average of 0.89 kcal/mol.[10] The overall performance

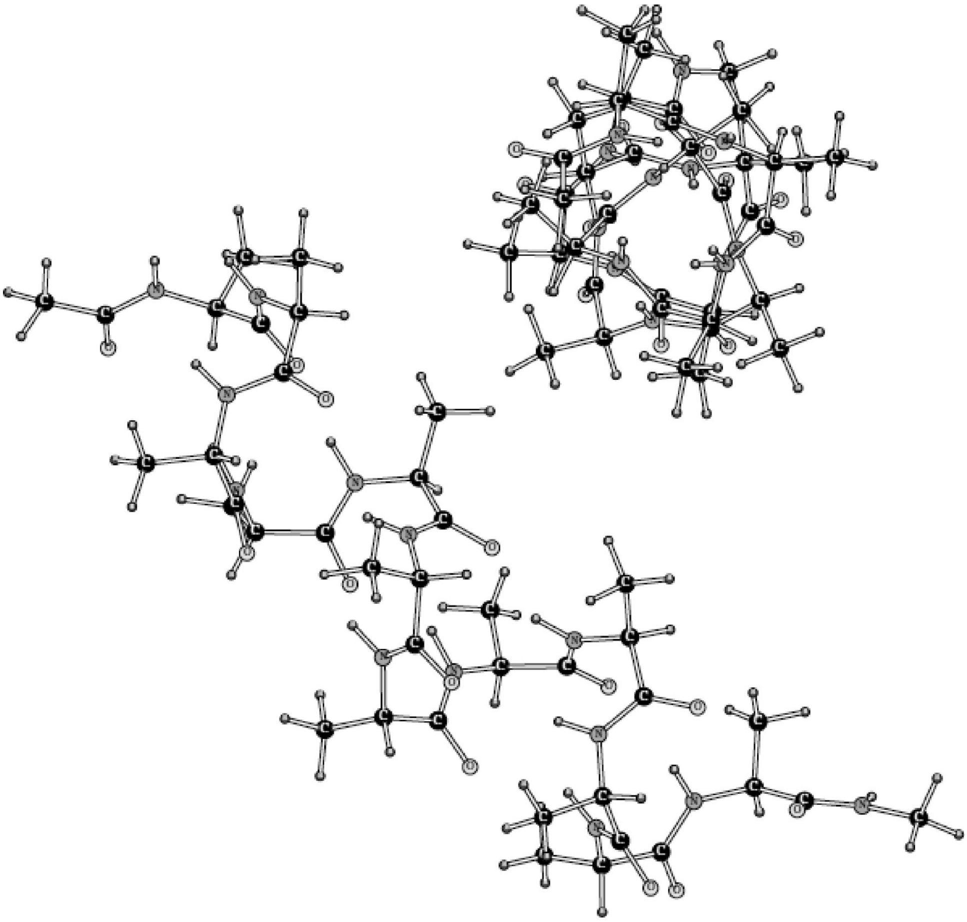


Figure 13. Structure of the ala-13 α-helix simulated with OPLS, in aqueous solution, after 25 × 10^6 Monte Carlo configurations. Water molecules are not shown for the sake of clarity.

of the NMA parameters was very good. This included reproducing liquid NMA heat of vaporization and density. Which lead us to the conclusion that our quantum mechanical NMA—water binding energies are probably somewhat overestimated. Therefore, a similar trend in the alanine dipeptide complex formation with water could have been expected and is not at all an indication of problems with the protein POSSIM force field. Moreover, it can be easily seen from the data in Table 5 that the POSSIM performed noticeably better than the OPLS in reproducing the hydrogen-bond lengths, which are probably given much more accurate than the energies by the quantum mechanics. The average errors in these lengths are 0.15 and 0.06 Å with the OPLS and POSSIM calculations, respectively.

It is also worth noting that the values of the ϕ and ψ backbone angles in this complex, as computed with the POSSIM, are much closer to the resulting quantum mechanical values of these angles than their OPLS counterparts, with the average error of only 5.3° vs 12.8°. This is so even though the POSSIM gives the lowest-energy monomer conformer (C7eq) ψ angle of only 34.4° vs the quantum mechanical 88.1° and the OPLS 61.8°. We believe that this fact confirms that: (i) the conformational energy surface is rather flat at that region, and so the precise location of the minimum is not entirely crucial; and (ii) the POSSIM force field is robust and adequate in reproducing important binding geometries.

We have further investigated the alanine dipeptide—water binding properties by running calculations, in which the values of ϕ and ψ were kept the same as in the fully optimized quantum mechanical dimers in all the cases (quantum mechanical, OPLS, and POSSIM monomers and also the OPLS and POSSIM dimers). The results are presented in Table 6. The structure B dimerization energy as computed with the POSSIM is slightly greater than the quantum mechanical one in this case (−12.2 vs −11.7 kca/mol), otherwise the trends are the same as in the fully relaxed geometry optimizations. The average errors in the dimerization energies with the POSSIM and OPLS are 0.76 and 0.39 kcal/mol, respectively. The POSSIM and OPLS errors in the hydrogen-bonding distances are 0.09 and 0.05 Å. Interestingly, the improvement in geometry achieved by fixing the backbone angles is greater with the OPLS than it is with the POSSIM. Once again, we believe this indicates that, even though the C7eq conformational geometry is better reproduced with the OPLS, the more important binding properties are better assessed with the POSSIM force field.

C. Gas-Phase and Hydrated Simulations of the Tridecaalanine Peptide (Ala-13). We have carried out Monte Carlo simulations of the ala-13 in order to test the robustness of the POSSIM force field by assessing stability of this experimentally known α-helical peptide. While quantum mechanical gas-phase alanine dipeptide conformational energies and

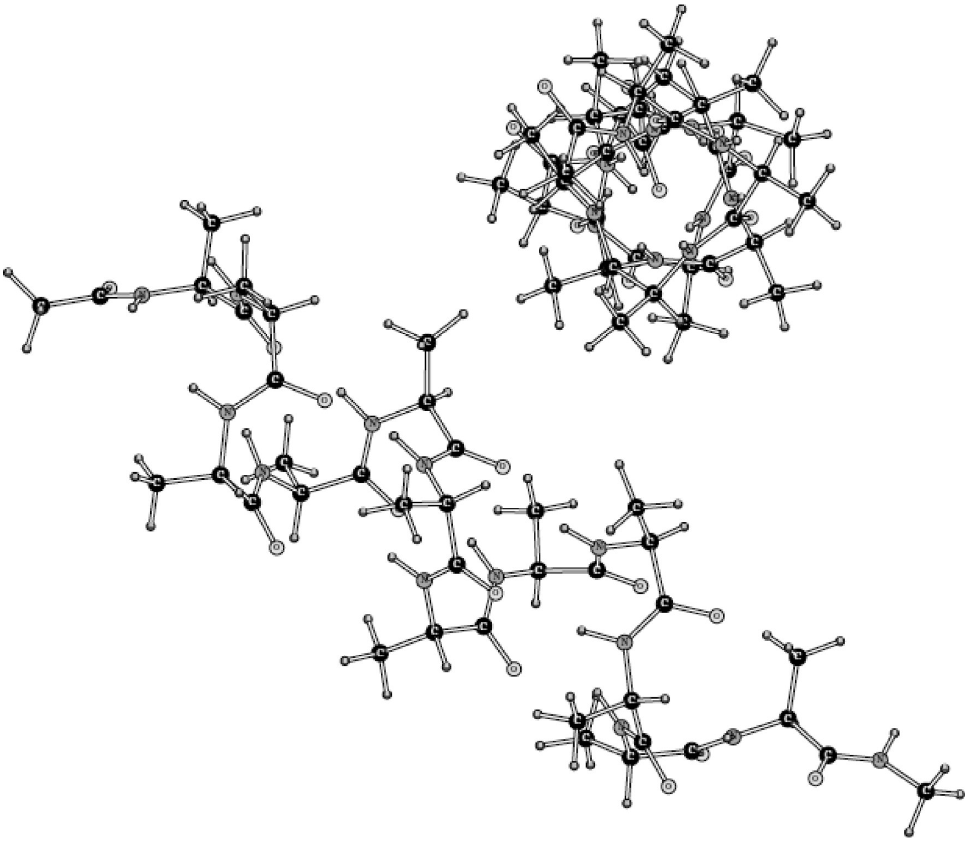


Figure 14. Structure of the ala-13 α-helix simulated with POSSIM, version tors.1, in aqueous solution, after 25 × 10^6 Monte Carlo configurations. Water molecules are not shown for the sake of clarity.

geometries are important in fitting, these simulations provided a direct comparison with the available experimental observations. In particular, we were assessing the general stability of the helix and the average values of the backbone ϕ and ψ angles. Figures 6 and 7 show graphs of the average values of these angles as a function of the simulation length (in millions of Monte Carlo configurations) for the OPLS force field, as well as with POSSIM, using both the tors.1 and tors.final torsional parameters. Each angle value represents averaging over the last 200 000 configurations before the indicated simulation length.

The experimental values of the backbone ϕ and ψ in an α-helix are 296° and 319°, respectively, with a 7° uncertainty.[16] In finding the average values of the backbone angles, we disregarded one residue on each end of the helix.

Two conclusions can be made from the presented results. First, the final versions of the POSSIM as well as the OPLS force field yield better agreement with the experimental data than the POSSIM version with the tors.1 parameters. Second, the ϕ values are more stable than those of the angle ψ with all the force fields tested.

But one should keep in mind that the experimental data represent crystallographic results, and thus the thermal motion allowed in the Monte Carlo calculations can cause oscillations beyond the ±7° experimental lines. Overall, we can conclude that the gas-phase simulations confirm that the newly developed POSSIM force field is stable and robust. They reproduce the experimentally observed α-helix gas-phase stability (see also Supporting Information). The stability of the simulated helixes can also be evaluated by studying the final structure of the system shown in Figures 8−10. One can see that, while the OPLS and POSSIM with tors.final produce a stable α-helix, the POSSIM/tors.1 helix denaturates. At the same time, the average ϕ and ψ angles in the tors.1 version of POSSIM are not extremely far from the experimental data, therefore the helicity of the structure is at least partially conserved.

Average values of the ϕ and ψ angles as a function of the simulation length for the ala-13 peptide in water are shown on Figures 11 and 12. In this case, as can be expected, the stability of the both angles is greater, and the deviations are smaller. It should be noted that the angle ϕ tends to be too low compared to the experimental crystallographic values, while the angle ψ is somewhat too high, thus their sum stays roughly at the same spot as the experimental one (255° or −105°), and the α-helicity of the structure for all the force fields employed is good.

Structures of the ala-13 peptide after 25 × 10^6 Monte Carlo configurations in water are given on Figures 13−15. Water molecules are not removed for clarity. It can be seen from the figures, in combination with the graphs and the table for the liquid-state simulations, that in this case (hydrated ala-13) all three force fields (OPLS and the two versions of POSSIM) perform adequately, and no denaturation of the tridecaalanine α-helix is observed.

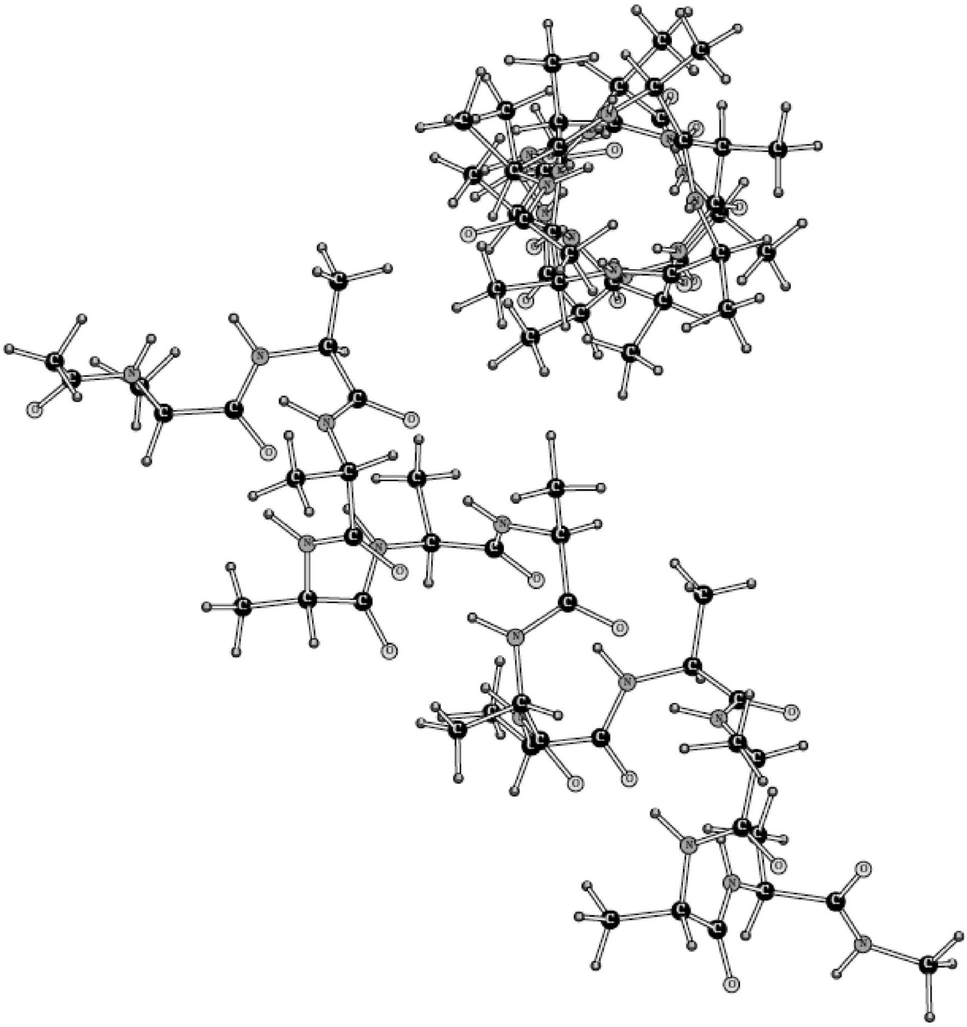


Figure 15. Structure of the ala-13 α-helix simulated with POSSIM, version tors.final, in aqueous solution, after 25 × 10^6 Monte Carlo configurations. Water molecules are removed for clarity.

IV. CONCLUSIONS

We have presented results of developing a fast polarizable POSSIM force field for alanine and protein backbones. The quantum mechanical data set used for fitting was streamlined and simplified as compared to the previous version of the complete polarizable force field for proteins, and a high degree of transferability of the potential energy parameters has been demonstrated.

We have included a previously unused step of calculating dipeptide dimerization energies with a water molecule as an additional proof of validity of the technique and the resulting force field. The POSSIM force field performs well in this test.

The torsional fitting procedure has been augmented by a new step, a direct optimization-type fitting of the torsional parameters to the quantum mechanical conformational energies and structures.

At the same time, we believe that quantum mechanical dipeptide conformers in themselves are not a sufficient tool in validation of a force field. One of the reasons for this assumption is that most of these conformers belong to parts of the total conformational space which are rarely found in experimentally known proteins. Therefore, we have included an additional step to further test the robustness of the POSSIM force field. We have simulated the tridecaalanine peptide (ala-13) in both gas phase and aqueous solution with the Monte Carlo technique. This peptide is experimentally known to form an α-helix under these conditions. The POSSIM ala-13 (and the OPLS used for benchmarking) was found to maintain a stable α-helical conformation as well.

We conclude that the resulting polarizable POSSIM force field is adequately accurate, and we will use this model for the alanine and protein backbones as the basis for further development of a complete polarizable POSSIM force field for proteins.

■ ASSOCIATED CONTENT

S Supporting Information. Tabulated values of ϕ and ψ angles of the α-helix in gas phase and solution as a function of simulation length. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: gkaminski@wpi.edu.

■ ACKNOWLEDGMENT

This project was supported by grant no. R01GM074624 from the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health. The authors express gratitude to Schrödinger, LLC, for the Jaguar and Impact software.

■ REFERENCES

(1) (1) See, for example: (a) Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 4177–4178. (b) Cieplak, P.; Caldwell, J.; Kollman, P. *J. Comput. Chem.* **2001**, *22*, 1048–1057. (c) Kaminski, G. A. *J. Phys. Chem. B* **2005**, *119*, 5884–5890. (d) Jiao, D.; Zhang, J. J.; Duke, R. E.; Li, G. H.; Schneiders, M. J.; Ren, P. Y. *J. Comput. Chem.* **2009**, *30*, 1701–1711. (e) Hernandez, G.; Anderson, J. S.; LeMaster, D. M. *Biochemistry* **2009**, *48*, 6482–6494. (f) Wang, X. Y.; Zhang, J. Z. H. *Chem. Phys. Lett.* **2011**, *501*, 508–512.

(2) (a) MacDermaid, C. M.; Kaminski, G. A. *J. Phys. Chem. B* **2007**, *111*, 9036–9044. (b) Click, T. H.; Kaminski, G. A. *J. Phys. Chem. B* **2009**, *113*, 7844–7850.

(3) Veluraja, K.; Margulis, C. J. *J. Biomol. Struct. Dyn.* **2005**, *23*, 101–111.

(4) (a) Ji, C.; Mei, Y.; Zhang, J. Z. H. *Biophys. J.* **2008**, *95*, 1080–1088. (b) Ji, C. G.; Zhang, J. Z. H. *J. Phys. Chem. B* **2009**, *113*, 16059–16064.

(5) For representative publications see: (a) Rick, S. W.; Stuart, S. J.; Berne, B. J. *J. Chem. Phys.* **1994**, *101*, 6141–6156. (b) Liu, Y. P.; Kim, K.; Berne, B. J.; Friesner, R. A.; Rick, S. W. *J. Chem. Phys.* **1998**, *108*, 4739–4755. (c) Ramon, J. M. H.; Rios, M. A. *Chem. Phys.* **1999**, *250*, 155–169. (d) Gonzalez, M. A.; Enciso, E.; Bermejo, F. J.; Bee, M. *J. Chem. Phys.* **1999**, *110*, 8045–8059. (e) Soetens, J. C.; Jansen, G.; Millot, C. *Mol. Phys.* **1999**, *96*, 1003–1012. (f) Dang, L. X. *J. Chem. Phys.* **2000**, *113*, 266–273. (g) Chen, B.; Xing, J. H.; Siepmann, J. I. *J. Phys. Chem. B* **2000**, *104*, 2391–2401. (h) Jedlovszky, P.; Vallauri, R. *J. Chem. Phys.* **2001**, *115*, 3750–3762. (i) Ribeiro, M. C. C. *Phys. Rev. B* **2001**, *6309*, 4205. (j) Rinker, S.; Gunsteren, W. F. *J. Chem. Phys.* **2011**, *134*, 084110. (k) Jiang, W.; Hardy, D. J.; Phillips, J. C.; MacKerrel, A. D.; Schulten, K.; Roux, B. *J. Phys. Chem. Lett.* **2011**, *2*, 87–92.

(6) Kaminski, G. A.; Zhou, R.; Friesner, R. A. *J. Comput. Chem.* **2003**, *24*, 267–276.

(7) Kaminski, G. A.; Stern, H. A.; Berne, B. J.; Friesner, R. A.; Cao, Y. X.; Murphy, R. B.; Zhou, R.; Halgren, T. *J. Comput. Chem.* **2002**, *23*, 1515–1531.

(8) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **2001**, *105*, 6474–6487.

(9) (a) Marqusee, S.; Robbins, V. H.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5286–5290. (b) Scholtz, J. M.; York, E. J.; Steward, J. M.; Baldwin, R. L. *J. Am. Chem. Soc.* **1991**, *113*, 5102–5104. (c) Scholtz, J. M.; Baldwin, R. L. *Annu. Rev. Biophys. Biomol. Struct.* **1992**, *21*, 95–119. (d) Kinnear, B. S.; Kaleta, D. T.; Kohtani, M.; Hudgins, R. R.; Jarrold, M. F. *J. Am. Chem. Soc.* **2000**, *122*, 9243–9256. (e) Wei, Y.; Nader, W.; Hansmann, U. H. E. *J. Chem. Phys.* **2007**, *126*, 204307.

(10) Kaminski, G. A.; Ponomarev, S. Y.; Liu, A. B. *J. Chem. Theory Comput.* **2009**, *5*, 2935–2943.

(11) (a) *Jaguar*, v3.5, Schrödinger, Inc.: Portland, OR, 1998; (b) *Jaguar*, v4.2, Schrödinger, Inc.: Portland, OR, 2000.

(12) Kaminski, G. A.; Maple, J. R.; Murphy, R. B.; Braden, D.; Friesner, R. A. *J. Chem. Theory Comput.* **2005**, *1*, 248–254.

(13) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W. *J. Chem. Phys.* **1983**, *79*, 926–935.

(14) Takekiyo, T.; Imai, T.; Kato, M.; Taniguchi, Y. *Biopolymers* **2004**, *73*, 283–290.

(15) Distasio, R. A.; Steele, R. P.; Rhee, Y. M.; Shao, Y.; Head-Gordon, M. *J. Comput. Chem.* **2007**, *28*, 839–856.

(16) Berndt K. D. *Protein Secondary Structure*, Birkbeck College, University of London: London; http://www.cryst.bbk.ac.uk/PPS2/course/section8/ss-960531_6.html. Accessed on December 10, 2010).

Vibrational Energy Levels via Finite-Basis Calculations Using a Quasi-Analytic Form of the Kinetic Energy

Juana Vázquez,*[†,‡] Michael E. Harding,*[†,‡] John F. Stanton,[†] and Jürgen Gauss[‡]

[†]Center for Theoretical Chemistry, Department of Chemistry and Biochemistry, University of Texas at Austin, Austin, Texas 78712, United States

[‡]Institut für Physikalische Chemie, Universität Mainz, D-55099 Mainz, Germany

S Supporting Information

ABSTRACT: A variational method for the calculation of low-lying vibrational energy levels of molecules with small amplitude vibrations is presented. The approach is based on the Watson Hamiltonian in rectilinear normal coordinates and characterized by a quasi-analytic integration over the kinetic energy operator (KEO). The KEO beyond the harmonic approximation is represented by a Taylor series in terms of the rectilinear normal coordinates around the equilibrium configuration. This formulation of the KEO enables its extension to arbitrary order until numerical convergence is reached for those states describing small amplitude motions and suitably represented with a rectilinear system of coordinates. A Gauss−Hermite quadrature grid representation of the anharmonic potential is used for all the benchmark examples presented. Results for a set of molecules with linear and nonlinear configurations, i.e., CO_2, H_2O, and formyl fluoride (HFCO), illustrate the performance of the method and the versatility of our implementation.

INTRODUCTION

The use of perturbation theory in second and higher orders has a long tradition in the theoretical prediction and interpretation of vibrational spectra[1,2] and is still extensively applied with new variations and extensions.[3−8] Nevertheless, variational methods constitute a more robust approach for solving the vibrational Schrödinger equation; they overcome the multiple degeneracy or resonance problems associated with the perturbational approach and provide an exact solution to the problem within the constraints imposed by the basis and the potential energy surface (PES) used. The earliest variational developments used a direct-product representation of the wave function in terms of an orthonormal set of basis functions (e.g., harmonic, Morse oscillators, etc.), which has been referred to as finite basis representation (FBR). Pioneer studies in this direction were carried out by Whitehead and Handy,[9,10] Carney et al.,[11−13] and others.[14−15] Other methods applying the FBR approach appeared later, for example, the vibrational self-consistent field (VSCF) method and different 'CI' schemes[16−21] as well as some modifications such as parallel vibrational multiple window configuration interaction (P-VMWCI) and the vibrational mean field configuration interaction (VMFCI) approaches.[23] In variational or pseudovariational methods it is also possible to construct a grid-based representation of the wave function, the so-called discrete variable representation (DVR).[23−28] After Light et al.[26−28] had shown the equivalence between FBR and DVR approaches, the latter soon started to be used extensively in almost all fields of nuclear motion theory.[28,29]

The definition and computation of the two components of the nuclear motion Hamiltonian, the kinetic energy and the potential energy operators, are aspects to consider in the theoretical prediction and analysis of vibrational spectra. A choice for representing the potential energy is a Taylor expansion around a reference configuration (usually the equilibrium structure). This has proven especially suitable for describing regions of the PES near minima, and it is well-suited to the framework of perturbational approaches.[1−5] This representation of the potential can be computed by fitting energy points to an analytic function or alternatively by numerical or analytical calculation of derivatives of the electronic energy.[30,31] Another possibility is to define the potential by a (semi)global PES; this idea provides a better description of areas of the PES other than those close to stationary points. Examples are the work of Braams and coworkers[32−35] as well as the so-called *n*-mode[36−38] and product[39] representations as well as a direct quadrature. The latter has the advantage of providing a diagonal form of the potential energy reducing significantly the computational demand. The choice of coordinates determines the form of the kinetic-energy operator (KEO) and consequently the total Hamiltonian, (assuming that if the coordinate system of the KEO and the potential are different, the transformation between them is known[40]). One option is to express the KEO in curvilinear internal coordinates, (e.g., valence, Jacobi, Radau, or polyspherical coordinates). Some of these coordinates can describe the complete range of eigenstates of a molecule and are suitable for 'floppy' molecules and in general for large amplitude vibrational motions. General formulations of the KEO in curvilinear coordinates leading to particularly simple structures have existed for a long time,[41,42] and research in this direction continues.[43−53] However, explicit algebraic expressions of KEOs for some choices of curvilinear coordinates may involve some complexity, especially when

Received: December 11, 2010
Published: April 22, 2011

increasing the molecular size. In these cases, it has been more common to build specific KEOs for molecules with a particular number of atoms (from three to five) and/or atomic arrangements.[54−69] A further possibility is to use rectilinear coordinates, in particular, normal coordinates. Even though they provide a poor description of large amplitude motion, the resulting KEO is unique and can be represented by a compact form.[70,71] These two features make the rectilinear representation attractive when dealing with (semirigid) medium- and large-size molecules and, in general, when attention is placed on low-lying vibrational energy levels and zero-point energy (ZPE), as is the usual case in thermochemical studies as well as the analysis of most infrared spectra. Because of these advantages several investigations have been devoted to Watson's simplified form of the nuclear Hamiltonian in rectilinear normal coordinates, for carrying out either a perturbative treatment[2−5,8] or a variational approach with finite basis[72−83] or discrete variable[40,84−88] representations. Nevertheless, because of complications in factorizing the KEO, the complete Watson Hamiltonian has not always been considered. In the past, this problem has been circumvented by including only the harmonic contribution of the kinetic energy[8,85,86,89−91] or by incorporating only some terms of the KEO beyond the harmonic approximation.[72−75,80−82] So far, the exact KEO has only been considered with numerical integration in the FBR or in the DVR, some examples of the use of these two representations can be found in refs 9−13 and 16−21 (FBR) as well as in refs 40, 84, and 88 (DVR).

In this work, we return to the original work of Whitehead and Handy[10] and describe a vibrational full configuration interaction (VFCI) method in a FBR to calculate energy levels using the complete vibrational Hamiltonian as given by Watson.[70,71] A quadrature grid representation of the anharmonic potential is combined with a quasi-analytic integration of the kinetic energy. The KEO beyond the harmonic oscillator treatment is represented by a Taylor series in rectilinear normal coordinates. We present a formulation of the KEO that allows the expansion to be extended to arbitrary order. Given a highly accurate PES, the applicability of our approach extends to the determination of accurate zero-point energies (ZPE) and low-lying vibrational energy levels. Results for a set of molecules with linear and nonlinear configuration, i.e., CO_2, H_2O, and HFCO are presented to illustrate the accuracy of the approach and the applicability of our implementation.

THEORY

Vibrational Hamiltonian for Nonlinear and Linear Configurations. The pure vibrational Hamiltonian (rotational angular momentum equal to zero, $J = 0$) in rectilinear dimensionless normal coordinates (q) for a nonlinear molecular system with N nuclei can be expressed in the following compact form and in units of cm^{-1}:[70,92,93]

$$\hat{\mathscr{H}} = \frac{\hbar^2}{2}\sum_{\alpha\beta}^{x,y,z}\hat{\pi}_\alpha\mu_{\alpha\beta}\hat{\pi}_\beta - \frac{\hbar^2}{8}\sum_{\alpha}^{x,y,z}\mu_{\alpha\alpha} + \frac{1}{2}\sum_{k=1}^{3N-6}\omega_k\hat{p}_k^2 + \hat{V}(q) \tag{1}$$

The three first terms of the Hamiltonian define the nuclear kinetic energy, $\hat{T}(q, \hat{p})$. The first represents the vibrational Coriolis term, and the second is the kinetic pseudopotential term (traditionally called Watson or $\hat{U}$ term). The latter has a quantum-mechanical origin and is a small mass-dependent correction to the vibrational energy that is essentially constant and therefore of little spectroscopic relevance. It can be, however, essential for predicting accurate ZPE necessary, for example, in thermochemical studies.[94] $\hat{V}(q)$ represents the potential energy surface as a function of the rectilinear dimensionless normal coordinates, ω_k is the harmonic frequency for the normal mode k, α and β denote the principal rotational axes (A, B, and C). The α-th component of the effective vibrational angular momentum operator, $\hat{\pi}_\alpha$, is defined by

$$\hat{\pi}_\alpha = \sum_{kl}^{3N-6}\zeta_{kl}^{\alpha}\left(\frac{\omega_l}{\omega_k}\right)^{1/2} q_k\hat{p}_l \tag{2}$$

The Coriolis constant, ζ_{kl}^{α}, describes the coupling between the rectilinear normal coordinates k and l along the α axis, and $\hat{p}_k$ is the linear vibrational momentum conjugate to the reduced dimensionless normal coordinate q_k and given by $\hat{p}_k = -i(\partial/\partial q_k)$. The $\alpha\beta$-component of the modified reciprocal inertia tensor μ is defined by $\mu_{\alpha\beta} \equiv (I'^{-1})_{\alpha\beta}$. Amat and Henry[95,96] derived an expression for $\mathbf{I}'$ which takes the following form for the $\alpha\beta$-component:

$$\mathrm{I}'_{\alpha\beta} = \mathrm{I}^{e}_{\alpha\beta} + \sum_{k}^{3N-6} a_k^{\alpha\beta}\gamma_k^{-1/2}q_k + \frac{1}{4}\sum_{kl}^{3N-6}\sum_{\gamma\delta}^{x,y,z} a_k^{\alpha\gamma}I_{\gamma\delta}^{e-1}a_l^{\delta\beta}\gamma_k^{-1/2}\gamma_l^{-1/2}q_kq_l \tag{3}$$

where

$$a_k^{\alpha\beta} = \left(\frac{\partial I_{\alpha\beta}}{\partial Q_k}\right)_e \tag{4}$$

Q_k denotes the mass-weighted rectilinear normal coordinate associated with the normal mode k and is related to q_k by $Q_k = (2\pi c\omega_k/\hbar)^{-1/2}q_k$, ($\gamma_k = 2\pi c\omega_k/\hbar$). Here, $\mathbf{I}^{\mathrm{e}}$ is the moment of inertia at the reference configuration (equilibrium geometry in the present case).

Watson also derived a similar quantum-mechanical expression for linear molecules[71] and avoided the difficulties with the relations between the angular momentum components by transforming the Hamiltonian to the isomorphic form introduced earlier by Hougen.[97] Assuming that the molecular reference configuration lies along the z axis, the resulting pure vibrational Hamiltonian for linear systems can be written as

$$\hat{\mathscr{H}} = \frac{\hbar^2}{2}\mu(\hat{\pi}_x^2 + \hat{\pi}_y^2) + \frac{1}{2}\sum_{k=1}^{3N-5}\omega_k\hat{p}_k^2 + \hat{V}(q) \tag{5}$$

where, $\boldsymbol{\mu} = (\mathbf{I}')^{-1}$. For simplicity, the notation is based on: $\mu = \mu_{xx} = \mu_{yy}$ and $\mathrm{I}' = \mathrm{I}'_{xx} = \mathrm{I}'_{yy}$. The components of I' are defined by

$$\mathbf{I}' = \mathbf{I}^{\mathrm{e}\,-1}\left(\mathbf{I}^{\mathrm{e}} + \frac{1}{2}\sum_{k}^{3N-5}\mathbf{a}_k\gamma_k^{1/2}q_k\right)^2 \tag{6}$$

where the relations $\boldsymbol{a}_k = a_k^{xx} = a_k^{yy}$ have been exploited. The vibrational Hamiltonian of eqs 1 and 5 can be regrouped in two parts: the harmonic-oscillator Hamiltonian and the remaining terms. For nonlinear molecules this is

$$\begin{aligned}\hat{\mathscr{H}} &= \hat{\mathscr{H}}_{\text{harm}} + \frac{\hbar^2}{2}\sum_{\alpha\beta}^{x,y,z} \hat{\pi}_\alpha \mu_{\alpha\beta}\hat{\pi}_\beta - \frac{\hbar^2}{8}\sum_{\alpha}^{x,y,z} \mu_{\alpha\alpha} + \hat{V}_{\text{anh}}(q) \\ &= \hat{\mathscr{H}}_{\text{harm}} + \hat{C}_{\text{nonlinear}}(q,\hat{p}) + \hat{U}(q) + \hat{V}_{anh}(q) \\ &= \hat{\mathscr{H}}_{\text{harm}} + \hat{K}_{\text{nonlinear}}(q,\hat{p}) + \hat{V}_{\text{anh}}(q) \qquad (7)\end{aligned}$$

where $\hat{K}_{\text{nonlinear}}$ contains the vibrational Coriolis and the Watson terms. Similarly, for linear configurations:

$$\begin{aligned}\hat{\mathscr{H}} &= \hat{\mathscr{H}}_{\text{harm}} + \frac{\hbar^2}{2}\mu(\hat{\pi}_x^2 + \hat{\pi}_y^2) + \hat{V}_{\text{anh}}(q) \\ &= \hat{\mathscr{H}}_{\text{harm}} + \hat{C}_{\text{linear}}(q,\hat{p}) + \hat{V}_{\text{anh}}(q) \\ &= \hat{\mathscr{H}}_{\text{harm}} + \hat{K}_{\text{linear}}(q,\hat{p}) + \hat{V}_{\text{anh}}(q) \qquad (8)\end{aligned}$$

In both cases, linear and nonlinear configurations, $\hat{\mathscr{H}}_{\text{harm}}$ is defined by

$$\hat{\mathscr{H}}_{\text{harm}} = \hat{\mathscr{H}}_{\text{o}} = \frac{1}{2}\sum_{k=1}^{3N-6(5)} \omega_k(q_k^2 + \hat{p}_k^2) \qquad (9)$$

The eigenfunctions of $\hat{\mathscr{H}}_{\text{harm}}$ are one-dimensional harmonic oscillator wave functions,[98] ϕ_{n_i} [$\phi_{n_i} = N_i e^{-q_i^2/2} H_{n_i}(q_i)$, where H_{n_i} is the Hermite polynomial associated to the i-th normal mode], and the total vibrational molecular wave function, Ψ, is given by a multidimensional product of harmonic oscillators basis functions (FBR). Use of this ansatz leads to an eigenvalue equation; elements of the corresponding Hamiltonian matrix $\mathbf{H}$ are given by:

$$H_{KL} = \underbrace{<\Psi_K|\hat{\mathscr{H}}_{harm}|\Psi_L>}_{\delta_{KL}\ \varepsilon_K} + \underbrace{<\Psi_K|\hat{V}_{anh}|\Psi_L>}_{anharmonic} + \underbrace{<\Psi_K|\hat{K}|\Psi_L>}_{kinetic} \qquad (10)$$

where ε_K represents the harmonic energy of the vibrational state K, and the second and third terms account for the anharmonic contribution and the remaining part of the kinetic energy, respectively. The anharmonic contribution is calculated by multidimensional numerical integration employing Gauss–Hermite quadratures in the FBR,[9−14] and the evaluation of the integrals associated with the kinetic energy operator is discussed in the next section.

Quasi-Analytic Integration of $\hat{K}(q,\hat{p})$. The KEO beyond the harmonic contribution, i.e., $\hat{K}_{\text{nonlinear}}(q,\hat{p})$ and $\hat{K}_{\text{linear}}(q,\hat{p})$, is defined in eqs 7 and 8 for nonlinear and linear configurations, respectively. These equations are a function of $\boldsymbol{\mu}$, the reciprocal of $\mathbf{I}'$, which is quadratic in the rectilinear normal coordinates, see eqs 3 and 6. However $\mathbf{I}'$ can also be written as a product of three terms, two of them linear in the rectilinear normal coordinate (q),[70] and for which Watson proposed the following general formulation:

$$\mu = \mathbf{I}_{\text{e}}^{-1/2}\left(1 + \frac{1}{2}\,\mathbf{b}\right)^{-2}\mathbf{I}_{\text{e}}^{-1/2} \quad \text{with} \quad \mathbf{b} \equiv \sum_k \mathbf{I}_{\text{e}}^{-1/2}\mathbf{a}_k\mathbf{I}_{\text{e}}^{-1/2}\gamma_k^{-1/2}q_k \qquad (11)$$

with the elements of $\mathbf{a}_k$ defined via eq 4. With the expansion of the binomial $(1 + 1/2\mathbf{b})^{-2}$,[99] the expression of $\mu_{\alpha\beta}$ derived earlier by Amat and Henry[95,96] is now reformulated as follows:

$$\begin{aligned}\mu &= \mathbf{I}_{\text{e}}^{-1/2}\left\{1 - \mathbf{b} + \frac{3}{4}\mathbf{b}^2 - \frac{1}{2}\mathbf{b}^3 + ...\right\}\mathbf{I}_{\text{e}}^{-1/2} \\ &= \mathbf{I}_{\text{e}}^{-1} - \mathbf{I}_{\text{e}}^{-1}\mathbf{a}\mathbf{I}_{\text{e}}^{-1} + \frac{3}{4}\mathbf{I}_{\text{e}}^{-1}\mathbf{a}\mathbf{I}_{\text{e}}^{-1}\mathbf{a}\mathbf{I}_{\text{e}}^{-1} - \frac{1}{2}\mathbf{I}_{\text{e}}^{-1}\mathbf{a}\mathbf{I}_{\text{e}}^{-1}\mathbf{a}\mathbf{I}_{\text{e}}^{-1}\mathbf{a}\mathbf{I}_{\text{e}}^{-1} + ... \qquad (12)\end{aligned}$$

where

$$\mathbf{a} = \sum_{k=1}^{3N-6(5)} \mathbf{a}_k\gamma_k^{-1/2}q_k \qquad (13)$$

For semirigid molecules, vibrations are small amplitude motions, and this allows the components of the reciprocal modified moment of inertia, $\mu_{\alpha\beta}$, to be represented by a Taylor expansion with respect to the rectilinear normal coordinates,[100] viz.:

$$\begin{aligned}\mu_{\alpha\beta} &= \mu_e^{\alpha\beta}\delta_{\alpha\beta} + \sum_k \mu_{(k)}^{(1)\alpha\beta}q_k + \frac{1}{2}\sum_{kl}\mu_{(k,\,l)}^{(2)\alpha\beta}q_kq_l \\ &+ \frac{1}{6}\sum_{klm}\mu_{(k,\,l,\,m)}^{(3)\alpha\beta}q_kq_lq_m + \frac{1}{24}\sum_{klmn}\mu_{(k,\,l,\,m,\,n)}^{(4)\alpha\beta}q_kq_lq_mq_n + ... \qquad (14)\end{aligned}$$

and $\mu_e^{\alpha\beta}$ is defined by $\mu_e^{\alpha\beta} = \delta_{\alpha\beta}(I_e^{-1})^{\alpha\beta}$. In the reduced dimensionless normal-coordinate representation derivatives of $\mu_{\alpha\beta}$ to *all* orders $[\mu_{(k)}^{(1)\alpha\beta}, \mu_{(k,l)}^{(2)\alpha\beta}, \mu_{(k,l,m)}^{(3)\alpha\beta}, ...]$ can be associated with the terms of eq 12 and expressed entirely in terms of the *first* derivatives of the inertia tensor with respect to the rectilinear normal coordinates, i.e.:

$$\begin{aligned}\mu_{(k)}^{(1)} &= \left(\frac{\partial\mu}{\partial Q_k}\right)_e = \mathbf{I}_e^{-1}\mathbf{a}\mathbf{I}_e^{-1} \Longrightarrow \mu_{(k)}^{(1)\alpha\beta}\left(\frac{\partial\mu_{\alpha\beta}}{\partial Q_k}\right)_e = -\frac{a_k^{\alpha\beta}}{\mathrm{I}_e^{\alpha}\mathrm{I}_e^{\beta}\gamma_k^{1/2}} \\ \mu_{(k,\,l)}^{(2)} &= \left(\frac{\partial^2\mu}{\partial Q_k\partial Q_l}\right)_e = \frac{3}{4}\mathbf{I}_e^{-1}\mathbf{a}\mathbf{I}_e^{-1}\mathbf{a}\mathbf{I}_e^{-1} \Longrightarrow \mu_{(k,\,l)}^{(2)\alpha\beta} = \left(\frac{\partial^2\mu_{\alpha\beta}}{\partial Q_k\partial Q_l}\right)_e = \frac{3}{4}\sum_{\gamma}\frac{a_k^{\alpha\gamma}a_l^{\gamma\beta} + a_l^{\alpha\gamma}a_k^{\gamma\beta}}{\mathrm{I}_e^{\alpha}\mathrm{I}_e^{\gamma}\mathrm{I}_e^{\beta}(\gamma_k\gamma_l)^{1/2}} \\ \mu_{(k,\,l,\,m)}^{(3)} &= \left(\frac{\partial^3\mu}{\partial Q_k\partial Q_l\partial Q_m}\right)_e = -\frac{1}{2}\mathbf{I}_e^{-1}\mathbf{a}\mathbf{I}_e^{-1}a\mathbf{I}_e^{-1}\mathbf{a}\mathbf{I}_e^{-1} \Longrightarrow \mu_{(k,\,l,\,m)}^{(3)\alpha\beta} = \left(\frac{\partial^3\mu_{\alpha\beta}}{\partial Q_k\partial Q_l\partial Q_m}\right)_e \\ &= -\frac{1}{2}\sum_{\gamma\delta}\left[\frac{a_k^{\alpha\gamma}a_l^{\gamma\delta}a_m^{\delta\beta} + a_k^{\alpha\gamma}a_m^{\gamma\delta}a_l^{\delta\beta} + a_m^{\alpha\gamma}a_l^{\gamma\delta}a_k^{\delta\beta} + a_l^{\alpha\gamma}a_k^{\gamma\delta}a_m^{\delta\beta} + a_m^{\alpha\gamma}a_k^{\gamma\delta}a_l^{\delta\beta} + a_l^{\alpha\gamma}a_m^{\gamma\delta}a_k^{\delta\beta}}{\mathrm{I}_e^{\alpha}\mathrm{I}_e^{\gamma}\mathrm{I}_e^{\delta}\mathrm{I}_e^{\beta}(\gamma_k\gamma_l\gamma_m)^{1/2}}\right] \qquad (15) \\ &\vdots \qquad \vdots \qquad \vdots\end{aligned}$$

The first step in the quasi-analytic evaluation of the kinetic energy is the calculation of the set of derivatives of $\mu_{\alpha\beta}$. Fortunately, this is easily affordable as only products of $a_k^{\alpha\beta}$ are involved, i.e., derivatives of the $\alpha\beta$-component of the inertia tensor, $I_{\alpha\beta}$, with respect to the rectilinear normal coordinate Q_k, see eq 4. Although the evaluation of integrals involving the coordinate and momentum operators is a trivial matter,[98] the processing and bookkeeping associated with sorting and linking these integrals and these contributions to the Hamiltonian is the most complicated and tedious element of the present approach particularly in the case of the Coriolis term.[101] When this difficulty is overcome, an accurate and efficient computation of the kinetic energy can be achieved.

The complexity associated with computing $\hat{K}(q,\hat{p})$ can be reduced by a formulation in terms of ladder operators. Initial work in this direction was carried out for nonlinear triatomic systems by Huber[102] and continued and extended by others in the context of variational and perturbational approaches.[6,7,83,103] Details of the ladder operator formalism are described in the next section.

Ladder Operator Formalism. Normalized raising, $\mathscr{L}_k^+$, and lowering, $\mathscr{L}_k^-$, operators for the vibrational normal mode k can be defined as

$$\mathscr{L}_k^+ = \mathscr{N}(-i\hat{p}_k + q_k); \qquad \mathscr{L}_k^- = \mathscr{N}(i\hat{p}_k + q_k) \tag{16}$$

The two ladder operators, $\mathscr{L}_k^+$ and $\mathscr{L}_k^-$, act only on the part of the total vibrational wave function, $\Psi = \phi_{n_i}(q_1)\phi_{n_j}(q_2)\phi_{n_k}(q_k)$ ... = $|n_1, n_2, ..., n_k, ...\rangle$, which depends on the k-th rectilinear normal coordinate, i.e., $\phi_{n_k}(q_k) = |n_k\rangle$. The normalization constant is chosen in such a way that:

$$\begin{aligned} \mathscr{L}_k^+|..., n_k, ...\rangle &= \left(\frac{n_k+1}{2}\right)^{1/2}|..., n_k+1, ...\rangle \\ \mathscr{L}_k^-|..., n_k, ...\rangle &= \left(\frac{n_k}{2}\right)^{1/2}|..., n_k-1, ...\rangle \end{aligned} \tag{17}$$

The vibrational quantum number associated with normal mode k is represented by n_k. Solving eqs 16 for $\hat{p}_k$ and q_k yields

$$q_k = (\mathscr{L}_k^+ + \mathscr{L}_k^-); \qquad \hat{p}_k = i(\mathscr{L}_k^+ - \mathscr{L}_k^-) \tag{18}$$

Using eq 18, the terms of the operator $\hat{K}(q,\hat{p})$ can be written in a ladder operator formalism. The components of the modified reciprocal moment of inertia are functions of the rectilinear normal coordinates and are thus also easily expressed in terms of ladder operators as

$$\begin{aligned} \mu_{\alpha\beta} = I_e^{-1} &+ \sum_k \mu_{(k)}^{(1)\alpha\beta}(\mathscr{L}_k^+ + \mathscr{L}_k^-) \\ &+ \frac{1}{2}\sum_{kl}\mu_{(k,l)}^{(2)\alpha\beta}(\mathscr{L}_k^+ + \mathscr{L}_k^-)(\mathscr{L}_l^+ + \mathscr{L}_l^-) \\ &+ \frac{1}{6}\sum_{klm}\mu_{(k,l,m)}^{(3)\alpha\beta}(\mathscr{L}_k^+ + \mathscr{L}_k^-)(\mathscr{L}_l^+ + \mathscr{L}_l^-)(\mathscr{L}_m^+ + \mathscr{L}_m^-) \\ &+ \frac{1}{24}\sum_{klmn}\mu_{(k,l,m,n)}^{(4)\alpha\beta}(\mathscr{L}_k^+ + \mathscr{L}_k^-)(\mathscr{L}_l^+ + \mathscr{L}_l^-)(\mathscr{L}_m^+ + \mathscr{L}_m^-) \\ &\qquad (\mathscr{L}_n^+ + \mathscr{L}_n^-) + \cdots \end{aligned} \tag{19}$$

The effective vibrational angular momentum $\hat{\pi}_\alpha$ defined in eq 2 involves the product $q_k\hat{p}_k$ and has a more complicated structure.[102] This operator is given by

$$\begin{aligned} \hat{\pi}_\alpha = i\sum_{k<l}\zeta_{kl}^{\alpha}[&\mathbf{R}_-^{kl}(\mathscr{L}_k^+\mathscr{L}_l^+ - \mathscr{L}_k^-\mathscr{L}_l^-) \\ &+ \mathbf{R}_+^{kl}(\mathscr{L}_k^-\mathscr{L}_l^+ - \mathscr{L}_k^+\mathscr{L}_l^-)] \end{aligned} \tag{20}$$

where the coefficients $\mathbf{R}_-^{kl}$ and $\mathbf{R}_+^{kl}$ are defined by

$$\begin{aligned} \mathbf{R}_-^{kl} &= \left(\frac{\omega_l}{\omega_k}\right)^{1/2} - \left(\frac{\omega_k}{\omega_l}\right)^{1/2}; \\ \mathbf{R}_+^{kl} &= \left(\frac{\omega_l}{\omega_k}\right)^{1/2} + \left(\frac{\omega_k}{\omega_l}\right)^{1/2} \end{aligned} \tag{21}$$

Based on the Taylor expansion of $\mu_{\alpha\beta}$, the Watson term for nonlinear configurations $(\hat{U}(q))$ and the vibrational Coriolis term $(\hat{C}(q,\hat{p}))$ for both linear and nonlinear systems can be written in terms of a hierarchy of contributions or "orders" as follows:

$$\begin{aligned} \hat{O} &= \sum_{i=0}^{\infty}\hat{O}^{(i)} \\ &= \hat{O}^{(0)} + \hat{O}^{(1)} + \hat{O}^{(2)} + \hat{O}^{(3)} + \hat{O}^{(4)} + \cdots \end{aligned} \tag{22}$$

where $\hat{O} = \hat{U}(q),\hat{C}(q,\hat{p})$ and $\hat{O}^{(i)} = \hat{U}^{(i)}(q),\hat{C}^{(i)}(q,\hat{p})$.

For $\hat{U}$ the specific terms in the expansion are:

$$\begin{aligned} \hat{U}^{(0)} &= -\frac{\hbar^2}{8}\mathrm{I}^{e-1} \\ \hat{U}^{(1)} &= -\frac{\hbar^2}{8}\sum_k \mu_{(k)}^{(1)\alpha\alpha}(\mathscr{L}_k^+ + \mathscr{L}_k^-) \\ \hat{U}^{(2)} &= -\frac{\hbar^2}{16}\sum_{kl}\mu_{(k,l)}^{(2)\alpha\alpha}(\mathscr{L}_k^+\mathscr{L}_l^+ + \mathscr{L}_k^+\mathscr{L}_l^- + \mathscr{L}_k^-\mathscr{L}_l^+ + \mathscr{L}_k^-\mathscr{L}_l^-) \\ \hat{U}^{(3)} &= -\frac{\hbar^2}{48}\sum_{klm}\mu_{(k,l,m)}^{(3)\alpha\alpha}(\mathscr{L}_k^+\mathscr{L}_l^+\mathscr{L}_m^+ + \mathscr{L}_k^-\mathscr{L}_l^+\mathscr{L}_m^+ + \mathscr{L}_k^+\mathscr{L}_l^-\mathscr{L}_m^+ \\ &\quad + \mathscr{L}_k^+\mathscr{L}_l^+\mathscr{L}_m^- + \mathscr{L}_k^+\mathscr{L}_l^-\mathscr{L}_m^- + \mathscr{L}_k^-\mathscr{L}_l^+\mathscr{L}_m^- + \mathscr{L}_k^-\mathscr{L}_l^-\mathscr{L}_m^+ \\ &\quad + \mathscr{L}_k^-\mathscr{L}_l^-\mathscr{L}_m^-) \\ &\qquad \vdots \qquad \vdots \qquad \vdots \end{aligned} \tag{23}$$

Considering the commutation relation between $\hat{\pi}_\alpha$ and μ, i.e., $\Sigma_\alpha[\hat{\pi}_\alpha,\mu_{\alpha\beta}] = 0$, terms in the expansion of the Coriolis term, $\hat{C}(q,\hat{p})$, (eqs 7, 8, and 22) take the following general form:

$$\begin{aligned} \hat{C}^{(0)} &= -\sum_{r<s}\sum_{t<u}\sum_{\alpha\beta}B_e^{\alpha}\zeta_{rs}^{\alpha}\zeta_{tu}^{\beta}(\mathbf{R}_{\mathbf{rs}}^-\mathbf{R}_{\mathbf{tu}}^-\hat{C}_{a,(rstu)}^{(0)} + \mathbf{R}_{\mathbf{rs}}^+\mathbf{R}_{\mathbf{tu}}^-\hat{C}_{b,(rstu)}^{(0)} \\ &\qquad + \mathbf{R}_{\mathbf{rs}}^-\mathbf{R}_{\mathbf{tu}}^+\hat{C}_{c,(rstu)}^{(0)} + \mathbf{R}_{\mathbf{rs}}^+\mathbf{R}_{\mathbf{tu}}^+\hat{C}_{d,(rstu)}^{(0)}) \\ \hat{C}^{(1)} &= -\sum_{r<s}\sum_{t<u}\sum_{k}\sum_{\alpha\beta}\mu_{(k)}^{(1)\alpha\beta}\zeta_{rs}^{\alpha}\zeta_{tu}^{\beta}(\mathbf{R}_{\mathbf{rs}}^-\mathbf{R}_{\mathbf{tu}}^-\hat{C}_{a,(k;rstu)}^{(1)} + \mathbf{R}_{\mathbf{rs}}^+\mathbf{R}_{\mathbf{tu}}^-\hat{C}_{b,(k;rstu)}^{(1)} \\ &\qquad + \mathbf{R}_{\mathbf{rs}}^-\mathbf{R}_{\mathbf{tu}}^+\hat{C}_{c,(k;rstu)}^{(1)} + \mathbf{R}_{\mathbf{rs}}^+\mathbf{R}_{\mathbf{tu}}^+\hat{C}_{d,(k;rstu)}^{(1)}) \\ \hat{C}^{(2)} &= -\sum_{r<s}\sum_{t<u}\sum_{kl}\sum_{\alpha\beta}\mu_{(k,l)}^{(2)\alpha\beta}\zeta_{rs}^{\alpha}\zeta_{tu}^{\beta}(\mathbf{R}_{\mathbf{rs}}^-\mathbf{R}_{\mathbf{tu}}^-\hat{C}_{a,(kl;rstu)}^{(2)} + \mathbf{R}_{\mathbf{rs}}^+\mathbf{R}_{\mathbf{tu}}^-\hat{C}_{b,(kl;rstu)}^{(2)} \\ &\qquad + \mathbf{R}_{\mathbf{rs}}^-\mathbf{R}_{\mathbf{tu}}^+\hat{C}_{c,(kl;rstu)}^{(2)} + \mathbf{R}_{\mathbf{rs}}^+\mathbf{R}_{\mathbf{tu}}^+\hat{C}_{d,(kl;rstu)}^{(2)}) \ldots \end{aligned} \tag{24}$$

The operators $\hat{C}_a^{(i)}$, $\hat{C}_b^{(i)}$, $\hat{C}_c^{(i)}$, and $\hat{C}_d^{(i)}$ $(i = 0, 1, 2, ..., \infty)$ represent sums of products of ladder operators. Explicit expressions for the two first sets of the series are given next:

$i = 0$

$$\begin{aligned} \hat{C}_{a,(rstu)}^{(0)} &= \mathscr{L}_r^+\mathscr{L}_s^+\mathscr{L}_t^+\mathscr{L}_u^+ - \mathscr{L}_r^+\mathscr{L}_s^+\mathscr{L}_t^-\mathscr{L}_u^- \\ &\quad - \mathscr{L}_r^-\mathscr{L}_s^-\mathscr{L}_t^+\mathscr{L}_u^+ + \mathscr{L}_r^-\mathscr{L}_s^-\mathscr{L}_t^-\mathscr{L}_u^- \\ \hat{C}_{b,(rstu)}^{(0)} &= \mathscr{L}_r^+\mathscr{L}_s^+\mathscr{L}_t^-\mathscr{L}_u^+ - \mathscr{L}_r^+\mathscr{L}_s^+\mathscr{L}_t^+\mathscr{L}_u^- \\ &\quad - \mathscr{L}_r^-\mathscr{L}_s^-\mathscr{L}_t^-\mathscr{L}_u^+ + \mathscr{L}_r^-\mathscr{L}_s^-\mathscr{L}_t^+\mathscr{L}_u^- \\ \hat{C}_{c,(rstu)}^{(0)} &= \mathscr{L}_r^-\mathscr{L}_s^+\mathscr{L}_t^+\mathscr{L}_u^+ - \mathscr{L}_r^-\mathscr{L}_s^+\mathscr{L}_t^-\mathscr{L}_u^- \\ &\quad - \mathscr{L}_r^+\mathscr{L}_s^-\mathscr{L}_t^+\mathscr{L}_u^+ + \mathscr{L}_r^+\mathscr{L}_s^-\mathscr{L}_t^-\mathscr{L}_u^- \\ \hat{C}_{d,(rstu)}^{(0)} &= \mathscr{L}_r^-\mathscr{L}_s^+\mathscr{L}_t^-\mathscr{L}_u^+ - \mathscr{L}_r^-\mathscr{L}_s^+\mathscr{L}_t^+\mathscr{L}_u^- \\ &\quad - \mathscr{L}_r^+\mathscr{L}_s^-\mathscr{L}_t^-\mathscr{L}_u^+ + \mathscr{L}_r^+\mathscr{L}_s^-\mathscr{L}_t^+\mathscr{L}_u^- \end{aligned} \tag{25}$$

$i = 1$

$$
\begin{aligned}
\hat{C}^{(1)}_{a,(k;rstu)} = {} & \mathscr{L}^{+}_{k}\mathscr{L}^{+}_{r}\mathscr{L}^{+}_{s}\mathscr{L}^{+}_{t}\mathscr{L}^{+}_{u} - \mathscr{L}^{+}_{k}\mathscr{L}^{+}_{r}\mathscr{L}^{+}_{s}\mathscr{L}^{-}_{t}\mathscr{L}^{-}_{u} \\
& - \mathscr{L}^{+}_{k}\mathscr{L}^{-}_{r}\mathscr{L}^{-}_{s}\mathscr{L}^{+}_{t}\mathscr{L}^{+}_{u} + \mathscr{L}^{+}_{k}\mathscr{L}^{-}_{r}\mathscr{L}^{-}_{s}\mathscr{L}^{-}_{t}\mathscr{L}^{-}_{u} \\
& + \mathscr{L}^{-}_{k}\mathscr{L}^{+}_{r}\mathscr{L}^{+}_{s}\mathscr{L}^{+}_{t}\mathscr{L}^{+}_{u} - \mathscr{L}^{-}_{k}\mathscr{L}^{+}_{r}\mathscr{L}^{+}_{s}\mathscr{L}^{-}_{t}\mathscr{L}^{-}_{u} \\
& - \mathscr{L}^{-}_{k}\mathscr{L}^{-}_{r}\mathscr{L}^{-}_{s}\mathscr{L}^{+}_{t}\mathscr{L}^{+}_{u} + \mathscr{L}^{-}_{k}\mathscr{L}^{-}_{r}\mathscr{L}^{-}_{s}\mathscr{L}^{-}_{t}\mathscr{L}^{-}_{u} \\
\hat{C}^{(1)}_{b,(k;rstu)} = {} & \mathscr{L}^{+}_{k}\mathscr{L}^{+}_{r}\mathscr{L}^{+}_{s}\mathscr{L}^{-}_{t}\mathscr{L}^{+}_{u} - \mathscr{L}^{+}_{k}\mathscr{L}^{+}_{r}\mathscr{L}^{+}_{s}\mathscr{L}^{+}_{t}\mathscr{L}^{-}_{u} \\
& - \mathscr{L}^{+}_{k}\mathscr{L}^{-}_{r}\mathscr{L}^{-}_{s}\mathscr{L}^{-}_{t}\mathscr{L}^{+}_{u} + \mathscr{L}^{+}_{k}\mathscr{L}^{-}_{r}\mathscr{L}^{-}_{s}\mathscr{L}^{+}_{t}\mathscr{L}^{-}_{u} \\
& + \mathscr{L}^{-}_{k}\mathscr{L}^{+}_{r}\mathscr{L}^{+}_{s}\mathscr{L}^{-}_{t}\mathscr{L}^{+}_{u} - \mathscr{L}^{-}_{k}\mathscr{L}^{+}_{r}\mathscr{L}^{+}_{s}\mathscr{L}^{+}_{t}\mathscr{L}^{-}_{u} \\
& - \mathscr{L}^{-}_{k}\mathscr{L}^{-}_{r}\mathscr{L}^{-}_{s}\mathscr{L}^{-}_{t}\mathscr{L}^{+}_{u} + \mathscr{L}^{-}_{k}\mathscr{L}^{-}_{r}\mathscr{L}^{-}_{s}\mathscr{L}^{+}_{t}\mathscr{L}^{-}_{u} \\
\hat{C}^{(1)}_{c,(k;rstu)} = {} & \mathscr{L}^{+}_{k}\mathscr{L}^{-}_{r}\mathscr{L}^{+}_{s}\mathscr{L}^{+}_{t}\mathscr{L}^{+}_{u} - \mathscr{L}^{+}_{k}\mathscr{L}^{-}_{r}\mathscr{L}^{+}_{s}\mathscr{L}^{-}_{t}\mathscr{L}^{-}_{u} \\
& - \mathscr{L}^{+}_{k}\mathscr{L}^{+}_{r}\mathscr{L}^{-}_{s}\mathscr{L}^{+}_{t}\mathscr{L}^{+}_{u} + \mathscr{L}^{+}_{k}\mathscr{L}^{+}_{r}\mathscr{L}^{-}_{s}\mathscr{L}^{-}_{t}\mathscr{L}^{-}_{u} \\
& + \mathscr{L}^{-}_{k}\mathscr{L}^{-}_{r}\mathscr{L}^{+}_{s}\mathscr{L}^{+}_{t}\mathscr{L}^{+}_{u} - \mathscr{L}^{-}_{k}\mathscr{L}^{-}_{r}\mathscr{L}^{+}_{s}\mathscr{L}^{-}_{t}\mathscr{L}^{-}_{u} \\
& - \mathscr{L}^{-}_{k}\mathscr{L}^{+}_{r}\mathscr{L}^{-}_{s}\mathscr{L}^{+}_{t}\mathscr{L}^{+}_{u} + \mathscr{L}^{-}_{k}\mathscr{L}^{+}_{r}\mathscr{L}^{-}_{s}\mathscr{L}^{-}_{t}\mathscr{L}^{-}_{u} \\
\hat{C}^{(1)}_{d,(k;rstu)} = {} & \mathscr{L}^{+}_{k}\mathscr{L}^{-}_{r}\mathscr{L}^{+}_{s}\mathscr{L}^{-}_{t}\mathscr{L}^{+}_{u} - \mathscr{L}^{+}_{k}\mathscr{L}^{-}_{r}\mathscr{L}^{+}_{s}\mathscr{L}^{+}_{t}\mathscr{L}^{-}_{u} \\
& - \mathscr{L}^{+}_{k}\mathscr{L}^{+}_{r}\mathscr{L}^{-}_{s}\mathscr{L}^{-}_{t}\mathscr{L}^{+}_{u} + \mathscr{L}^{+}_{k}\mathscr{L}^{+}_{r}\mathscr{L}^{-}_{s}\mathscr{L}^{+}_{t}\mathscr{L}^{-}_{u} \\
& + \mathscr{L}^{-}_{k}\mathscr{L}^{-}_{r}\mathscr{L}^{+}_{s}\mathscr{L}^{-}_{t}\mathscr{L}^{+}_{u} - \mathscr{L}^{-}_{k}\mathscr{L}^{-}_{r}\mathscr{L}^{+}_{s}\mathscr{L}^{+}_{t}\mathscr{L}^{-}_{u} \\
& - \mathscr{L}^{-}_{k}\mathscr{L}^{+}_{r}\mathscr{L}^{-}_{s}\mathscr{L}^{-}_{t}\mathscr{L}^{+}_{u} + \mathscr{L}^{-}_{k}\mathscr{L}^{+}_{r}\mathscr{L}^{-}_{s}\mathscr{L}^{+}_{t}\mathscr{L}^{-}_{u}
\end{aligned} \tag{26}
$$

IMPLEMENTATION

The implementation follows the algebraic expressions outlined here and keeps the number of basis functions per degree of freedom independent from the number of sampling points used for the quadrature grid. Three main features characterize the present implementation: (a) determination of the prefactors resulting from the expansion of the modified reciprocal moment of inertia tensor, $\boldsymbol{\mu}$; (b) evaluation of the associated integrals (matrix elements); and (c) the procedure used to obtain eigenvalues of the Hamiltonian matrix. These are discussed in the following paragraphs.

First, the prefactors in eqs 14 and 19 are calculated by an efficient permutation matrix multiplication algorithm where it is essential to use loop unrolling as well as to make full use of restricted summations. Regarding the second point, the integrals depend only on a sum of ladder operator products (see, for example eqs 25 and 26). These integrals are represented by lists of integers where 1 and 0 identify raising and lowering operators, respectively. The latter lists are regrouped by normal modes and written as a string of bits. Each string can be processed in a recursive scheme. By systematically applying the commutation relations

$$[\mathscr{L}^{+}_{i},\mathscr{L}^{+}_{j}]_{-} = 0; \quad [\mathscr{L}^{-}_{i},\mathscr{L}^{-}_{j}]_{-} = 0; \quad [\mathscr{L}^{-}_{i},\mathscr{L}^{+}_{j}]_{-} = \delta_{ij} \tag{27}$$

where i, j, ... represent normal modes or vibrational degrees of freedom, and following Wick's theorem for bosons,[104] the raising operators can be clustered on the left (normal ordering) or the right (antinormal ordering).[105] For example, in the case of the sum products of three ladder operators associated with the vibrational mode i, an antinormal ordering leads the next results:

$$
\begin{aligned}
& \mathscr{L}^{+}_{i}\mathscr{L}^{+}_{i}\mathscr{L}^{+}_{i} + \mathscr{L}^{-}_{i}\mathscr{L}^{+}_{i}\mathscr{L}^{+}_{i} + \mathscr{L}^{+}_{i}\mathscr{L}^{-}_{i}\mathscr{L}^{+}_{i} + \mathscr{L}^{+}_{i}\mathscr{L}^{+}_{i}\mathscr{L}^{-}_{i} + \mathscr{L}^{-}_{i}\mathscr{L}^{-}_{i}\mathscr{L}^{+}_{i} \\
& \quad + \mathscr{L}^{-}_{i}\mathscr{L}^{+}_{i}\mathscr{L}^{-}_{i} + \mathscr{L}^{+}_{i}\mathscr{L}^{-}_{i}\mathscr{L}^{-}_{i} + \mathscr{L}^{-}_{i}\mathscr{L}^{-}_{i}\mathscr{L}^{-}_{i} \\
& = 2(\mathscr{L}^{-}_{i}\mathscr{L}^{+}_{i}\mathscr{L}^{+}_{i} + \mathscr{L}^{-}_{i}\mathscr{L}^{-}_{i}\mathscr{L}^{+}_{i}) - 3(\mathscr{L}^{+}_{i} + \mathscr{L}^{-}_{i})
\end{aligned} \tag{28}
$$

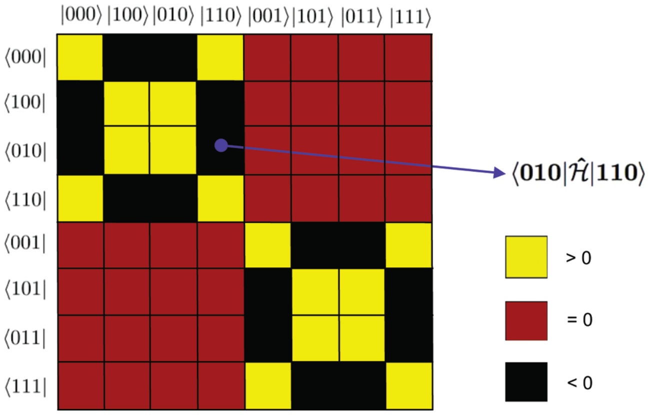


Figure 1. Schematic structure of the Hamiltonian matrix, **H**, for a XY$_2$ molecule (C_{2v} symmetry) in the FBR including only the anharmonic potential energy and harmonic oscillator, i.e., $\hat{\mathscr{H}}_0 + \hat{V}_{\text{anh}}$. (Two basis functions were used per vibrational degree of freedom).

According to the relations above, the eight possible sequences of three lowering and raising operators are reduced to four combinations of three operators, a contribution from $\mathscr{L}^{-}_{i}$ and another from $\mathscr{L}^{+}_{i}$. Within this framework, and for a given order n (see Section II), the number of integrals involving a product of n operators, $[n(n-1)+2]$, is reduced by expressing them in terms of normal- or antinormal-ordered products of n, $n-2$, $n-4$, $n-6$, ..., p operators, where p = 1 or 0 for n odd or even, respectively. In the present work we have chosen an antinormal ordering, but we note that the choice of normal or antinormal ordering is arbitrary.[106]

For an odd order n, all contributions of antinormal products of n, $n-2$, $n-4$, ..., 1 operators associated with a vibrational mode i, $\Omega_{\text{odd},i}$, are obtained by the expression:

$$
\begin{aligned}
\Omega_{\text{odd},\,i} \Longrightarrow & \sum_{l=0}^{n} \binom{n}{l} (\mathscr{L}^{-}_{i})^{l}(\mathscr{L}^{+}_{i})^{n-l} \\
& + \sum_{m=1}^{(n-1)/2} \left\{\Theta_m (\mathscr{L}^{-}_{i})^{l}(\mathscr{L}^{+}_{i})^{n-2m-l}\right\}
\end{aligned} \tag{29}
$$

and in the case of an even order n, $\Omega_{\text{even},i}$, the corresponding expression is

$$
\begin{aligned}
\Omega_{\text{even},\,i} \Longrightarrow & \sum_{l=0}^{n} \binom{n}{l} (\mathscr{L}^{-}_{i})^{l}(\mathscr{L}^{+}_{i})^{n-l} \\
& + \sum_{m=1}^{(n-2)/2} \left\{\Theta_m (\mathscr{L}^{-}_{i})^{l}(\mathscr{L}^{+}_{i})^{n-2m-l}\right\} + (-1)^{r+n}(n-1)!!
\end{aligned} \tag{30}
$$

where

$$\Theta_m = \frac{(-1)^m}{m!2^m}\left[\prod_{k=1}^{m}\frac{(n-2k+2)!}{(n-2k)!}\right]\left[\sum_{l=0}^{n-2m} n-2m\binom{n-2m}{l}\right] \tag{31}$$

and

$$r = 1 \text{ for } n = 2, 6, 10, 14\ , \ ...$$
$$\text{and } r = 0 \text{ for } n = 4, 8, 12, 16\ , \ ...$$

Thus, the $n(n-1)+2$ combinations of products of n raising ($\mathscr{L}_i^+$) and lowering ($\mathscr{L}_i^-$) operators that appear at n-th order can be reduced to only $(n+1)$ combinations of products of n operators, $(n-1)$ products of $(n-2)$ operators, $(n-3)$ products of $(n-4)$ operators, etc. In order to maintain the antinormal ordering and exploit eqs 29−31 computing the vibrational Coriolis term, the following relations were taken into account for a particular vibration i and for each order of the expansions defined by eqs 22 and 24:

$$\begin{array}{lcl}
\mathscr{L}_i^-(\mathscr{L}_i^-)^l(\mathscr{L}_i^+)^{n-l} & = & (\mathscr{L}_i^-)^{l+1}(\mathscr{L}_i^+)^{n-l} \\
\mathscr{L}_i^+(\mathscr{L}_i^-)^l(\mathscr{L}_i^+)^{n-l} & = & (\mathscr{L}_i^-)^l(\mathscr{L}_i^+)^{n-l+1}-l(\mathscr{L}_i^-)^{l-1}(\mathscr{L}_i^+)^{n-l} \\
\mathscr{L}_i^-\mathscr{L}_i^-(\mathscr{L}_i^-)^l(\mathscr{L}_i^+)^{n-l} & = & (\mathscr{L}_i^-)^{l+2}(\mathscr{L}_i^+)^{n-l} \\
\mathscr{L}_i^-\mathscr{L}_i^+(\mathscr{L}_i^-)^l(\mathscr{L}_i^+)^{n-l} & = & (\mathscr{L}_i^-)^{l+1}(\mathscr{L}_i^+)^{n-l+1}-l(\mathscr{L}_i^-)^l(\mathscr{L}_i^+)^{n-l} \\
\mathscr{L}_i^+\mathscr{L}_i^-(\mathscr{L}_i^-)^l(\mathscr{L}_i^+)^{n-l} & = & (\mathscr{L}_i^-)^{l+1}(\mathscr{L}_i^+)^{n-l+1}-(l+1)(\mathscr{L}_i^-)^l(\mathscr{L}_i^+)^{n-l} \\
\mathscr{L}_i^+\mathscr{L}_i^+(\mathscr{L}_i^-)^l(\mathscr{L}_i^+)^{n-l} & = & (\mathscr{L}_i^-)^l(\mathscr{L}_i^+)^{n-l+2}-2l(\mathscr{L}_i^-)^{l-1}(\mathscr{L}_i^+)^{n-l+1}+l(l-1)(\mathscr{L}_i^-)^{l-2}(\mathscr{L}_i^+)^{n-l}
\end{array} \tag{32}$$

where l is an integer and its values ranges over the interval $1 \leq l \leq n-1$. This scheme allows the analytic solution of integrals involving an arbitrary number of ladder operators. This kernel for evaluating the Hamiltonian matrix elements resulting from the $\hat{K}(q, \hat{p})$ operator was checked by implementing the lower-order Hamiltonian contributions explicitly in terms of dimensionless rectilinear normal coordinates (q) and its conjugate momentum ($\hat{p}$).

Finally, concerning the diagonalization procedure, a few remarks should be made in relation to the structure of the Hamiltonian matrix, $\mathbf{H}=\langle\Psi|\hat{\mathscr{H}}|\Psi\rangle$. In FBR the matrix elements of **H** can be schematically represented by distinguishing positive, negative and zero values (Figures 1 and 2). In particular, Figure 1 illustrates the structure of **H** for a XY_2 molecule (C_{2v} symmetry) using two basis functions per degree of freedom and ordering the vibrational modes by increasing value of their harmonic frequency, i.e., first the two totally symmetric vibrations (a_1 symmetry) and next the asymmetric stretching mode (b_2 symmetry). Only the harmonic oscillator operator and the anharmonic potential were included as part of $\hat{\mathscr{H}}$, i.e., $\hat{\mathscr{H}}_o + \hat{V}_{anh}(q)$. The matrix is block diagonal due to symmetry constraints, all diagonal contributions are positive, and about half of the elements have values larger than 1 cm^{-1}. Increasing the number of basis functions per vibrational degree of freedom preserves the block diagonal profile but obviously increases the density of nonzero matrix elements (see first column of Figure 2). The structure of **H** due to the contributions from different orders of the pseudopotential term ($\hat{U}$) is presented in Figure SI of the Supporting Information. As in the case of $\hat{V}(q)$, $\hat{U}$ is a function of the coordinate operator, and consequently the structure of the contributions to **H** is similar to that of $\langle\Psi|\hat{V}_{harm}+\hat{V}_{anh}|\Psi\rangle$ (second column of Figure 2,[107] with the proviso that at zeroth-order ($\hat{U}^{(0)}$) only about 1% of them have a value larger than 1 cm^{-1} and at tenth-order ($\hat{U}^{(10)}$) no contribution is larger than 1 cm^{-1} and only 2% of them have values larger than 10^{-2} cm^{-1}. The nonzero contributions to **H** resulting from the vibrational Coriolis term, $\hat{C}$, have a more peculiar structure, in the present example, mostly because of symmetry considerations,[108] (Figure SII, Supporting Information and third column of Figure 2). Nevertheless, at zeroth-order ($\hat{C}^{(0)}$) only 5% of these contributions have values larger than 1 cm^{-1}, while at tenth-order ($\hat{C}^{(10)}$) there are no elements larger than 1 cm^{-1} and only 2% of them are larger than 10^{-5} cm^{-1}. For Coriolis and pseudopotential contributions, the number of nonzero matrix elements increases rapidly with order, see Figure SIII of the Supporting Information. The complete structure of the Hamiltonian matrix, including all terms, is illustrated in the last column of Figure 2. The sequence of orders in this Figure shows a dense **H**, which is characterized by two aspects. First, the contribution from the potential energy quantitatively dominates the rest, and second the scarcity of zero matrix elements becomes larger with the increase of basis functions per degree of freedom and with the increase of order in the $\hat{U}$ and $\hat{C}$ contributions. These characteristics had to be considered in the implementation.

For triatomic molecules the eigenvalue problem is solved by direct diagonalization of **H**, however for larger systems, a complete or partial diagonalization is prohibitively expensive (even if the Hamiltonian matrix still can be kept in main memory), and an iterative Lanczos procedure[109,110] is applied. Here, the key step is the multiplication of the Hamiltonian matrix with the Lanczos vector, $\mathbf{V_L}$: **H** does not need to be constructed and stored explicitly, and only its product with $\mathbf{V_L}$ is required. In order to construct the resulting product−vector, $\mathbf{V_T}$, a process of sequential additions of the different contributions from the Hamiltonian matrix is performed, i.e., first the harmonic oscillator energies are multiplied with the trial vector and added to $\mathbf{V_T}$. This is followed by an efficient contraction of the vector matrix multiplication and the numerical integration of the anharmonic potential. The anharmonic potential is multiplied point-wise with the corresponding weights, while $\mathbf{V_L}$ is sequentially contracted with the normalized Hermite polynomials from the right side (this step has been implemented using matrix−matrix multiplications). The transformed vector is then contracted with Hermite polynomials from the left side where each of these operations yields the complete contribution from the anharmonic potential to $\mathbf{V_T}$. Next, the contributions from the $\hat{U}$ and $\hat{C}$ terms to $\mathbf{V_T}$ are constructed by multiplication of $\mathbf{V_L}$ with the corresponding one mode integrals, for $\hat{U}$, and the one and two mode integrals, for $\hat{C}$. These steps are repeated several times depending on the order of the expansion and the number of vibrational degrees of freedom. An additional complication arises with the $\hat{C}$ term as the resulting combinations of integrals have to be multiplied with different elements of $\mathbf{V_L}$ and contribute to several elements of $\mathbf{V_T}$. Once the matrix−vector multiplication is done, the algorithm continues. After the new Lanczos vector is constructed every other iteration, a complete reorthogonalization of the Lanczos vectors is carried out[111,112] in order to avoid spurious eigenvalues.[113]

■ RESULTS AND DISCUSSION

The approach outlined here has been tested for two triatomic molecules, the normal isotopic species of H_2O and CO_2. For the

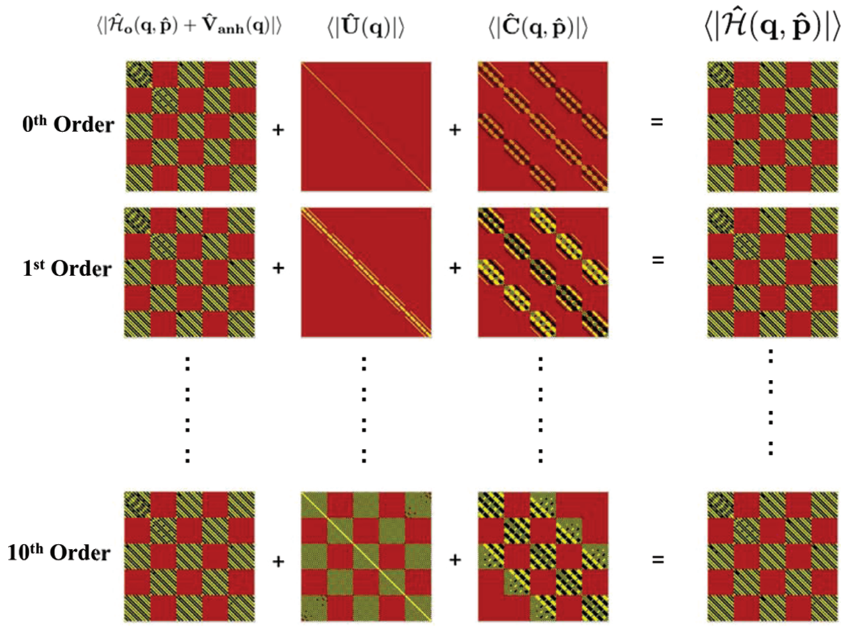

$\langle|\hat{\mathcal{H}}_o(\mathbf{q},\hat{\mathbf{p}})+\hat{\mathbf{V}}_{\mathrm{anh}}(\mathbf{q})|\rangle$
$\langle|\hat{\mathbf{U}}(\mathbf{q})|\rangle$
$\langle|\hat{\mathbf{C}}(\mathbf{q},\hat{\mathbf{p}})|\rangle$
$\langle|\hat{\mathcal{H}}(\mathbf{q},\hat{\mathbf{p}})|\rangle$
0th Order
1st Order
10th Order
+
=

Figure 2. Structure of the full Hamiltonian matrix in the FBR for a XY_2 (C_{2v} symmetry) molecule adding different contributions to $\mathscr{H}$. (All the data was obtained using five basis functions per vibrational mode, and all the pictorial representations contain the contribution from the harmonic oscillator operator in their diagonal).

Table 1. Convergence of the Anharmonic Contribution to the Zero-Point Energy and the Low-Lying Vibrational Energy Levels of $H_2^{16}O$ as Function of Number of Grid Points Used in the Numerical Integration (in cm^{-1})[a,b]

states[c]	number of grid points per mode										
(n_1 n_2 n_3)	7	11	17	20	25	27	29	31	33	35	37[d]	ref 40
(0 0 0)	4657.55	**4649.22**	4649.22	4649.22	4629.22	4629.22	4629.22	4629.22	4629.22	4629.22	**4649.22**	4649.22
(0 1 0)	1593.01	**1582.46**	1582.46	1582.46	1582.46	1582.46	1582.46	1582.46	1582.46	1582.46	**1582.46**	1582.46
(0 2 0)	3127.22	3126.71	**3126.70**	3126.70	3126.70	3126.70	3126.70	3126.70	3126.70	3126.70	**3126.70**	3126.70
(1 0 0)	3657.86	**3656.95**	3656.95	3656.95	3656.95	3656.95	3656.95	3656.95	3656.95	3656.95	**3656.95**	3656.95
(0 0 1)	3742.66	**3742.57**	3742.57	3742.57	3742.57	3742.57	3742.57	3742.57	3742.57	3742.57	**3742.57**	3742.57
(0 3 0)	4636.84	4628.87	**4628.80**	4628.80	4628.80	4628.80	4628.80	4628.80	4628.80	4628.80	**4628.80**	4628.80
(1 1 0)	5224.42	5223.39	**5223.38**	5223.38	5223.38	5223.38	5223.38	5223.38	5223.38	5223.38	**5223.38**	5223.39
(0 1 1)	5281.41	**5281.31**	5281.31	5281.31	5281.31	5281.31	5281.31	5281.31	5281.31	5281.31	**5281.31**	5281.31
(0 4 0)	6147.68	6083.88	6082.54	6082.54	**6082.53**	6082.53	6082.53	6082.53	6082.53	6082.53	**6082.53**	6082.54
(1 2 0)	6756.61	6751.64	**6751.55**	6751.55	6751.55	6751.55	6751.55	6751.55	6751.55	6751.55	**6751.55**	6751.56
(0 2 1)	6784.44	**6783.89**	6783.89	6783.89	6783.89	6783.89	6783.89	6783.89	6783.89	6783.89	**6783.89**	6783.89
(2 0 0)	7218.12	7198.30	**7198.09**	7198.09	7198.09	7198.09	7198.09	7198.09	7198.09	7198.09	**7198.09**	7198.09
(1 0 1)	7241.67	7236.36	**7236.31**	7236.31	7236.31	7236.31	7236.31	7236.31	7236.31	7236.31	**7236.3*1***	7236.32
(0 0 2)	7425.65	7421.14	**7421.10**	7421.10	7421.10	7421.10	7421.10	7421.10	7421.10	7421.10	**7421.10**	7421.10
(0 5 0)	7774.51	7492.79	7477.54	7477.42	7477.37	7477.37	7477.37	7477.36	7477.37	7477.36	7477.3*7*	7477.38[e]
(1 3 0)	8253.40	8239.20	8237.88	8237.86	8237.85	8237.85	8237.85	8237.89	8237.85	8237.86	8237.8*5*	8237.86
(0 3 1)	8273.14	8246.75	**8246.69**	8246.69	8246.69	8246.69	8246.69	8246.69	8246.69	8246.69	**8246.69**	8246.69
(2 1 0)	8756.47	8740.00	8739.74	**8739.73**	8739.73	8739.73	8739.73	8739.73	8739.73	8739.73	**8739.73**	8739.73
(1 1 1)	8764.15	8758.91	**8758.87**	8758.87	8758.87	8758.87	8758.87	8758.87	8758.87	8758.87	**8758.87**	8758.87
(0 6 0)	9893.84	8892.33	8796.29	8793.26	8792.19	8791.95	8791.34	8792.55	8791.65	8792.40	879*1.84*	8792.77[e]
(0 1 2)	8930.91	8925.12	**8925.07**	8925.07	8925.07	8925.07	8925.07	8925.07	8925.07	8925.07	**8925.07**	8925.07

[a] Results obtained using the CVRQD potential energy surface of refs 40, 115, and 116. [b] Bold font was used when two digits convergence (in cm^{-1}) was reached. [c] The ordering of the vibrational modes is consistent with the standard spectroscopic criteria[174] i.e., ν_1: symmetric stretching; ν_2: bending; ν_3: asymmetric stretching. [d] Bold font was used in this column with states that converge with 37 or less number of grid points. Italic underlined refers to those states whose converged values were slightly different from ref 40 or were unconverged in the present study as well. [e] Eigenvalues with low convergence rates in the discrete variable representation approach of Mátyus et al.[40] Problems of convergence in this numerical method are underlined. Compare footnote of Table 1 in ref 40.

Table 2. Corrections to the Harmonic ZPE and Low-Lying Vibrational Transitions of $H_2^{16}O$ Resulting from Different Orders of the Expansion of the $\hat{U}$ term (in cm^{-1})[a]

states[b]		order of the $\hat{U}$ term contribution[c,d]							
$(n_1\ n_2\ n_3)$	harm	$\hat{U}^{(0)}$	$+\hat{U}^{(2)}$	$+\hat{U}^{(4)}$	$+\hat{U}^{(6)}$	$+\hat{U}^{(8)}$	$+\hat{U}^{(10)}$	$+\hat{U}^{(12)}$	$+\hat{U}^{(14)}$
(0 0 0)	4714.58	−12.88	−13.31	−13.33	**−13.34**	−13.34	−13.34	−13.34	−13.34
(0 1 0)	1649.20	–	−0.52	−0.59	**−0.60**	−0.60	−0.60	−0.60	−0.60
(0 2 0)	3298.40	–	−1.04	−1.23	−1.26	**−1.27**	−1.27	−1.27	−1.27
(1 0 0)	3834.44	–	−0.19	**−0.21**	−0.21	−0.21	−0.21	−0.21	−0.21
(0 0 1)	3945.53	–	−0.15	−0.16	**−0.17**	−0.17	−0.17	−0.17	−0.17
(0 3 0)	4947.59	–	−1.55	−1.92	−2.00	**−2.02**	−2.02	−2.02	−2.02
(1 1 0)	5483.63	–	−0.71	−0.82	−0.84	**−0.85**	−0.85	−0.85	−0.85
(0 1 1)	5594.73	–	−0.67	−0.76	−0.77	**−0.78**	−0.78	−0.78	−0.78
(0 4 0)	6596.79	–	−2.07	−2.66	−2.82	−2.86	**−2.88**	−2.88	−2.88
(1 2 0)	7132.83	–	−1.23	−1.49	−1.54	−1.54	**−1.56**	−1.56	−1.56
(0 2 1)	7243.93	–	−1.19	−1.41	−1.45	**−1.46**	−1.46	−1.46	−1.46
(2 0 0)	7668.87	–	−0.38	**−0.43**	−0.43	−0.43	−0.43	−0.43	−0.43
(1 0 1)	7779.97	–	−0.34	−0.38	**−0.39**	−0.39	−0.39	−0.39	−0.39
(0 0 2)	7891.07	–	−0.30	−0.33	−0.33	**−0.34**	−0.34	−0.34	−0.34
(0 5 0)	8245.99	–	−2.59	−3.45	−3.73	−3.82	−3.85	**−3.86**	−3.86
(1 3 0)	8782.03	–	−1.74	−2.20	−2.32	−2.36	**−2.37**	−2.37	−2.37
(0 3 1)	8893.13	–	−1.70	−2.11	−2.20	**−2.23**	−2.23	−2.23	−2.23
(2 1 0)	9318.07	–	−0.90	−1.06	**−1.10**	−1.10	−1.10	−1.10	−1.10
(1 1 1)	9429.17	–	−0.86	−1.00	−1.03	−1.03	**−1.04**	−1.04	−1.04
(0 6 0)	9895.19	–	−3.11	−4.30	−4.74	−4.90	−4.97	−5.00	**−5.00**
(0 1 2)	9540.26	–	−0.82	−0.94	**−0.96**	−0.96	−0.96	−0.96	−0.96

[a] Results obtained using 20 grid points and basis functions per degree of freedom. [b] The ordering of the vibrational modes is consistent with the standard spectroscopic criteria,[174] i.e., ν_1: symmetric stretching; ν_2: bending; ν_3: asymmetric stretching. [c] Odd orders have been omitted because their contributions are almost negligible. [d] Bold font was used when two digits convergence (in cm^{-1}) was reached.

first, a high-accuracy potential energy surface is available, and its vibrational level structure has been extensively studied (a starting point in the literature can be found in ref 114), while the second, CO_2, represents a prototype for linear systems. Additionally, the applicability for four-atom and low-symmetry molecules is illustrated by the prediction of the low-lying vibrational spectrum of the normal isotopic species of formyl fluoride (HFCO).

Water. The investigation of the low-lying vibrational structure of the most abundant isotopologue of water, $H_2^{16}O$, was carried out with the semiglobal PES developed by Polyansky et al. (CVRQD).[115,116] The convergence of the ZPE and the 20 lowest vibrational transitions as a function of the number of quadrature points per degree of freedom used to describe the anharmonic contribution is shown in Table 1. As expected, fast convergence is observed for the ZPE and for the three fundamental vibrations. The ZPE and fundamentals converge using only 9−11 grid points per degree of freedom. Transitions involving higher excitation of the bending mode (ν_2, $1_0 2_n 3_0$, $n = 3, 4, 5, ...$) were more difficult to converge due to the somewhat large amplitude character of this motion. With 19−20 grid points per degree of freedom all states are converged except for the third, fourth, and fifth overtones of ν_2, i.e., (0,4,0), (0,5,0), and (0,6,0). The first, (0,4,0), required 27 basis functions per degree of freedom to converge, and the other two [(0,5,0) and (0,6,0)] could not be converged even with 37 basis functions per mode. The (0,6,0) state still shows a final uncertainty of 0.28 cm^{-1} (see Table 1). Our results compare well with those obtained in a DVR by Mátyus et al.[40] The largest discrepancy is observed for the (0,6,0) transition with a difference between the DEWE approach,[40,117] and our results of the order of 1.5 cm^{-1} using the same number of basis functions as Mátyus et al.,[40] i.e., 20 quadrature points per degree of freedom (see Table 1).

Table 2 shows the corrections to the harmonic ZPE and vibrational transitions resulting from the different expansion orders of $\hat{U}$ (i.e.; $\hat{U}^{(0)}$,$\hat{U}^{(2)}$,$\hat{U}^{(4)}$, ...), see eqs 22 and 23. It can be observed that all corrections decrease the energies and, as expected, that the zeroth-order contribution ($\hat{U}^{(0)}$) is a constant correction to all the vibrational levels with a quantitatively relevant effect on the ZPE and, of course, no impact for transition energies. The correction to the ZPE due to the $\hat{U}$ term converges at sixth-order ($\hat{U}^{(n)}$, $n = 0−6$) and differs by only −0.46 cm^{-1} from the corresponding zeroth-order correction usually applied in VPT2.[94,118−120] The corrections to the fundamentals also converge at sixth-order and their magnitudes range between −0.60 and −0.17 cm^{-1}. States involving two and higher quanta excitations of the bending motion exhibit the largest corrections, i.e., 2−5 cm^{-1} and require higher orders in the expansions because of the slow convergence (which is caused, in part, by both the large amplitude character of these states and the incipient singularity in μ at linear geometries).[121−123] This behavior is typical for vibrations with large amplitude character[124,125] and can be used to identify this type of motion. The kinetic energy contributions resulting from $\hat{C}$ are presented in Table 3; these corrections are characterized by positive contributions to the energies and relatively fast convergence. The $\hat{C}^{(i)}$ corrections to the ZPE are significantly smaller than those due to $\hat{U}$. However, their quantitative influence on the vibrational states in general gains importance with increasing level of excitation. As expected, the zeroth-order correction, $\hat{C}^{(0)}$, is the leading contribution.

Table 3. Corrections to the Harmonic ZPE and Low-Lying Vibrational Transitions of $H_2^{16}O$ Resulting from Different Orders of the Expansion of the $\hat{C}$ Term (in cm^{-1})[a]

states[b]		order of the $\hat{C}$ term contribution[c,d]			
(n_1 n_2 n_3)	harm	$\hat{C}^{(0)}$	$+\hat{C}^{(0)}$	$+\hat{C}^{(4)}$	$+\hat{C}^{(6)}$
(0 0 0)	4714.58	1.91	1.91	**1.92**	1.92
(0 1 0)	1649.20	13.30	**13.40**	13.40	13.40
(0 2 0)	3298.40	26.45	**26.64**	26.64	26.64
(1 0 0)	3834.44	**0.00**	0.00	0.00	0.00
(0 0 1)	3945.53	13.26	**13.36**	13.36	13.36
(0 3 0)	4947.59	39.45	**39.73**	39.73	39.73
(1 1 0)	5483.63	13.31	13.63	**13.64**	13.64
(0 1 1)	5594.73	53.02	53.40	**53.41**	53.41
(0 4 0)	6596.79	52.30	52.66	**52.67**	52.67
(1 2 0)	7132.83	26.45	27.05	**27.07**	27.07
(0 2 1)	7243.93	92.32	92.96	**92.97**	92.97
(2 0 0)	7668.87	0.00	**0.06**	0.06	0.06
(1 0 1)	7779.97	13.26	13.59	**13.60**	13.60
(0 0 2)	7891.07	26.53	**26.73**	26.73	26.73
(0 5 0)	8245.99	65.01	**65.46**	65.46	65.46
(1 3 0)	8782.03	39.45	40.31	**40.33**	40.33
(0 3 1)	8893.13	131.16	132.04	**132.05**	132.05
(2 1 0)	9318.07	13.31	13.86	**13.87**	13.87
(1 1 1)	9429.17	53.02	54.23	**54.26**	54.26
(0 6 0)	9895.19	77.59	78.11	**78.12**	78.12
(0 1 2)	9540.26	92.77	93.45	**93.46**	93.46

[a] Results obtained using 20 grid points and basis functions per degree of freedom. [b] The ordering of the vibrational modes is consistent with the standard spectroscopic criteria,[174] i.e., ν_1: symmetric stretching; ν_2: bending; ν_3: asymmetric stretching. [c] Odd orders have been omitted because their contributions are almost negligible. [d] Bold font was used when two digits convergence (in cm^{-1}) was reached.

Fundamentals and ZPE converge at fourth order ($\hat{C}^{(4)}$) and all the 20 states at sixth order ($\hat{C}^{(6)}$). When both contributions to the kinetic energy ($\hat{K}(q,\hat{p})$), i.e., the pseudopotential, $\hat{U}^{(i)}$, and the Coriolis terms, $\hat{C}^{(i)}$, i = 1, 2, 3, ..., are simultaneously taken into account the quantitative results are dominated by the $\hat{C}^{(i)}$ corrections, while the convergence is slower for the $\hat{U}^{(i)}$ contributions (see Table 4). Once the convergence of the kinetic energy contribution is achieved by reaching an appropriate order in the expansions, which is expected always for small amplitude motions, it can be used in conjunction with the contribution from the anharmonic potential to estimate the final vibrational levels, see Table 5. Our results agree well with those from the numerical procedure of Mátyus et al.,[40] with the exception of the states (0,5,0) and (0,6,0), for which numerical results show a very slow convergence and are affected by singularities in the pseudopotential contribution.[121−123]

CO$_2$. Since the resonance between the symmetric stretching mode (ν_1) of carbon dioxide and the first overtone of its degenerate bending motion ($2\nu_2$) was identified by Fermi,[126] this molecule became a famous problem in molecular spectroscopy. The isotopologue $^{12}C^{16}O_2$ and some others were extensively studied experimentally; as a result of that it was possible to experimentally infer quartic[127] and sextic[128−132] force fields for this species. Ab initio calculations of the quartic force field of $^{12}C^{16}O_2$ have been carried out at different levels of theory.[133−139] Recently, Rodriguez-Garcia et al.[139] calculated the PES of $^{12}C^{16}O_2$ from a set of energy points computed using coupled-cluster methods with partial and full inclusion of triple excitation effects together with correlation-consistent basis sets and an extrapolation technique to converge to the basis set limit. They presented the PES in terms of a fourth-order Taylor expansion and by means of numerical values on a Gauss−Hermite quadrature grid.

Concerning the prediction of the vibrational spectrum, most theoretical studies were based on vibrational perturbation theory in second order (VPT2),[133−137] and fundamental transitions were the main focus of interest. ZPE and 13 vibrational states in the range of 2000−4900 cm^{-1} were calculated within DVR using different quartic and sextic force fields[131,132,138] and a Hamiltonian expressed in terms of orthogonal Jacobi and Radau coordinates.[140] DVR was also used with a Hamiltonian in rectilinear normal coordinates,[40] and in this occasion, the ZPE and a total of 13 states between 600 and 2800 cm^{-1} were predicted using the experimental sextic force field of Chédin.[131] Rodriguez-Garcia et al.[140] applied the VSCF and VCI approaches with their extrapolated PES computing the lowest eight states of $^{12}C^{16}O_2$ and analyzed the classic resonance $\nu_1 \approx 2\nu_2$.

In the present work we use Chédin's experimental quartic force field expressed in rectilinear normal coordinates.[131] Quintic and sextic force fields were excluded for two reasons: First, inconsistencies were previously found in the sign of some force constants in internal coordinates,[138] which have been confirmed in this research when expressing the force field in rectilinear normal coordinates;[141] and second, it has been demonstrated that the influence of the quintic and sextic potential in the low-energy region of the vibrational spectrum of $^{12}C^{16}O_2$ is essentially negligible.[140] Considering the truncation of the potential expansion at the fourth order, we restrict our analysis to the ZPE and the eight lowest energy levels (see Table 6). The present results are compared with experimental values, from which the PES was derived and to results obtained with the VCI approach and the extrapolated PES,[139] as well as with the values calculated within DVR using Chédin's sextic force field in internal coordinates.[131] The third column of Table 6 contains the contribution from the vibrational Coriolis term, $\hat{C}_{\text{linear}} = \Sigma_i \hat{C}_{\text{linear}}^{(i)}$ (i = 1, 2, 3, ...). In contrast to $H_2^{16}O$, the Coriolis contribution converges quickly and is essentially converged at zeroth-order ($\hat{C}_{\text{linear}}^{(0)}$).[142] It is also noteworthy that the first and higher excitations of the symmetric stretch, ν_1 [(n 0^0; 0), n = 1,2, ...] do not have contributions from this part of the KEO due to the fact that the Coriolis constants involving this mode ($\zeta^x_{\nu_1\nu_{(2,3)}}, \zeta^y_{\nu_1\nu_{(2,3)}}$) vanish. Our results are close to the experimental values (mean deviation of about 1.0 cm^{-1}); the largest difference of about 3.0 cm^{-1} for the (1 1^1 0) state probably is due to an insufficient inclusion of anharmonic effects and the computation of only "*pure*" vibrational states. Comparing with other theoretical predictions, the DVR calculations of Mátyus et al.[40] have a mean deviation of 1.1 cm^{-1} with respect to experiment which, together with our results, confirms the small influence of higher-order force fields on the low-lying vibrational levels; Rodriguez-Garcia et al.[139] obtained values with a mean deviation of about 3.0 cm^{-1} from experiment using VCI and an extrapolated PES; their larger discrepancies might be partially due to the neglect of the vibrational Coriolis term in the kinetic energy operator. These contributions amount to 0.75−2.24 cm^{-1}.

HFCO. Several theoretical studies on the unimolecular dissociation and rearrangement reactions of HFCO have been carried out over the years,[143−149] although to the best of our knowledge only two analytic representations of its global PES were constructed.[148,149] The first[148] was obtained by fitting 3855 energy

Table 4. Corrections to the Harmonic ZPE and Low-Lying Vibrational Transitions of $H_2^{16}O$ Resulting from Different Orders of the Expansion of the $\hat{U}$ and $\hat{C}$ Terms (in cm^{-1})[a]

states[b]		order of the $\hat{O} = \hat{U} + \hat{C}$ contribution[c,d]							
(n_1 n_2 n_3)	harm	$\hat{O}^{(0)}$	$+\hat{O}^{(2)}$	$+\hat{O}^{(4)}$	$+\hat{O}^{(6)}$	$+\hat{O}^{(8)}$	$+\hat{O}^{(10)}$	$+\hat{O}^{(12)}$	$+\hat{O}^{(14)}$
(0 0 0)	4714.58	−10.97	−11.38	**−11.41**	−11.41	−11.41	−11.41	−11.41	−11.41
(0 1 0)	1649.20	13.30	12.89	12.82	**12.81**	12.81	12.81	12.81	12.81
(0 2 0)	3298.40	26.45	25.61	25.43	**25.39**	25.39	25.39	25.39	25.39
(1 0 0)	3834.44	0.00	−0.16	**−0.18**	−0.18	−0.18	−0.18	−0.18	−0.18
(0 0 1)	3945.53	13.26	13.21	**13.20**	13.20	13.20	13.20	13.20	13.20
(0 3 0)	4947.59	39.45	38.19	37.83	37.75	**37.73**	37.73	37.73	37.73
(1 1 0)	5483.63	13.31	12.92	12.82	12.80	**12.79**	12.79	12.79	12.79
(0 1 1)	5594.73	53.02	52.75	52.66	**52.65**	52.65	52.65	52.65	52.65
(0 4 0)	6596.79	52.30	50.61	50.03	49.87	49.83	49.82	**49.81**	49.81
(1 2 0)	7132.83	26.45	25.84	25.59	25.53	25.52	25.52	**25.51**	25.51
(0 2 1)	7243.93	92.32	91.79	91.59	91.55	**91.54**	91.54	91.54	91.54
(2 0 0)	7668.87	0.00	−0.32	**−0.37**	−0.37	−0.37	−0.37	−0.37	−0.37
(1 0 1)	7779.97	13.26	13.25	13.22	**13.21**	13.21	13.21	14.21	14.21
(0 0 2)	7891.07	26.53	26.43	**26.41**	26.41	26.41	26.41	26.41	26.41
(0 5 0)	8245.99	65.01	62.89	62.04	61.77	61.68	61.65	**61.64**	61.64
(1 3 0)	8782.03	39.45	38.58	38.14	38.02	37.99	37.98	**37.97**	37.97
(0 3 1)	8893.13	131.16	130.37	129.99	129.90	129.88	129.88	**129.87**	129.87
(2 1 0)	9318.07	13.31	12.96	12.81	12.78	12.78	**12.77**	12.77	12.77
(1 1 1)	9429.17	53.02	53.39	53.27	53.25	**53.24**	53.24	54.24	54.24
(0 6 0)	9895.19	77.59	75.03	73.86	73.43	73.26	73.20	**73.18**	73.18
(0 1 2)	9540.26	92.77	92.65	92.55	**92.53**	92.53	92.53	92.53	92.53

[a] Results obtained using 20 grid points and basis functions per degree of freedom. [b] The ordering of the vibrational modes is consistent with the standard spectroscopic criteria,[174] i.e., ν_1: symmetric stretching; ν_2: bending; ν_3: asymmetric stretching. [c] Odd orders have been omitted because their contributions are almost negligible. [d] Bold font was used when two digits convergence (in cm^{-1}) was reached.

Table 5. Corrections to the ZPE and Low-Lying Vibrational Transitions of $H_2^{16}O$ Resulting from the $\hat{U}$ and $\hat{C}$ Contributions (in cm^{-1})[a,b]

states[c]	$\hat{\mathscr{H}}_o + \hat{V}_{anh}$		$\hat{\mathscr{H}}_o + \hat{V}_{anh} + \hat{U}^{conv}$		$\hat{\mathscr{H}}_o + \hat{V}_{anh} + \hat{U}^{conv}$		$\hat{\mathscr{H}}_o + \hat{V}_{anh} + (\hat{U} + \hat{C})^{conv}$	
(n_1 n_2 n_3)	this work	ref 40	this work	ref 40	this work	ref 40	this work	ref 40
(0 0 0)	4649.22	4629.22	4636.30	4636.30	4651.23	4651.23	4638.21	4638.21
(0 1 0)	1582.46	1582.46	1581.58	1581.58	1595.94	1595.94	1595.08	1595.08
(0 2 0)	3126.70	3126.70	3124.64	3124.64	3154.22	3154.22	3152.20	3152.20
(1 0 0)	3656.95	3656.95	3657.23	3657.23	3656.77	3656.77	3657.05	3657.05
(0 0 1)	3742.57	3742.57	3742.97	3742.98	3755.32	3755.32	3755.73	3755.73
(0 3 0)	4628.80	4628.80	4625.02	4625.02	4671.26	4671.26	4667.58[f]	4667.57
(1 1 0)	5223.38	5223.39	5222.82	5222.82	5236.05	5236.05	5235.49	5235.49
(0 1 1)	5281.31	5281.31	5280.90	5280.90	5331.88	5331.88	5331.51	5331.51
(0 4 0)	6082.5$\underline{3}$	6082.54	6075.99	6075.9$\underline{4}$[e]	6141.44	6141.44	6135.1$\underline{6}$[f]	6135.1$\underline{0}$[e]
(1 2 0)	6751.55	6751.56	6749.85	6749.85	6777.62	6777.63	6775.96[f]	6775.97
(0 2 1)	6783.89	6783.89	6782.41	6782.41	6873.53	6873.53	6872.15	6872.15
(2 0 0)	7198.09	7198.09	7198.67	7198.68	7200.62	7200.62	7201.19	7201.19
(1 0 1)	7236.31	7236.32	7236.99	7236.99	7248.54	7248.55	7249.22	7249.22
(0 0 2)	7421.10	7421.10	7421.86	7321.86	7444.12	7444.12	7444.88	7444.88
(0 5 0)	7477.3$\underline{7}$	7477.3$\underline{8}$[d]	7465.45	7465.$\underline{40}$[e]	7555.49	7555.4$\underline{8}$[d]	7544.4$\underline{0}$[f]	7444.$\underline{19}$[e]
(1 3 0)	8237.8$\underline{5}$	8237.86	8234.43	8234.4$\underline{3}$[e]	8278.39	8278.39	8275.1$\underline{0}$[f]	8275.1$\underline{0}$[e]
(0 3 1)	8246.69	8246.69	8243.69	8243.69	8377.55	8377.55	8374.77	8374.77
(2 1 0)	8739.73	8739.73	8739.41	8739.45	8762.14	8762.14	8761.92	8761.92
(1 1 1)	8758.87	8758.87	8758.77	8758.77	8807.10	8807.10	8807.03	8807.03
(0 6 0)	879$\underline{3.26}$	8792.$\underline{77}$[d]	8767.57	877$\underline{1.13}$[e]	8896.33	8896.$\underline{06}$[d]	887$\underline{3.56}$[f]	887$\underline{5.62}$[e]
(0 1 2)	8925.07	8925.07	8925.06	8925.06	9000.34	9000.34	9000.39	9000.40

[a] Results obtained using 20 grid points and basis functions per degree of freedom and with the expansion of the $\hat{U}$ and $\hat{C}$ terms up to the twelfth-order, i.e., conv = 12. [b] Digits underlined present problems of convergence when using 20 sampling points per degree of freedom in the anharmonic potential. [c] The ordering of the vibrational modes is consistent with the standard spectroscopic criteria,[174] i.e., ν_1: symmetric stretching; ν_2: bending; ν_3: asymmetric stretching. [d] Eigenvalues with low convergence rate in ref 40. [e] Digits underlined and extracted from ref 40 (third, fifth, seventh, and ninth columns) did not converge due to the singularity present in the operator $\hat{U}$. [f] Using 37 grid points/basis functions per degree of freedom these transitions are predicted at 4667.58, 6135.16, 6775.96, 7544.37, 8275.10, and 8870.99 cm^{-1} for (0,3,0), (0,4,0), (1,2,0), (0,5,0), and (0,6,0), respectively.

points calculated at the MP4[150] level of theory, and the second[149] was based on approximately 4000 energy points calculated at the MP2 level[151] using a (10s5p2d/5s2p)/[4s3p2d/3s2p] and a correlation consistent polarized valence triple-ζ (cc-pVTZ) basis sets.[152] The MP2 PES was constructed by splitting and treating the energy surface in three different parts, each one describing a

Table 6. Final Results for the Prediction of the ZPE and Low-Lying Vibrational Transitions of $^{12}C^{16}O_2$ (in cm^{-1})[a]

states[b]	this work[c,d]			exp.	calcd[c]	
$(n_1\, n_2^{\lvert l \rvert}\, n_3)$	$\hat{\mathscr{H}}_o$	$\hat{\mathscr{H}}_o + \hat{V}_{anh}$	total	ref 131	ref 40	ref 139
$(0\ 0^0\ 0)$	2548.05	2535.80	2536.15		2535.45	
$(0\ 1^1\ 0)$	672.89	666.73	667.47 (*−0.09*)	667.38	667.68 (*−0.30*)	669.1 (*−1.72*)
$(0\ 1^1\ 0)$	672.89	666.73	667.47 (*−0.09*)	667.38	667.68 (*−0.30*)	669.1 (*−1.72*)
$(1\ 0^0\ 0)$	1353.78	1284.30	1285.10 (*−0.93*)	1284.17	1284.98 (*−0.81*)	1288.9 (*−4.73*)
$(0\ 2^2\ 0)$	1345.79	1334.48	1335.95 (*−0.82*)	1335.13	1336.48 (*−1.35*)	1339.6 (*−4.47*)
$(0\ 2^2\ 0)$	1345.79	1334.48	1335.95 (*−0.82*)	1335.13	1336.48 (*−1.35*)	1339.6 (*−4.47*)
$(0\ 2^0\ 0)$	1345.79	1387.29	1387.93 (*0.26*)	1388.19	1387.46 (*0.73*)	1389.3 (*−1.11*)
$(1\ 1^1\ 0)$	2026.67	1928.03	1929.56 (*2.91*)	1932.47	1932.37 (*0.10*)	1938.0 (*−5.53*)
$(1\ 1^1\ 0)$	2026.67	1928.03	1929.56 (*2.91*)	1932.47	1932.37 (*0.10*)	1938.0 (*−5.53*)
$(0\ 3^3\ 0)$	2018.68	2003.07	2005.25 (*−1.97*)	2003.28	2006.43 (*−3.15*)	2011.4 (*−8.12*)
$(0\ 3^3\ 0)$	2018.68	2003.07	2005.25 (*−1.97*)	2003.28	2006.43 (*−3.15*)	2011.4 (*−8.12*)
$(0\ 3^1\ 0)$	2018.68	2076.81	2078.15 (*−1.29*)	2076.86	2076.27 (*0.59*)	2080.0 (*−3.14*)
$(0\ 3^1\ 0)$	2018.68	2076.81	2078.15 (*−1.29*)	2076.86	2076.27 (*0.59*)	2080.0 (*−3.14*)
$(0\ 0^0\ 1)$	2396.53	2347.87	2349.38 (*−0.18*)	2349.20	2347.32 (*1.88*)	2349.2 (*0.00*)

[a] Results obtained with 14 grid points and basis functions per degree of freedom. [b] The ordering of the normal modes is: symmetric stretching, ν_1; degenerate bending motion, ν_2; asymmetric stretching, ν_3. [c] Numbers in parentheses and italic are the differences from the experimental values. [d] According to eqs 10, 11, and 15, the total pure vibrational Hamiltonian can be written as a sum of three terms, i.e., $\hat{\mathscr{H}} = \hat{\mathscr{H}}_o + \hat{C}_{linear} + \hat{V}_{anh}$. "Total" denotes the complete Hamiltonian including all three contributions. The $\hat{C}_{linear}$ term was converged in zeroth-order.

Table 7. Corrections to the ZPE and Low-Lying Vibrational Transitions of HFCO Resulting from the $\hat{U}$ and $\hat{C}$ Contributions (in cm^{-1})[a,b]

states[c]	$\hat{\mathscr{H}}_o + \hat{V}_{anh}$				states[c]	$\hat{\mathscr{H}}_o + \hat{V}_{anh}$			
$(n_1\, n_2\, n_3\, n_4\, n_5\, n_6)$		$+\hat{U}^{conv}$	$+\hat{C}^{conv}$	$+(\hat{U}+\hat{C})^{conv}$	$(n_1\, n_2\, n_3\, n_4\, n_5\, n_6)$		$+\hat{U}^{conv}$	$+\hat{C}^{conv}$	$+(\hat{U}+\hat{C})^{conv}$
(0 0 0 0 0 0)	4541.5	4540.6	4542.6	4541.7					
(0 0 0 0 1 0)	657.7	657.7	658.1	658.1	(0 1 0 1 0 0)	2862.4	2862.4	2863.9	2863.9
(0 0 0 0 0 1)	1014.2	1014.2	1019.2	1019.2	(0 0 0 0 3 1)	2979.3	2979.3	2986.7	2986.7
(0 0 0 1 0 0)	1048.9	1048.9	1049.5	1049.5	(1 0 0 0 0 0)	2999.4	2999.4	3003.2	3003.2
(0 0 0 0 2 0)	1314.0	1313.9	1314.8	1314.8	(0 0 0 1 3 0)	2994.1	2994.1	2996.4	2996.4
(0 0 1 0 0 0)	1369.5	1369.5	1370.3	1370.3	(0 0 0 0 0 3)	3021.5	3021.5	3036.8	3036.8
(0 0 0 0 1 1)	1670.7	1670.7	1676.5	1676.5	(0 0 1 0 1 1)	3042.3	3042.3	3049.5	3049.5
(0 0 0 1 1 0)	1698.8	1698.8	1699.9	1699.9	(0 0 1 1 1 0)	3060.8	3060.8	3063.2	3063.2
(0 1 0 0 0 0)	1820.7	1820.8	1821.3	1821.3	(0 0 0 1 0 2)	3062.7	3062.7	3074.4	3074.4
(0 0 0 0 3 0)	1968.7	1968.7	1970.0	1970.0	(0 0 0 2 0 1)	3090.4	3090.4	3098.0	3098.0
(0 0 1 0 1 0)	2027.3	2027.3	2028.3	2028.3	(0 0 0 3 0 0)	310$\underline{7.8}$	310$\underline{7.9}$	310$\underline{9.6}$	310$\underline{9.6}$
(0 0 0 0 0 2)	2021.2	2021.2	2031.4	2031.5	(0 1 0 0 2 0)	3124.1	3124.1	3126.2	3126.2
(0 0 0 1 0 1)	2059.1	2059.1	2065.4	2065.4	(0 1 1 0 0 0)	3189.3	3189.3	3190.7	3190.7
(0 0 0 2 0 0)	2084.2	2084.2	2085.3	2085.3	(0 0 0 0 5 0)	327$\underline{3.8}$	3273.8	3275.9	3275.9
(0 0 0 0 2 1)	2325.8	2325.7	2332.4	2332.4	(0 0 0 0 2 2)	3330.2	3330.2	3343.3	3343.3
(0 0 0 1 2 0)	2347.2	2347.2	2349.0	2348.9	(0 0 1 0 3 0)	3337.1	3337.1	3338.6	3338.6
(0 0 1 0 0 1)	2385.8	2385.8	2392.2	2392.2	(0 0 0 1 2 1)	3354.9	3354.9	3363.2	3363.2
(0 0 1 1 0 0)	2411.5	2411.5	2412.9	2412.9	(0 0 0 2 2 0)	3367.1	3367.1	3369.7	3369.7
(0 1 0 0 1 0)	2473.1	2473.1	2474.4	2474.4	(0 0 2 0 1 0)	3373.0	3373.0	3375.3	3375.3
(0 0 0 0 4 0)	2621.7	2621.8	2623.5	2623.5	(0 0 1 0 0 2)	3394.6	3394.6	3406.4	3406.4
(0 0 1 0 2 0)	2683.1	2683.1	2684.4	2684.4	(0 0 1 1 0 1)	3423.6	3423.6	3431.4	3431.4
(0 0 0 0 1 2)	2676.4	2676.4	2688.1	2688.1	(0 0 1 2 0 0)	3440.0	3440.1	3441.9	3442.0
(0 0 0 1 1 1)	2707.7	2707.7	2715.0	2715.0	(0 1 0 0 1 1)	3478.4	3478.4	3485.6	3485.6
(0 0 2 0 0 0)	2715.0	2715.0	2716.7	2716.7	(0 1 0 1 1 0)	3507.2	3507.2	3509.6	3509.6
(0 0 0 2 1 0)	2726.4	2726.3	2728.2	2728.2	(0 2 0 0 0 0)	3622.4	3622.4	3623.7	3623.7
(0 1 0 0 0 1)	2827.1	2827.1	2833.3	2833.3	(0 0 0 0 4 1)	3631.1	3631.1	3639.3	3639.3

[a] Results obtained using 10 grid points and basis functions per degree of freedom and with expansion of the $\hat{U}$ and $\hat{C}$ terms up to second order, i.e., conv = 2. Digits underlined present problems of convergence. [b] Potential energy surface from ref 149. [c] The ordering of the vibrational normal modes follow the standard spectroscopic criteria,[174] i.e., the first five vibrations are those with a' symmetry and the last one is the out-of-plane bending motion (a'' symmetry). The vibrational characterization is: ν_1(CH-str); ν_2(CO-str); ν_3(CH bend); ν_4(CF-str); ν_5(FCO bend).

particular region of the PES, i.e., equilibrium, transition state, and asymptotic regions. In the present study, we use Yamamoto and Kato's PES[149] which enables comparison of our results with those obtained by other methods[153−160] that employed the same potential.

Theoretical predictions of the bound region of the vibrational spectrum of HFCO have been carried out using both perturbational[161,162] and variational or pseudovariational[153,154,158−160] methods. However, the studies applying vibrational perturbation theory in second order (VPT2)[163] calculated only fundamentals[161,162] and a few two-quanta transitions ($2\nu_2$, $2\nu_4$, and $\nu_4 + \nu_5$).[161] Leforestier et al.[153,154] were probably among the first who applied other types of approaches for computing vibrational states of HFCO up to 5000 cm^{-1}. Their initial study[153] was based on a six-dimensional Hamiltonian expressed in Jacobi coordinates using a DVR of the Hamiltonian matrix contracted by a pseudospectral method.[164] In a second paper, they proposed the Jacobi–Wilson (JW) method[154] using HFCO as a test molecule. Later, the same authors employed the JW approach in combination with a modified Davidson algorithm based on a prediagonalization–perturbation step and calculated the lowest 350 vibrational states of a' symmetry[158] and the high-energy overtones of the out-of-plane mode (ν_6) (6400–10 900 cm^{-1}).[157] The latter spectral region was also investigated by means of the multiconfiguration time-dependent Hartree (MCTDH) method.[159] By contrast, the approach presented here is expected to provide a suitable description of the low energy range, and its compact form facilitates its use with medium and large molecules as well as different bonding arrangements.

Results obtained in the present research for the ZPE and the 50 lowest vibrational transitions of formyl fluoride are presented in Table 7. For this molecule, expansions of the pseudopotential and Coriolis terms converge very rapidly; both contributions are already converged at the second order of the Taylor expansion.[142] It can also be observed that the effect of the $\hat{U}$ term is essentially negligible, except for the zero-point energy which is lowered by 0.9 cm^{-1} by this correction. This indicates that HFCO is a quite rigid molecule. By contrast, the Coriolis term has a larger quantitative influence; for most of the states its contribution ranges from 1 to 10 cm^{-1} and for some higher quanta transitions involving the out-of-plane mode, ν_6, the Coriolis term contributes 10–15 cm^{-1}; see in Table 7 cases as, for example, $2\nu_6$, $\nu_5 + 2\nu_6$, $3\nu_6$, $\nu_4 + 2\nu_6$, $2\nu_5 + 2\nu_6$, and $\nu_3 + 2\nu_6$.

The transition energies calculated in this work were compared with results obtained using other theoretical approaches, see Table 8. Nevertheless the earliest studies[153,154] have been omitted because some problems have subsequently been found in those calculations.[158,160] Thus, Table 8 contains the ZPE and transition energies calculated by the JW method coupled to a modified Davidson scheme,[158] those obtained with the MCTDH approach,[159] and finally the values recently computed by Wang et al.[160] in their rovibrational study of HFCO based on a direct product of rotational Wigner functions and a DVR of the vibrational part. In references 158 and 160 the KEO was expressed in terms of Jacobi vectors, while valence polyspherical coordinates were used with the MCTDH method.[159] Our results agree well with those obtained for the a$'$-transitions calculated by the JW approach[158] and with the fundamental frequencies estimated with the MCTDH method;[159] the largest discrepancies are 1.0 and 1.7 cm^{-1} for the CH stretching and the ZPE, respectively. The fundamental vibrations predicted by DVR and FBR in the rovibrational study of Wang et al.[160] are slightly higher than the results of this work; these differences might be related with the constraints on the bond angles made in ref 160 in order to eliminate artifacts of the Yamamoto et al.'s surface.[149] Concerning the experimental values, although several experiments were devoted to the highly excited vibrational levels of HFCO (above 14 000 cm^{-1}),[165–167] no intensive research has been carried out on the low-energy region of the vibrational spectrum of this molecule. The earliest and almost only infrared and Raman experimental studies assigned the five a$'$ fundamentals and a few higher quanta transitions ($2\nu_2$, $2\nu_4$, $3\nu_4$, $4\nu_4$, $\nu_3 + \nu_4$, $\nu_3 + \nu_5$, and $\nu_4 + \nu_5$) but could only provide tentative assignments of the out-of-plane bending mode, ν_6.[168–170] It was not until 1978,[171] and later confirmed by stimulated emission pumping spectroscopy,[165–167] that this a''-symmetry motion was assigned to a wavenumber of 1011.2 cm^{-1}. Beside these experiments, some high-resolution laser Stark measurements of ν_2 (C=O stretch) were done,[172] and the only new experimental data for two and higher quanta transitions come from indirect sources.[153,154] Thus, comparing our results with the scarce experimental information in the spectral range of 0–3650 cm^{-1} (see Table 8), it can be seen that the calculated CO and CF stretchings (ν_2 and ν_4, respectively) are about 15 cm^{-1} lower than the observed frequencies, and the CH bend (ν_3) and stretch (ν_1) are overestimated by 28 and 23 cm^{-1}, respectively. The differences in comparison with experiment become even larger for most of the nine two-quanta transition energies included in Table 8. These large discrepancies to experiments are essentially due of deficiencies in the MP2 PES.[173] Nevertheless the consistent level of agreement of our results with a wide variety of theoretical approaches confirms the reliability of our implementation.

Table 8. Calculated ZPE and Experimental and Calculated Low-Lying Vibrational Transitions of HFCO in the Spectral Range Between 0 and 3660 cm^{-1} (in cm^{-1})[a]

	states[b]	calcd					
no.[c]	(n_1 n_2 n_3 n_4 n_5 n_6)	JW[d]	MCTDH[e]	DVR[f]	FBR[f]	This work[g]	exp.[h]
0	(0 0 0 0 0 0)		4540.0	4542.58	4542.56	4541.7	
1	(0 0 0 0 1 0)	658.1	658.1	659.39	659.37	658.1	662.6
2	(0 0 0 0 0 1)		1019.1	1019.43	1019.43	1019.2	1011.2
3	(0 0 0 1 0 0)	1049.5	1049.5	1050.45	1050.42	1049.5	1064.9
4	(0 0 0 0 2 0)	1314.8				1314.8	1324.1
5	(0 0 1 0 0 0)	1370.2	1370.3	1370.34	1370.33	1370.3	1342.3
7	(0 0 0 1 1 0)	1699.9				1699.9	1719.3
8	(0 1 0 0 0 0)	1821.3	1821.4	1822.17	1822.14	1821.4	1836.8
13	(0 0 0 2 0 0)	2085.4				2085.3	2115.6
17	(0 0 1 1 0 0)	2412.8				2412.9	2412.0
18	(0 1 0 0 1 0)	2474.3				2474.4	2494.2
25	(0 1 0 0 0 1)					2833.3	2841.0
26	(0 1 0 1 0 0)	2863.8				2863.9	2895.0
28	(1 0 0 0 0 0)	3003.1	3003.2	3004.97	3005.08	3003.2	2981.2
37	(0 1 0 0 2 0)	3126.1				3126.2	3150.6
49	(0 2 0 0 0 0)	3623.5				3623.6	3652.8

[a] Results obtained using 10 grid points per degree of freedom and with expansion of the $\hat{U}$ and $\hat{C}$ terms up to second order, i.e., conv = 2. [b] The ordering of the vibrational normal modes follows the standard spectroscopic criteria,[174] i.e., the first five vibrations are those with a' symmetry and the last one is the out-of-plane bending motion (a'' symmetry). The vibrational characterization is: ν_1(CH-str); ν_2(CO-str); ν_3(CH bend); ν_4(CF-str); ν_5(FCO bend). [c] Ordering of the transitions in increasing value of energy. [d] JW stands for Jacobi–Wilson method used in ref [158]. [e] MCTDH stands for multiconfiguration time-dependent Hartree method used in ref 159. [f] Ref 160. [g] The potential energy surface from ref 149 was employed. [h] Experimental values extracted from refs 153 and 154.

■ SUMMARY

In this paper we have presented a full configuration interaction method for calculating low-lying vibrational energy levels of semirigid molecules based on Watson's pure vibrational Hamiltonian that is suitable for small amplitude motions. It is characterized by a finite basis representation in conjunction with a general quasi-analytic scheme for the evaluation of the kinetic energy terms, $\hat{U}$ and $\hat{C}$, expressed as Taylor expansions with respect to the rectilinear normal coordinates around the equilibrium configuration. The generality of this vibrational Hamiltonian allows application to a large range of molecules, and the quasi-analytic treatment, although being restricted to small amplitude motions, avoids the explicit use of geometries close to linearity and problems associated with that region by the singularity of the Watson term. Our approach has been tested for a set of prototype molecules, i.e.,

H_2O, CO_2, and HFCO; the results confirm the applicability and accuracy of the presented method, for example, for determining accurate ZPEs, which are of importance to thermochemistry.[94] In addition, our code also offers a mechanism for analyzing the performance of perturbational approaches based on Taylor expansions of the potential and kinetic energy operators up to arbitrary order.

■ ASSOCIATED CONTENT

S Supporting Information. Figures SI−SIII present the structure of the Hamiltonian matrix, **H**, due to the contributions from different orders of the pseudopotential term ($\hat{U}$), the Coriolis term ($\hat{C}$), and both of them ($\hat{U}+ \hat{C}$), respectively. This material is available free of charge via the Internet at http://pubs.acs.org/.

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: juana@mail.cm.utexas.edu, harding@mail.utexas.edu.

■ ACKNOWLEDGMENT

The authors thank Professor Gregor Diezemann for his involvement during the early stage of this project, Edit Mátyus and Professor Attila G. Császár for providing the potential energy surface of water, and Professor Shigeki Kato for providing the potential energy surface of HFCO. This work has been supported by the United States National Science Foundation (Grant CHE0710146); the United States Department of Energy, Basic Energy Sciences (Contract FG02-07ER15884), the Robert A. Welch Foundation under grant F-1283, the Deutsche Forschungsgemeinschaft (DFG GA 370/5-1), and the Fonds der Chemischen Industrie.

■ REFERENCES

(1) Nielsen, H. H. In *Handbuch der Physik, Part. 1*; ; Flügge, S., Ed.; Spring-Verlag: Berlin, 1959; Vol. 37, p 171.

(2) Papoušek, D.; Aliev, M. R. *Molecular Vibrational-Rotational Spectra*, 1st ed.; Elsevier Scientifics: New York, 1982.

(3) Konen, I. M.; Li, E. X. J.; Lester, M. I.; Vázquez, J.; Stanton, J. F. *J. Chem. Phys.* **2006**, *125*, 074310.

(4) Matthews, D. A.; Vázquez, J.; Stanton, J. F. *Mol. Phys.* **2007**, *105*, 2659.

(5) Vázquez, J.; Stanton, J. F. *Mol. Phys.* **2007**, *105*, 101.

(6) McCoy, A. B.; Sibert, E. L., III *J. Chem. Phys.* **1991**, *95*, 3476.

(7) McCoy, A. B.; Sibert, E. L., III In *Dynamics of Molecules and Chemical Reactions*; Wyatt, R. E., Zhang, J. Z. H., Eds.; Marcel Dekker: New York, 1996; p 151.

(8) Norris, L. S.; Ratner, M. A.; Roitberg, A. E.; Gerber, R. B. *J. Chem. Phys.* **1996**, *105*, 11261.

(9) Bucknell, M. G.; Handy, N. C.; Boys, S. F. *Mol. Phys.* **1974**, *28*, 759.

(10) Whitehead, R. J.; Handy, N. C. *J. Mol. Spectrosc.* **1975**, *55*, 356.

(11) Carney, G. D.; Kern, C. W. *Int. J. Quant. Chem. Suppl.* **1975**, *Y-9*, 317.

(12) Carney, G. D.; Langhoff, S. R.; Curtiss, L. A. *J. Chem. Phys.* **1977**, *66*, 3724.

(13) Carney, G. D.; Sprandel, L. L.; Kern, C. W. *Adv. Chem. Phys.* **1978**, *37*, 305.

(14) Dunn, K. M.; Boggs, J. E.; Pulay, P. *J. Chem. Phys.* **1986**, *85*, 5838.

(15) Dunn, K. M.; Boggs, J. E.; Pulay, P. *J. Chem. Phys.* **1987**, *86*, 5088.

(16) Bowman, J. M. *J. Chem. Phys.* **1978**, *68*, 608.

(17) Bowman, J. M.; Christoffel, K. M; Tobin, F. *J. Phys. Chem.* **1979**, *83*, 905.

(18) Tobin, F.; Bowman, J. M. *Chem. Phys.* **1980**, *47*, 151.

(19) Christoffel, K. M.; Bowman, J. M. *Chem. Phys. Lett.* **1982**, *85*, 220.

(20) Carter, S.; Bowman, J. M.; Harding, L. B. *Spectrochim. Acta* **1997**, *53A*, 1179.

(21) Carter, S.; Bowman, J. M. *J. Chem. Phys.* **1998**, *108*, 4397.

(22) Bégué, D.; Gohaud, N.; Pouchan, C.; Cassan-Chenaï, P.; Liévin, J. *J. Chem. Phys.* **2007**, *127*, 164115.

(23) Harris, D. O.; Engerholm, G. G.; Gwinn, W. D. *J. Chem. Phys.* **1965**, *43*, 1515.

(24) Endres, P. F. *J. Chem. Phys.* **1967**, *47*, 798.

(25) Dickinson, A. S.; Certain, P. R. *J. Chem. Phys.* **1968**, *49*, 4209.

(26) Lill, J. V.; Parker, G. A.; Light, J. C. *Chem. Phys. Lett.* **1982**, *89*, 483.

(27) Heather, R. W.; Light, J. C. *J. Chem. Phys.* **1983**, *79*, 147.

(28) Light, J. C.; Hamilton, I. P.; Lill, J. V. *J. Chem. Phys.* **1985**, *82*, 1400.

(29) Bowman, J. M.; Carrington, T., Jr.; Meyer, H. D. *Mol. Phys.* **2008**, *106*, 2145.

(30) Stanton, J. F.; Gauss, J. *Int. Rev. Phys. Chem.* **2000**, *19*, 61.

(31) Ruden, T. A.; Taylor, P. R.; Helgaker, T. *J. Chem. Phys.* **2003**, *119*, 1951.

(32) Park, S. C.; Braams, B. J.; Bowman, J. M. *J. Theor. Comput. Chem.* **2005**, *4*, 163.

(33) Huang, X.; Braams, B. J.; Bowman, J. M. *J. Chem. Phys.* **2005**, *122*, 044308.

(34) Xie, Z.; Braams, B. J.; Bowman, J. M. *J. Chem. Phys.* **2005**, *122*, 224307.

(35) Sharma, A. R.; Braams, B. J.; Carter, S.; Shepler, B. C.; Bowman, J. M. *J. Chem. Phys.* **2009**, *130*, 174301.

(36) Li, G. Y.; Wang, S.-W.; Rosenthal, C.; Rabitz, H. *J. Math. Chem.* **2001**, *30*, 1.

(37) Carter, S.; Culik, S. J.; Bowman, J. M. *J. Chem. Phys.* **1997**, *107*, 10458.

(38) Jung, J. O.; Gerber, R. B. *J. Chem. Phys.* **1996**, *105*, 10332.

(39) Beck, M. H.; Jäckle, A.; Worth, G. A.; Meyer, H. D. *Phys. Rep.* **2000**, *324*, 1.

(40) Mátyus, E.; Czakó, G.; Sutcliffe, B. T.; Császár, A. G. *J. Chem. Phys.* **2007**, *127*, 084102.

(41) Meyer, R.; Günthard, H. H. *J. Chem. Phys.* **1968**, *49*, 1510.

(42) Pickett, H. M. *J. Chem. Phys.* **1972**, *56*, 1715.

(43) Lukka, T. J. *J. Chem. Phys.* **1995**, *102*, 3945.

(44) Schwenke, D. W. *J. Chem. Phys.* **2003**, *118*, 10431.

(45) Watson, J. K. G. *J. Mol. Spectrosc.* **2004**, *228*, 645.

(46) Fehrensen, B.; Luckhaus, D.; Quack, M. *Chem. Phys. Lett.* **1999**, *300*, 312.

(47) Luckhaus, D. *J. Chem. Phys.* **2000**, *113*, 1329.

(48) Luckhaus, D. *J. Chem. Phys.* **2003**, *118*, 8797.

(49) Lauvergnat, D.; Nauts, A. *J. Chem. Phys.* **2002**, *116*, 8560.

(50) Lauvergnat, D.; Baloïtcha, E.; Dive, G.; Desouter-Lecomte, M. *Chem. Phys.* **2006**, *326*, 500.

(51) Yurchenko, S. N.; Thiel, W.; Jensen, P. *J. Mol. Spectrosc.* **2007**, *245*, 126.

(52) Makarewicz, J. In *Computational Molecular Spectroscopy*; Jensen, P., Bunker, P. R., Eds.; Wiley: Chichester, U.K., 2000; p 391.

(53) Mátyus, E.; Czakó, G.; Császár, A. G. *J. Chem. Phys.* **2009**, *130*, 134112.

(54) Nauts, A.; Chapuisat, X. *Mol. Phys.* **1985**, *55*, 1287.

(55) Chapuisat, X.; Belafhal, A.; Nauts, A. *J. Mol. Spectrosc.* **1991**, *149*, 274.

(56) Gatti, F.; Iung, C.; Leforestier, C.; Chapuisat, X. *J. Chem. Phys.* **1999**, *111*, 7236.

(57) Gatti, F. *J. Chem. Phys.* **1999**, *111*, 7225.

(58) Gatti, F.; Muñoz, C.; Iung, C. *J. Chem. Phys.* **2001**, *114*, 8275.

(59) Mladenović, M. *J. Chem. Phys.* **2000**, *112*, 1070.

(60) Mladenović, M. *J. Chem. Phys.* **2000**, *112*, 1082.

(61) Wang, X.-G.; Carrington, T., Jr. *J. Chem. Phys.* **2004**, *121*, 2937.

(62) Sutcliffe, B. T. In *Conceptual Trends in Quantum Chemistry*; Kryachko, S., Calais, J. L., Eds.; Kluwer: Dordrecht The Netherlands, 1994; p 53; and references therein.

(63) Sutcliffe, B. T.; Tennyson, J. *Int. J. Quantum Chem.* **1991**, *39*, 183.

(64) Tennyson, J.; Sutcliffe, B. T. *J. Chem. Phys.* **1982**, *77*, 4061.

(65) Brocks, G.; Avoird, A. v. D.; Sutcliffe, B. T. *Mol. Phys.* **1983**, *50*, 1025.

(66) Handy, N. C. *Mol. Phys.* **1987**, *61*, 207.

(67) Császár, A. G.; Handy, N. C. *J. Chem. Phys.* **1995**, *102*, 3962.

(68) Vendrell, O.; Gatti, F.; Meyer, H.-D. *J. Chem. Phys.* **2007**, *127*, 184303.

(69) Gatti, F.; Iung, C. *Phys. Rep.* **2009**, *484*, 1.

(70) Watson, J. K. G. *Mol. Phys.* **1968**, *15*, 479.

(71) Watson, J. K. G. *Mol. Phys.* **1970**, *19*, 465.

(72) Webster, F.; Huang, M.-J.; Wolfsberg, M. *J. Chem. Phys.* **1981**, *75*, 2306.

(73) Maessen, B.; Wolfsberg, M. *J. Chem. Phys.* **1984**, *80*, 4651.

(74) Maessen, B.; Wolfsberg, M. *J. Phys. Chem.* **1984**, *88*, 6420.

(75) Chen, C.-L.; Maessen, B.; Wolfsberg, M. *J. Chem. Phys.* **1985**, *83*, 1795.

(76) Searles, D. J.; von Nagy-Felsobuki, E. I. *J. Chem. Phys.* **1991**, *95*, 1107.

(77) Wang, F.; von Nagy-Felsobuki, E. I. *Mol. Phys.* **1992**, *77*, 1197.

(78) Searles, D.; von Nagy-Felsobuki, E. I. *Ab Initio Variational Calculations of Molecular Vibration-Rotation Spectra. Lecture Notes in Chemistry*, 1st ed.; Springer-Verlag: Berlin, Germany, 1993; No. 61.

(79) Page, A. J.; von Nagy-Felsobuki, E. I. *Mol. Phys.* **2007**, *105*, 2527.

(80) Carter, S.; Bowman, J. M.; Handy, N. C. *Theor. Chem. Acc.* **1998**, *100*, 191.

(81) Carter, S.; Bowman, J. M. *J. Phys. Chem. A* **2000**, *104*, 2355.

(82) Bowman, J. M.; Huang, X. C.; Carter, S. *Spectrochim. Acta* **2002**, *58A*, 839.

(83) Aoyagi, M.; Gray, S. K. *J. Chem. Phys.* **1991**, *94*, 195.

(84) Seideman, T.; Miller, W. H. *J. Chem. Phys.* **1992**, *97*, 2499.

(85) Balint-Kurti, G. G.; Pulay, P. *J. Mol. Struct. (THEOCHEM)* **1995**, *341*, 1.

(86) Yonehara, T.; Yamamoto, T.; Kato, S. *Chem. Phys. Lett.* **2004**, *393*, 98.

(87) Wang, F.; McCourt, F. R. W.; von Nagy-Felsobuki, E. I. *J. Mol. Struct. (THEOCHEM)* **2000**, *497*, 227.

(88) Mátyus, E.; Šimunek, J.; Császár, A. G. *J. Chem. Phys.* **2009**, *131*, 074106.

(89) Rauhut, G. *J. Chem. Phys.* **2004**, *121*, 9313.

(90) Rauhut, G.; Hrenar, T. *Chem. Phys.* **2008**, *346*, 160.

(91) Heislbetz, S.; Rauhut, G. *J. Chem. Phys.* **2010**, *132*, 124102.

(92) Louck, J. D. *J. Mol. Spectrosc.* **1976**, *61*, 107.

(93) Makushkin, Y. S.; Ulenikov, O. N. *J. Mol. Spectrosc.* **1977**, *68*, 1.

(94) Harding, M. E.; Vázquez, J.; Ruscic, B.; Wilson, A. K.; Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **2008**, *128*, 114111.

(95) Amat, G.; Henry, L. *Cah. Phys.* **1958**, *12*, 273.

(96) Amat, G.; Henry, L. *Cah. Phys.* **1961**, *12*, 472.

(97) Hougen, J. T. *J. Chem. Phys.* **1962**, *36*, 519.

(98) Wilson, E. B., Jr.; Decius, J. C.; Cross, P. C. *Molecular Vibration. The Theory of Infrared and Raman Vibrational Spectra*; Dover Publications: New York, 1955.

(99) *Handbook of Mathematical Functions*; Abramowitz, M., Stegun, I. A., Eds. Dover Publications: New York, 1972.

(100) Aliev, M. R. *Opt. Spectrosc.* **1969**, *26*, 463.

(101) Mladenović, M. *Spectrochim. Acta* **2002**, *58A*, 795.

(102) Huber, D. *Int. J. Quantum Chem.* **1985**, *28*, 245.

(103) Aliev, M. R.; Watson, J. K. G. In *Molecular Spectroscopy: Modern Research.* ; Rao, K. N., Ed.; Academic Press: Orlando, FL, 1985; Vol. III, p 1.

(104) Wick, G. C. *Phys. Rev.* **1950**, *80*, 268.

(105) Mandl, F.; Shaw, G. *Quantum Field Theory*; Wiley: New York, 1984.

(106) The ladder operator formalism used here does not correspond to a complete second quantization formalism for creation and annihilation operators. Concepts and premises, as for example, vacuum and creation/annihilation of normal modes are not defined.

(107) The contributions from the potential ($\hat{V}$) and pseudopotential ($\hat{U}$) could be integrated jointly numerically or analytically, which would lead to a small reduction of the computational cost. However, the main purpose of this study is to analyze the magnitude of each contribution independently and to point out the advantages of the quasi-analytic calculation of KEO expressed as a Taylor expansion around the equilibrium configuration.

(108) Nemes, L. In *Vibrational Spectra and Structure*; During, J. R., Ed.; Elsevier: Amsterdam, The Netherlands, 1981; Vol. 10, p 395.

(109) Lanczos, C. *J. Res. Natl. Bur. Stand.* **1950**, *45*, 255.

(110) Cullum, J. K.;Willoughby, R. A. *Lanczos Algorithms for Large Symmetry Eigenvalue Computations*; Birkhäuser: Boston, MA, 1985; Vol. 1-2.

(111) Simon, H. D. *Math. Comput.* **1984**, *42*, 115.

(112) Golub, G. H.; Van Loan, C. F. *Matrix Computations*; Johns Hopkins University Press: Baltimore, MD, 1996.

(113) The most efficient way for the reorthogonalization of a Hermitian matrix was found to be a formulation of the modified Gram−Schmidt method, which heavily involves highly optimized BLAS and LAPACK routines (see http://www.netlib.org/lapack/). A complete Cholesky decomposition of the overlap matrix of the complete set of Lanczos vectors is performed, and the Lanczos vectors are then multiplied in place with the inverse triangular matrix resulting from the Cholesky procedure.

(114) Furtenbacher, T.; Császár, A. G.; Tennyson, J. *J. Mol. Spectrosc.* **2007**, *245*, 115.

(115) Polyansky, O. L.; Császár, A. G.; Shirin, S. V.; Zobov, N. F.; Barletta, P.; Tennyson, J.; Schwenkey, D. W.; Knowles, P. J. *Science* **2003**, *299*, 539.

(116) Barletta, P.; Shirin, S. V.; Zobov, N. F.; Polyansky, O. L.; Tennyson, J.; Valeev, E. F.; Császár, A. G. *J. Chem. Phys.* **2006**, *125*, 204307.

(117) DEWE stands for discrete variable representation (D) of the Eckart−Watson (EW) Hamiltonian with exact inclusion of an arbitrary potential energy function (E).

(118) Schuurman, M. S.; Muir, S. R.; Allen, W. D.; Schaefer, H. F., III *J. Chem. Phys.* **2004**, *120*, 11586.

(119) Wheeler, S. E.; Robertson, K. A.; Allen, W. D.; Schaefer, H. F., III; Bomble, Y. J.; Stanton, J. F. *J. Chem. Phys. A* **2007**, *111*, 3819.

(120) Because of its small magnitude, the $\hat{U}$ term has traditionally been ignored in the perturbational approaches except for the leading term, $-1/4\Sigma_\alpha B_e^\alpha$, which contributes to the zero-point energy of nonlinear polyatomic molecules.

(121) Bartholomae, R.; Martin, D.; Sutcliffe, B. T. *J. Mol. Spectrosc.* **1981**, *87*, 367.

(122) Carter, S.; Handy, N. C. *J. Mol. Spectrosc.* **1982**, *95*, 9.

(123) In the case of numerical approaches this problem is addressed in different ways, see for example ref 40 and also: Scivetti, I.; Kohanoff, J.; Gidopoulos, N. I. *Int. J. Quantum Chem.* **2011**, *111*, 307.

(124) Sarka, K.; Bunker, P. R. *J. Mol. Spectrosc.* **1987**, *122*, 259.

(125) Epa, V. C.; Bunker, P. R. *J. Mol. Spectrosc.* **1991**, *150*, 511.

(126) Fermi, E. *Z. Physik* **1931**, *71*, 250.

(127) Suzuki, I. *J. Mol. Spectrosc.* **1968**, *25*, 479.

(128) Cihla, Z.; Chédin, A. *J. Mol. Spectrosc.* **1971**, *40*, 337.

(129) Chédin, A.; Cihla, Z. *J. Mol. Spectrosc.* **1973**, *47*, 554.

(130) Jobard, I.; Chédin, A. *J. Mol. Spectrosc.* **1975**, *57*, 464.

(131) Chédin, A. *J. Mol. Spectrosc.* **1979**, *76*, 430.

(132) Lacy, M. *Mol. Phys.* **1982**, *45*, 253.

(133) Steele, D.; Person, W. B.; Brown, K. G. *J. Phys. Chem.* **1981**, *85*, 2007.

(134) Allen, W. D.; Yamaguchi, Y.; Császár, A. G.; Clabo, D. A., Jr.; Remington, R. B.; Schaefer, H. F., III *Chem. Phys.* **1990**, *145*, 427.

(135) Maslen, P. E.; Jayatilaka, D.; Colwell, S. M.; Amos, R. D.; Handy, N. C. *J. Chem. Phys.* **1991**, *95*, 7409.

(136) Martins Filho, H. P. *Spectroschim. Acta.* **2002**, *58A*, 2621.

(137) Martin, J. M. L.; Taylor, P. R.; Lee, T. J. *Chem. Phys. Lett.* **1993**, *205*, 535.

(138) Császár, A. G. *J. Phys. Chem.* **1992**, *96*, 7898.

(139) Rodriguez-Garcia, V.; Hirata, S.; Yagi, K.; Hirao, K.; Taketsugu, T.; Schweigert, I.; Tasumi, M. *J. Chem. Phys.* **2007**, *126*, 124303.

(140) Czakó, G.; Furtenbacher, T.; Császár, A. G.; Szalay, V. *Mol. Phys.* **2004**, *102*, 2411.

(141) A grid of 4153 energy points calculated at the fc-CCSD(T)/ANO1 level of theory was used to fit a Taylor expansion expression of the potential energy in terms of dimensionless normal coordinates. Three of the sextic force constants resulted to have different sign from the experimental values provided by Chédin. The results we obtained for those constants: Φ_{113333} = 2.3 cm^{-1}, Φ_{111133} = 4.9 cm^{-1}, and Φ_{333333} = 10.5 cm^{-1}; (fc) stands for "frozen core", i.e., that only the valence electrons were correlated in the post Hartree−Fock treatment.

(142) In the present context, values are considered converged when a *n*th order and a $(n + 1)$th order in the expansion of $\hat{U}$ and $\hat{C}$ give differences smaller than 10^{-2} and 10^{-1} cm^{-1} for CO_2 and HFCO, respectively.

(143) Morokuma, K.; Kato, S.; Hirao, K. *J. Chem. Phys.* **1980**, *72*, 6800.

(144) Morokuma, K.; Kato, S. In *Potential energies surfaces and dynamics calculations for chemical reactions and molecular energy transfer*; Truhlar, D. G., Ed.; Plenum: New York, 1981; p 243.

(145) Goddard, J. D.; Schaefer, H. F., III *J. Chem. Phys.* **1990**, *93*, 4907.

(146) Kamiya, K.; Morokuma, K. *J. Chem. Phys.* **1991**, *94*, 7287.

(147) Francisco, J. S.; Zhao, Y. *J. Chem. Phys.* **1992**, *96*, 7587.

(148) Wei, T.-G.; Wyatt, R. E. *J. Phys. Chem.* **1993**, *97*, 13580.

(149) Yamamoto, T.; Kato, S. *J. Chem. Phys.* **1997**, *107*, 6114.

(150) Bartlett, R. J. *Annu. Rev. Phys. Chem.* **1981**, *32*, 359.

(151) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.

(152) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007.

(153) Viel, A.; Leforestier, C. *J. Chem. Phys.* **2000**, *112*, 1212.

(154) Leforestier, C.; Viel, A.; Gatti, F.; Mũnoz, C.; Iung, C. *J. Chem. Phys.* **2001**, *114*, 2099.

(155) Iung, C.; Ribeiro, F. *J. Chem. Phys.* **2005**, *123*, 174105.

(156) Iung, C.; Ribeiro, F.; Sibert, E. L., III *J Phys. Chem. A* **2006**, *110*, 5420.

(157) Ribeiro, F.; Iung, C.; Leforestier, C. *J. Chem. Phys.* **2005**, *123*, 054106.

(158) Ribeiro, F.; Iung, C.; Leforestier, C. *J. Theor. Comput. Chem.* **2003**, *2*, 609.

(159) Pasin, G.; Gatti, F.; Iung, C.; Meyer, H. D. *J. Chem. Phys.* **2006**, *124*, 194304.

(160) Wang, X.-G.; Carrington, T., Jr. *J. Chem. Phys.* **2009**, *130*, 094101.

(161) Green, W. H.; Jayatilaka, D.; Willetts, A.; Amos, R. D.; Handy, N. C. *J. Chem. Phys.* **1990**, *93*, 4965.

(162) Vázquez, J.; Stanton, J. F. *Mol. Phys.* **2006**, *104*, 377.

(163) Mills, I. M. In *Modern Spectroscopy: Modern Research*; Rao, K. N., Matthews, C. W., Eds.; Academic Press: New York, 1972; Vol. I, p 115.

(164) Friesner, R. A.; Bentley, M.; Menou, M.; Leforestier, C. *J. Chem. Phys.* **1993**, *99*, 324.

(165) Choi, Y. S.; Moore, C. B. *J. Chem. Phys.* **1991**, *94*, 5414.

(166) Choi, Y. S.; Moore, C. B. *J. Chem. Phys.* **1992**, *97*, 1010.

(167) Choi, Y. S.; Moore, C. B. *J. Chem. Phys.* **1995**, *103*, 9981.

(168) Morgan, H. W.; Staats, P. A.; Goldstein, J. H. *J. Chem. Phys.* **1956**, *25*, 337.

(169) Stratton, R. F.; Nielsen, A. H. *J. Mol. Spectrosc.* **1960**, *4*, 373.

(170) Kattenberg, H. W.; Elst, R.; Oskam, A. *J. Mol. Spectrosc.* **1971**, *39*, 29.

(171) Mizuno, M.; Saëki, D. *Spectrochim. Acta* **1978**, *34A*, 407.

(172) Wong, M.; Johns, J. W. C.; Mckellar, A. R. W. *J. Mol. Spectrosc.* **1982**, *94*, 79.

(173) To explore this possibility, the ZPE and the 50 lowest vibrational energy levels were recomputed using geometric, harmonic, and kinetic energy parameters calculated at the fc-CCSD(T)/ANO2 level of theory and taking only the anharmonic potential from ref 149. In this case the agreement with experiment is improved with maximum difference of only 16.7 cm^{-1} for the CH stretching, ν_1, while most of the transitions are predicted with discrepancies smaller than 10 cm^{-1}. In addition, the same states were calculated with vibrational second-order perturbation theory (VPT2) using harmonic, cubic, and (semi)diagonal quartic force fields as well as kinetic energy parameters computed at the fc-CCSD(T)/ANO2 level. These values are in excellent agreement with experiment considering the incompleteness in the description of the anharmonic potential (only quartic diagonal and semidiagonal force constants are included in a VPT2 treatment); states are predicted with less than 5.0 cm^{-1} difference from experiment, except for the first overtone of the C−F stretching, $2\nu_4$, for which this difference is 14.1 cm^{-1}.

(174) Herzberg, G. *Molecular Spectra and Molecular Structure*; van Nostrand: New York, 1950; Vol. I−III.

JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

First Principles Simulations of the Infrared Spectrum of Liquid Water Using Hybrid Density Functionals

Cui Zhang,[†] Davide Donadio,[†,⊥] François Gygi,[‡,§] and Giulia Galli*[,†,∥]

[†]Department of Chemistry, University of California, Davis, California 95616, United States
[‡]Department of Applied Science, University of California, Davis, California 95616, United States
[§]Department of Computer Science, University of California, Davis, California 95616, United States
[∥]Department of Physics, University of California, Davis, California 95616, United States

S *Supporting Information*

ABSTRACT: We show that first principles hybrid functional (PBE0) simulations of the infrared spectrum of liquid water yields a much better agreement with experimental results than a semilocal functional description; in particular, the quantitative accord with measured stretching and bending bands is very good. Such an improved description stems from two effects: a more accurate account, at the PBE0 level of theory, of the vibrational properties of the monomer and dimer and an underlying structural model for the liquid with a smaller number of hydrogen bonds and oxygen coordination than those obtained with semilocal functionals. The average electronic gap of the liquid is increased by 60% with respect to the PBE value, when computed at the PBE0 level of theory, and is in fair agreement with experimental results.

1. INTRODUCTION

Providing a theoretical description of the properties of liquid water and acquiring the ability to simulate them have been central topics in physical chemistry and the subject of intense activities for many decades. However, a description of hydrogen bonding in liquid water is still the subject of debate,[1−3] with controversies recently stirred, for example, by spectroscopic measurements[1] pointing at an oxygen coordination and the number of hydrogen bonds substantially smaller than those in ice.

Building on the pioneering work of Rahman and Stillinger,[4] molecular dynamics (MD) and Monte Carlo (MC) techniques have been extensively used to simulate water in different thermodynamic states, by using classical interatomic potentials fitted to specific sets of experimental data. A number of force fields has been employed, including simple point charge (SPC) models[5] (either rigid[6] or flexible[7−9]) or more sophisticated parametrizations such as TIP4P[10] and TIP5P.[11] Although useful and successful in describing a number of structural and dynamical properties of aqueous systems, classical potentials often lack transferability to thermodynamic conditions and environments different from those for which they were fitted; in addition, they cannot describe processes where bond breaking and formation occur, and they may not be used to analyze spectroscopic measurements where electronic effects play an important role, for example, vibrational spectra.[12] Developing theoretical frameworks capable of accounting not only for structural and thermodynamic properties but also for spectroscopic data is of the greatest importance to interpreting complex measurements and thus gaining insight into hydrogen bonding in the liquid.

The ability to account for electronic effects and carry out MD simulations where the electronic structure of liquid water, treated as a condensed system, is computed at each step of the dynamics came with the advent of the Car−Parrinello (CP) method.[13] This framework also opened the way to addressing spectroscopic properties from first principles. Within this approach, the electronic structure of the liquid is described with density functional theory (DFT), and different levels of approximations for the exchange correlation functional have been used. Among semilocal functionals, the most widely adopted are the gradient corrected functionals BLYP[14,15] and PBE.[16,17] The first simulation using BLYP appeared in 1993,[18] and it gave important, qualitative information about a DFT-based description of the structure of liquid water under ambient conditions. It has now been established by several authors[19−24] that if the electronic structure of the liquid is properly converged (that is, numerical inaccuracies present in some early simulations are eliminated), both PBE and BLYP functionals yield overstructured pair correlation functions $g_{OO}(r)$ under ambient conditions, as compared to experimental results, a self-diffusion constant that is greatly underestimated, and a number of hydrogen bonds that is most likely overestimated. The errors of PBE and BLYP on both structural and self-diffusion properties can be artificially "corrected" by a temperature shift of ∼50 to 100 K, depending on whether rigid or flexible water models[25,26] are employed.[27] Therefore, a DFT description with BLYP and PBE gradient corrected functionals can give a reasonably good, qualitative account of water structure, although the accord with experimental results is not fully quantitative. In addition, both PBE and BLYP functionals yield a reasonably good description of the vibrational properties of the liquid,[28−33] although a quantitative discrepancy with experimental results remains, for example, a sizable red shift of about 200 cm^{-1} of the infrared (IR) stretching

Received: February 9, 2011
Published: March 21, 2011

band, indicating that improvements in the description of hydrogen bonds in the liquid are necessary.

Such improvements may come from the use of a higher level of theory than provided by local and semilocal density functionals. Hybrid functionals have been used by several authors[34−38] to investigate the structural properties of the liquid. Substantial differences have been found between the PBE and BLYP descriptions and that provided by empirical functionals such as HCTH.[34] Two studies using approximate forms of the hybrid functionals PBE0[39] and B3LYP[40] have also appeared in the literature.[35,36] Todorova et al.[35] found a more diffusive and less structured liquid, at the PBE0 level of theory, than with PBE, in qualitative agreement with the report of Li et al.;[37] Guidon et al.[36] reported instead negligible differences between pair correlation functions computed with PBE and PBE0, indicating negligible changes in the number and character of hydrogen bonds found in the system. The study of water clusters reported in refs 41 and 42 is consistent with the findings of a less structured liquid within PBE0; indeed it was shown that PBE0 yields smaller binding energies[41,42] and smaller polarizabilities[43] than PBE. However, in ref 35, larger binding energies are reported when using PBE0. We note that approximations in the implementation of the exact exchange operator have been adopted in both refs 35 and ref 36 and that the two studies were carried out in different ensembles.[44]

The great majority of first principles simulations have so far neglected the quantum nature of the proton. The significance of treating the proton as a quantum mechanical particle has been recently investigated by Morrone and Car[45] using path integral (PI) simulations and a DFT description of the electronic structure with the BLYP functional. These authors showed that the agreement with experimental results for the liquid structural properties is improved when proton quantum effects are included, consistent with the results obtained with PI calculations and empirical potentials.[46] PI simulations have also been carried out using *ab initio* based potentials, finding an effect similar to that reported in ref 45, that is a softening of the oxygen−oxygen pair correlation function corresponding to less structured hydrogen bonded networks, when treating protons quantum mechanically.[47,48]

Most of the theoretical investigations at the DFT level (and all of those using nonlocal functionals) have focused on the structural properties of the liquid, with few studies of vibrational[22,29,31−33] and electronic[2,3,49−52] properties appearing in the past several years. In order to gain a better understanding of water and in particular of hydrogen bonding, it is important to develop tools to compute accurate spectroscopic quantities and thus compare directly with a wealth of available measurements, in addition to structure factors and pair correlation functions. The IR spectra of water were first studied at the PBE level by Sharma et al.,[29] yielding a qualitative agreement with experimental results. Subsequent studies[32,33] outlined the importance of intermolecular dipolar correlations in shaping the IR spectrum of the liquid, and in particular of the complex stretching band. However quantitative discrepancies between theory and experiments remain, for example, on the position of the IR stretching band, and the origin of these discrepancies is not well understood. For example, it is yet unknown how the inaccuracies in the description of the liquid structural properties and thus of hydrogen bonding found in first principles simulations impact our understanding of vibrational measurements.

In this paper, we focus on the vibrational spectroscopy of the liquid,[53] and we report IR spectra computed using the nonlocal functional PBE0. So far, no vibrational study of liquid water or ice using nonlocal functionals has appeared in the literature. We find a much better agreement with experimental results, especially in the description of the stretching and bending bands, than obtained with semilocal functionals (at the GGA level); two main reasons are responsible for our improved description: a better account of the vibrational properties of the monomer and dimer and an underlying structural model for the liquid with a smaller number of hydrogen bonds and a smaller effective molecular dipole than those obtained at the GGA level. These results have important implications for the study of vibrational properties of water in contact with surfaces. The rest of the paper is organized as follows: our methodological approach is presented in section 2, and our results are discussed in section 3, for both structural and vibrational properties. Section 4 contains our conclusions.

2. METHODS

We performed first principles molecular dynamics simulations of heavy water D_2O using the *Qbox* code,[54] with cubic cells containing 32 molecules, and all of our results for vibrational spectra have been obtained at a fixed density of 1.108 g/cm^3. We employed both semilocal (PBE) and hybrid (PBE0) exchange and correlation functionals. In the case of simulations using PBE, we compared our results with those obtained with a 96 molecule cell. The equilibrium density of water using PBE has been predicted to be 0.85−0.90 g/cm^3,[55,56] that is, ≃10−15% smaller than in experiments. We carried out simulations at this lower density with the PBE functional at ~400 K, and we found a worse agreement with neutron diffraction data than at the experimental density. In particular, the liquid turned out to be more structured than at higher density, as indicated, e.g., by an analysis of the pair correlation function $g_{OO}(r)$ (see Figure 1 in the Supporting Information). Given these results and given that the equilibrium density of water at the PBE0 level of theory has not yet been determined, we carried out simulations of vibrational spectra at the experimental equilibrium density of deuterated water, and we defer investigations as a function of density to a later study. Although it is in principle possible to compute the equilibrium density of the liquid using the PBE0 functional, this would be computationally very intensive, as PBE0 calculations are substantially heavier, from a computational standpoint, than those using PBE.[57]

In all liquid simulations, we adopted a plane wave (PW) basis set and norm-conserving pseudopotentials (PP)[58] with a kinetic energy cutoff of 85 Ry. The computation of the Hartree−Fock exchange operator included in the PBE0 functional was carried out without truncation of the range of the Coulomb interaction, unlike in previous studies,[35,36] and in such a way to ensure quadratic convergence with respect to Brillouin zone integration. In particular, calculations performed at the Γ point of the Brillouin zone correctly include divergent terms stemming from the long range of the Coulomb potential. An efficient parallelization scheme implemented in *Qbox* mitigates the high computational cost of the Hartree−Fock exchange energy when using PW basis sets.[59] Simulations were carried out with a time step of 10 a.u. in the NVE ensemble at several temperatures (T) ranging from 370 to 470 K, within a Born−Oppenheimer (BO) framework. At each T, the system was equilibrated for at least 10 ps and up to 20 ps, and trajectories were collected for 17 ps. The electronic contributions to the molecular dipole moment were computed using maximally localized Wannier functions[60] (MLWFs), evaluated at each MD step with the algorithm proposed in ref 61. The IR absorption coefficient per unit length was obtained within linear response theory from the Fourier transform of the time correlation function of the system's dipole moment:[62]

$$\alpha(\omega) = \frac{2\pi\omega^2\beta}{3cVn(\omega)}\int_{-\infty}^{\infty} \mathrm{d}t e^{-i\omega t}\langle \sum_{ij} \vec{\mu}_i(0)\cdot\vec{\mu}_j(t)\rangle \qquad (1)$$

where $n(\omega)$ is the refractive index, V is the volume, $\beta = 1/k_BT$ is the inverse temperature, and $\vec{\mu}_i$ is the molecular dipole moment. Before discussing the IR spectra of liquid water, we present results obtained for the vibrational properties of the water monomer and dimer using the PBE and PBE0 functionals, and we discuss the effect of the basis set (energy cutoff) and PP on our results.

3. RESULTS AND DISCUSSION

3.1. Vibrational Properties of the Water Molecule and the Water Dimer. We first analyzed the effect of the PP used in our calculations by comparing results at the PBE and PBE0 levels of theory with all electron (AE) and experimental results for the H_2O molecule and the H_2O dimer (see details of the calculations in the Supporting Information). AE results are not available for the deuterated water molecule and dimer. We found that for converged basis sets (200 Ry) at the PBE level, PP and AE results differ by a few wavenumbers, showing the excellent performance of converged PP calculations. Deviation from AE results appears to be bigger in the case of PBE0 (up to 34 cm^{-1}), most likely because we have used PP generated within PBE.

Table 1. Vibrational Frequencies (cm^{-1}) of the Deuterated Water Monomer D_2O[a,b]

	Ecut (Ry)	pseudopotentials	ν_1	ν_2	ν_3
PBE	85	Hamann	2605	1164	2735
PBE	200	Hamann	2616	1165	2746
PBE	85	HSCV	2655	1164	2779
PBE	200	HSCV	2656	1165	2780
PBE0	85	Hamann	2717	1195	2850
PBE0	200	Hamann	2719	1199	2851
PBE0	85	HSCV	2763	1197	2890
PBE0	200	HSCV	2758	1199	2885
expt. harm.[65]			2764	1206	2889
expt. anharm.[65]			2671	1178	2788

[a] Simulations of the water molecule were carried out in a cubic cell with L = 30 Bohr. [b] ν_1, symmetric stretching; ν_2, bending; ν_3, asymmetric stretching.

In Tables 1 and 2, we show calculated vibrational frequencies of the D_2O molecule and the dimer obtained with two different norm-conserving PPs: Hamann[58] and Hamann–Schlüter–Chiang–Vanderbilt (HSCV).[63,64] As for H_2O, we compare our results to experimental harmonic frequencies, because calculations of vibrational frequencies at T = 0 by, e.g., finite differences, do not include anharmonic effects, unlike our MD simulation results at finite T, which account for classical anharmonic effects. For converged basis sets (200 Ry), all of the modes of the monomer are underestimated, with respect to experimental results, by up to 5.4% with the PBE functional, while converged PBE0 results are in excellent agreement with harmonic frequencies extracted from measured spectra.[65] Experimental harmonic frequencies are not available for low frequency modes of the dimers; however, from a comparison with anharmonic frequencies, and given the shape of the potential energy curve of the dimer, it is likely that both PBE and PBE0 slightly overestimate the low frequency modes (see Tables 3 and 4 in the Supporting Information).

In the case of intramolecular frequencies of $(D_2O)_2$ within PBE0, the comparison with harmonic frequencies extracted from experimental data[66] shows that stretching frequencies are slightly underestimated when using Hamann PP with a converged basis set (200 Ry), with the errors on the donor being slightly larger than those on the acceptor. HSCV PP yields instead a slight overestimate of the stretching frequencies. Bending frequencies are consistently underestimated by 9–10 cm^{-1} with either type of PP.

The computed binding energy of the dimer is 5.12 kcal/mol with PBE0 and 5.24 kcal/mol with PBE, in good agreement with the data obtained with AE calculations and Gaussian basis sets.[68] We find that the difference between the binding energies obtained with the two functionals, $\Delta E_{\mathrm{binding}} = E_{\mathrm{binding}}^{\mathrm{PBE}} - E_{\mathrm{binding}}^{\mathrm{PBE0}}$, is 0.12 kcal/mol, to be compared with 0.15 kcal/mol from ref 68. Overall, our results for the vibrational frequencies of the water monomer and dimer show that the PBE0 functional very much improves the description obtained at the PBE level, and thus it appears to be a promising exchange correlation functional to

Table 2. Vibrational Frequencies (cm^{-1}) of the Deuterated Water Dimer $(D_2O)_2$[a,b]

	Ecut (Ry)	pseudopotentials	ν_1	ν_2	ν_3	ν_4	ν_5	ν_6
PBE	85	Hamann	2601	1164	2698	2492	1176	2729
PBE	200	Hamann	2613	1165	2710	2510	1177	2741
PBE	85	HSCV	2650	1164	2772	2549	1175	2744
PBE	200	HSCV	2652	1165	2774	2549	1176	2745
PBE0	85	Hamann	2714	1196	2845	2624	1208	2817
PBE0	200	Hamann	2716	1199	2846	2631	1212	2819
PBE0	85	HSCV	2759	1196	2885	2673	1207	2859
PBE0	200	HSCV	2754	1199	2879	2666	1211	2853
expt. harm.[66]			2738	1209	2857	2689	1221	2838
expt. anharm.[67]			2650	1182	2757	2599	1193	2738

[a] Simulations of the water dimer were carried out in a cubic cell with L = 30 Bohr. [b] ν_1, symmetric stretching of acceptor; ν_2, bending of acceptor; ν_3, asymmetric stretching of acceptor; ν_4, symmetric stretching of donor; ν_5, bending of donor; ν_6, asymmetric stretching of donor.

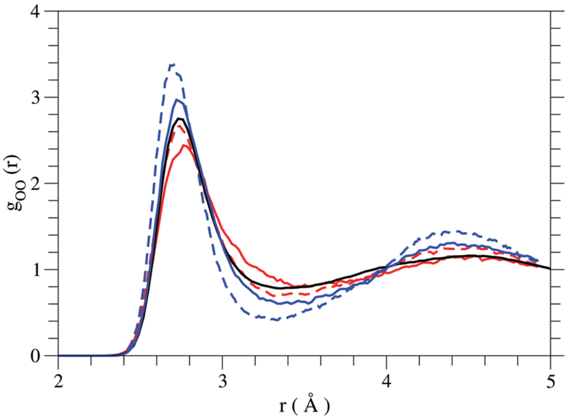


Figure 1. Comparison of oxygen−oxygen pair correlation functions for systems consisting of 32 water molecules, obtained with the PBE0 functional at 438 ± 29 K (solid red) and 374 ± 27 K (solid blue), and the PBE functional at 439 ± 29 K (dash red) and 367 ± 25 K (dash blue). The experimental result at room temperature is displayed by the black line.[70]

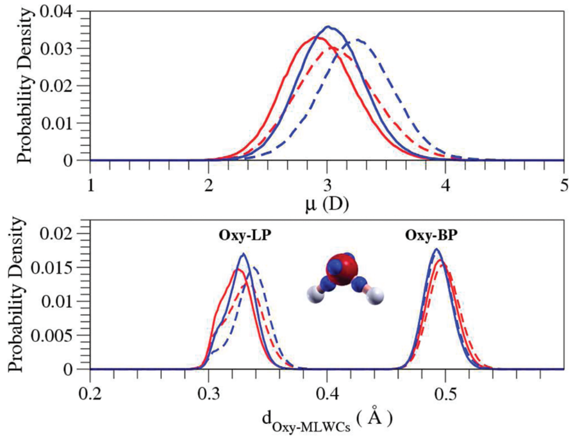


Figure 2. Distributions of molecular dipole moments and distances between oxygen and maximally localized Wannier centers (MLWCs), obtained with the PBE0 functional at 438 ± 29 K (solid red) and 374 ± 27 K (solid blue), and the PBE functional at 439 ± 29 K (dash red) and 367 ± 25 K (dash blue). The inset shows the positions of the centers of maximally localized Wannier functions in a water molecule. The two orbitals centered close to the OD bonds are bond pair (BP) orbitals; the other two are lone pair (LP) orbitals.

study the vibrational properties of liquid water. Our results for the liquid are presented in the next section using an 85 Ry cutoff and the Hamann PP,[69] after comparing structural properties obtained with PBE and PBE0 functionals.

3.2. Vibrational Properties and Hydrogen Bonding of Liquid Deuterated Water. Figure 1 shows the oxygen−oxygen pair correlation functions $g_{OO}(r)$ at two different temperatures slightly below and above 400 K, obtained with the PBE and PBE0 functionals. (We recall that within PBE, at the experimental density, we obtain a good agreement with the measured pair correlation function at room temperature by shifting the simulation temperature to about 400 K). The result of PBE calculations carried out with the cell containing 96 molecules yields a pair correlation function $g_{OO}(r)$ which, within error bars,[20] is the same as that obtained with 32 molecules (see Figure 2 in the Supporting Information). Pair correlation functions computed using the Hamann and HSCV PPs and 32 molecule cells are also the same, within error bars (see Figure 3 in the Supporting Information). A comparison between PBE and PBE0 results at the two temperatures reported in Figure 1 clearly shows that the PBE0 functional improves over the PBE description, yielding a much less structured pair correlation function. Our findings are consistent (though not in close agreement) with those of ref 35, showing differences between PBE and PBE0 results for $g_{OO}(r)$, but they are at variance with the results of ref 36, which found hardly any change when carrying out calculations of structural properties at the PBE and PBE0 levels of theory.[71] The structural differences found here are consistent with a decrease of the molecular dipole moment found when using PBE0 (see Figure 2, upper panel). The calculated average molecular dipole moments with PBE0 are 2.88 ± 0.30 D (at 471 K), 2.94 ± 0.30 D (at 438 K), and 3.06 ± 0.29 D (at 374 K), and the corresponding ones with PBE are 3.03 ± 0.35 D (at 470 K), 3.09 ± 0.34 D (at 439 K), and 3.24 ± 0.32 D (at 367 K). An effective dipole moment of 2.9 ± 0.6 D has been derived from X-ray measurements,[72] which is consistent with both our calculations and the value of the dipole extracted for water clusters with six molecules (2.7 D).[73] Our findings for the liquid are also consistent with our results on the binding energy of the dimer, which is smaller with PBE0 than PBE.

The calculated structural differences between PBE0 and PBE stem from a decrease in the average number of hydrogen bonds when using the hybrid functional: 3.26 vs 3.43 at ∼438 K and 3.60 vs 3.79 at ∼374 K. Hydrogen bonds are defined using a geometrical criterion: two molecules are regarded as hydrogen bonded if the OO distance is less than 3.35 Å, and the O−OD angle is less than 30° (see Table 5 in the Supporting Information for further details). Differences in hydrogen bonding are also shown by the distribution of distances between oxygen and the centers of the MLWFs, displayed in Figure 2, lower panel. Two MLWFs are centered along OD bonds, and we call them bond pair (BP) orbitals. The other two are approximately centered on symmetric tetrahedral sites, and we call these lone pair (LP) orbitals. The four centers of the MLWFs of a water molecule are shown in the inset of Figure 2, lower panel. It is seen that the distributions of Oxy−BP distances are very similar when using PBE and PBE0, while those of Oxy−LP distances show marked differences.

In particular, the average distance between oxygen atoms and Wannier centers of the LP orbitals is shorter when computed with PBE0 than with PBE; in addition, the shoulder on the left-hand side of the Oxy−LP distance distribution, associated with single acceptor hydrogen bonds, is more pronounced. This indicates that at the PBE0 level, liquid water exhibits a more distorted hydrogen bonded network, with a larger number of broken hydrogen bonds. We also note that in addition to improving structural properties, calculations with the PBE0 functional yield differences in the electronic structure of the fluid, with an average electronic band gap of 6.73 eV, to be compared with the PBE value of 4.23 eV. However, also within PBE0, we obtain a value which is underestimated compared with experimental results (8.7 eV).[74]

We now turn to discussing our results for the vibrational spectrum of the liquid, as a function of temperature. The IR spectra of liquid deuterated water computed with the PBE and PBE0 functionals are compared with experimental results in Figure 3. The improvement of the PBE0 description with respect

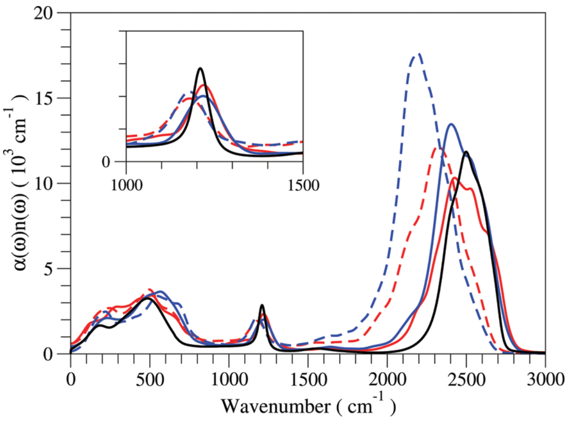


Figure 3. Calculated IR spectra of liquid D_2O with the PBE0 functional at 438 ± 29 K (solid red) and 374 ± 27 K (solid blue), compared with the ones calculated with the PBE functional at 439 ± 29 K (dash red) and 367 ± 25 K (dash blue). The experimental spectrum at room temperature is displayed by the black line.[75] The inset shows the spectra in the range 1000 cm^{-1} to 1500 cm^{-1}.

to PBE over the entire spectrum is apparent, especially so for the positions of the stretching and bending bands (see inset in Figure 3). This improvement comes from the combined effect of a better description of the frequencies of the monomer and dimer (see section 3.1) and an improved description of the structure of the liquid, with a smaller average molecular dipole moment corresponding to reduced strength of the hydrogen bonds. Weaker hydrogen bonding leads to stiffer intramolecular covalent bonds and higher stretching and bending frequencies. The positions of band maxima corresponding to hindered translations, librations, bending, and stretching modes computed with the PBE0 functional at 438 K are 180 (186, 246), 503 (486, 497), 1219 (1209, 1179), and 2426 (2498, 2322) cm^{-1}, respectively, where wavenumbers in parentheses are experimental values at room temperature[75] and results calculated with the PBE functional at 439 K, respectively. The corresponding values with PBE0 at 374 K are are 211 (186, 219), 579 (486, 549), 1211 (1209, 1179), and 2388 (2498, 2192) cm^{-1}, respectively, where wavenumbers in parentheses have the same meaning as those at 438 K. With the PBE0 functional, the red shift of the calculated stretching band with respect to experimental results is reduced to 72 cm^{-1} at ~438 K and 110 cm^{-1} at ~374 K, compared with 176 cm^{-1} and 306 cm^{-1} obtained with the PBE functional, respectively. We expect that the use of a PBE0 PP would bring our PBE0 results in closer agreement with experimental results, based on the comparison shown in Tables 1 and 2 in the Supporting Information. The comparison of our results to experimental ones for the low frequency band is more delicate as our sample is rather small. We note that a comparison of calculations at the PBE level of theory carried out with 32 and 96 molecules (see Figure 4) shows modest size effects for the stretching and bending bands.

The intensity of the stretching band calculated with the PBE0 functional is significantly smaller compared to the one computed at the PBE level. The ratio between the experimental intensities of ice[76] and water[75] stretching bands, after rescaling *T*, is approximately 1.3, suggesting that the less hydrogen bonded the system, the less intense the IR stretching band. This is consistent with the PBE0 intensity being smaller than the PBE one, at the same *T*. In addition to a blue-shifted main peak, the stretching band obtained with PBE0 shows a more pronounced shoulder at higher frequencies, compared with the spectrum obtained with PBE at the same temperature. This is again due to the increased number of broken hydrogen bonds in the liquid. As discussed in ref 33, within PBE, the contribution to the IR stretching band from molecules with broken hydrogen bonds gives rise to a shoulder at higher frequencies but not to a distinctive peak. This shoulder is more pronounced when using the hybrid functional, as expected from a less structured and less hydrogen bonded fluid. However, we emphasize that in the liquid, irrespective of which description is used (semilocal or hybrid functionals), hydrogen bonds are broken only for a short time, thus a clear IR signal associated to steadily broken bonds is not present in vibrational spectra. Also, the line shape of the bending modes is significantly improved within PBE0, consistent with the improved description of bending modes in the water molecule and dimer. Finally, we note that the analysis of the relative contributions of inter- and intramolecular correlations to the IR stretching band reported in ref 33 holds at the PBE0 level as well.

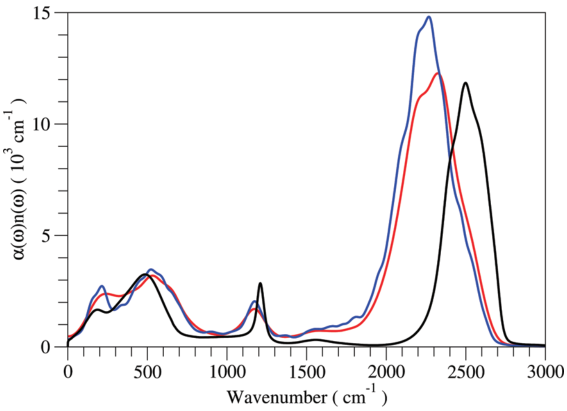


Figure 4. Calculated IR spectra of liquid D_2O with the PBE functional for systems consisting of 32 water molecules at 408 ± 27 K (red) and 96 water molecules at 407 ± 16K (blue), compared with the experimental spectrum at room temperature[75] (black).

The results reported in Figure 3 show that the positions of the bending and stretching peaks are both blue-shifted when *T* is increased. An increase of temperature may mimic, to some extent, the effect one would find by including a quantum mechanical description of the deuterons. More delocalized deuterons may weaken hydrogen bonds and thus enhance the signature of molecules with broken or distorted hydrogen bonds, which yield a blue-shifted IR signal, compared to that of perfectly hydrogen bonded molecules. The centroid MD simulations reported in refs 47 and 77 found instead a red shift in the position of the stretching band, with respect to classical MD. However, these results appear to be controversial;[78,79] indeed, it was recently pointed out that the observed red shift may be an artifact of the centroid MD technique.

4. CONCLUSIONS

We have shown that when using the PBE0 functional to describe the electronic structure of liquid water within DFT, one obtains structural and vibrational properties in better agreement with experimental results than calculations done with PBE and, in general, with semilocal functionals. In particular, we find a less structured fluid, with a lower dipole moment associated with each molecule and a higher number of fleetingly broken or distorted hydrogen bonds. We also find the IR spectrum to be in closer agreement with recent measurements,[75] showing a more pronounced shoulder corresponding to temporarily broken hydrogen bonds. The improved agreement with experimental results found for the spectrum stems from an improved description of both the structure of the liquid and the monomer and dimer vibrational frequencies. We believe that our findings represent a significant step forward in the theoretical modeling of water from first principles, as one may now extend the framework adopted here to other aqueous environments, such as water in contact with surfaces and simple aqueous solutions, thus opening the way to complement and interpret many spectroscopic measurements in an accurate fashion. In addition, encouraging results were found for electronic properties as well, and work is in progress to investigate in detail the electronic structure of the liquid using hybrid functionals.

However, several issues in the description of both structural and vibrational properties of water remain to be addressed, including the determination of the equilibrium density of water and its melting temperature when using hybrid functionals. In addition, it would be interesting to quantitatively assess the significance of proton quantum effects on vibrational spectra, when using a DFT-PBE0 description of the fluid. On the basis of the findings of ref 45, we expect the inclusion of proton quantum effects on our PBE0 results would lead to a less structured fluid, bringing the pair correlation function in even better agreement with experimental results. Our results on changes of IR spectra as a function of T, showing a blue shift of bending and stretching peaks as T is increased, indicate that the inclusion of proton quantum effects may improve the description of vibrations as well, although a detailed analysis is clearly needed to draw any firm conclusion.

While the qualitative effect on structural properties of treating the proton quantum mechanically is relatively straightforward to predict, an understanding of the influence of a more accurate description of van der Waals and dispersion forces is still under debate. Investigations of water clusters have so far shown critical dependence of results (e.g., order of energetically favored configurations) on the so-called dispersion corrections used[42,80] on top of simulations within DFT-PBE. Although no firm conclusion has yet been reported for liquid water, it appears that adding dispersion contributions to semilocal density functionals leads to a less structured liquid;[55,56,81] however, several open problems remain, e.g., in the description of the second solvation shell (next-nearest neighbor structural properties).

ASSOCIATED CONTENT

Supporting Information. Vibrational frequencies, average number of hydrogen bonds, and calculated oxygen−oxygen pair correlation functions. This material is available free of charge via the Internet at http://pubs.acs.org/.

AUTHOR INFORMATION

Corresponding Author

*E-mail: gagalli@ucdavis.edu.

Present Addresses

⊥Max Planck Institute for Polymer Research, 55128 Mainz, Germany

ACKNOWLEDGMENT

We would like to thank Ivan Duchemin and Roberto Car for helpful discussions. This work was supported by NSF Grant OCI-0749217. Use of ALCF computational resources at Argonne National Laboratory is gratefully acknowledged.

REFERENCES

(1) Wernet, Ph.; Nordlund, D.; Bergmann, U.; Cavalleri, M.; Odelius, M.; Ogasawara, H.; Näslund, L. Å.; Hirsch, T. K.; Ojamäe, L.; Glatzel, P.; Pettersson, L. G. M.; Nilsson, A. *Science* **2004**, *304*, 995–999.

(2) Prendergast, D.; Galli, G. *Phys. Rev. Lett.* **2006**, *96*, 215502.

(3) Chen, W.; Wu, X.; Car, R. *Phys. Rev. Lett.* **2010**, *105*, 017802.

(4) Rahman, A.; Stillinger, F. H. *Phys. Rev. A* **1974**, *10*, 368–378.

(5) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces*; Pullman, B., Ed.; Reidel: Dordrecht, The Netherlands, 1981; p 331.

(6) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269–6271.

(7) Toukan, K.; Rahman, A. *Phys. Rev. B* **1985**, *31*, 2643–2648.

(8) Dang, L. X.; Pettitt, B. M. *J. Phys. Chem.* **1987**, *91*, 3349–3354.

(9) Wu, Y.; Tepper, H. L.; Voth, G. A. *J. Chem. Phys.* **2006**, *124*, 024503.

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

(11) Mahoney, M. W.; Jorgensen, W. L. *J. Chem. Phys.* **2000**, *112*, 8910–8922.

(12) Donadio, D.; Cicero, G.; Schwegler, E.; Sharma, M.; Galli, G. *J. Phys. Chem. B* **2009**, *113*, 4170–4175.

(13) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471–2474.

(14) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.

(15) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.

(16) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

(17) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1997**, *78*, 1396.

(18) Laasonen, K.; Sprik, M.; Parrinello, M.; Car, R. *J. Chem. Phys.* **1993**, *99*, 9080–9089.

(19) Asthagiri, D.; Pratt, L. R.; Kress, J. D. *Phys. Rev. E* **2003**, *68*, 041505.

(20) Grossman, J. C.; Schwegler, E.; Draeger, E. W.; Gygi, F.; Galli, G. *J. Chem. Phys.* **2004**, *120*, 300–311.

(21) Schwegler, E.; Grossman, J. C.; Gygi, F.; Galli, G. *J. Chem. Phys.* **2004**, *121*, 5400–5409.

(22) Sit, P. H.-L.; Marzari, N. *J. Chem. Phys.* **2005**, *122*, 204510.

(23) Fernández-Serra, M. V.; Artacho, E. *J. Chem. Phys.* **2004**, *121*, 11136–11144.

(24) Kühne, T. D.; Krack, M.; Parrinello, M. *J. Chem. Theory Comput.* **2009**, *5*, 235–241.

(25) Allesch, M.; Schwegler, E.; Gygi, F.; Galli, G. *J. Chem. Phys.* **2004**, *120*, 5192–5198.

(26) Leung, K.; Rempe, S. B. *Phys. Chem. Chem. Phys.* **2006**, *8*, 2153–2162.

(27) In a so-called rigid water model, the O−H bond length and the H−O−H bond angle are constrained to be fixed during the simulation; in a flexible water model, all ionic degrees of freedom are allowed to vary.

(28) Silvestrelli, P. L.; Bernasconi, M.; Parrinello, M. *Chem. Phys. Lett.* **1997**, *277*, 478–482.

(29) Sharma, M.; Resta, R.; Car, R. *Phys. Rev. Lett.* **2005**, *95*, 187401.

(30) Iftimie, R.; Tuckerman, M. E. *J. Chem. Phys.* **2005**, *122*, 214508.

(31) Lee, H.-S.; Tuckerman, M. E. *J. Chem. Phys.* **2007**, *126*, 164501.

(32) Chen, W.; Sharma, M.; Resta, R.; Galli, G.; Car, R. *Phys. Rev. B* **2008**, *77*, 245114.

(33) Zhang, C.; Donadio, D.; Galli, G. *J. Phys. Chem. Lett.* **2010**, *1*, 1398–1402.

(34) VandeVondele, J.; Mohamed, F.; Krack, M.; Hutter, J.; Sprik, M.; Parrinello, M. *J. Chem. Phys.* **2005**, *122*, 014515.

(35) Todorova, T.; Seitsonen, A. P.; Hutter, J.; Kuo, I.-F. W.; Mundy, C. J. *J. Phys. Chem. B* **2006**, *110*, 3685–3691.

(36) Guidon, M.; Schiffmann, F.; Hutter, J.; VandeVondele, J. *J. Chem. Phys.* **2008**, *128*, 214104.

(37) Li, Z.; Wu, X.; Car, R. http://meetings.aps.org/Meeting/MAR10/Event/119032 (accessed March 2011).

(38) Guidon, M.; Hutter, J.; VandeVondele, J. *J. Chem. Theory Comput.* **2010**, *6*, 2348–2364.

(39) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158–6170.

(40) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(41) Santra, B.; Michaelides, A.; Scheffler, M. *J. Chem. Phys.* **2007**, *127*, 184104.

(42) Santra, B.; Michaelides, A.; Fuchs, M.; Tkatchenko, A.; Filippi, C.; Sheffler, M. *J. Chem. Phys.* **2008**, *129*, 194111.

(43) Xu, X.; Goddard, W. A. *J. Phys. Chem. A* **2004**, *108*, 2305–2313.

(44) Reference 35 reports simulations within the NVT ensemble, and ref 36 reports simulations within the NVE ensemble. In ref 36, a multiple time step scheme was used.

(45) Morrone, J. A.; Car, R. *Phys. Rev. Lett.* **2008**, *101*, 017801. The results presented in this paper differ from earlier PI simulations within DFT reported by Chen et al.[82] However, they are consistent with a large body of PI simulations carried out with classical or ab initio derived potentials. It is possible that the results of Chen et al. may suffer from inaccuracies in the comparison between classical and quantum simulations deriving from the choice of technical parameters in their simulations.

(46) Kuharski, R. A.; Rossky, P. J. *J. Chem. Phys.* **1985**, *82*, 5164–5177.

(47) Paesani, F.; Iuchi, S.; Voth, G. A. *J. Chem. Phys.* **2007**, *127*, 074506.

(48) Fanourgakis, G. S.; Xantheas, S. S. *J. Chem. Phys.* **2008**, *128*, 074506.

(49) Hetényi, B.; Angelis, F. D.; Giannozzi, P.; Car, R. *J. Chem. Phys.* **2004**, *120*, 8632–8637.

(50) Prendergast, D.; Grossman, J. C.; Galli, G. *J. Chem. Phys.* **2005**, *123*, 014501.

(51) Fernández-Serra, M. V.; Artacho, E. *Phys. Rev. Lett.* **2006**, *96*, 016404.

(52) Kulik, H. J.; Marzari, N.; Correa, A. A.; Prendergast, D.; Schwegler, E.; Galli, G. *J. Phys. Chem. B* **2010**, *114*, 9594–9601.

(53) Bakker, H. J.; Skinner, J. L. *Chem. Rev.* **2010**, *110*, 1498–1517.

(54) Qbox code. http://eslab.ucdavis.edu/software/qbox (accessed March 2011).

(55) Schmidt, J.; VandeVondele, J.; Kuo, I.-F. W.; Sebastiani, D.; Siepmann, J. I.; Hutter, J.; Mundy, C. J. *J. Phys. Chem. B* **2009**, *113*, 11959–11964.

(56) Wang, J.; Román-Pérez, G.; Soler, J. M.; Artacho, E.; Fernández-Serra, M.-V. *J. Chem. Phys.* **2011**, *134*, 024516.

(57) For 32 molecule cells, the ratio of the CPU time per MD step using PBE0 and PBE is about a factor of 25.

(58) Hamann, D. R. *Phys. Rev. B* **1989**, *40*, 2980–2987.

(59) Duchemin, I.; Gygi, F. *Comput. Phys. Commun.* **2010**, *181*, 855–860.

(60) Marzari, N.; Vanderbilt, D. *Phys. Rev. B* **1997**, *56*, 12847–12865.

(61) Gygi, F.; Fattebert, J. L.; Schwegler, E. *Comput. Phys. Commun.* **2003**, *155*, 1–6.

(62) Ramírez, R.; López-Ciudad, T.; Kumar-P, P.; Marx, D. *J. Chem. Phys.* **2004**, *121*, 3973–3983.

(63) Pseudopotential Table. http://fpmd.ucdavis.edu/potentials (accessed March 2011).

(64) Vanderbilt, D. *Phys. Rev. B* **1985**, *32*, 8412–8415.

(65) Benedict, W. S.; Gailar, N.; Plyler, E. K. *J. Chem. Phys.* **1956**, *24*, 1139–1165.

(66) Fredin, L.; Nelander, B.; Ribbegard, G. *J. Chem. Phys.* **1977**, *66*, 4065–4072.

(67) Tursi, A. J.; Nixon, E. R. *J. Chem. Phys.* **1970**, *52*, 1521–1528.

(68) Santra, B.; Michaelides, A.; Scheffler, M. *J. Chem. Phys.* **2009**, *131*, 124509.

(69) Although the HSCV PP yields results in slightly better agreement with AE calculations than the Hamann PP, we used the latter here, in order to be able to compare them to our previous results[33] for IR spectra, obtained with the Hamann PP.

(70) Soper, A. K. *Chem. Phys.* **2000**, *258*, 121–137.

(71) In refs 35 and 36, a cutoff function of the type $erfc(\alpha r)/r$ was adopted for the screened Coulomb operator; different values of α were used in the two papers. The functionals used in refs 35 and 36 are therefore different from the PBE0 one as commonly defined. Another difference between our simulation and those of refs 35 and 36 is the length of the trajectories (10(5) ps and 7 ps in refs 35 and 36, respectively, and 17 ps in our studies) and the equilibration procedure. The simulations of ref 35 contain 32 water molecules and are performed within the NVT ensemble at 350 K, while those of ref 36 contain 64 water molecules and are carried out within the NVE ensemble at 325 K.

(72) Badyal, Y. S.; Saboungi, M.-L.; Price, D. L.; Shastri, S. D.; Haeffner, D. R.; Soper, A. K. *J. Chem. Phys.* **2000**, *112*, 9206–9208.

(73) Gregory, J. K.; Clary, D. C.; Liu, K.; Brown, M. G.; Saykally, R. J. *Science* **1997**, *275*, 814–817.

(74) Bernas, A.; Ferradini, C.; Jay-Gerin, J.-P. *Chem. Phys.* **1997**, *222*, 151–160.

(75) Max, J.-J.; Chapados, C. *J. Chem. Phys.* **2009**, *131*, 184505.

(76) Bergren, M. S.; Schuh, D.; Sceats, M. G.; Rice, S. A. *J. Chem. Phys.* **1978**, *69*, 3477–3482.

(77) Fanourgakis, G. S.; Xantheas, S. S. *J. Chem. Phys.* **2008**, *128*, 074506.

(78) Ivanov, S. D.; Witt, A.; Shiga, M.; Marx, D. *J. Chem. Phys.* **2010**, *132*, 031101.

(79) Paesani, F.; Voth, G. A. *J. Chem. Phys.* **2010**, *132*, 014105.

(80) Kelkkanen, A. K.; Lundqvist, B. I.; Nørskov, J. K. *J. Chem. Phys.* **2009**, *131*, 046102.

(81) Lin, I.-C.; Seitsonen, A. P.; Coutinho-Neto, M. D.; Tavernelli, I.; Rothlisberger, U. *J. Phys. Chem. B* **2009**, *113*, 1127–1131.

(82) Chen, B.; Ivanov, I.; Klein, M. L.; Parrinello, M. *Phys. Rev. Lett.* **2003**, *91*, 215503.

■ NOTE ADDED AFTER ASAP PUBLICATION

This paper was published on the Web on March 21, 2011, with minor errors in references 15 and 72. The corrected version was reposted on March 31, 2011.

JCTC Journal of Chemical Theory and Computation
ARTICLE
pubs.acs.org/JCTC

Combination of RISM and Cheminformatics for Efficient Predictions of Hydration Free Energy of Polyfragment Molecules: Application to a Set of Organic Pollutants

Ekaterina L. Ratkova and Maxim V. Fedorov*

The Max Planck Institute for Mathematics in the Sciences, Inselstrasse 22, Leipzig, 04103, Germany

S *Supporting Information*

ABSTRACT: Here, we discuss a new method for predicting the hydration free energy (HFE) of organic pollutants and illustrate the efficiency of the method on a set of 220 chlorinated aromatic hydrocarbons. The new model is computationally inexpensive, with one HFE calculation taking less than a minute on a PC. The method is based on a combination of a molecular integral equations theory, one-dimensional reference interaction site model (1D RISM), with the cheminformatics approach. We correct HFEs obtained by the 1D RISM with a set of empirical corrections. The corrections are associated with the partial molar volume and structural descriptors of the molecules. We show that the introduced corrections can significantly improve the quality of the 1D RISM HFE predictions obtained by the partial wave free energy expression [Ten-no, S. J. *Chem. Phys.* **2001**, *115*, 3724] and the Kovalenko–Hirata closure [Kovalenko, A.; Hirata, F. *J. Chem. Phys.* **1999**, *110*, 10095]. We also show that the quality of the model can be further improved by the reparametrization using QM-derived partial charges instead of the originally used OPLS-AA partial charges. The final model gives good results for polychlorinated benzenes (the mean and standard deviation of the error are 0.02 and 0.36 kcal/mol, correspondingly). At the same time, the model gives somewhat worse results for polychlorobiphenyls (PCBs) with a systematic bias of −0.72 kcal/mol but a small standard deviation equal to 0.55 kcal/mol. We note that the error remains the same for the whole set of PCBs, whereas errors of HFEs predicted with continuum solvation models (data were taken from Phillips, K. L. et al. *Environ. Sci. Technol.* **2008**, *42*, 8412) increase significantly for higher chlorinated PCB congeners. In conclusion, we discuss potential future applications of the model and several avenues for its further improvement.

INTRODUCTION

Chlorinated aromatic hydrocarbons (CAHs) are a group of compounds that belong to the category of "persistent organic pollutants" (POPs). This class of pollutants is characterized by (i) long-term persistence, (ii) long-range atmospheric transport and deposition, (iii) bioaccumulation, and (iv) adverse effects on biota.[1,2] For a long time, in many countries, CAHs (such as polychlorobiphenyls, hexachlorobenzene, etc.) were used in agriculture as pesticides, fungicides, and agents controlling arthropods.[1] Although CAHs have been banned from further use and production,[3] their persistence in biological compartments (e.g., soil, water, plants, and sediment) means that they still pose a significant environmental hazard. Understanding and clarifying the global fate of CAHs is one of the most important environmental and ecological problems.[1,2,4,5] The semivolatile nature of CAHs allows them to evaporate from soil and water into the atmosphere, where they can exist both in gaseous and particle-absorbed forms (these can be atmospheric aerosol particles, e.g., cloud droplets, as well as dust particles). Several dominant mechanisms that determine the distribution of CAHs between atmosphere and water are shown in Figure 1.

There are several physical/chemical properties of CAHs that determine their global fate: vapor pressure; aqueous solubility; partition coefficients between different media; and half-lives in the air, solids, and water. These parameters are intensively used in mathematical models describing the global fate and long-range transport of CAHs.[6−10] One of the most important parameters in these models is the flux of a compound across surfaces, which characterizes the exchange of the compound between the corresponding compartments.[2,11] As an example, the flux of molecules *i* between two compartments 1 and 2 can be modeled as

$$F_{1\rightarrow 2} = K_{1/2(i)}\left(C_{1(i)} - \frac{C_{2(i)}}{P_{i,\,\mathrm{eq}}}\right) \tag{1}$$

where $F_{1\rightarrow 2}$ is the flux from compartment 1 to compartment 2, $K_{1/2(i)}$ is the kinetic parameter represented by the mass transfer coefficient on the molecules *i*, $C_{1(i)}$ and $C_{2(i)}$ are equilibrium molecular concentrations of the molecules *i* in compartments 1 and 2, respectively, and $P_{i,\mathrm{eq}}$ is the equilibrium partition coefficient of the molecules *i* between the two compartments.

Thus, accurate data for the partition coefficients are of a high importance for modeling CAH exchange between compartments. In the case of the air–water flux, the widely used partition coefficient is the Henry's law constant (K_H), which shows the distribution of a compound between gaseous and aqueous phases:

$$K_H = \frac{C_{\mathrm{aq}(i)}}{C_{\mathrm{g}}(i)} \tag{2}$$

Received: November 12, 2010
Published: April 22, 2011

dx.doi.org/10.1021/ct100654h | *J. Chem. Theory Comput.* 2011, 7, 1450–1457

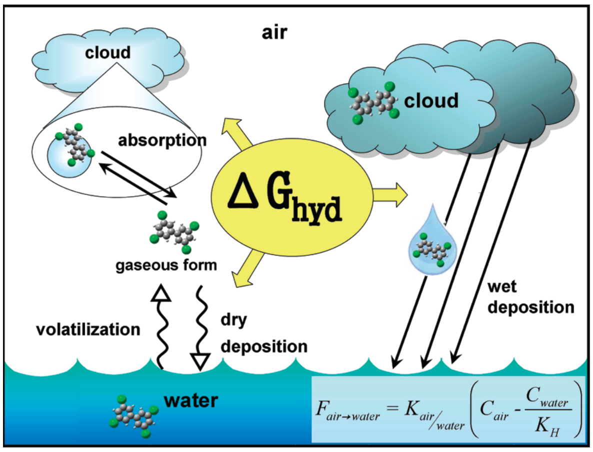


Figure 1. Dominant mechanisms that determine the distribution of CAHs between atmosphere and water. Hydration free energy (ΔG_{hyd}) is an important thermodynamic parameter used to describe the main processes of CAH distribution between atmosphere and water. It is closely related to the Henry's law constant (K_{H}) as $\Delta G_{\text{hyd}} = -RT \ln(K_{\text{H}})$. In turn, K_{H} is widely used to model the flux of a molecule from air to water, $F_{\text{air}\rightarrow\text{water}}$ (see the inset equation and eq 1 for the notation).

where $C_{\text{aq}(i)}$ and $C_{\text{g}(i)}$ are equilibrium molecular concentrations of the molecules i in aqueous and gaseous phases, respectively.

We note that the Henry's law constant is closely related with the HFE as[12]

$$\Delta G_{\text{hyd}} = -RT \ln(K_{\text{H}}) \qquad (3)$$

where ΔG_{hyd} is the hydration free energy, K_{H} is the Henry's law constant, R is the ideal gas constant, and T is the temperature.

Recently, we reported a novel computational method for accurate estimations of the HFEs of organic molecules—the structural descriptors correction (SDC) model.[13] The method is based on a combination of the computationally inexpensive one-dimensional RISM (1D RISM) with several corrections that can be obtained in a straightforward manner from the molecular structure. The main advantage of the model is a small number of chemical descriptors associated with main structural features of solutes: partial molar volume (PMV), aromatic rings, electron-donating/withdrawing substituents, etc. We have shown that the 1D RISM-SDC model with the OPLS-AA partial charges[14,15]—1D RISM-SDC(OPLSq) model—allows one to obtain HFEs for monofragment solutes with high accuracy.[13] In the case of polyfragment solutes, the 1D RISM-SDC(OPLSq) model is more sensitive to the chemical nature of solutes. Thus, the model allows one to predict HFEs with an accuracy of about 1 kcal/mol for chlorinated benzenes with fewer than three chlorine atoms, but it provides worse results for chlorinated benzenes with a larger number of chlorine atoms.[13] The main reason for this deviation is the fact that OPLS-AA partial charges are not sensitive to the mesomeric effect of aromatic polyfragment solutes.[13]

Here, we show that the quality of the 1D RISM-SDC model can be further improved by the model reparametrization using QM-derived partial charges (1D RISM-SDC(QMq) model) instead of the originally used OPLS-AA partial charges. In this paper, we would like to demonstrate the efficiency of the 1D RISM-SDC(QMq) model for two classes of CAHs: (i) polychlorinated benzenes and (ii) polychlorobiphenyls. Other classes of CAHs will be considered in our forthcoming publications.

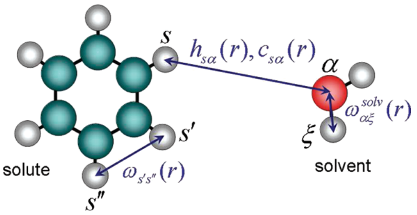

s
$h_{s\alpha}(r), c_{s\alpha}(r)$
α
$\omega^{solv}_{\alpha\xi}(r)$
ξ
s'
solute
solvent
$\omega_{s's''}(r)$
s''

Figure 2. Representations of solute and solvent molecules and correlation functions in the 1D RISM approach. Both molecules are modeled as sets of sites (atoms). The molecules structures are described with site−site intramolecular correlation functions: $\omega_{ss'}(r)$ and $\omega^{\text{solv}}_{\alpha\xi}(r)$. Solvent density distributions around the solute molecule are described with intermolecular total $h_{s\alpha}(r)$ and direct $c_{s\alpha}(r)$ correlation functions.

METHODS

1D RISM Approach. We use here the 1D RISM approach,[16] where the solute and solvent molecules are modeled as sets of sites (atoms) interacting via pairwise spherically symmetric potentials (Figure 2).[16] We use the common form of the interaction potential represented by the long-range electrostatic term and short-range Lennard-Jones (LJ) term.[17] The 1D RISM operates with site−site correlation functions: intramolecular correlation functions $\omega_{ss'}(\mathbf{r})$, $\omega^{\text{solv}}_{\alpha\xi}(\mathbf{r})$, total correlation functions $h_{s\alpha}(\mathbf{r})$, and direct correlation functions $c_{s\alpha}(\mathbf{r})$ (where s and s' are solute atoms, and α and ξ are solvent atoms; Figure 2).[16] In general, these are 3D-functions, but due to the spherical symmetry, we consider only their 1D-radial parts $\omega_{ss'}(r)$, $\omega^{\text{solv}}_{\alpha\xi}(r)$, $h_{s\alpha}(r)$, and $c_{s\alpha}(r)$, which depend only on the radial distance r. Direct correlation functions are connected with the total correlation functions via the set of 1D RISM integral equations:[16]

$$h_{s\alpha}(|\mathbf{r}_1 - \mathbf{r}_2|) = \sum_{s'=1}^{N_{\text{solute}}} \sum_{\xi=1}^{N_{\text{solvent}}} \int_{R^3}\int_{R^3} \omega_{ss'}(|\mathbf{r}_1 - \mathbf{r}'|)\, c_{s'\xi}(|\mathbf{r}' - \mathbf{r}''|)\, \chi_{\alpha\xi}(|\mathbf{r}'' - \mathbf{r}_2|)\, d\mathbf{r}'\, d\mathbf{r}'' \qquad (4)$$

where $\chi_{\alpha\xi}(r) = \omega^{\text{solv}}_{\alpha\xi}(r) + \rho h^{\text{solv}}_{\alpha\xi}(r)$ are the bulk solvent susceptibility functions and N_{solute} and N_{solvent} are the numbers of sites in the solute and solvent, correspondingly. We note that in the current work we expressed the intramolecular correlation functions in terms of Dirac δ functions considering molecules as rigid objects. However, molecules under investigation are almost rigid, and this simplification does not lead to considerable changes in the description of the hydration process. In general, for more flexible compounds, the changes in molecular conformations upon hydration have to be taken into account (e.g., with a coupled RISM/MD or RISM/MC simulation methodology[18−20]).

To make eq 4 complete, $N_{\text{solute}} \times N_{\text{solvent}}$ site−site *closure* relations are introduced:

$$\begin{aligned} &h_{s\alpha}(r) = \exp(-\beta u_{s\alpha}(r) + h_{s\alpha}(r) - c_{s\alpha}(r) + B_{s\alpha}(r)) - 1 \\ &s = 1, ..., N_{\text{solute}};\ \alpha = 1, ..., N_{\text{solvent}} \end{aligned} \qquad (5)$$

where $u_{s\alpha}(r)$ is a pair interaction potential between the sites s and α, $B_{s\alpha}(r)$ are site−site bridge functions, and $\beta = 1/k_{\text{B}}T$, where k_{B} is the Boltzmann constant and T is the temperature. In general, the exact bridge functions are practically uncomputable, and one needs to use some approximation.[16,21,22] The most straightforward and widely used model is the HNC approximation, which sets the bridge functional $B_{s\alpha}(r)$ to zero.[23] However, due to the uncontrolled growth of the argument of the exponent (eq 5), use of the HNC closure can lead to a slow convergence rate, and in many cases even divergence of the numerical solution of 1D RISM equations. One way to overcome this problem is to linearize the exponent when its argument is larger than a certain threshold constant C:

$$h_{s\alpha}(r) = \begin{cases} \exp(\Xi_{s\alpha}(r)) - 1 & \text{when} \quad \Xi_{s\alpha}(r) < C \\ \Xi_{s\alpha}(r) + \exp(C) - C - 1 & \text{when} \quad \Xi_{s\alpha}(r) > C \end{cases} \tag{6}$$

where $\Xi_{s\alpha}(r) =- \beta u_{s\alpha}(r) + h_{s\alpha}(r) - c_{s\alpha}(r)$.

The linearized HNC closure for the case $C = 0$ was proposed by Hirata and Kovalenko in ref 24, where it has been called KH closure. More details of the theoretical and computational background behind the 1D RISM equations can be found in refs 16, 25, and 26. In the current work, we performed 1D RISM calculations with the KH closure. Previously, we showed[27,28] that the efficiency of HFE calculations with a set of semiempirical corrections is almost not sensitive to the choice of closure relation (KH, HNC). However, the KH closure allows one to perform the quickest and the most stable RISM calculations. That is why we used the KH closure in this work rather than the HNC.

Within the framework of the 1D RISM theory, there are several free energy expressions which allow one to obtain values of the HFE from the total and direct correlation functions: HNC,[16,23] GF,[29] KH,[30] PW,[31] HNCB,[32] and PWC.[33] Comparisons of these expressions[13,31,33−38] show that the PW free energy expression has better performance than the KH, HNC, and HNCB free energy expressions. Therefore, as in our previous work,[13] we use here the PW free energy expression to calculate HFE values:

$$\Delta G_{\text{hyd}}^{\text{PW}} = 2\pi\rho k_{\text{B}} T \sum_{s=1}^{N_{\text{solute}}} \sum_{\alpha=1}^{N_{\text{solvent}}} \int_0^\infty [-2c_{s\alpha}(r) - c_{s\alpha}(r)\, h_{s\alpha}(r) + \tilde{h}_{s\alpha}(r)\, h_{s\alpha}(r)] r^2 \, \mathrm{d}r \tag{7}$$

where $r = |\mathbf{r}_1 - \mathbf{r}_2|$ and

$$\tilde{h}_{s\alpha}(|\mathbf{r}_1 - \mathbf{r}_2|) = \sum_{s'=1}^{N_{\text{solute}}} \sum_{\xi=1}^{N_{\text{solvent}}} \int_{R^3} \int_{R^3} \tilde{\omega}_{ss'}(|\mathbf{r}_1 - \mathbf{r}'|)\, h_{s'\xi}(|\mathbf{r}' - \mathbf{r}''|)\, \tilde{\omega}_{\alpha\xi}^{\text{solv}}(|\mathbf{r}'' - \mathbf{r}_2|)\, \mathrm{d}\mathbf{r}' \, \mathrm{d}\mathbf{r}''$$

$\tilde{\omega}_{ss'}$ and $\tilde{\omega}_{\alpha\xi}^{\text{solv}}(r)$ are the elements of matrices $\mathbf{W}^{-1}$ and $\mathbf{W}_{\text{solv}}^{-1}$, which are inverses to the matrices $\mathbf{W} = [\omega_{ss'}(r)]_{N_{\text{solute}} \times N_{\text{solute}}}$ and $\mathbf{W}_{\text{solv}} = [\omega_{\alpha\xi}^{\text{solv}}(r)]_{N_{\text{solvent}} \times N_{\text{solvent}}}$ built from the solute and solvent intramolecular correlation functions, respectively.

SDC Model. We define the *modeling error* (ε) of the RISM-based HFEs calculations for a solute as the difference between the calculated and experimental values:

$$\varepsilon = \Delta G_{\text{hyd}}^{\text{model}} - \Delta G_{\text{hyd}}^{\text{exp}} \tag{8}$$

where $\Delta G_{\text{hyd}}^{\text{exp}}$ is the experimental value of HFE and $\Delta G_{\text{hyd}}^{\text{model}}$ is the HFE calculated by the RISM approach (superscript model denotes the RISM-based HFE expression, e.g., PW).

The main idea behind the SDC model is that we parametrize the modeling error ε with a set of descriptors $\{D_i\}$ associated with specific features of the chemical structure of solutes such as partial molar volume (PMV), aromatic rings, electron-donating/withdrawing substituents, etc. (Figure 3). The model contains the assumption that different structural properties of the solute molecule contribute independently to the error in HFE calculations. We use here the multilinear regression model where the impact of the selected chemical properties on the HFE is linearly proportional to the values of the corresponding descriptors $\{D_i\}$ with empirical coefficients $\{a_i^{\text{model}}\}$:

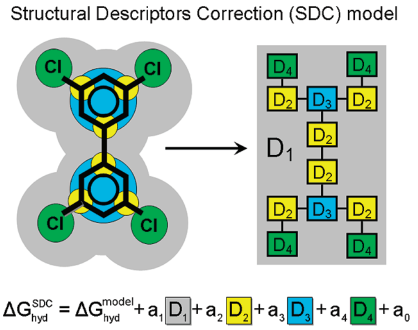

Structural Descriptors Correction (SDC) model
Cl
Cl
Cl
Cl
D4
D4
D2
D3
D2
D1
D2
D2
D2
D3
D2
D4
D4
ΔG^SDC_hyd = ΔG^model_hyd + a1 D1 + a2 D2 + a3 D3 + a4 D4 + a0

Figure 3. Schematic representation of a molecule (3,3′,5,5′-tetrachlorobiphenyl) as a combination of fragment counts. The SDC model equation as a linear combination of the corresponding structural corrections: a_1D_1 is the correction on dimensionless partial molar volume, a_2D_2 is the correction on branches, a_3D_3 is the correction on the benzene ring, a_4D_4 is the correction on the halogen atom, and a_0 is a constant (solute-independent systematic error; see eq 9).

$$\Delta G_{\text{hyd}}^{\text{SDC}} = \Delta G_{\text{hyd}}^{\text{model}} + \sum_i a_i^{\text{model}} D_i + a_0^{\text{model}} \tag{9}$$

where the second term is the set of structural corrections and a_0^{model} is a systematic error.[13]

Training and Test Sets. Another basic idea behind the SDC model is to calibrate the empirical coefficients $\{a_i^{\text{model}}\}$ on a set of "simple" solutes. They can be represented as an alkyl chain (linear or branched) which can contain only *one* substituent (e.g., benzene ring or chlorine atom). In the present work for the 1D RISM-SDC(QMq) model, we used a training set of 46 small neutral organic solutes: 22 alkanes, 17 alkylbenzenes, and 7 monochloroalkanes (see the Supporting Information). The modeling error for alkanes can be parametrized with corrections on PMV and branches. The set of alkylbenzenes requires an additional correction on the benzene ring; in turn, the modeling error for chloroalkanes can be parametrized with a linear combination of corrections on PMV, branches, and chlorine atoms. Experimental HFEs for all solutes from the training set were taken from ref 13, where HFEs were collected from several literature sources and then averaged.

We tested the calibrated 1D RISM-SDC model on a set of 220 chlorinated aromatic hydrocarbons (CAHs): 11 polychlorinated benzenes (from chlorobenzene to hexachlorobenzene, Table 2) and 209 polychlorinated biphenyls, PCBs (see the Supporting Information). The set of experimental HFEs for CAHs was compiled from different literature sources: (i) HFEs were taken from ref 13; (ii) log P(water/gas) values were collected from ref 39 and recalculated to HFEs with eq 10 ; (iii) K_{H} constants were taken from refs 40−44 and recalculated to HFEs with eq 3.

$$\Delta G_{\text{hyd}} = -(\ln 10) RT \log P(\text{water/gas}) \tag{10}$$

where ΔG_{hyd} is the hydration free energy, log P(water/gas) is the logarithm of the partition coefficient between the gaseous phase and water, R is the ideal gas constant, and T is the temperature.

Table 1. Descriptors and Corresponding Multilinear Regression Coefficients of the 1D RISM-SDC(QMq) Model for the Training Set of Solutes

descriptor		coefficient (kcal/mol)
		$a_0^{PW} = -4.19$
dimensionless partial molar volume	$(D_1 = \rho\overline{V})^a$	$a_1^{PW} = -1.48$
number of branches	$(D_2 = N_{br})$	$a_2^{PW} = 0.98$
number of benzene rings	$(D_3 = N_{benz})$	$a_3^{PW} = -3.11$
number of halogen atoms	$(D_4 = N_{hal})$	$a_4^{PW} = -1.30$

[a] $\overline{V}$ is the partial molar volume of the solute; $\rho = 0.0337$ Å^{-3}.

Computational Details. The HFEs were calculated with the 1D RISM method using the collection of numerical routines developed by our group.[45−47] Calculations were performed for the case of infinitely diluted aqueous solutions at $T = 300$ K. We used the Lue and Blankschtein version of the modified SPC/E model of water (MSPC/E),[48] proposed earlier by Pettitt and Rossky.[49] It differs from the original SPC/E water model[50] by the addition of Lennard-Jones (LJ) potential parameters for the water hydrogen ($\sigma_{Hw}^{LJ} = 0.8$ Å and $\varepsilon_{Hw}^{LJ} = 0.046$ kcal/mol). We took the MSPC/E bulk solvent correlation functions $h_{\alpha\beta}^{solv}(r)$ from ref 37.

To perform the calculations, one needs three sets of input solute data: (1) coordinates of atoms, (2) partial charges on atoms, and (3) atom LJ potential parameters. Coordinates of atoms for each molecule were optimized using the Gaussian 03 quantum chemistry software[51] at the B3LYP/6-31G(d,p) level of theory. The initial configurations for the solutes from the training set were taken from ref 13. In the case of the test set, atomic coordinates for several PCBzs and PCBs were taken from the Cambridge Structural Database.[52] Due to the fact that the hydrogen positions determined by standard X-ray methods can be inadequate,[53] we optimized the length of the C−H bonds with constrained C−C and C−Cl bonds. The geometrical parameters of all other CAHs (not presented in the Cambridge Structural Database) were found by structural optimization at the same level of theory without constrained bonds. Partial charges for all molecules were obtained by with the CHELPG procedure[54] at the B3LYP/6-31G(d,p) level of theory using the Gaussian 03 software.[51] We modeled all compounds with OPLS-AA (optimized potential for liquid simulations−all atom) LJ potential parameters[14,15,55] which were assigned to each atom automatically by the Maestro software (Schroedinger Inc.). The set of structural descriptors (eq 9) was assigned to each molecule automatically by the computer program "checkmol"[56] with the use of Python scripts.

RESULTS AND DISCUSSION

The 1D RISM-SDC(QMq) Model Calibration. Values of coefficients $\{a_i^{PW}\}$ of 1D RISM-SDC(QMq) model eq 9 with the considered set of descriptors were obtained using multilinear regression[57] against a training set of 46 solutes. The regression analysis was performed with the function *regress* from the Matlab Statistics Toolbox (MATLAB, version 7.8.0.347(R2009a), The MathWorks Inc., 2009). As one can see (Table 1), coefficients a_2^{PW}, a_3^{PW}, and a_4^{PW} have the same order of magnitude, indicating that each structural descriptor from the considered set is significant.

HFEs calculated by the 1D RISM-SDC(QMq) model for the training set of solutes are shown in Figure 4. Correlation coefficient r between the calculated and experimental HFEs was obtained with the function *corrcoef* from the same Matlab Statistics Toolbox and equals 0.92. It shows that the 1D RISM-SDC(QMq) model with four structural descriptors describes HFEs of 46 solutes from different chemical classes with high accuracy (the standard deviation of the error is 0.64 kcal/mol).

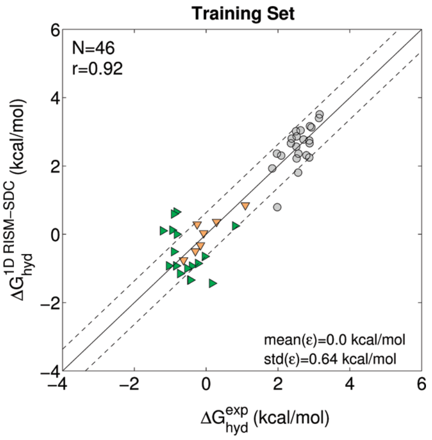


Figure 4. Correlation between the calculated and experimental HFEs for the training sets of solutes (gray circles are alkanes, orange triangles are alkylbenzenes, green triangles are chlorinated alkanes). The inset data show the statistical profile of the error $\varepsilon = \Delta G_{hyd}^{1DRISM\text{-}SDC} - \Delta G_{hyd}^{exp}$. Solid line illustrates the ideal correlation. Dashed lines indicate the standard deviation of the error.

Predictive Ability of the 1D RISM-SDC(QMq) Model for CAHs. The predictive ability of the 1D RISM-SDC(QMq) model for HFE calculations was analyzed on the test set of 220 CAHs (see the section Training and Test Sets) and *the same* set of coefficients from Table 1 as for the training set. A comparison of the predicted and experimental HFEs is discussed below. We note that the reliable experimental data are very important for estimations of the accuracy of predicted results. Due to that, before the analysis of calculated data, we performed an estimation of reliability of experimentally obtained HFE values for each class of compounds from the test set.

Polychlorinated Benzenes (PCBzs). First of all, we analyzed the difference between the experimental HFEs for PCBzs obtained by different sources (see Table 2). Despite the fact that for several solutes (1,2,3-trichlorobenzene, 1,3,5-trichlorobenzene, and hexachlorobenzene) the HFE values differ by 0.5−0.6 kcal/mol (see Table 2), on average, HFE values obtained with different techniques deviate from the mean value by 0.2−0.3 kcal/mol. Thus, we concluded that experimental data for polychlorinated benzenes are sufficiently accurate and can be used for the estimation of the accuracy of the predicted data.

The comparison of the predicted and experimental HFE values is shown in Figure 5. To quantify the accuracy, we calculated statistical parameters of the error $\varepsilon = \Delta G_{hyd}^{1DRISM-SDC} - \Delta G_{hyd}^{exp}$ for the test set of polychlorinated benzenes (Figure 5, inset data). As one can see, results obtained with the 1D RISM-SDC(QMq) model are nonbiased (mean of the difference equals 0.02 ± 0.11 kcal/mol), and the standard deviation of the error is in the range of the deviation between different sources of the corresponding experimental data (∼0.4 kcal/mol).

Polychlorobiphenyls (PCBs). For PCBs, experimental values of neither hydration free energy nor log P(water/gas) are available in the literature. However, since the 1980s, there have been

Table 2. Descriptors of the 1D RISM-SDC Model (eq 9) and Hydration Free Energies (ΔG_{hyd}) for Polychlorinated Benzenes[a]

					ΔG_{hyd} (kcal mol $^{-1}$)			
name	D_1 ($\rho\overline{V}$)	D_2 (N_{br})	D_3 (N_{benz})	D_4 (N_{hal})	1D RISM-PW	1D RISM-SDC	$\exp_{average}$	$\exp_{\|max\|-\|min\|}$
1,2,3,4-tetrachlorobenzene	5.24	4	1	4	14.75	−1.83	−1.32[39,40,43]	0.07
1,2,3-trichlorobenzene	4.85	3	1	3	13.77	−1.79	−1.49[39,43]	0.50
1,2,4,5-tetrachlorobenzene	5.30	4	1	4	15.40	−1.27	−1.34[39,43]	0.00
1,2,4-trichlorobenzene	4.89	3	1	3	14.23	−1.40	−1.22[39,40,43]	0.29
1,2-dichlorobenzene	4.45	2	1	2	12.85	−1.69	−1.47[13,39−41,43]	0.27
1,3,5-trichlorobenzene	4.93	3	1	3	14.71	−1.97	−1.09[39,43]	0.63
1,3-dichlorobenzene	4.48	2	1	2	13.24	−1.34	−1.13[39,41,43]	0.29
1,4-dichlorobenzene	4.49	2	1	2	13.15	−1.44	−1.15[39,41,43]	0.21
2-chlorotoluene	4.52	2	1	1	12.76	−0.55	−1.14[39]	
chlorobenzene	4.04	1	1	1	11.98	−1.51	−1.07[13,39−41,43]	0.22
hexachlorobenzene	5.95	6	1	6	16.33	−2.17	−2.26[43,42]	0.50

[a] Hydration free energies (ΔG_{hyd}) predicted by the uncorrected PW free energy expression and the 1D RISM-SDC model with QM-derived partial charges. Experimental values were averaged over different sources ($\exp_{average}$); $\exp_{|max|-|min|}$ shows the difference between the maximum and minimum values from different literature sources.

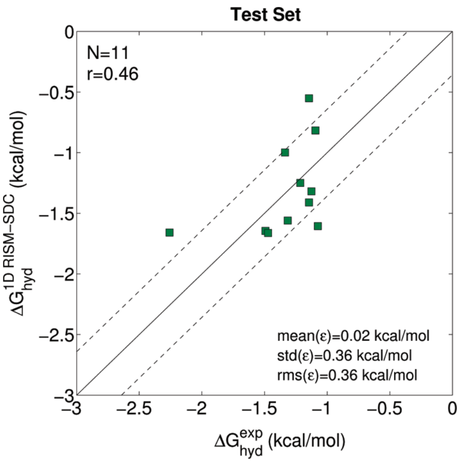


Figure 5. Correlation between the experimental HFE and values predicted by the 1D RISM-SDC(QMq) model for the test set of polychlorinated benzenes. The inset data show the statistical profile of the error $\varepsilon = \Delta G_{hyd}^{calc} - \Delta G_{hyd}^{exp}$. Solid line illustrates the ideal correlation. Dashed lines indicate the std(ε).

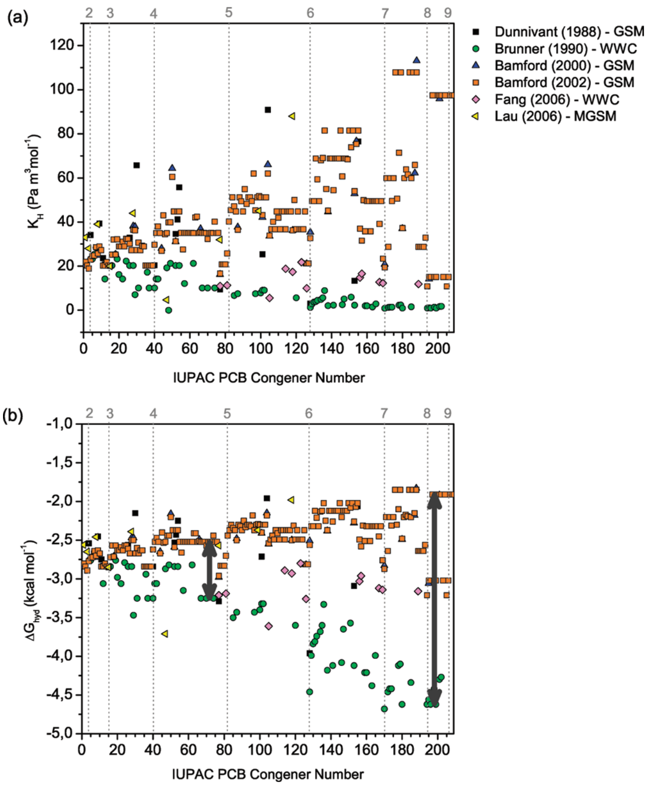


Figure 6. Experimental data for PCB congeners: (a) Henry's law constants, K_H, obtained with wetted-wall column (WWC), gas stripping method (GSM), or modified GSM (MGSM). (b) Hydration free energies (ΔG_{hyd}) recalculated from K_H. Black arrows show the deviation of experimental data obtained by the different techniques. Dashed lines show the separation of the whole set of PCBs with respect to the number of chlorine atoms (shown on the top).

several experimental investigations of K_H values of PCBs reported, where the experiments were carried out with two dynamic techniques: (i) the gas stripping method (GSM)[58−61] and (ii) the "wetted-wall column" (WWC) or the concurrent flow technique.[44,62,63] All values are presented in Figure 6a; corresponding HFEs recalculated with eq 3 are presented in Figure 6b. One can see that the experimental K_H values are presented mainly by two sets of data obtained by the GSM (Bamford[59]) and the WWC technique (Brunner et al.[44]). Other sets of K_H values are not very large and contain about 20−30 values from 209 possible. Figure 6 shows that, for the same solutes, experimental K_H values from the GSM and WWC sets can differ considerably. The difference increases with the increase in the number of chlorine atoms in a solute. In terms of HFE, the difference varies from 1 kcal/mol for lighter PCB congeners (PCB with 4−5 chlorine atoms) to up to 3 kcal/mol for heavier congeners (higher chlorinated PCBs) (Figure 6b).

Recently, it was found that the GSM overestimates K_H values for highly chlorinated biphenyls .[64,65] The problem is hidden in the technical implementation of the GSM. Within the method, the K_H of a compound is determined as a ratio of the equilibrium concentrations of the compound in aqueous solution and vapor, accordingly. The compound is stripped from the aqueous phase into a gaseous phase using a bubble column apparatus (see the Supporting Information).[66] It was found that the sorption of the solute molecules to the surface of gas bubbles leads to a higher compound concentration in the gaseous phase, which, in turn, results in the overestimated K_H value. With the WWC technique, one can avoid this drawback. The technical implementation of the method consists of the equilibration of a compound between a thin layer of water and a concurrent flow of gas within the contact region.[66] Due to that, we accepted the experimental data obtained by the WWC method[44] as the most reliable set. Unfortunately, the total number of experimental values published in ref 44 is only 57 from 209 possible.

A comparison of HFEs, predicted by the 1D RISM-SDC-(QMq) model, with the experimental data (Table 3) shows that the calculated values are biased with respect to experimental ones, mean(ε) = -0.72 ± 0.07 kcal/mol, but have a small standard deviation of error. Figure 7 shows that the error remains the same for the whole set of PCBs and does not increase for the heavier PCB congeners.

Table 3. Statistical Profiles of Errors for Results Obtained by the Implicit Solvent Models for the Test Set of Polychlorobiphenyls (N = 57): Mean Value, Standard Deviation (std), and Root Mean Square (rms) of the Error $\varepsilon = \Delta G_{\text{hyd}}^{\text{calc}} - \Delta G_{\text{hyd}}^{\text{exp}}$ (kcal/mol)[a]

	model			
	1D RISM-PW	1D RISM-SDC	SM_6[67]	COSMO-SAC[67]
mean(ε)	20.35	−0.72	1.28	1.15
std(ε)	1.62	0.55	0.78	0.94
rms(ε)	20.42	0.91	1.50	1.49
r	−0.80	0.65	−0.35	−0.70

[a] r is the correlation coefficient. Results obtained by the SM_6 and COSMO-SAC methods were collected from ref 67.

Also, we performed a comparison of our results with HFEs obtained by other implicit models, SM_6 and COSMO-SAC (the data were taken from ref 67). Both of them treat the solvent as a homogeneous medium characterized by its dielectric constant (continuum solvent methods). Statistical analysis of the literature results is shown in Table 3. As one can see (Figure 7), HFEs obtained by these models are in good agreement with each other. However, the models allow predictions of HFE with high accuracy only for light congeners, whereas for the heavier PCBs, the error of HFE increases with the increase in the number of chlorine atoms. In the case of the highly chlorinated biphenyls (N_{Cl} = 8−9), the error is ~3 kcal/mol. We explain these results as follows. In the case of lighter PCB congeners, the chlorine atoms are well-separated from each other. Thus, the total effect of chlorine atoms interactions with the solvent molecules can be presented as a sum of single chlorine atoms' contributions. Increasing the number of chlorine atoms in biphenyl leads to the interference of the chlorine atoms' interactions with the solvent molecules and, as a result, to a nonlinearity of the solvent response in the process of hydration. We underline that the 1D RISM approach considers these effects in a proper way, even in the case of highly chlorinated compounds. In turn, the continuum solvent models (SM_6 and COSMO-SAC) are not sensitive to the nonlinear solvent response. We note that using the RISM model for solvent is essential for the efficiency of the SDC model. As such, we tested our correction scheme with the use of the PBSA solvent model instead of the 1D RISM (see the Supporting Information). The results show that the RISM-based SDC model is superior to the PBSA-based model. Thus, for the test set of polychlorinated benzenes, the mean of error and the standard deviation of error of the PBSA-SDC model are ~11.3 kcal/mol and ~6.7 kcal/mol, accordingly; that is much worse than the 1D RISM-SDC results.

The results of this work show the potential of the 1D RISM-SDC(QMq) approach for the description of a hydration/solvation process for a wide range of chemical solutes. It makes the model a good candidate for use in large-scale environmental modeling of hydration pathways of organic pollutants.

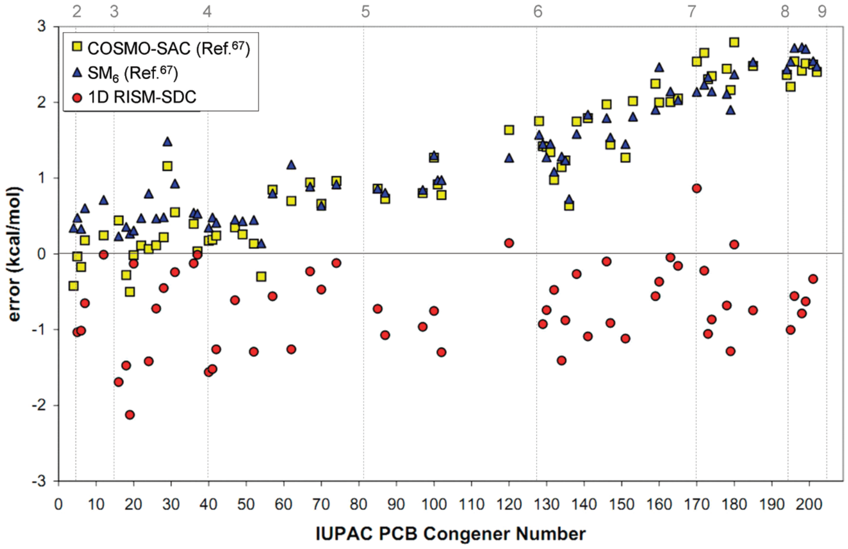


Figure 7. Errors for HFE predictions by the 1D RISM-SDC(QMq) model proposed in this work for the test set of polychlorobiphenyls (PCBs). The errors are compared with the corresponding literature data for SM_6 and COSMO-SAC (taken from ref 67). The HFE prediction error increases for SM_6 and COSMO-SAC with the increase in IUPAC number. At the same time, the 1D RISM-SDC(QMq) error remains the same for all PCBs. Dashed lines show the separation of the whole set of PCBs with respect to the number of chlorine atoms (shown on the top).

CONCLUSIONS

Here, we discussed a new method for predicting the hydration free energy (HFE) of organic pollutants and illustrated the efficiency of the method on a set of 220 chlorinated aromatic hydrocarbons that are in the list of persistent organic pollutants. The model is computationally inexpensive, and one HFE calculation takes only a minute on a standard PC (3.3 GHz). The method provides good accuracy for the test set of organic pollutant molecules. However, analysis of the results shows that the model performs better for polychlorinated benzenes than for polychlorobiphenyls. On one hand, that means that the SDC model might still require some improvement. That can be done in two directions: (i) one can use more sophisticated molecular theories, such as 3D RISM[16,68,69] [we note, however, that the 3D approach is significantly more computationally expensive than the 1D RISM approach used here (roughly by 2 orders of magnitude) and that might limit its application for large-scale screening of pollutants]; (ii) one can also work on the improvement of the theoretical part of the model by developing new, more efficient forms of the HFE functional. This is the subject of our future research.

On the other hand, the observed ~1 kcal/mol bias of the model results from experimental data for PCBs may be attributed to the differences in the quality of experimental data for polychlorinated benzenes and polychlorobiphenyls. We note that the sources of experimental data for these two classes of pollutants are different. However, as shown in Figure 6, the PCB congeners HFEs obtained from different sources can vary by several kilocalories per mole. We note that the problem of the lack of reliable experimental data for pollutants was highlighted in refs 2 and 65. Computational and theoretical scientists can do very little to improve the situation in that respect, but we hope that our results and analysis of the available experimental data will provoke experimentalists to revisit the question and, hopefully, to make additional independent measurements of HFE for CAHs. Such new experimental data would be very valuable in creating and testing new models for environmental modeling with high predictive ability.

ASSOCIATED CONTENT

Supporting Information. The composition of the training set and test subset of polychlorobiphenyls together with the corresponding experimental and calculated hydration free energies. Performance of the SDC model with the continuum PBSA model as the initial approximation. This material is available free of charge via the Internet at http://pubs.acs.org/.

AUTHOR INFORMATION

Corresponding Author

*Phone: +49 341 9959 804. Fax: +49 341 9959 999. E-mail: fedorov@mis.mpg.de.

ACKNOWLEDGMENT

This work was partially funded by Deutsche Forschungsgemeinschaft (DFG)—German Research Foundation, Research Grant FE 1156/2-1 and REA—Research Executive Agency, Grant No. 247500 "BioSol", Programme: FP7-PEOPLE-2009-IRSES. The authors would like to acknowledge the use of the Chemical Database Service at Daresbury. We are thankful to Andriy Kovalenko for his insightful comments on the history of RISM and to Sergey Gusarov for useful discussions on hybrid RISM/QM models. We thank David S. Palmer and Andrey I. Frolov for critical reading of the manuscript and useful discussions. We acknowledge Katarzyna Baier and Evelyn Weiser for their help with finding references for this paper.

REFERENCES

(1) Vallack, H. W.; *Environ. Toxicol. Pharmacol.* **1998**, *6*, 143–175.
(2) Valsaraj, K. T.; Thibodeaux, L. J. *J. Phys. Chem. Lett.* **2010**, *1*, 1694–1700.
(3) Aarhus Protocol on Persistent Organic Pollutants (POPs). http://www.unece.org/env/lrtap/pops_h1.htm (accessed December 18, 2009).
(4) Jones, K. C.; de Voogt, P. *Environ. Pollut.* **1999**, *100*, 209–221.
(5) Wine, P. H. *J. Phys. Chem. Lett.* **2010**, *1*, 1749–1751.
(6) Mackay, D. In *Multimedia Environmental Models: The Fugacity Approach*, 2nd ed.; CRC Press: Boca Raton, FL, 2001; Chapter Environmental Chemicals and Their Properties, pp 29—54.
(7) Beyer, A.; Biziuk, M. *Rev. Environ. Contam. Toxicol.* **2009**, *201*, 137–158.
(8) Scheringer, M.; Wegmann, F.; Fenner, K.; Hungerbuhler, K. *Environ. Sci. Technol.* **2000**, *34*, 1842–1850.
(9) Wania, F.; Mackay, D. *The Global Distribution Model. A Non-Steady State Multicompartment Mass Balance Model of the Fate of Persistent Organic Pollutants in the Global Environment*; University of Toronto: Scarborough, Canada, 2000
(10) Strand, A.; Hov, O. *Water, Air, Soil Pollut.* **1996**, *86*, 283–316.
(11) Liss, P. S.; Slater, P. G. *Nature* **1974**, *247*, 181–184.
(12) Modarresi, H.; Modarress, H.; Dearden, J. C. *SAR QSAR Environ. Res.* **2005**, *16*, 461–482.
(13) Ratkova, E. L.; Chuev, G. N.; Sergiievskyi, V. P.; Fedorov, M. V. *J. Phys. Chem. B* **2010**, *114*, 12068–12079.
(14) Jorgensen, W. L.; Maxwell, D. S.; TiradoRives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.
(15) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B* **2001**, *105*, 6474–6487.
(16) *Molecular Theory of Solvation*; Hirata, F., Ed.; Kluwer Academic Publishers: Dordrecht, Netherlands, 2003.
(17) Frenkel, D.; Smit, B. In *Understanding Molecular Simulation*; Academic Press: New York, 2002; Chapter Basics, pp 7—23.
(18) Kinoshita, M.; Okamoto, Y.; Hirata, F. *J. Am. Chem. Soc.* **1998**, *120*, 1855–1863.
(19) Kinoshita, M.; Okamoto, Y.; Hirata, F. *J. Chem. Phys.* **1999**, *110*, 4090–4100.
(20) Freedman, H.; Truong, T. N. *Chem. Phys. Lett.* **2003**, *381*, 362–367.
(21) Monson, P. A.; Morriss, G. P. *Adv. Chem. Phys.* **1990**, *77*, 451–550.
(22) Duh, D. M.; Haymet, A. D. J. *J. Chem. Phys.* **1995**, *103*, 2625–2633.
(23) Singer, S. J.; Chandler, D. *Mol. Phys.* **1985**, *55*, 621–625.
(24) Kovalenko, A.; Hirata, F. *J. Chem. Phys.* **1999**, *110*, 10095–10112.
(25) Chandler, D.; Andersen, H. C. *J. Chem. Phys.* **1972**, *57*, 1930–1937.
(26) Hirata, F.; Rossky, P. J. *Chem. Phys. Lett.* **1981**, *83*, 329–334.
(27) Palmer, D. S.; Sergiievskyi, V. P.; Jensen, F.; Fedorov, M. V. *J. Chem. Phys.* **2010**, *133*, 044104.
(28) Karino, Y.; Fedorov, M. V.; Matubayasi, N. *Chem. Phys. Lett.* **2010**, *496*, 351–355.
(29) Chandler, D.; Singh, Y.; Richardson, D. M. *J. Chem. Phys.* **1984**, *81*, 1975–1982.
(30) Kovalenko, A.; Hirata, F. *J. Chem. Phys.* **2000**, *112*, 10403–10417.
(31) Ten-no, S. *J. Chem. Phys.* **2001**, *115*, 3724–3731.
(32) Kovalenko, A.; Hirata, F. *J. Chem. Phys.* **2000**, *113*, 2793–2805.
(33) Chuev, G. N.; Fedorov, M. V.; Crain, J. *Chem. Phys. Lett.* **2007**, *448*, 198–202.
(34) Lee, P. H.; Maggiora, G. M. *J. Phys. Chem.* **1993**, *97*, 10175–10185.
(35) Ten-no, S.; Jung, J.; Chuman, H.; Kawashima, Y. *Mol. Phys.* **2010**, *108*, 327–332.
(36) Sato, K.; Chuman, H.; Ten-no, S. *J. Phys. Chem. B* **2005**, *109*, 17290–17295.

(37) Chuev, G. N.; Fedorov, M. V. *J. Comput. Chem.* **2004**, *25*, 1369–1377.

(38) Chuev, G. N.; Fedorov, M. V.; Chiodo, S.; Russo, N.; Sicilia, E. *J. Comput. Chem.* **2008**, *29*, 2406–2415.

(39) Abraham, M. H.; Andonianhaftvan, J.; Whiting, G. S.; Leo, A.; Taft, R. S. *J. Chem. Soc., Perkin Trans. 2* **1994**, 1777–1791.

(40) Ryu, S. A.; Park, S. J. *Fluid Phase Equilib.* **1999**, *161*, 295–304.

(41) Ashworth, R. A.; Howe, G. B.; Mullins, M. E.; Rogers, T. N. *J. Hazard. Mater.* **1988**, *18*, 25–36.

(42) Jantunen, L. M.; Bidleman, T. F. *Chemosphere* **2006**, *62*, 1689–1696.

(43) Rounds, S. A.; Pankow, J. F. *J. Chromatogr.* **1993**, *629*, 321–327.

(44) Brunner, S.; Hornung, E.; Santl, H.; Wolff, E.; Piringer, O. G.; Altschuh, J.; Brueggemann, R. *Environ. Sci. Technol.* **1990**, *24*, 1751–1754.

(45) Fedorov, M. V.; Kornyshev, A. A. *Mol. Phys.* **2007**, *105*, 1–16.

(46) Fedorov, M. V.; Flad, H. J.; Chuev, G. N.; Grasedyck, L.; Khoromskij, B. N. *Computing* **2007**, *80*, 47–73.

(47) Sergiievskyi, V. P.; Hackbusch, W.; Fedorov, M. V. *J. Comput. Chem.* **2011**, in press, DOI: 10.1002/jcc.21783.

(48) Lue, L.; Blankschtein, D. *J. Phys. Chem.* **1992**, *96*, 8582–8594.

(49) Pettitt, B. M.; Rossky, P. J. *J. Chem. Phys.* **1982**, *77*, 1451–1457.

(50) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269–6271.

(51) Frisch, M. J. *Gaussian 03*; Gaussian, Inc.: Wallingford, CT, 2004.

(52) Fletcher, D. A.; McMeeking, R. F.; Parkin, D. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 746–749.

(53) Hope, H.; Ottersen, T. *Acta Crystallogr., Sect. B: Struct. Sci.* **1978**, *34*, 3623–3626.

(54) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361–373.

(55) Jacobson, M. P.; Kaminski, G. A.; Friesner, R. A.; Rapp, C. S. *J. Phys. Chem. B* **2002**, *106*, 11673–11680.

(56) Feldman, H.; Dumontier, M.; Ling, S.; Haider, N.; Hogue, C. *FEBS Lett.* **2005**, *579*, 4685–4691.

(57) *Handbook of Chemoinformatics. 4 Bde. From Data to Knowledge*, 1st ed.; Gasteiger, J., Ed.; Wiley-VCH: New York, 2003.

(58) Dunnivant, F. M.; Coates, J. T.; Elzerman, A. W. *Environ. Sci. Technol.* **1988**, *22*, 448–453.

(59) Bamford, H. A.; Poster, D. L.; Baker, J. E. *J. Chem. Eng. Data* **2000**, *45*, 1069–1074.

(60) Bamford, H. A.; Poster, D. L.; Huie, R. E.; Baker, J. E. *Environ. Sci. Technol.* **2002**, *36*, 4395–4402.

(61) Lau, F. K.; Charles, M. J.; Cahill, T. M. *J. Chem. Eng. Data* **2006**, *51*, 871–878.

(62) Fendinger, N.; Glotfelty, D. *Environ. Toxicol. Chem.* **1990**, *9*, 731–735.

(63) Fang, F.; Chu, S. G.; Hong, C. S. *Anal. Chem.* **2006**, *78*, 5412–5418.

(64) Goss, K.-U.; Wania, F.; McLachlan, M. S.; Mackay, D.; Schwarzenbach, R. P. *Environ. Sci. Technol.* **2004**, *38*, 1626–1628.

(65) Shunthirasingham, C.; Lei, Y. D.; Wania, F. *Environ. Sci. Technol.* **2007**, *41*, 3807–3814.

(66) Bamford, H.; Baker, J. In *Review of Methods and Measurements of Selected Hydrophobic Organic Contaminant Aqueous Solubilities, Vapor Pressures, and Air-Water Partition Coefficients*; National Institute of Standards and Technology: Gaithersburg, MD, 1998; Chapter Physical and Chemical properties.

(67) Phillips, K. L.; Sandler, S. I.; Greene, R. W.; Di Toro, D. M. *Environ. Sci. Technol.* **2008**, *42*, 8412–8418.

(68) Palmer, S.; Frolov, A. I.; Ratkova, E. L.; Fedorov, M. V. *J. Phys.: Condens. Matter* **2010**, *22*, 492101.

(69) Frolov, A. I.; Ratkova, E. L.; Palmer, D. S.; Fedorov, M. V. *J. Phys. Chem. B* **2011**, doi: 10.1021/jp111271c.

Using Theory to Reconcile Experiment: The Structural and Thermodynamic Basis of Ligand Recognition by Phenylethanolamine *N*-Methyltransferase (PNMT)

Pramod C. Nair,[†] Alpeshkumar K. Malde,[†] and Alan E. Mark*,[†,‡]

[†]School of Chemistry and Molecular Biosciences (SCMB) and [‡]Institute for Molecular Bioscience (IMB), The University of Queensland (UQ), St. Lucia Campus, Brisbane, QLD 4072 Australia

ABSTRACT: A fundamental challenge in computational drug design is the availability of reliable and validated experimental binding and structural data against which theoretical calculations can be compared. In this work a combination of molecular dynamics (MD) simulations and free energy calculations has been used to analyze the structural and thermodynamic basis of ligand recognition by phenylethanolamine *N*-methyltransferase (PNMT) in an attempt to resolve uncertainties in the available binding and structural data. PNMT catalyzes the conversion of norepinephrine into epinephrine (adrenaline), and inhibitors of PNMT are of potential therapeutic importance in Alzheimer's and Parkinson's disease. Excellent agreement between the calculated and recently revised relative binding free energies to human PNMT was obtained with the average deviation between the calculated and the experimentally determined values being only 0.8 kJ/mol. In this case, the variation in the experimental data over time is much greater than the uncertainties in the theoretical estimates. The calculations have also enabled the refinement of structure−activity relationships in this system, to understand the basis of enantiomeric selectivity of substitution at position three of tetrahydroisoquinoline and to identify the role of specific structural waters. Finally, the calculations suggest that the preferred binding mode of *trans*-(1*S*,2*S*)-2-amino-1-tetralol is similar to that of its epimer *cis*-(1*R*,2*S*)-2-amino-1-tetralol and that the ligand does not adopt the novel binding mode proposed in the pdb entry 2AN5. The work demonstrates how MD simulations and free energy calculations can be used to resolve uncertainties in experimental binding affinities, binding modes, and other aspects related to X-ray refinement and computational drug design.

■ INTRODUCTION

The primary challenge in rational drug design is to understand how a protein recognizes a specific ligand. X-ray structures of protein−ligand complexes can provide detailed information regarding the location of the ligand within the complex and information on specific ligand-protein interactions. They do not, however, provide information on how these interactions may contribute to the net binding free energy. Thus, the question of why a specific ligand binds better than another is frequently open to speculation.[1−3] An additional difficulty is that the binding mode (the position, the orientation, and the conformation) of small ligands can be uncertain in medium-(0.2−0.3 nm) to low-(>0.3 nm) resolution structures where alternative binding modes cannot be distinguished based solely on the electron density.[4,5] In such cases the combined use of molecular dynamics (MD) simulations and free energy calculations (FE) has proved to be a powerful approach to identify the thermodynamically stable binding mode. For example, Malde and Mark[5] recently discussed a number of examples where the stereochemistry, conformation, and orientation and the protonation and tautomeric states of specific ligand structures were ambiguous and where the published structures were possibly inappropriate. A case in point was the binding of noradrenochrome and tetrahydroisoquinoline-7-sulphonamide to phenyl ethanolamine *N*-methyltransferase (PNMT). In particular, the novel binding mode proposed in the case of noradrenochrome and the proposed orientation of the sulphonamide group in the ligand tetrahydroisoquinoline-7-sulphonamide were unstable, suggesting that the crystallographic models represented high-energy states.

PNMT catalyzes the formation of epinephrine (adrenaline) from norepinephrine.[6] In the central nervous system (CNS) epinephrine is linked to the control of blood pressure and respiration as well as the secretion of pituitary hormones.[7,8] CNS specific PNMT inhibitors are of potential therapeutic importance as the progression of diseases, such as Alzheimer's and Parkinson's diseases, is associated with increased levels of epinephrine in the CNS.[9−12] Several classes of PNMT inhibitors have been identified. These include derivatives of phenylethylamine and α-methylphenethylamine (amphetamine) as well as sulfhydryl-binding agents and benzamidine-based compounds.[9,10] While many of these compounds are potent inhibitors of PNMT in vitro, the design of inhibitors that are active in vivo and in particular within the CNS remains a significant challenge. For example, phenylethylamine- and benzamidine-based PMNT inhibitors are of little use clinically as they show cross reactivity with the α-adrenergic receptor.[13] Tetrahydroisoquinoline (THIQ) (Figure 1, molecule **1**) derivatives are also potent inhibitors of PNMT in vitro and exhibit good selectivity.[14,15] It has also been claimed that derivatives of THIQ should be active within the CNS.[14,16] THIQ is a structural analog of norepinephrine, the main substrate of PMNT. Based on structure−activity relationships (SAR), it has been reported that the combination of an electron-withdrawing substituent at position 7 and an alkyl substituent at position 3 on the THIQ scaffold (see Figure 1) leads to enhanced PNMT inhibition with good selectivity over the α-adrenergic receptor.[17−19] A number of X-ray crystal structures of PNMT:THIQ complexes have been

Received: December 16, 2010
Published: March 30, 2011

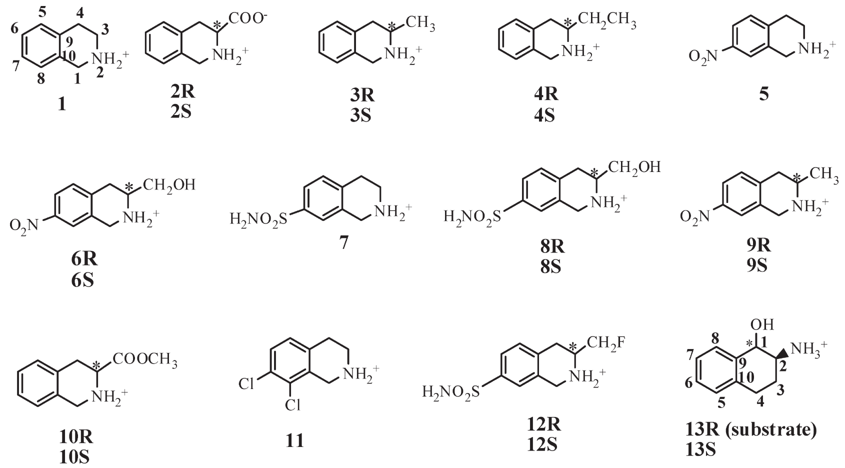


Figure 1. Three-, seven-, and eight-substituted tetrahydroisoquinoline (THIQ) derivatives and 2-amino-1-tetralol used in the study. The numbering scheme for THIQ and 2-amino-1-tetralol is given for ligands **1** and **13**, respectively; * indicates the chiral center.

reported. In these structures the bulkier substituents at position 7 of the THIQ ring were found to bind within a pocket which is not evident in the ligand-free protein, suggesting that the protein has a high degree of conformational plasticity complicating structure-based ligand design.[20,21] Detailed analysis of this system is further complicated by the fact that there is uncertainty regarding the experimental binding data.[21−24] This is illustrated in Table 1, which lists experimental binding data for the PNMT inhibitors shown in Figure 1. As can be seen, differences of between 10- and a 1000-fold in the value of K_i, which corresponds to an uncertainty of between 5 and 11 kJ/mol in the free energy of binding, have been reported by the same authors using different assay conditions.[23] In addition, binding data is only available for racemic mixtures in several cases or only for one of the potential isomers in others.

In the present study a combination of MD simulations and FE calculations have been used in order to understand in detail the structural and thermodynamic basis of ligand recognition by PNMT. A series of THIQ derivatives with substitutions at positions three, seven, and eight that shows a wide range of binding free energies (−48 to > −15 kJ/mol) and the potential enantiomeric selectivity have been examined.[19,21,23−27] Excellent agreement between the calculated and the recent experimental values for the FE of binding of these THIQ derivatives to human PNMT was obtained with the variation in the experimental data over time being much greater than the uncertainties in the theoretical estimates. In addition, alternative binding modes of *trans*-(1*S*,2*S*)-2-amino-1-tetralol (Figure 1, **13S**) to human PNMT have been considered and the role specific water molecules play in stabilizing the binding of *cis*-(1*R*,2*S*)-2-amino-1-tetralol (Figure 1, **13R**) to PNMT was examined.

■ METHODS

MD Simulations. All MD simulations were performed using the GROMOS96 simulation package in conjunction with the GROMOS 53A6 force field.[28,29] The initial structure of human PNMT complexed with the cofactor *S*-adenosyl-L-homocysteine (SAH) and 1,2,3,4-tetrahydro-isoquinoline-7-sulphonamide (inhibitor 7) taken from the pdb entry 1HNN was used for all studies involving THIQ derivatives. The initial structures for the studies involving the binding of **13R** and **13S** were taken from pdb entries 2AN3 and 2AN5, respectively. The topologies of the ligands (Figure 1) were generated using the 'Automated Topology Builder' (ATB, http://compbio.biosci.uq.edu.au/atb/), version 2009-06-10.[30] Missing parameters were manually assigned where possible based on comparable groups within the GROMOS force field.[29] The parameters used for the $-NO_2$ and $-SO_2NH_2$ groups are shown in Tables 2a and 2b. Simulations of the ligands free in solution were performed by placing the ligand in a periodic rectangular box containing 975 simple point charge (SPC) water molecules.[31] For the systems in which the ligand was bound to the protein, the configuration of the solvent was relaxed by performing a steepest descent minimization in which the protein and ligand atoms were positionally restrained to their initial positions using a harmonic interaction potential with a force constant of 2×10^3 kJ/mol/nm^2. The system was then further equilibrated by performing a 200 ps simulation, with the heavy atoms of the protein positionally restrained, before a series of unrestrained MD simulations were commenced. All the simulations were performed at constant temperature (298 K) and pressure (1 atm). This was achieved using a Berendsen thermostat[32] with a coupling time of 0.1 ps and a Berendsen barostat with a coupling time of 0.5 ps. The isothermal compressibility was set to 4.575×10^{-4} kJ/mol/nm^3. Nonbonded interactions were calculated using a twin-range cutoff. Interactions within the short-range cutoff of 0.8 nm were updated every time step. Interactions within the longer-range cutoff of 1.4 nm were updated every 5 time steps together with the pairlist. To correct for the truncation of electrostatic interactions beyond the 1.4 nm long-range cutoff, a reaction field correction was applied using an effective dielectric (ε) of 54.[28] The equations of motion were integrated using the leapfrog scheme with a 2 fs time step. Initial velocities at a given temperature were taken from a Maxwell−Boltzmann distribution. The lengths of all bonds were constrained to ideal values using the SHAKE algorithm with a geometric tolerance of 0.0001.[33]

System Setup. The initial structure of the different protein: ligand complexes were derived from the crystal structure of PNMT complexed with 1,2,3,4-tetrahydroisoquinoline-7- sulphonamide (inhibitor 7) and the cofactor *S*-adenosyl-L-homocysteine, pdb code 1HNN.[34] The GROMOS force field treats aliphatic hydrogen atoms as united atoms together with the carbon atom to which they are attached. The coordinates of polar hydrogen atoms (bound to nitrogen, oxygen, or sulfur atoms)

Table 1. Experimental (Human and/or Bovine) and Calculated (Human) Relative Gibbs Free Energies of Binding (Calculated Relative to Inhibitor 5) with PNMT[a]

	inhibitor	K_i (μM)	ref	ΔG_{expt} (kJ/mol)	relative $\Delta\Delta G_{expt}$ (kJ/mol)	relative $\Delta\Delta G_{calcd}$ (kJ/mol) [d]	$\Delta\Delta G_{expt} - \Delta\Delta G_{calcd}$ (kJ/mol)
1	**1**	15.0 ± 1	22	−27.8 ± 0.2	13.1 ± 0.6	9.7 ± 1.8	3.4
2		***5.8 ± 0.5***	24	***−30.1 ± 0.2***	***10.8 ± 0.6***	***9.7 ± 1.8***	***1.1***
3	**1**[b]	10.0 ± 0.9	19, 25	−28.8 ± 0.2	12.1 ± 0.6	9.7 ± 1.8	2.4
4	**2**[b,c]	>2000	25	> −15.5	>25.4	–	–
5	**2R**	–	–	–	–	66.9 ± 2.5	–
6	**2S**	–	–	–	–	99.5 ± 1.6	–
7	**3R**[b]	38.0 ± 2	19	−25.4 ± 0.1	15.5 ± 0.5	14.0 ± 2.4	1.5
8	**3S**[b]	1.0 ± 0.1	19	−34.5 ± 0.2	6.4 ± 0.6	4.8 ± 2.0	1.6
9	**4**[b,c]	24.0 ± 1	19, 25	−26.6 ± 0.1	14.3 ± 0.5	–	–
10	**4R**	–	–	–	–	8.6 ± 2.6	–
11	**4S**	–	–	–	–	4.2 ± 2.1	–
12	**5**	0.078 ± 0.014	21	−40.9 ± 0.4	0.0	0.0	0.0
13	**6R**	***0.017 ± 0.01***	21	***−44.7 ± 0.1***	***−3.8 ± 0.5***	***−4.6 ± 2.2***	***0.8***
14	**6R**[b]	0.24 ± 0.04	19	−38.1 ± 0.4	2.8 ± 0.8	−4.6 ± 2.2	7.4
15	**6S**[b]	0.9 ± 0.03	19	−34.8 ± 0.1	6.1 ± 0.5	3.1 ± 2.1	3.0
16	**7**	0.58 ± 0.04	22	−35.9 ± 0.2	5.0 ± 0.6	−1.8 ± 2.5	6.8
17		0.28 ± 0.02	23, 27	−37.7 ± 0.1	3.2 ± 0.5	−1.8 ± 2.5	5.0
18		***0.12 ± 0.02***	21, 26	***−39.8 ± 0.3***	***1.1 ± 0.7***	***−1.8 ± 2.5***	***2.9***
19	**7**[b]	0.56 ± 0.04	22	−35.8 ± 0.1	4.9 ± 0.5	−1.8 ± 2.5	6.7
20	**8R**	2.1 ± 0.1	22	−32.7 ± 0.1	8.2 ± 0.5	−0.8 ± 2.3	9.0
21		***0.052 ± 0.004***	24	***−41.9 ± 0.1***	***−1.0 ± 0.5***	***−0.8 ± 2.3***	***−0.2***
22	**8R**[b]	0.34 ± 0.06	22	−37.2 ± 0.3	3.7 ± 0.7	−0.8 ± 2.3	4.5
23	**8S**	–	–	–	–	1.0 ± 2.9	–
24	**9R**[b]	1.3 ± 0.1	19	−33.9 ± 0.2	7.0 ± 0.6	10.2 ± 2.1	−3.2
25	**9S**[b]	0.25 ± 0.02	19	−38.0 ± 0.2	2.9 ± 0.6	−2.1 ± 1.3	5.0
26	**10**[b,c]	69.5 ± 6	19, 25	−24.0 ± 0.3	16.9 ± 0.7	–	–
27	**10R**	–	–	–	–	12.7 ± 2.8	–
28	**10S**	–	–	–	–	22.3 ± 2.8	–
29	**11**	0.3 ± 0.04	22	−37.5 ± 0.2	3.4 ± 0.6	−8.3 ± 2.2	11.7
30		***0.0031 ± 0.0006***	16, 23	***−48.9 ± 0.3***	***−8.0 ± 0.7***	***−8.3 ± 2.2***	***0.3***
31	**11**[b]	0.22 ± 0.05	22	−38.3 ± 0.5	2.6 ± 0.9	−8.3 ± 2.2	10.9
32	**12R**[b]	0.15 ± 0.01	16	−39.3 ± 0.2	1.6 ± 0.6	−1.7 ± 2.5	3.3
33	**12S**	–	–	–	–	9.3 ± 2.3	–

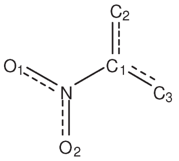
[a] The values in the bold italics correspond to the most recent experimental data for binding to human PNMT at the time of publication. [b] Data for binding to bovine PNMT. [c] Racemic mixture. [d] The standard error for the relative free energy involving two legs were calculated by formula $[(s_1)^2 + (s_2)^2]^{1/2}$, where $s_1$and s_2 are standard errors of two different legs.

and aromatic hydrogen atoms were generated based on ideal geometries. In the chain A of pdb 1HNN, the first 21 N-terminal residues were not resolved. In addition atoms were missing in seven other residues. The missing N-terminal residues were not included in the model as they lie far from the active site. The coordinates for the missing atoms in residues Arg33, Lys136, Arg145, Gln163, Glu241, Arg245, and Leu282 were generated as follows: All atoms in the protein for which the coordinates were available were positionally restrained using a harmonic restraining potential and a force constant of 2 × 10^3 kJ/mol/nm^2. The coordinates for the other atoms were arbitrarily set to zero. A series of minimizations were then performed in which the bond length, the bond angle, the dihedral, the improper dihedral, and finally the nonbonded terms for the missing atoms were added progressively. After that the restraints were removed, and the whole protein was minimized. The charges of the ionizable groups were chosen to correspond to a pH of 7, resulting in a net charge of −2e. No counterions were added. The histidine residues were assigned appropriate tautomeric configurations based on the local environment of these residues. The protein was placed at the center of a periodic truncated octahedral box, which was filled with 7239 SPC water molecules. In this procedure, the minimum distance between water oxygen atoms and nonhydrogen protein atoms was 0.23 nm, and the minimum distance between the protein and the wall of the box was 0.9 nm.

Free Energy Calculations. The change in Gibbs FE (ΔG), associated with different mutations of the ligand in water and in the protein was determined using the coupling parameter approach in conjunction with the thermodynamic integration eq 1:

$$\Delta G_{0\rightarrow 1} = \int_{\lambda=0}^{\lambda=1} \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_{\lambda} d\lambda \qquad (1)$$

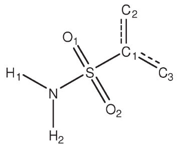
Table 2a. Bonded and Nonbonded Parameters for the Functional Group $-NO_2$



atom[a]	atom type	$[C6(i,i)]^{1/2}$ $[(kJ/mol\,nm^6)^{1/2}]^b$	$[C12(i,i)]^{1/2}$ $[10^{-3}$ $(kJ/mol\,nm^{12})^{1/2}]^b$	partial atomic charge (e)
N	7	0.04936	1.523, 2.250	0.700
O1	1	0.04756	1.000, 1.130	0.350
O2	1	0.04756	1.000, 1.130	0.350
C1	12	0.04838	2.222	0.000
bond	bond type	bond length b_0 (nm)	force constant K_b (10^6kJ/mol/nm^4)	
C1−N	21	0.147	8.71	
N−O1, N−O2	5	0.123	16.6	
angle	angle type	bond angle θ_0 (°)	force constant K_θ (kJ/mol)	
C1−N−O1, C1−N−O2	22	117	635	
O1−N−O2	36	126	575	
dihedral	dihedral type	phase shift cos(δ)	force constant K_φ (kJ/mol)	multiplicity m
C2−C1−N−O1	14	−1	33.5	2
improper dihedral angle	improper dihedral type	improper dihedral angle ξ_0 (deg)	force constant K_ξ (kJ/mol/deg^2)	
N−O1−O2−C1	1	0.0	0.0510	

[a] The carbon C2 and C3 have standard parameters for aromatic carbons in the GROMOS 53A6 parameter set. [b] Lennard-Jones parameters C6-(i,j) and C12(i,j) were obtained using the following combination rules: $C_6(i,j) = [C_6(i,i)^{1/2}\ C_6(j,j)\ ^{1/2}]$ and $C_{12}(i,j) = [C_{12}(i,i)^{1/2}\ C_{12}(j,j)\ ^{1/2}]$. The atom, bond, angle, and dihedral types are from the GROMOS 53A6 parameter set.[29]

Table 2b. Bonded and Nonbonded Parameters for the Functional Group $-SO_2NH_2$



atom[a]	atom type	$[C6(i,i)]^{1/2}$ $[(kJ/mol\,nm^6)^{1/2}]^b$	$[C12(i,i)]^{1/2}$ $[10^{-3}$ $(kJ/mol\,nm^{12})^{1/2}]^b$	partial atomic charge (e)
S	42	0.10277	4.6366	1.157
O1	44	0.047652	0.86686, 1.1250	−0.550
O2	44	0.047652	0.86686, 1.1250	−0.561
N	6	0.04936	1.523, 1.943	−0.832
H1	21	0.0	0.0	0.393
H2	21	0.0	0.0	0.393
C1	12	0.04838	2.222	0.000
bond	bond type	bond length b_0 (nm)	force constant K_b (10^6 kJ/mol/nm^4)	
C1−S	31	0.178	5.94	
S−O	25	0.150	8.37	
S−N	41	0.153	8.04	
angle	angle type	bond angle θ_0 (°)	force constant K_θ (kJ/mol)	
C1−S−N, C1−S−O1, O1−S−O2, O1−S−N	13	109.5	520	
dihedral	dihedral type	phase shift cos(δ)	force constant K_φ (kJ/mol)	multiplicity m
C2−C1−S−N	43[c]	+1	0.75	2
C1−S−N−H1	40	+1	1.0	6
improper dihedral angle	improper dihedral type	improper dihedral angle ξ_0 (°)	force constant K_ξ (kJ/mol/deg^2)	
N−H1−H2-S	1	0.0	0.0510	

[a] The carbon C2 and C3 have standard parameters for aromatic carbons in the GROMOS 53A6 parameter set. [b] Lennard-Jones parameters C6-(i,j) and C12(i,j) were obtained using the following combination rules: $C_6(i,j) = [C_6(i,i)^{1/2}\ C_6(j,j)\ ^{1/2}]$ and $C_{12}(i,j) = [C_{12}(i,i)^{1/2}\ C_{12}(j,j)\ ^{1/2}]$. The atom, bond, angle, and dihedral types are from the GROMOS 53A6 parameter set.[29] [c] Nonstandard dihedral type derived by fitting the molecular mechanics (MM) dihedral profile to the one obtained from quantum mechanical (QM) calculations (data not shown).

where $\lambda = 0$ corresponded to the initial state of the system, and $\lambda = 1$ corresponded to the final state of the system. H is the Hamiltonian of the system, and the brackets $<...>_\lambda$ correspond to an average over an equilibrium ensemble at λ. The relative FE of binding $\Delta\Delta G$ was determined from the difference in the change in FE of performing the same mutation free in solution and bound to the protein. Equation 1 was integrated by performing separate simulations at a series of 15 (0.0, 0.1, 0.2, ..., 0.8, 0.9, 1.0) λ points, including 0.025, 0.05, 0.95, 0.975 (to smooth the integrand) in both the bound and unbound states. For the mutations in water, a 1 ns simulation was performed at each λ value. For the mutations in the protein, the system was first equilibrated for 0.2 ns, and 1.8 ns of sampling used to provide an initial estimate of $<\partial H/\partial\lambda>_\lambda$. In cases where the value of $<\partial H/\partial\lambda>_\lambda$ had not converged, the simulations were extended to 3 ns. To determine the degree of convergence, thermodynamic cycles wherein the molecules were transformed from one to another in circular path in water, and when bound to the protein, were constructed. The mutations were chosen in order to maximize the number of closed thermodynamic cycles that could be generated with a limited number of mutations. The degree of convergence was also checked by performing the forward and backward mutations. To prevent numerical instabilities as atoms were created or destroyed, the soft-core potential as described by Beutler et al.[35,36] was used with $\alpha_{LJ} = 0.5$ and $\alpha_C = 0.5$ nm^2. The area beneath the curve in (1) was estimated using a trapezoidal approximation. The error in $<\partial H/\partial\lambda>_\lambda$ was estimated using a block averaging procedure at each λ-point.[37] The individual errors were then integrated to yield an estimate of the error in ΔG. All mutations performed as part of this work are summarized in Table 3. All possible thermodynamic cycles that could be constructed from these mutations are shown diagrammatically in Figure 2.

Table 3. Change in the Gibbs FE for Mutations of Pairs of Inhibitors Listed in Figure *1* in Water and in PNMT

	ΔG(kJ/mol)				ΔG(kJ/mol)				ΔΔG (kJ/mol)		
	water				PNMT				PNMT−water		
mutation	forward[a]	backward[a]	|hysteresis|	average[b]	forward[a]	backward[a]	|hysteresis|	average[b]	$\Delta\Delta G_{calcd}$[c]	$\Delta\Delta G_{exp}$[d]	$\Delta\Delta G_{expt} - \Delta\Delta G_{calcd}$
1−2R	−372.4(1.2)	371.7(1.4)	0.7	−372.1 ± 1.4	−255.2(2.5)	268.1(2.2)	12.9	−261.7 ± 2.4	110.4 ± 2.8	−	−
1−2S	−371.9(1.6)	371.3(2.3)	0.6	−371.6 ± 2.0	−288.1(2.1)	297.3(1.7)	9.2	−292.7 ± 1.9	78.9 ± 2.8	−	−
1−3R	−5.5 (0.3)	5.1 (0.5)	0.4	−5.3 ± 0.4	−0.7(1.9)	1.2(1.2)	0.5	−1.0 ± 1.5	4.3 ± 1.6	−	−
1−3S	−5.5 (0.2)	5.3 (0.4)	0.2	−5.4 ± 0.3	−10.1(0.9)	10.6 (1.1)	0.5	−10.3 ± 1.0	−4.9 ± 1.0	−	−
1−4R	−3.2(1.3)	3.5(1.1)	0.3	−3.4 ± 1.2	−4.7(1.9)	4.3(1.0)	0.4	−4.5 ± 1.5	−1.1 ± 1.9	−	−
1−4S	−4.4(1.4)	4.7(0.4)	0.3	−4.6 ± 0.9	−9.1(0.4)	11.1(0.6)	2.0	−10.1 ± 0.5	−5.5 ± 1.0	−	−
1−5	6.2(0.6)	−6.1(0.9)	0.1	6.2 ± 0.8	−3.1(1.8)	4.0(1.5)	0.9	−3.5 ± 1.6	−9.7 ± 1.8	−10.8 ± 0.6	−1.1
1−6R	−39.2(0.8)	39.4(1.0)	0.2	−39.3 ± 0.9	−55.0(0.8)	52.2(1.1)	2.8	−53.6 ± 0.9	−14.3 ± 1.3	−14.6 ± 0.3	−0.3
1−6S	−37.2(0.7)	38.3(1.4)	1.1	−37.8 ± 1.0	−47.2(1.4)	48.9(2.4)	1.7	−48.1 ± 1.9	−10.3 ± 2.1	−	−
1−7	−342.8(1.4)	343.1(0.7)	0.3	−343.0 ± 1.1	−353.8(1.1)	355.2(1.7)	1.4	−354.5 ± 1.4	−11.5 ± 1.8	−9.7 ± 0.5	1.8
1−8R	−388.6(0.7)	388.7(0.9)	0.1	−388.7 ± 0.8	−397.9(1.3)	400.5(1.1)	2.7	−399.2 ± 1.2	−10.5 ± 1.4	−11.8 ± 0.3	−1.3
1−8S	−388.2(1.4)	387.9(0.8)	0.3	−388.1 ± 1.1	−397.8(0.8)	395.7(1.2)	2.1	−396.8 ± 2.0	−8.7 ± 2.3	−	−
1−9R	0.2(0.2)	−0.3(0.4)	0.5	0.2 ± 0.3	1.0(0.8)	−1.4(0.9)	0.4	1.2 ± 0.9	1.0 ± 1.0	−	−
1−9S	0.2(0.4)	−0.1(0.3)	0.3	0.2 ± 0.4	−12.4(1.8)	13.8(1.5)	1.4	−13.0 ± 1.7	−13.2 ± 1.7	−	−
1−10R	−124.2(0.9)	123.8(1.5)	0.4	−124.0 ± 1.2	−119.8(1.5)	122.2(2.1)	2.5	−121.0 ± 1.8	3.0 ± 2.2	−	−
1−10S	−123.2(1.4)	123.7(1.5)	0.5	−123.5 ± 1.5	−109.5(1.2)	112.2(1.8)	2.7	−110.9 ± 1.5	12.6 ± 2.1	−	−
1−11	−17.8(0.5)	17.9(0.6)	0.1	−17.8 ± 0.6	−33.9(0.7)	37.6(1.3)	3.7	−35.8 ± 1.0	−18.0 ± 1.2	−18.8 ± 0.4	−0.8
1−12R	−351.4(0.6)	351.8(0.8)	0.4	−351.6 ± 0.7	−362.3(2.3)	363.6(1.1)	1.3	−363.0 ± 1.7	−11.6 ± 1.8	−	−
1−12S	−351.2(0.7)	351.5 (0.3)	0.3	−351.6 ± 0.5	−353.1(1.2)	350.8(1.7)	2.3	−352.0 ± 1.4	−0.4 ± 1.5	−	−
2R−5	377.2(1.2)	−378.1(1.9)	0.9	377.7 ± 1.6	310.8(1.9)	−	−	−	−66.9 ± 2.5	−	−
2S−5	377.4(0.7)	−377.5(1.7)	0.1	377.5 ± 1.2	274.2(0.6)	−281.8 (1.6)	7.6	278.0 ± 1.1	−99.5 ± 1.6	−	−
2R−6R	332.0(1.6)	−331.8(0.9)	0.2	331.9 ± 1.3	286.6(2.3)	−	−	−	−46.3 ± 2.6	−	−
2S−6S	332.2(0.9)	−332.4(1.3)	0.2	332.3 ± 1.1	236.7(1.7)	−205.0(2.4)	31.7	220.9 ± 2.1	−111.4 ± 2.4	−	−
2R−9S	371.8(0.7)	−372.9(1.9)	1.1	372.4 ± 1.3	315.0(2.3)	−	−	−	−57.4 ± 2.6	−	−
2S−9R	372.2(0.4)	−372.7(1.4)	0.5	372.5 ± 0.9	272.2(2.3)	−252.0(1.4)	20.2	262.1 ± 1.9	−110.4 ± 2.1	−	−
3S−8R	−382.2(3.4)	382.1(1.4)	0.1	−382.2 ± 2.4	−398.2(2.6)	394.1(1.8)	4.1	−396.2 ± 2.2	−14.0 ± 3.2	−	−
3S−9S	6.2(0.6)	−6.1(1.1)	0.1	6.2 ± 0.9	−8.6(2.6)	8.0(1.2)	0.6	−8.3 ± 1.9	−14.5 ± 2.1	−	−
4R−7	−340.6(0.9)	340.2(1.3)	0.4	−340.4 ± 1.1	−353.7(2.1)	357.5(1.8)	3.8	−355.6 ± 2.0	−15.2 ± 2.3	−	−
4R−10S	−120.0(1.3)	120.0(0.8)	0.0	−120.0 ± 1.1	−103.3(1.5)	−	−	−	16.7 ± 1.8	−	−
5−6S	−44.5(1.5)	44.2(0.5)	0.3	−44.4 ± 1.0	−41.3(1.3)	41.2(2.2)	0.1	−41.3 ± 1.8	3.1 ± 2.1	−	−
5−9S	−6.2(0.5)	6.3(0.5)	0.1	−6.2 ± 0.5	−8.1(1.1)	8.5(1.3)	0.4	−8.3 ± 1.2	−2.1 ± 1.3	−	−
6S−10S	−84.2(1.3)	85.1(0.6)	0.9	−84.7 ± 1.0	−69.8(2.2)	65.8(1.7)	4.0	−67.5 ± 2.0	17.2 ± 2.2	−	−
7−8S	−44.7(0.4)	45.2(0.8)	0.5	−45.0 ± 0.6	−45.8(1.9)	−	−	−	−0.8 ± 2.0	−	−

[a] The values in the parentheses show the standard error estimate obtained by block averaging. [b] The standard error in the average column were calculated by $[(s_1)^2 + (s_2)^2/2]^{1/2}$, where s_1 and s_2 are the standard error for the forward and backward mutations, respectively. [c] The error for the $\Delta\Delta G_{calcd}$ (PNMT−water) was calculated by formula $[(s_1)^2 + (s_2)^2]^{1/2}$, where s_1 and s_2 are the standard error for the mutations in water and PNMT, respectively. [d] Where multiple values for the FE of binding are available only the values obtained from refs 16, 21, 23, 24, and 26 shown as bold italics in Table 1 were used as these were the most recently available at the time of publication.

■ RESULTS AND DISCUSSION

MD Simulation of Ligand−PNMT Complexes. To examine if the system was equilibrated and if the ligand−PNMT complexes were stable in the force field used, the atomic positional root-mean-square deviation (rmsd) of the protein bound to inhibitor 7 with respect to initial X-ray crystal structure (pdb code 1HNN) was calculated. The average rmsd value with respect to the starting structure equilibrates after approximately 2 ns of simulation at 0.2 nm for the backbone atoms and 0.25 nm for all atoms, respectively, with the interactions between ligand and binding site remaining essentially the same as that proposed in the X-ray crystal structure. The aromatic ring of the THIQ nucleus is involved in a π-stacking arrangement with the side chain of Phe181. The ring N of THIQ formed a salt bridge with the side chain carboxylate of Glu219 and the water-mediated hydrogen bond with the side chain O of Asn39 and side chain O of Asp267. Almost all of the other ligands showed a similar degree of stability and similar interactions with PNMT. The exceptions were the molecules **2R** and **2S**. Molecules **2R** and **2S**, which are very weak inhibitors (>2000 μM), cannot adopt the same binding mode as the other ligands due to the fact molecule **2** has a charged carboxylate group and will be discussed in detail later.

Convergence of the Free Energy Calculations in Water and Bound to PNMT. The degree of convergence in the FE calculations was monitored in two ways. Out of the 33 pairs of mutations investigated, 28 pairs were performed in both the forward and the backward directions (Table 3). In addition, thermodynamic cycles in water and in the protein were constructed as described in the Methods Section. As the FE is a state function, the FE for the forward and backward mutations should be identical except for the sign. In addition, the FE for any closed cycle should be zero. Figure 2 shows all possible thermodynamic cycles for the different mutations performed in water and PNMT. All possible three-membered thermodynamic cycles that can be constructed from Table 3 are listed in Table 4. Using the mean of the FE in the forward and backward directions all three-membered cycles in water close to within 1.1 kJ/mol. The convergence of the forward and backward transformations in

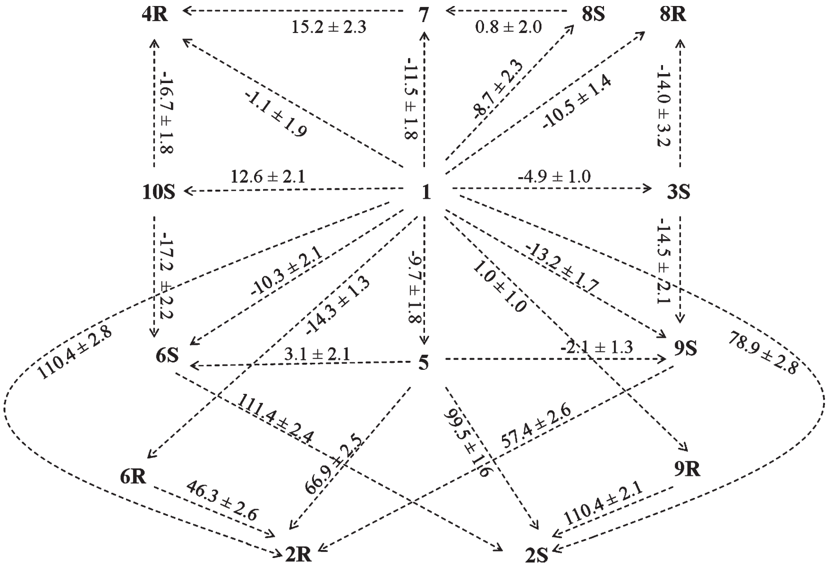


Figure 2. Schematic showing all possible closed thermodynamic cycles can be built from the mutations listed in Table 3. Note that the arrows refer the mutation performed in a given direction. The values listed on each line correspond to the difference in the average FE values between PNMT and water for these mutations.

Table 4. All Possible Three-Memebered Thermodynamic Cycles That Can Be Constructed from Table 3

thermodynamic cycles	water ΔG^a (kJ/mol)^b	PNMT ΔG^a (kJ/mol)^b	PNMT−water $\Delta\Delta G^a$ (kJ/mol)^c
1→**2R**→**5**→**1**	−0.6 ± 2.3	186.4 ± 3.5	187.0 ± 4.1
1→**2R**→**6R**→**1**	−0.9 ± 2.1	77.5 ± 3.4	78.4 ± 4.0
1→**2S**→**5**→**1**	−0.3 ± 2.5	−11.2 ± 2.7	−10.9 ± 3.7
1→**2S**→**9R**→**1**	0.7 ± 2.2	−31.8 ± 2.8	−32.5 ± 3.6
1→**3S**→**8R**→**1**	1.1 ± 2.5	−7.3 ± 2.7	−8.4 ± 3.7
1→**3S**→**9S**→**1**	0.6 ± 1.0	−5.6 ± 2.7	−6.2 ± 2.9
1→**4R**→**10S**→**1**	0.3 ± 2.2	3.3 ± 2.6	3.0 ± 3.4
1→**5**→**9S**→**1**	−0.2 ± 1.0	1.2 ± 2.6	1.4 ± 2.8
1→**6S**→**5**→**1**	0.4 ± 1.4	−3.3 ± 3.1	−3.7 ± 3.7
1→**7**→**4R**→**1**	0.8 ± 2.0	5.6 ± 2.9	4.8 ± 3.5
1→**7**→**8S**→**1**	0.0 ± 1.7	−3.6 ± 3.1	−3.6 ± 3.5
1→**9S**→**2R**→**1**	−0.1 ± 2.0	−66.3 ± 3.7	−66.2 ± 4.2
1→**10S**→**6S**→**1**	−1.0 ± 2.1	4.7 ± 3.1	5.7 ± 3.8
5→**6S**→**2S**→**5**	0.8 ± 1.8	15.8 ± 3.0	15.0 ± 3.5
5→**9S**→**2R**→**5**	−0.9 ± 2.1	121.3 ± 3.2	122.2 ± 3.9

[a] The residual FE averaged for the forward and backward mutations for each leg in water and in PNMT and for the difference between water and PNMT is shown. [b] The errors were calculated by formula $[(s_1)^2 + (s_2)^2 + (s_3)^2]^{1/2}$ where s_1, s_2, and s_3 are the standard errors for the three different legs for mutations in water and PNMT. [c] The errors for $\Delta\Delta G$ were calculated by $[(s_1)^2 + (s_2)^2]^{1/2}$, where s_1 and s_2 are the standard errors for cycles in water and PNMT, respectively.

water is illustrated in Figure 3, which show plots of <∂H/∂λ> versus λ for selected mutations. Figure 3a shows the forward and backward mutations for the transformation of molecule **1** to **3R** and molecule **1** to **3S**. As shown in Figure 3a, there is an almost perfect overlap in the value of <∂H/∂λ> for all λ values for the forward and backward mutations for both **1** to **3R** and **1** to **3S** as expected. Comparable results were obtained in all other cases with the difference in FE for the forward and backward mutations (hysteresis) in water being ≤1.1 kJ/mol, demonstrating that the calculations in water were well converged.

From Table 4 it can be seen that taking the average between the forward and backward mutations, all cycles not involving **2R** and **2S** in PNMT converged to within 7.3 kJ/mol, with the average residual being 4.3 kJ/mol. While clearly the results in PNMT are not as well converged as in water, the intrinsic error in most cases is still low. The convergence of the forward and backward transformations is illustrated in Figure 3b, which shows a plot of <∂H/∂λ> versus λ for the forward and backward mutations for the transformation of molecule **1** to **3R** and molecule **1** to **3S** in PNMT. Again there is an almost perfect overlap in the value of <∂H/∂λ> for all λ values for the forward and backward mutations in both cases. Note, the change in FE for the mutations **1**−**3R** and **1**−**3S** in water (Figure 3a) are essentially identical as required. From Figure 3b it can be seen there is a significant difference between the mutations **1**−**3R** and **1**−**3S** when bound to PNMT, reflecting enantiomeric selective

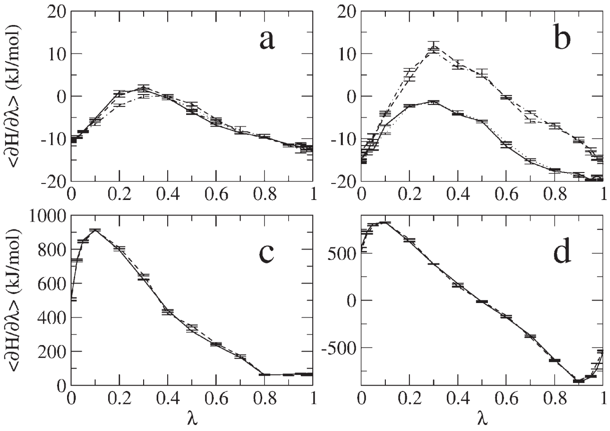


Figure 3. FE profiles for the mutation of specific inhibitors in water and in PNMT. Each graph shows the value of the integrand $<\partial H/\partial\lambda>_\lambda$ at each λ-value. The error bars correspond to the standard error of $\partial H/\partial\lambda$ at each λ-value. (a) Mutation of **1**→**3R** (solid line), **3R**→**1** (dashed line), **1**→**3S** (dots), and **3S**→**1** (dots and dashed line) in water; (b) mutation of **1**→**3R** (solid line) and **1**→**3S** (dashed line) in PNMT; (c) mutation of **8R**→**3S** (solid line) and **3S**→**8R** (dashed line) in PNMT; (d) mutation of the inhibitor **13S** from the binding mode proposed in pdb 2AN5 to the binding mode of the substrate **13R** in pdb 2AN3. All values are in kJ/mol.

binding. The failure of the cycle **1**−**3S**−**8R**−**1** to close within 7.3 ± 4.4 kJ/mol is primarily due to the high hysteresis between the forward and backward calculations for the mutation **3S**−**8R** (Table 3). The mutation **3S**−**8R** has a hysteresis of 4.1 kJ/mol. Figure 3c shows a plot of $<\partial H/\partial\lambda>$ versus λ for the forward and backward mutations for **3S**−**8R**. Even in this case there is still almost perfect overlap between the forward and backward mutations for each λ value, with the intrinsic error being less than 0.5% of the value of $<\partial H/\partial\lambda>_\lambda$ at $\lambda = 0.1$.

The three-membered thermodynamic cycles in PNMT involving the molecules **2R** and **2S** have an error ranging from ~11 kJ/mol to as high as ~186 kJ/mol based on the average of the forward and backward mutations for each leg. Clearly the calculations involving molecules **2R** and **2S** have not converged. This is also reflected in the high hysteresis between the forward and backward mutations involving **2R** and **2S** (Table 3). Experimental binding data are only available for a racemic mixture of molecule **2** against the bovine PNMT. This suggests that both **2R** and **2S** bind only weakly. In the MD simulations of the human PNMT−**2R** and PNMT−**2S** complexes, the binding modes of **2R** and **2S** are unstable, suggesting that they have very low affinity for human PNMT and explaining why the calculations are poorly converged.

Comparison to Experiment. In order to compare the calculated differences in the FE of binding between the different THIQ derivatives to the available experimental data, a series of thermodynamic cycles were constructed. Note in this system, the validation of the results from the simulations by comparison to experiment is complicated by several factors. Table 1 shows there are large differences in the experimental estimates reported by different authors with a range of values for the binding affinity of specific inhibitors having been published even by the same group.[22,23] For example, the variation in the experimental estimate of the binding FE is in the order of ~2 kJ/mol for molecule **1**, ~4 kJ/mol for molecule **7**, ~9 kJ/mol for molecule **8R**, and ~11 kJ/mol for molecule **11**. This is despite the fact that the error in any of the individual values was claimed to be less than 1.0 kJ/mol. Also in some cases experimental binding data are only available for the bovine PNMT and usually for one isomer in the case of human PNMT.

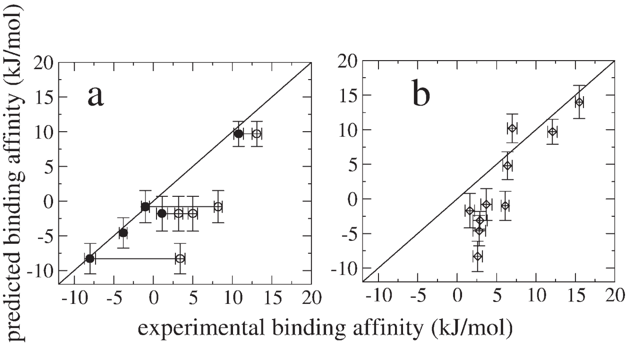


Figure 4. Plot of the experimental versus the calculated binding affinities in kJ/mol (relative to inhibitor **5**). The straight diagonal line has a slope of 1.0 and corresponds to a perfect correlation between the calculated and the experimental values. The vertical lines show the error in the calculated FE values. (a) Comparison with human PNMT, the horizontal lines connect different experimental values for one compound, the filled circles show the recent experimental binding data, the open circle shows the earlier experimental binding data; and (b) comparison with bovine PNMT.

Table 3 shows a direct comparison between the mutations performed and the experimental relative binding FE. Specifically, column 12 of Table 3 lists the difference between the FE of binding determined experimentally and the FE of binding estimated from the FE calculations for individual mutations for which a one-to-one comparison in human PNMT can be made. As can be seen, there is an almost exact correlation between the calculated and the experimental free energies for the inhibitors binding to human PNMT based on the most recent human PNMT data highlighted in bold italics in Table 1. The average absolute deviation is 1.0 kJ/mol with the maximum deviation of 1.8 kJ/mol in the case of molecule **7**.

In order to compare the calculated relative binding free energies to all the available experimental data for both human and bovine PNMT, the relative binding free energies with respect to inhibitor **5** were calculated. Inhibitor **5**, which is achiral and conformationally rigid, was selected as a reference as only one experimental K_i value for human PNMT has been published. The last column of Table 1 shows the difference between the calculated and the experimental values for the FE of binding relative to inhibitor **5** ($\Delta\Delta G_{\mathrm{expt}} - \Delta\Delta G_{\mathrm{calcd}}$). Again it can be seen there is an almost one-to-one correspondence between the calculated and the most recent experimental estimates of the inhibitors binding to human PNMT. Figure 4a shows a plot of the calculated and the various experimental values for the relative binding FE of the available THIQ derivatives to human PNMT. In Figure 4a, the filled circles indicate the most recent experimental estimates, whereas the open circles connected by the horizontal lines correspond to earlier experimental estimates. Note, even in the case of molecule **7**, there is a progressive convergence of the experimental estimate of binding affinity toward the calculated value with time. The binding affinity of molecule **7** for human PNMT was estimated to be 0.58 ± 0.04 μM (−35.9 ± 0.2 kJ/mol) by Grunewald et al.[22] in 2001, 0.28 ± 0.02 μM (−37.7 ± 0.1 kJ/mol) by Wu et al.[23,27] in 2004 (quoting earlier values of Pendleton et al.),[27] and 0.12 ± 0.02 μM (−39.8

± 0.3 kJ/mol) by Wu et al.[21,26] in 2005. A partial explanation for this variation is that earlier estimates were obtained using low concentrations of the cofactor and PNMT and also the possible contamination of the cofactor.[23] This not only highlights the predictive power of the FE calculations but also the pressing need for reliable experimental data against which computational models can be validated. In this case any comparison to data published earlier this decade would have led to the incorrect assumption that the FE calculations were not predictive. In reality the uncertainty in the calculations in this case is much less than the variation in the experimental values over time.

Another point worth noting is that much of the analysis of the binding of THIQ inhibitors to PNMT is based on an analysis of binding data obtained using the bovine enzyme. In fact for many of the compounds listed in Table 1, experimental data are only publically available for the bovine form. Despite the very high sequence identity between the human and bovine forms of the enzyme of 84%, it is known that, while certain compounds may have similar binding affinities between the human and the bovine forms of the enzyme in other cases such as molecules **6R**, **8R** and **11**, the values differ markedly.[19,22,24,25] Figure 4b shows a plot of the calculated relative binding free energies in human PNMT versus the values for bovine PMNT reported in the literature. Based on the calculations we would predict that molecules **1**, **2R**, **2S**, **3R**, **3S**, **7**, **9R**, and **12R** will have similar binding affinities (<4 kJ/mol) to both the human and the bovine forms of the enzyme, whereas molecules **4R**, **4S**, **6R**, **8R**, **10R**, **10S**, and **11** would have a significantly higher affinity (>6 kJ/mol) for the human as compared to the bovine form. Note, no X-ray crystallographic structure of bovine PNMT is currently available.

Enantiomeric Selectivity. Experimental binding data related to the enantiomeric selectivity of the THIQ derivatives considered in this study toward PNMT are only available for the bovine form of the enzyme. The calculations nevertheless correctly predict the relative binding FE of the enantiomers in all cases for which experimental data is available. The FE calculations predict that PNMT preferentially binds **3S** over **3R** by ~9 kJ/mol. The calculations also correctly predict the preferential binding of **6R** over **6S** by ~8 kJ/mol and preferential binding of **9S** over **9R** by ~12 kJ/mol. For the other molecules in the test set there is no data for both enantiomers, and for three molecules (**2**, **4**, and **10**), experimental binding data for bovine PNMT are only available for racemic mixtures. The calculations predict that human PNMT would preferentially bind molecule **4S** over **4R** by ~4 kJ/mol and **8R** over **8S** by ~2 kJ/mol, and thus the enzyme shows only weak enantiomeric selectivity to these compounds. However, the calculation would predict that human PNMT would preferentially bind molecule **10R** over **10S** and molecule **12R** over **12S** by more than 10 kJ/mol in each case.

Structure−Activity Relationships (SAR). In the current work, the FE of binding for 20 analogues of THIQ interacting with human PNMT has been derived with enantiomeric selective data being derived for eight of these compounds. This provides an opportunity to analyze possible structure−activity relationships. As noted previously, electron-withdrawing polar substituents ($-NO_2$, $-SO_2NH_2$, and $-Cl$) at the seven position of the THIQ ring enhance binding affinity toward human PNMT at least by ~10 kJ/mol (compared to THIQ, molecule **1**). An additional $-Cl$ substituent at position eight of THIQ (**11**) improves the binding affinity by a further 10 kJ/mol. Based on this we would predict that the addition of a small electron-withdrawing group at position eight may be an effective means to increase the affinity of THIQ derivatives toward human PNMT. It is also clear that the stereochemistry at position three plays an important role in determining the binding affinity. From Table 1 it can be seen that for nonpolar substituents, such as $-CH_3$ (**3**, **9**) and $-C_2H_5$ (**4**), the '*S*' enantiomer binds preferentially, while in the case of the polar substituents, such as $-CH_2OH$ (**6**, **8**), $-COOCH_3$ (**10**), and $-CH_2F$ (**12**), the '*R*' enantiomer is preferred. It is important to note, however, that in terms of absolute stereochemistry the preferred compounds are those in which substituent at the three position lies equatorial with respect to the piperidine ring projecting toward Tyr222 with the nitrogen of the ring forming a salt bridge with the side chain of Glu219. In the other enantiomer, if the salt bridge to Glu219 is maintained, then the substituent at position three would lie axial to the ring and project toward the backbone of Phe182, which could explain the lower affinity. However, the overall affinity depends on competing interactions that cannot easily be reduced to a simple structure−activity relationship, such as the size or the hydrophobicity of the substituent at position three. This is illustrated by the fact that while compounds **3**, **9**, and **12**, which have only a small substituent at position three ($-CH_3$ or $-CH_2F$), and molecule **10**, which has the largest substituent at position three ($-COOCH_3$), show good enantiomeric selectivity, and intermediate size substituents ($-C_2H_5$ and $-CH_2OH$) are predicted to show only weak selectivity.

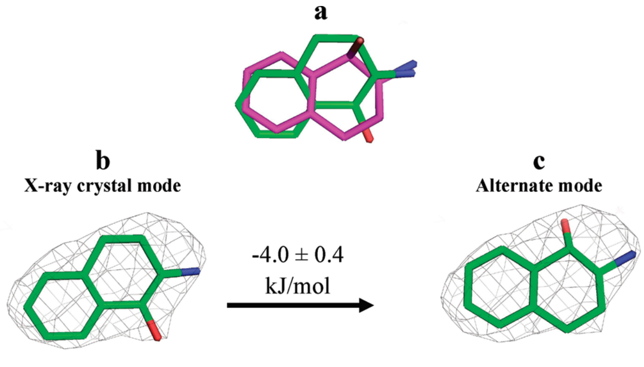


Figure 5. (a) The binding mode of substrate **13R** (magenta) inhibitor **13S** (green) as proposed in the X-ray crystal structure pdb code 2AN3 and 2AN5, respectively. 2Fo−Fc map contoured at 1.0σ for inhibitor **13S** in (b) pdb code 2AN5 and (c) alternate binding mode, which is similar to substrate **13R**.

Binding of Chiral 2-Amino-1-tetralol to PNMT. Stereochemistry also plays an important role in the recognition of other PNMT ligands.[38] For example, compound **13R** and **13S**, which are epimers (diastereoisomers with the opposite stereochemistry at one of the chiral centers), behave very differently when bound to PMNT. The compound **13R** is a substrate for PNMT, whereas **13S** acts as an inhibitor. The X-ray crystallographic models of Gee et al. suggested that the two compounds have different modes of binding (Figure 5a, pdb code 2AN3 and 2AN5) and used this to explain their different biochemical behavior.[38] Specifically it was proposed that the inhibitor **13S** binds in an orientation that is rotated by 180° to the long axis of the fused ring when compared to the substrate **13R** (Figure 5a).

In order to investigate the validity of this proposal, MD simulations of **13R** and **13S** in both the proposed binding modes were performed beginning from pdb structures 2AN3 and 2AN5. In case of **13R** the proposed X-ray binding mode was stable, whereas the alternate binding mode resulted in the disruption of

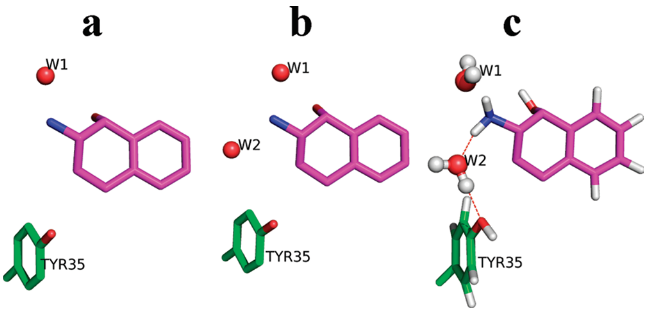


Figure 6. The crystal structure of substrate **13R** complexed with PNMT: (a) molecule A of the asymmetric unit with water, W1 and (b) molecule B of the asymmetric unit with water W1 and W2. (c) The $-NH_3^+$ of substrate **13R** shows water-mediated (W2) hydrogen-bond interaction (red, dotted line) with Tyr35 observed in MD simulation.

active site. In case of **13S**, while the proposed X-ray crystal mode was stable, the alternate binding mode was equally stable (Figure 5c). In fact, as shown in Figure 5b and c, both orientations fit equally well within the experimental electron density. In such cases one cannot easily distinguish the preferred binding mode based on the density, geometry, or energetic criteria or on the global indicators of quality, such as R and R_{free}.[5] In such cases one must instead turn to FE approaches to determine which of the two modes is the more thermodynamically stable. The calculations suggest that the preferred binding mode of the inhibitor **13S** is the same as that of the substrate **13R** with the alternative binding mode proposed in the X-ray structure being higher in FE by ~4 kJ/mol. Again as can be seen from Figure 3d, there was an almost perfect overlap of forward and backward transformations between X-ray crystal and alternate modes of **13S** in PNMT, indicating the calculations were very well converged. Again, the difference in FE, which corresponds to a factor of 5 in the binding affinity, is much greater than the uncertainty in the calculations.

The Role of Water in Ligand Binding. The crystallographic model of the **13R**–PNMT complex (pdb code 2AN3, resolution 2.20 Å) has two molecules (A and B) in the asymmetric unit. Molecule A contains a single structural water molecule (W1, Figure 6a), while molecule B contains two structural water molecules (W1 and W2, Figure 6b) in the binding pocket. To determine whether these water molecules were critical to maintain the stability of the active site, separate simulations of both A and B were performed. In the case of molecule A, an additional water molecule entered the pocket forming hydrogen bonds with the $-NH_3^+$ group of the ligand and the $-OH$ group of Tyr35, as shown in Figure 6c, and remained stable in this position throughout the simulation. In the case of molecule B, the second structural water molecule, W2, occupied at the same position (Figure 6c) as described earlier. This suggests that the water molecule W2 is required to maintain the interaction between PNMT and **13R** as observed in the crystal structure.

Experimentally the mutation of Tyr35Phe shows a substantial decrease in the binding affinity for **13R** and the cofactor *S*-adenosyl-L-methionine (SAM).[38] The reduction in the binding affinity for the cofactor can be understood in terms of the interaction of the Tyr35 hydroxyl group with the carboxylate oxygen of the amino acid fragment of the cofactor analogue SAH used in crystallization. However, it was not clear why the binding affinity of the substrate **13R** was also affected by Tyr35Phe mutation.[38] In the simulations, W2 forms a water-mediated interaction between **13R** and the hydroxyl group of Tyr35, and it is likely that this accounts for the reduction in the binding affinity of the substrate when Tyr35 is mutated to Phe.

■ CONCLUSIONS

In this work the stereospecific binding affinity for a series of 20 analogues of THIQ and the stereospecific binding mode of 2-amino-1-tetralol to human PNMT have been investigated. Specifically, molecular dynamics simulations and free energy calculations have been used to understand in detail the structural and thermodynamic basis of ligand recognition in this system. This is an important case study as the binding affinities proposed by different sets of workers in different studies differ significantly and revised values based on new assay conditions have been recently published. For those THIQ analogues for which recent experimental data are available, excellent agreement between calculated and measured relative binding free energies to human PNMT was obtained. The average deviation between the calculated and the experimentally determined values for these compounds using molecule **5** as a reference was only 0.8 kJ/mol, showing that the calculations can easily distinguish between the data sets. This highlights the fundamental challenge when attempting to compare theoretical calculations to measured binding affinities and the critical need for reliable and validated experimental data.[39] Clearly, the variation in the published experimental data (despite the small errors claimed for each of the individual experimental observations) is much greater than the intrinsic uncertainty in the theoretical estimates.

The calculations have also enabled a detailed analysis of the structure–activity relationships of these THIQ analogues. In particular, the addition of a small electron-withdrawing group at position eight is predicted to be an effective means to increase the affinity of THIQ derivatives toward human PNMT. It is also evident that the relative orientation of the substituents rather than absolute stereochemistry at position three of THIQ appears to govern enantiomeric selectivity. The size of the group at position three of THIQ also plays a nontrivial role in enantiomeric selectivity with small substituents, such as a methyl or fluoromethyl group, showing greater enantiomeric selectivity than slightly larger groups, such as ethyl and hydroxmethyl. In case of **13R**, the importance of the role specific structural waters can play in ligand recognition has been illustrated. For example, interactions between the substrate **13R** and Tyr35 involving a specific structural water (W2) can explain the effect of mutations at position 35 on enzymatic activity. Finally, the thermodynamically stable binding mode in case of the inhibitor **13S** is predicted to be similar to that of the substrate **13R** as opposed to the novel binding mode proposed in the pdb entry 2AN5. Overall the work highlights the power of MD simulations and the free energy calculations to resolve uncertainties in experimental binding affinities, binding modes, and other aspects related to X-ray refinement and computational drug design.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: a.e.mark@uq.edu.au.

■ ACKNOWLEDGMENT

P.C.N. acknowledges the Australian government and the University of Queensland for the award of an Endeavour

International Postgraduate Research Scholarship (IPRS), A.K.M. acknowledges Australian Research Council (ARC; grant no. DP 0987043) for the award of an Australian Post-Doctoral (APD) fellowship. Computational resources were provided by the National Computational Infrastructure (NCI, Australia) projects m72 and n63. The authors would like to thank the reviewers for their comments and suggestions to improve the quality of the paper.

REFERENCES

(1) Verlinde, C. L. M. J.; Hol, W. G. J. Structure-based drug design: progress, results and challenges. *Structure* **1994**, *2*, 577–587.

(2) Karplus, P. A.; Faerman, C. Ordered water in macromolecular structure. *Curr. Opin. Struct. Biol.* **1994**, *4*, 770–776.

(3) Kollman, P. Free energy calculations: Applications to chemical and biochemical phenomena. *Chem. Rev.* **1993**, *93*, 2395–2417.

(4) Malde, A. K.; Mark, A. E. Binding and enantiomeric selectivity of threonyl-tRNA synthetase. *J. Am. Chem. Soc.* **2009**, *131*, 3848–3849.

(5) Malde, A. K.; Mark, A. E. Challenges in the determination of the binding modes of non-standard ligands in X-ray crystal complexes. *J. Comput.-Aided. Mol. Des.* **2011**, *25*, 1–12.

(6) Ziegler, M. G.; Kennedy, B.; Elayan, H. Extraadrenal adrenaline formation by two separate enzymes. *Cell. Mol. Life Sci.* **1989**, *45*, 718–720.

(7) Hökfelt, T.; Fuxe, K.; Goldstein, M.; Johansson, O. Immunohistochemical evidence for the existence of adrenaline neurons in the rat brain. *Brain Res.* **1974**, *66*, 235–251.

(8) Crowley, W. R.; Terry, L. C.; Johnson, M. D. Evidence for the involvement of central epinephrine systems in the regulation of luteinizing hormone, prolactin, and growth hormone release in female rats. *Endocrinology* **1982**, *110*, 1102–1107.

(9) Fuller, R. W.; Roush, B. W. Substrates and inhibitors of phenyletholamine n-methyl transferase from human adrenal glands. *Int. J. Biochem.* **1972**, *3*, 225–228.

(10) Fuller, R. W.; Roush, B. W.; Snoddy, H. D.; Day, W. A.; Molloy, B. B. Norepinephrine N-methyltransferase inhibition by benzamidines, phenylacetamidines, benzylguanidines, and phenylethylguanidines. *J. Med. Chem.* **1975**, *18*, 304–307.

(11) Grunewald, G. L.; Caldwell, T. M.; Li, Q.; Criscione, K. R. Synthesis and evaluation of 3-trifluoromethyl-7-substituted- 1,2,3,4-tetrahydroisoquinolines as selective inhibitors of phenylethanolamine N-methyltransferase versus the α2-Adrenoceptor. *J. Med. Chem.* **1999**, *42*, 3315–3323.

(12) Grunewald, G. L.; Dahanukar, V. H.; Ching, P.; Criscione, K. R. Effect of ring size or an additional heteroatom on the potency and selectivity of bicyclic benzylamine-type inhibitors of phenylethanolamine N-methyltransferase. *J. Med. Chem.* **1996**, *39*, 3539–3546.

(13) Toomey, R. E.; Horng, J. S.; Hemrick-Luecke, S. K.; Fuller, R. W. α2-Adrenoceptor affinity of some inhibitors of norepinephrine N-methyltransferase. *Life Sci.* **1981**, *29*, 2467–2472.

(14) Grunewald, G. L.; Romero, F. A.; Criscione, K. R. Nanomolar inhibitors of CNS epinephrine biosynthesis: (R)-(+)-3-fluoromethyl-7-(N-substituted aminosulfonyl)-1,2,3,4-tetrahydroisoquinolines as potent and highly selective inhibitors of phenylethanolamine N-methyltransferase. *J. Med. Chem.* **2004**, *48*, 1806–1812.

(15) Grunewald, G. L.; Romero, F. A.; Criscione, K. R. 3-Hydroxymethyl-7-(N-substituted aminosulfonyl)-1,2,3,4-tetrahydroisoquinoline Inhibitors of phenylethanolamine N-methyltransferase that display remarkable potency and selectivity. *J. Med. Chem.* **2004**, *48*, 134–140.

(16) Romero, F. A.; Vodonick, S. M.; Criscione, K. R.; McLeish, M. J.; Grunewald, G. L. Inhibitors of phenylethanolamine N-methyltransferase that are predicted to penetrate the blood-brain barrier: design, synthesis, and evaluation of 3-fluoromethyl-7-(N-substituted aminosulfonyl)-1,2,3,4-tetrahydroisoquinolines that possess low affinity toward the α2-adrenoceptor. *J. Med. Chem.* **2004**, *47*, 4483–4493.

(17) Grunewald, G. L.; Caldwell, T. M.; Li, Q.; Criscione, K. R. 1,3-Dimethyl-7-substituted-1,2,3,4-tetrahydroisoquinolines as probes for the binding orientation of tetrahydroisoquinoline at the active site of phenylethanolamine N-methyltransferase. *Bioorg. Med. Chem.* **1999**, *7*, 869–880.

(18) Grunewald, G. L.; Dahanukar, V. H.; Criscione, K. R. Effects of a 3-alkyl-, 4-hydroxy- and/or 8-aromatic-substituent on the phenylethanolamine N-methyltransferase inhibitor potency and [alpha]2-adrenoceptor affinity of 2,3,4,5-tetrahydro-1H-2-benzazepines. *Bioorg. Med. Chem.* **2001**, *9*, 1957–1965.

(19) Grunewald, G. L.; Dahanukar, V. H.; Teoh, B.; Criscione, K. R. 3,7-Disubstituted-1,2,3,4-tetrahydroisoquinolines display remarkable potency and selectivity as inhibitors of phenylethanolamine N-methyltransferase versus the α2-Adrenoceptor. *J. Med. Chem.* **1999**, *42*, 1982–1990.

(20) Begun, J.; McLeish, M. J.; Caine, J. M.; Palant, E.; Grunewald, G. L.; Martin, J. L. Crystallization of PNMT, the adrenaline-synthesizing enzyme, is critically dependent on a high protein concentration. *Acta Cryst. D* **2002**, *58*, 314–315.

(21) Gee, C. L.; Drinkwater, N.; Tyndall, J. D. A.; Grunewald, G. L.; Wu, Q.; McLeish, M. J.; Martin, J. L. Enzyme adaptation to inhibitor binding: A cryptic binding site in phenylethanolamine N-methyltransferase. *J. Med. Chem.* **2007**, *50*, 4845–4853.

(22) Grunewald, G. L.; McLeish, M. J.; Criscione, K. R. Phenylethanolamine N-methyltransferase kinetics: bovine versus recombinant human enzyme. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1579–1582.

(23) Wu, Q.; Criscione, K. R.; Grunewald, G. L.; McLeish, M. J. Phenylethanolamine N-methyltransferase inhibition: re-evaluation of kinetic data. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 4217–4220.

(24) Grunewald, G. L.; Seim, M. R.; Regier, R. C.; Martin, J. L.; Gee, C. L.; Drinkwater, N.; Criscione, K. R. Comparison of the binding of 3-fluoromethyl-7-sulfonyl-1,2,3,4-tetrahydroisoquinolines with their isosteric sulfonamides to the active site of phenylethanolamine N-methyltransferase. *J. Med. Chem.* **2006**, *49*, 5424–5433.

(25) Grunewald, G. L.; Sall, D. J.; Monn, J. A. Synthesis and evaluation of 3-substituted analogs of 1,2,3,4-tetrahydroisoquinoline as inhibitors of phenylethanolamine N-methyltransferase. *J. Med. Chem.* **1988**, *31*, 824–830.

(26) Wu, Q.; Gee, C. L.; Lin, F.; Tyndall, J. D.; Martin, J. L.; Grunewald, G. L.; McLeish, M. J. Structural, mutagenic, and kinetic analysis of the binding of substrates and inhibitors of human phenylethanolamine N-methyltransferase. *J. Med. Chem.* **2005**, *48*, 7243–7252.

(27) Pendleton, R. G.; Gessner, G.; Weiner, G.; Jenkins, B.; Sawyer, J.; Bondinell, W.; Intoccia, A. Studies on SK&F 29661, an organ-specific inhibitor of phenylethanolamine N-methyltransferase. *J. Pharmacol. Exp. Ther.* **1979**, *208*, 24–30.

(28) van Gunsteren, W. F.; Billeter., S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular Simulations: The GROMOS96 Manual and User Guide*; Swiss Federal Institute of Technology Zurich: Zurich, Switzerland, 1996.

(29) Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* **2004**, *25*, 1656–1676.

(30) Malde, A. K.; Zuo, L.; Breeze, M.; Stroet, M.; Poger, D.; Nair, P. C.; Oostenbrink, C.; Mark, A. E. An Automated force field Topology Builder (ATB) and repository: version 1.0. *J. Chem. Theory Comput.*, manuscript submitted.

(31) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. Interaction models for water in relation to protein hydration. In *Intermolecular Forces*, Pullman, B., Ed. Reidel: Dordrecht, 1981; pp 331−342.

(32) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

(33) Ryckaert, J.; Ciccotti, G.; Berendsen, H. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341.

(34) Martin, J. L.; Begun, J.; McLeish, M. J.; Caine, J. M.; Grunewald, G. L. Getting the adrenaline going: Crystal structure of the adrenaline-synthesizing enzyme PNMT. *Structure* **2001**, *9*, 977–985.

(35) Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. *Chem. Phys. Lett.* **1994**, *222*, 529–539.

(36) Zacharias, M.; Straatsma, T. P.; McCammon, J. A. Separation-shifted scaling, a new scaling method for Lennard-Jones interactions in thermodynamic integration. *J. Chem. Phys.* **1994**, *100*, 9025–9031.

(37) Allen, M. P.; Tildesley, D. J. *Computer Simulations of Liquids*. Clarendon Press: Oxford, U.K., 1987.

(38) Gee, C. L.; Tyndall, J. D. A.; Grunewald, G. L.; Wu, Q.; McLeish, M. J.; Martin, J. L. Mode of binding of methyl acceptor substrates to the adrenaline-synthesizing enzyme phenylethanolamine N-methyltransferase: Implications for catalysis. *Biochemistry* **2005**, *44*, 16875–16885.

(39) van Gunsteren, W. F.; Dolenc, J.; Mark, A. E. Molecular simulation as an aid to experimentalists. *Curr. Opin. Struct. Biol.* **2008**, *18*, 149–153.

JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

On the Calculation of the Dielectric Permittivity and Relaxation of Molecular Models in the Liquid Phase

Sereina Riniker, Anna-Pitschna E. Kunz, and Wilfred F. van Gunsteren*

Laboratory of Physical Chemistry, Swiss Federal Institute of Technology, ETH, 8093 Zürich, Switzerland

ABSTRACT: Methodology to compute the relative static dielectric permittivity and dielectric relaxation time of molecular liquids is reviewed and explicit formulas are given for the external field method in the case of simulations using a spherical cutoff, in which the background dielectric permittivity (ε_{cs}) can be larger than one, in combination with a Poisson−Boltzmann reaction-field approximation for long-range electrostatic interactions. The external field method is simple to implement and computationally efficient. It is particularly suitable for polarizable molecular models with zero permanent dipole moment and for coarse-grained molecular models with $\varepsilon_{cs} > 1$. The dielectric permittivities and relaxation times of water (H_2O), dimethylsulfoxide (DMSO), methanol (MeOH), and chloroform ($CHCl_3$), which range from 2 to 80 and from 5 ps to 50 ps, respectively, were calculated as an illustration.

1. INTRODUCTION

The interactions between molecules in the liquid phase, such as biomolecules in aqueous solution, are basically (i.e., at the quantum-chemical level) governed by electrostatic interactions. These manifest themselves in the form of Coulombic interactions between parts of molecules that have a nonzero net charge density, or of polarization interactions between electronically polarizable parts of molecules, or in the form of van der Waals interactions originating from mutual interactions between dipolar fluctuations of atoms. This implies that the dielectric properties of a molecular model for a given compound are of central importance when using such a model in a biomolecular simulation. For this reason, it is standard practice to compute and report the dielectric permittivity of molecular models for compounds that are used in such simulations. This property can be obtained from a simulation of the compound in the liquid phase.

Different methods are available to compute the dielectric permittivity (ε) of a molecular liquid via computer simulation, e.g., molecular dynamics (MD) simulation:[1]

(1) In the limit of low field strength, the fluctuations of the electric dipole moment $\vec{M}$ of the simulated system in an equilibrium simulation can be related to the polarization $\vec{P}$ induced in the system if an external electric field $\vec{E}^{ext}$ would be applied to it,[2] and thus to the dielectric permittivity ε. Technically, this relationship between the variance of the distribution of $\vec{M}$ and ε can be employed in different ways, using simulations in which $\vec{M}$ is unrestrained, or restrained by a biasing potential energy term, or constrained to a particular range of $\vec{M}$ values.[3]

(2) The dipole moment $\vec{M}$ that is induced by the application of an external field $\vec{E}^{ext}$ to the system can be analyzed to yield ε. The external field can be constant and homogeneous[4] or sinusoidal in space[5] in order to probe a range of $\vec{E}^{ext}$ values in one simulation. A variant of the external field method would be the coupling of the polarization of the system $\vec{P}^{syst}$ to an external polarization bath $\vec{P}^{bath}$, using a first-order weak coupling equation.[6]

When considering polarizable models for molecules that have zero permanent dipole moment, the fluctuating dipole moment methodology can obviously not be used, because no such fluctuations will occur in a simulation. Thus, an external field technique is the method of choice in such cases. Although this method seems simple to implement—just apply a homogeneous external field $\vec{E}^{ext}$ of a chosen value and calculate the polarization $\vec{P}$ induced in the system—the precise expressions relating $\vec{E}^{ext}$ to $\vec{P}$ and ε are dependent on the way the electrostatic interactions in the system are calculated, e.g., using a cutoff sphere, a minimum-image cutoff cube, a rectangular or oblique infinite lattice sum, or a spherical cutoff with continuum reaction field beyond the cutoff sphere. Since the first two schemes to calculate the electrostatic interactions induce sizable distortions of the configurational distribution of the system, only the latter two schemes are currently used in molecular simulations.

Each technique has advantages and disadvantages. Using a lattice-sum technique to compute long-range electrostatic interactions, one simulates an infinitely extended system in atomic detail, but at the cost of artificially enhancing periodic order in the system, which distorts the forces inside the system.[7−12] Using a spherical cutoff with a continuum reaction-field approximation representing the long-range electrostatic interactions, a distortive periodicity is avoided at the expense of a loss of atomic details beyond the cutoff sphere, which are modeled as an isotropic, homogeneous mean-field dielectric response.[13−15] In our view, this approximation is in regard to noncrystalline condensed-phase biomolecular systems less distortive than the imposition of artificial small-scale periodicity.

In the reaction-field approximation of the long-range electrostatic interactions, it is usually assumed that the dielectric continuum outside the cutoff sphere will react instantaneously to changes of the partial charges inside the cutoff sphere. In general, however, there will be a delay in the response caused by

Received: October 26, 2010
Published: April 15, 2011

classical Stokes and dielectric frictional effects in the surroundings, which can be modeled using a generalized Langevin equation for the reaction field.[13] We shall not consider this approximation of the long-range electrostatic forces, because the method is computationally expensive and an instantaneous treatment of the reaction-field turned out to be sufficiently accurate.

The original formulation of the dipolar reaction-field force induced by a dipole moment $\vec{M}$ inside a cutoff sphere through the surrounding dielectric continuum of permittivity ε_{rf}[14] was generalized[15] to the case of a dielectric continuum characterized by ε_{rf} and an ionic strength I represented by an inverse Debye screening length κ, leading to a so-called Poisson–Boltzmann reaction-field force, which reduces to the original dipolar one for $\kappa = 0$.

Here, we generalize the external field method to compute the static relative dielectric permittivity of a molecular model[4] to the case in which the background permittivity of the cutoff sphere (ε_{cs}) may be larger than one and the dielectric medium outside the cutoff sphere can have a nonzero ionic strength. We give expressions for the calculation of the dielectric permittivity ε from MD simulations in which a spherical truncation with $\varepsilon_{cs} \geq 1$ and a Poisson–Boltzmann continuum reaction field with $\kappa \geq 0$ outside the cutoff sphere is applied in combination with a constant homogeneous external field $\vec{E}^{ext}$. The external field method can also be used to obtain the Debye dielectric relaxation time τ_D from nonequilibrium MD simulations in which a homogeneous electric field is switched on. By performing simulations at different field strengths and measuring the polarization response, precise values of ε and τ_D can be obtained. The expressions for the static (i.e., zero frequency), relative dielectric permittivity $\varepsilon(0)$ and for the Debye relaxation time of the model, as a function of the magnitude of $\vec{E}^{ext}$, are tested via application to models of different molecular liquids (chloroform ($CHCl_3$), methanol (MeOH), dimethylsulfoxide (DMSO), and water (H_2O)) that cover a wide range of $\varepsilon(0)$ values (i.e., 2–80) and of τ_D values (i.e., 5–50 ps), and comparison to the $\varepsilon(0)$ values and τ_D values obtained for these models using the dipole moment fluctuation methodology, as reported in the literature.[3,16–20]

2. METHODS

2.1. Application of an External Electric Field. We consider a system of N particles, atoms, or beads in case a coarse-grained molecular model is used, in a computational box with or without periodic boundary conditions. The interaction between the particles consists of two components: the so-called "bonded interactions" between atoms covalently bound to each other in a molecule and "nonbonded interactions". In a typical biomolecular force field, the bonded interaction terms represent the local interactions between atoms that are separated by one, two, three, or possibly four bonds in a molecule. All other atom pairs interact through nonbonded interactions (generally electrostatic and van der Waals pairwise additive interactions). In addition, the molecular model can be polarizable, i.e., it contains polarizable atoms or sites, leading to a nonpairwise additive electrostatic interaction. When calculating the electrostatic interactions and forces between the particles of a molecule, a given set of particle pairs—the so-called "excluded neighbors" along the covalently bound chain of atoms—is excluded from the nonbonded interaction calculation, because the interaction of the pair, generally determined via quantum chemical or other methods, would be poorly represented as electrostatic, van der Waals, or polarization interactions. The set of particle pairs that is contributing to the nonbonded interaction is additionally limited when a so-called "cutoff sphere" for nonbonded interactions is applied. Since electrostatic interactions are spatially long-ranged, such a cutoff sphere restriction must be supplemented with a continuum or other approximation of the electrostatic interactions for the particle at the center of the cutoff sphere with those beyond it.

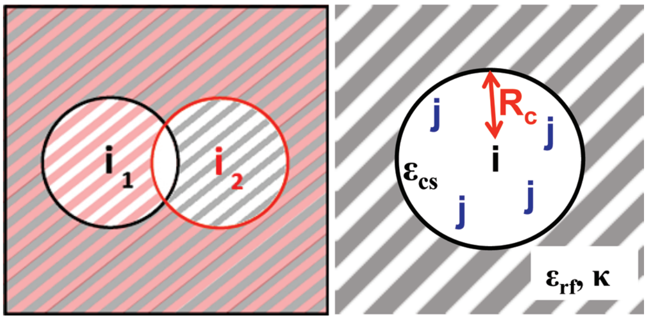


Figure 1. (Left) Schematic picture of the cutoff spheres of two particles i_1 and i_2, and the corresponding overlapping continuum reaction-field regions. (Right) Schematic picture of the interactions between particle i and all particles j in the cutoff sphere with radius $R_c = R_{rf}$.

We consider an electrostatic nonbonded interaction cutoff sphere with radius R_c around a particle i (see Figure 1). The particles j inside the cutoff sphere have charges q_j, and the static relative dielectric permittivity of the medium inside the cutoff sphere is ε_{cs}. For atomic particles, one generally has $\varepsilon_{cs} = 1$ (i.e., a vacuum background). In a coarse-grained molecular simulation, however, one may have $\varepsilon_{cs} > 1$ (i.e., a dielectric medium that represents the dielectric response of the degrees of freedom that are not considered in the coarse-grained model, via a mean-field response). The dielectric continuum that is intended to represent the electrostatic interactions beyond the cutoff sphere is modeled as a dielectric continuum outside a sphere of radius R_{rf} around particle i, characterized by a relative static permittivity ε_{rf} and an ionic strength I or inverse Debye screening length κ:

$$\kappa^2 = \frac{2IF^2}{\varepsilon_0 \varepsilon_{rf} RT} \tag{1}$$

where F is Faraday's constant, ε_0 the electric permittivity of vacuum, R the gas constant, and T the temperature. Generally, one would choose $R_c = R_{rf}$, but slightly different values for these radii might represent the true interactions better.[7]

The direct Coulomb force on particle i by particle j in the cutoff sphere is

$$\vec{f}_{ij}(r_{ij}) = \frac{q_i q_j}{4\pi\varepsilon_0\varepsilon_{cs}}\left(\frac{\vec{r}_{ij}}{r_{ij}^3}\right) \tag{2}$$

with $\vec{r}_{ij} = \vec{r}_i - \vec{r}_j$. The generalized Poisson–Boltzmann reaction-field force on particle i by particle j can be obtained by solving the Poisson equation inside the cutoff sphere and the Poisson–Boltzmann equation outside it, using the boundary condition of continuous radial dielectric displacement at the boundary $r = R_c = R_{rf}$[15] and zero potential at $r = \infty$,

$$\vec{f}_{ij}(r_{ij}) = -\frac{q_i q_j}{4\pi\varepsilon_0\varepsilon_{cs}}\left(\frac{1}{R_{rf}^3}\right)\left[\frac{(2\varepsilon_{rf} - 2\varepsilon_{cs})(1 + \kappa R_{rf}) + \varepsilon_{rf}(\kappa R_{rf})^2}{(2\varepsilon_{rf} + \varepsilon_{cs})(1 + \kappa R_{rf}) + \varepsilon_{rf}(\kappa R_{rf})^2}\right]\vec{r}_{ij} \tag{3}$$

The direct Coulomb force on particle i exerted by all particles j in the cutoff sphere, except those that are nearest neighbors and are therefore excluded, is then

$$\vec{\mathrm{f}}_i = \sum_{\substack{j \neq i \\ (i,j)\ \text{notexcluded}}}^{N_{cs}} \frac{q_i q_j}{4\pi\varepsilon_0\varepsilon_{cs}} \left(\frac{\vec{\mathrm{r}}_{ij}}{r_{ij}^3}\right) \tag{4}$$

and the Poisson—Boltzmann reaction-field force on particle i exerted by all particles j in the cutoff sphere is

$$\vec{\mathrm{f}}_i = \frac{q_i}{4\pi\varepsilon_0\varepsilon_{cs}} \left(\frac{1}{R_{rf}^3}\right) \left[\frac{(2\varepsilon_{rf} - 2\varepsilon_{cs})(1 + \kappa R_{rf}) + \varepsilon_{rf}(\kappa R_{rf})^2}{(2\varepsilon_{rf} + \varepsilon_{cs})(1 + \kappa R_{rf}) + \varepsilon_{rf}(\kappa R_{rf})^2}\right] \sum_{j \neq i}^{N_{cs}} q_j \vec{\mathrm{r}}_{ji} \tag{5}$$

where N_{cs} is the number of particles in the cutoff sphere around particle i. Using the short-hand notation,[21]

$$C_{rf} = \frac{(2\varepsilon_{cs} - 2\varepsilon_{rf})(1 + \kappa R_{rf}) - \varepsilon_{rf}(\kappa R_{rf})^2}{(\varepsilon_{cs} + 2\varepsilon_{rf})(1 + \kappa R_{rf}) + \varepsilon_{rf}(\kappa R_{rf})^2} \tag{6}$$

the reaction-field force reads

$$\vec{\mathrm{f}}_i = -\frac{q_i}{4\pi\varepsilon_0\varepsilon_{cs}} \left(\frac{C_{rf}}{R_{rf}^3}\right) \vec{\mathrm{M}}_i^{cs} = q_i \vec{\mathrm{E}}_i^{par,\, rf} \tag{7}$$

where the dipole moment of the cutoff sphere around particle i is denoted by

$$\vec{\mathrm{M}}_i^{cs} = \sum_{j \neq i}^{N_{cs}} q_j \vec{\mathrm{r}}_{ji} \tag{8}$$

and the corresponding dipolar reaction field in the cutoff sphere induced by the charges of the particles in it is given as

$$\vec{\mathrm{E}}_i^{par,\, rf} = -\frac{1}{4\pi\varepsilon_0\varepsilon_{cs}} \left(\frac{C_{rf}}{R_{rf}^3}\right) \vec{\mathrm{M}}_i^{cs} \tag{9}$$

We note that the summation in eq 8 may be extended to include the term for $j = i$, because $\vec{r}_{ii} = 0$.

The potential energy terms corresponding to the forces described by eqs 2 and 3 are, for $i \neq j$,

$$V_{ij}(r_{ij}) = \frac{q_i q_j}{4\pi\varepsilon_0\varepsilon_{cs}} \left\{\frac{1}{r_{ij}} - \left[\frac{(1/2)C_{rf}}{R_{rf}^3}\right] r_{ij}^2 - \frac{1 - (1/2)C_{rf}}{R_{rf}}\right\} \tag{10}$$

where the constant term ensures that $V_{ij}(R_{rf}) = 0$. The total electrostatic potential energy of the system then, using the notation $\vec{r}^{\,N} = (\vec{r}_1, \vec{r}_2, ..., \vec{r}_N)$, is given as

$$V(\vec{\mathrm{r}}^N) = \sum_{i=1}^{N-1} \sum_{\substack{j>i \\ j\ \text{inside cut-off}\ i \\ (i,j)\ \text{not excluded}}}^{N} \frac{q_i q_j}{4\pi\varepsilon_0\varepsilon_{cs}} \left(\frac{1}{r_{ij}}\right)$$

$$-\sum_{i=1}^{N-1} \sum_{\substack{j>i \\ j\ \text{insidecut-off}\ i}}^{N} \frac{q_i q_j}{4\pi\varepsilon_0\varepsilon_{cs}}$$

$$\left[\frac{(1/2)C_{rf}}{R_{rf}^3} r_{ij}^2 + \frac{1 - (1/2)C_{rf}}{R_{rf}}\right] - \sum_{i=1}^{N} \frac{q_i^2}{4\pi\varepsilon_0\varepsilon_{cs}} \left(\frac{1}{2}\right) \left[\frac{1 - (1/2)C_{rf}}{R_{rf}}\right] \tag{11}$$

where the third summation is a constant that is added to represent the self-interaction of the charged particles.

If a constant homogeneous external field $\vec{E}^{ext}$ is applied to the system, e.g., along the z-axis,

$$\vec{\mathrm{E}}^{ext} = E^{ext} \vec{\mathrm{e}}_z \tag{12}$$

an additional force on particle i is present,

$$\vec{\mathrm{f}}_i^{ext,\, cs} = \frac{q_i}{4\pi\varepsilon_0} \vec{\mathrm{E}}^{ext,\, cs} = \frac{q_i}{4\pi\varepsilon_0} \left(\vec{\mathrm{E}}^{ext} + \vec{\mathrm{E}}^{ext,\, rf}\right) \tag{13}$$

where $\vec{E}^{ext,rf}$ describes the contribution to the electric field $\vec{E}^{ext,cs}$ inside the cutoff sphere due to the polarization $\vec{P}^{ext,rf}$ in the dielectric continuum outside the cutoff sphere that is induced by the difference in dielectric permittivity inside the cutoff sphere (ε_{cs}) and outside the cutoff sphere (ε_{rf}). This contribution is dependent on the shape of the cutoff region.[22] For a spherical region, we have

$$\begin{aligned} \vec{\mathrm{E}}^{ext,\, cs} &= \vec{\mathrm{E}}^{ext} - \left(\frac{\varepsilon_{cs} - \varepsilon_{rf}}{\varepsilon_{cs} + 2\varepsilon_{rf}}\right) \vec{\mathrm{E}}^{ext} \\ &= \vec{\mathrm{E}}^{ext} + \left(\frac{4\pi}{\varepsilon_{cs} + 2\varepsilon_{rf}}\right) \vec{\mathrm{P}}^{ext,\, rf} \\ &= \left(\frac{3\varepsilon_{rf}}{\varepsilon_{cs} + 2\varepsilon_{rf}}\right) \vec{\mathrm{E}}^{ext} \end{aligned} \tag{14}$$

where the polarization induced around the cutoff cavity in the dielectric continuum by the external field is given as

$$\vec{\mathrm{P}}^{ext,\, rf} = \left(\frac{\varepsilon_{rf} - \varepsilon_{cs}}{4\pi}\right) \vec{\mathrm{E}}^{ext} \tag{15}$$

Thus, the electric field in the cutoff sphere has four components:[4] direct and reaction-field components that are due to the particles in the cutoff sphere,

$$\vec{\mathrm{E}}_i^{par,\, cs} = \sum_{\substack{j \neq i \\ \text{exclusions}}}^{N_{cs}} \left(\frac{q_i}{4\pi\varepsilon_0\varepsilon_{cs}}\right) \frac{\vec{\mathrm{r}}_{ij}}{r_{ij}^3} \tag{16}$$

$$\vec{\mathrm{E}}_i^{par,\, rf} = -\frac{1}{4\pi\varepsilon_0\varepsilon_{cs}} \left(\frac{C_{rf}}{R_{rf}^3}\right) \vec{\mathrm{M}}_i^{cs} \tag{17}$$

the direct external field $\vec{E}^{ext}$ and the reaction field induced by it,

$$\vec{\mathrm{E}}^{ext,\, rf} = -\left(\frac{\varepsilon_{cs} - \varepsilon_{rf}}{\varepsilon_{cs} + 2\varepsilon_{rf}}\right) \vec{\mathrm{E}}^{ext} \tag{18}$$

The electric field $\vec{E}^{ext}$ that is applied to the system will induce a polarization $\vec{P}$ in the system:[23,24]

$$\vec{\mathrm{P}} = \frac{1}{4\pi}(\varepsilon(0) - 1) \vec{\mathrm{E}}^{ext} \tag{19}$$

where $\varepsilon(0)$ is the zero-frequency or static dielectric permittivity of the molecular model. The polarization $\vec{P}$ can be calculated from the total dipole moment $\vec{M}$ of the system and its volume V:

$$\vec{\mathrm{P}} = \frac{\vec{\mathrm{M}}}{V} = \frac{1}{V}\sum_{i=1}^{N} q_i \vec{\mathrm{r}}_i \tag{20}$$

if $\Sigma_{i=1}^{N} q_i = \underline{0}$. If the latter condition is not satisfied, $\vec{M}$, and, therefore, $\vec{P}$ will be origin-dependent. For a homogeneous external field (see eq 12) along the z-axis, we then find

$$\varepsilon(0) = 1 + 4\pi\frac{\langle M_z\rangle}{VE_z^{ext}} \tag{21}$$

where $\langle M_z\rangle$ is the average dipole moment of the volume V in the z-direction in the simulation.

The method has been implemented into the GROMOS simulation software[25] using the expressions given. However, we note that the electric field eqs 13−21 could also be formulated differently, using

$$\vec{\mathrm{P}} = \varepsilon_0(\varepsilon(0) - 1)\,\vec{\mathrm{E}}^{ext} \tag{22}$$

instead of eq 19, which would imply

$$\vec{\mathrm{f}}_i^{\,ext,\,cs} = q_i\,\vec{\mathrm{E}}^{ext,\,cs} = q_i(\,\vec{\mathrm{E}}^{ext} + \vec{\mathrm{E}}^{ext,\,rf}) \tag{23}$$

$$\vec{\mathrm{E}}^{ext,\,cs} = \vec{\mathrm{E}}^{ext} + \frac{1}{\varepsilon_0}\left(\frac{1}{\varepsilon_{cs} + 2\varepsilon_{rf}}\right)\vec{\mathrm{P}}^{ext,\,rf} \tag{24}$$

$$\vec{\mathrm{P}}^{ext,\,rf} = \varepsilon_0(\varepsilon_{rf} - \varepsilon_{cs})\,\vec{\mathrm{E}}^{ext} \tag{25}$$

and

$$\varepsilon(0) = 1 + \left(\frac{1}{\varepsilon_0}\right)\frac{\langle M_z\rangle}{VE_z^{ext}} \tag{26}$$

instead of eqs 13−15 and 21, respectively.

2.2. Calculation of the Dielectric Relaxation Time.

The Debye dielectric relaxation time (τ_D) of a molecular liquid can be calculated from an equilibrium MD simulation of the liquid by evaluating the autocorrelation function $\langle\vec{M}(t_0)\cdot\vec{M}(t)\rangle_{t0}$ of the total dipole moment $\vec{M}$ of the system.[26] The averaging is over initial times t_0, i.e., $t \geq t_0$.

It is also possible to obtain a value of τ_D by averaging over a set of nonequilibrium MD simulations that start from a Boltzmann-distributed set of initial configurations and velocities and in which a homogeneous static external electric field $\vec{E}^{ext}$ is switched on at $t = t_0$. This is illustrated in Figure 2. Upon switching on $\vec{E}^{ext}$ along the z-axis at $t = t_0$, the z-component M_z of $\vec{M}$ will increase from its initial value $M_z(t_0)$, the values of which are Gaussian-distributed around $M_z = 0$, to a steady-state value $M_z(t = \infty)$. For a Debye dielectric medium, this buildup will be exponential:

$$\langle M_z(t)\rangle_{t_0} = \langle M_z(t = \infty)\rangle_{t_0}\left[1 - \exp\left(-\frac{t - t_0}{\tau_M}\right)\right] \tag{27}$$

The value of $\langle M_z(t = \infty)\rangle_{t0}$ will be larger for larger E_z^{ext}, but different field strengths E_z^{ext} should yield the same τ_M, as long as E_z^{ext} is not too small and not too large. The relationship between τ_M and τ_D is given as

$$\tau_D = \left[\frac{\varepsilon(0) + 2 + C_{rf}(\varepsilon(0) - 1)}{3}\right]\tau_M \tag{28}$$

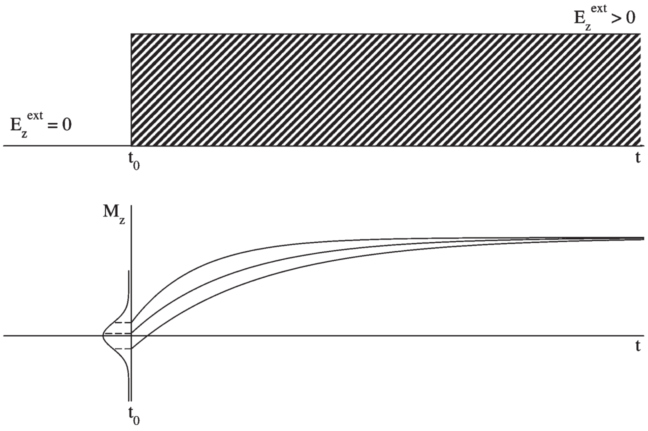


Figure 2. Sketch of the relaxation of the total dipole moment $M_z(t)$, as a function of time upon switching on a homogeneous static electric field E_z^{ext} at time t_0, for three different nonequilibrium molecular dynamics (MD) simulations.

which is a generalization for $\kappa \geq 0$ of the relationship for $\kappa = 0$ given by Neumann.[27] The value of $\varepsilon(0)$ can be calculated using eq 21, in which $\langle M_z\rangle = \langle M_z(t = \infty)\rangle_{t0}$.

2.3. Simulation Details.

The molecular models for water (H_2O), dimethylsulfoxide (DMSO), methanol (MeOH), and chloroform ($CHCl_3$) that have been chosen to test the external field methodology were the models for which the dielectric properties calculated with the dipole-fluctuation method had been reported in the literature.[16−20] The geometry of these rigid models was maintained using the SHAKE algorithm,[28] with a relative accuracy of 10^{-4}. All simulations were performed under NpT conditions using a modified version of the GROMOS05 package of programs.[25] The temperature was kept to a reference value by weak coupling to a temperature bath with a relaxation time of 0.1 ps,[6] and the pressure was maintained at 1.013 bar (1 atm) using the same type of algorithm, with a relaxation time of 0.5 ps and the experimental isothermal compressibility of the corresponding solvent. The integration time step was 2 fs. For the nonbonded interactions, a twin-range method was used with cutoff radii of 0.8 nm (short-range) and 1.4 nm (long-range). Outside the long-range cutoff, a reaction-field correction[15] with a relative dielectric permittivity (ε_{rf}) of 78.5[20] for SPC water, 46[18] for DMSO, 32.63[17] for methanol, and 5[16] for chloroform was applied. Values of $\varepsilon_{cs} = 1$, $\kappa = 0$, and $R_{rf} = 1.4$ nm were used. The pair list for pairs within the short-range cutoff and the energies and forces for long-range pairs were updated every 10 fs (5 steps). Cubic boxes of 5384 SPC water molecules (initial box length = 5.49 nm), 429 DMSO molecules (initial box length = 3.69 nm), 661 methanol molecules (initial box length = 3.60 nm), and 1000 chloroform molecules (initial box length = 5.11 nm) were used, together with periodic boundary conditions. To calculate $\varepsilon(0)$ for each electric field strength, an equilibration simulation of 100 ps and, subsequently, a 500 ps production run was performed at 298 K. The box dipole moment and the volume were saved every 0.4 ps (200 steps), and the atom positions every 2.0 ps (1000 steps) for analysis. The values of τ_D were obtained from 100 short (30 ps) nonequilibrium simulations at two different electric field strengths. The box dipole moment and the volume were saved every step for analysis. The 100 starting configurations were taken from an equilibrium simulation of 1-ns length, where the configuration was saved every 10 ps. The two

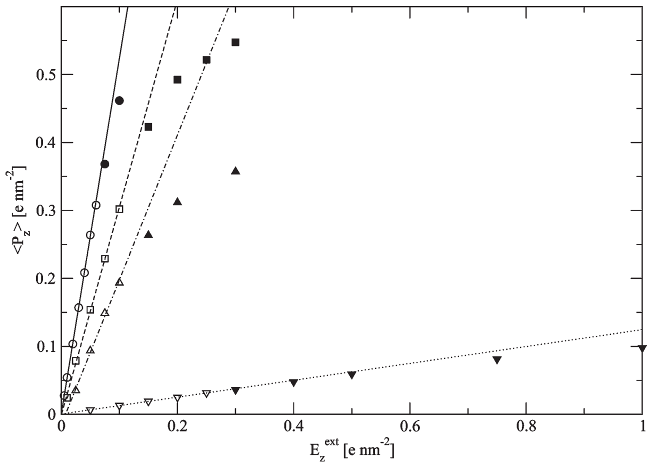


Figure 3. Dependence of the polarization averaged over time $\langle P_z \rangle$ on the applied electric field E_z^{ext} for different solvents: H_2O (circles), DMSO (squares), MeOH (triangles), and $CHCl_3$ (inverted triangles). Lines represent results obtained from linear regression: (—) H_2O, (- - -) DMSO, (- · -) MeOH, and (· · ·) $CHCl_3$. When $\mu_i E_z^{ext}/(3k_B) > 50$ K,[29,30] where μ is the molecular dipole moment, the data (full symbols) were excluded from linear regression.

electric field strengths were chosen for each solvent, such that they are as large as possible while being in the linear-response regime. With regard to DMSO, no value for τ_D obtained from the time correlation function of the dipole moment in an equilibrium simulation was reported in the literature, it was therefore calculated from a 1-ns equilibrium simulation, where the configurations were saved every 0.2 ps (100 steps) for analysis.

3. RESULTS

3.1. Static Dielectric Permittivity $\varepsilon(0)$. In short simulations for the models of four different molecular liquids (H_2O, DMSO, MeOH, and $CHCl_3$), an external electric field of varying strength was applied in the z-direction and the resulting polarization of the box was measured. In Figure 3, the polarization in z-direction is shown as a function of the strength of the applied field. For small fields, the energy of the molecular dipole $\vec{\mu}$ in the field is much smaller than k_BT, i.e., $\mu_i E_z^{ext}/3k_B \ll T$,[29] the response of $\vec{P}$ to $\vec{E}^{ext}$ is linear and only these data points (open symbols) were considered in the calculation of the dielectric permittivity $\varepsilon(0)$. For high field strengths, the relationship becomes nonlinear because of saturation effects (full symbols). The slope of the fitted linear function was used in eq 21, and the resulting $\varepsilon(0)$ values are shown in Table 1, together with the experimental values and those obtained using the dipole moment fluctuation methodology, as reported in the literature. For H_2O, DMSO, and $CHCl_3$, the values for $\varepsilon(0)$ resulting from the external field method agree well with values from the fluctuation method. For MeOH, the values differ, which could be due to the factors used to set up the various simulations: size of the system, constant volume versus constant pressure condition, cutoff radii and update frequency of the nonbonded interaction pairlist, equilibration and simulation time periods. Since the dielectric permittivity is a global property of the system, it converges slowly, especially for high values. We note that all simulations from which the dielectric permittivity was calculated, using the dipole moment fluctuation method,[16−18,20] were of rather short length. To obtain well-converged values with this method, simulations of periods of many nanoseconds should be performed, while 500 ps per field strength E^{ext} are sufficient using an applied external field. A 10-ns equilibrium simulation of the MeOH system, i.e., E^{ext} = 0, was analyzed using the dipole moment fluctuation formula and gave $\varepsilon(0)$ = 24.4.

Table 1. Experimental and Calculated Values for the Relative Static Dielectric Permittivity $\varepsilon(0)$ at 298 K and 1 atm for Water (H_2O), Dimethylsulfoxide (DMSO), Methanol (MeOH), and Chloroform ($CHCl_3$)[a]

	$\varepsilon(0)$		
solvent	expt	fluctuation formula	applied field method
H_2O	78.5[31]	66.6[20]	66.7
DMSO	46[32]	38[18]	39.5
MeOH	32.63[31]	19.8[17]	27.8
$CHCl_3$	4.81[31]	2.4[16]	2.6

[a] The lengths of the simulations for which the dipole moment fluctuation methodology was used were 3 ns,[20] 2 ns,[18] 2 ns,[17] and 1.2 ns.[16]

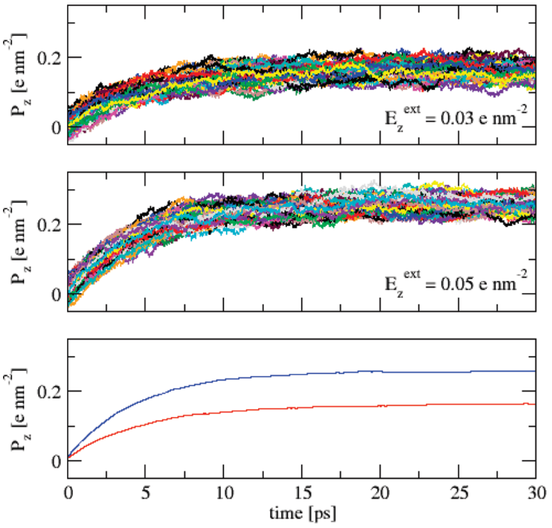


Figure 4. Polarization $P_z(t)$ for 100 nonequilibrium MD simulations of liquid water (5384 SPC molecules) after switching on an electric field E_z^{ext} at $t_0 = 0$, for $E_z^{ext} = 0.03\ e\ \mathrm{nm}^{-2}$ (upper panel) and $E_z^{ext} = 0.05\ e\ \mathrm{nm}^{-2}$ (middle panel). The averages over the 100 trajectories are shown in red ($E_z^{ext} = 0.03\ e\ \mathrm{nm}^{-2}$) and blue ($E_z^{ext} = 0.05\ e\ \mathrm{nm}^{-2}$) in the lower panel.

3.2. Debye Dielectric Relaxation Time (τ_D). In Figure 4, the relaxation of $P_z(t)$ and $\langle P_z(t) \rangle_{t0}$ toward $\langle P_z(t = \infty) \rangle_{t0}$ for liquid water is shown for two different electric field strengths. With increasing field strength, $\langle P_z(t = \infty) \rangle_{t0}$ increases and its variation decreases. However, both field strengths yield similar τ_D values, which are close to those obtained using the equilibrium time correlation function of $\vec{M}$, as reported in the literature. The values of τ_D derived from the relaxation for all four test solvents are given in Table 2. The value obtained using the fluctuation formula, τ_D = 12.3 ps, is not very precise, because of the short length of 1 ns of this simulation.

4. DISCUSSION

The models that were used for H_2O, DMSO, MeOH, and $CHCl_3$ were rigid (i.e., they did not possess internal degrees of

Table 2. Experimental and Calculated Values for the Debye Dielectric Relaxation Time τ_D at 298 K and 1 atm for Water (H_2O), Dimethylsulfoxide (DMSO), Methanol (MeOH), and Chloroform ($CHCl_3$)[a]

	τ_D[ps]		Applied Field Method			
		fluctuation	E_z^{ext}	τ_D	E_z^{ext}	τ_D
solvent	expt	formula	[e nm^{-2}]	[ps]	[e nm^{-2}]	[ps]
H_2O	8.3[33]	6.2[20]	0.03	6.6	0.05	6.0
DMSO	18.5[34]	12.3	0.05	9.6	0.075	9.7
MeOH	49[35]	14[17]	0.075	14.7	0.1	14.1
$CHCl_3$	5.4[36]	3.8[16]	0.2	4.2	0.4	4.3

[a] The electric field strengths were chosen such that they were as large as possible while within the linear-response regime.

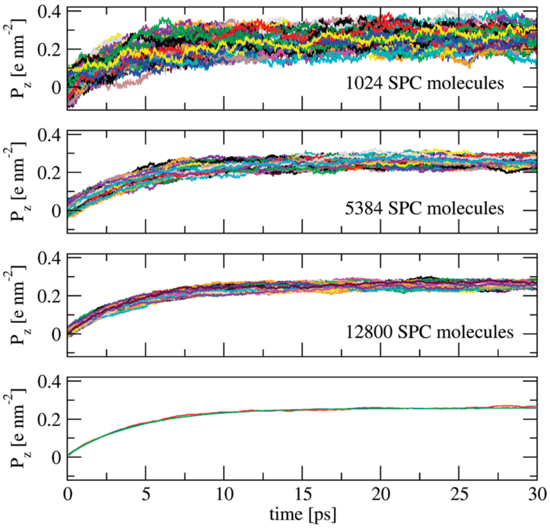


Figure 5. Polarization $P_z(t)$ for 100 nonequilibrium MD simulations of liquid water using three different system sizes (1024, 5384, and 12800 SPC molecules), after switching on an electric field $E_z^{ext} = 0.05$ e nm^{-2} at $t_0 = 0$. The averages over the 100 trajectories are shown in red (1024 molecules), blue (5384 molecules), and green (12800 molecules) in the lowest panel.

freedom). This meant that no vibrational contributions of such degrees of freedom to $\varepsilon(0)$ or τ_D were present.

Recently, the correlations in the total dipole moment $\vec{M}$ of a Stockmayer liquid were analyzed in equilibrium MD simulations using periodic boundary conditions and either a lattice-sum (Ewald) or a reaction-field method to approximate the long-range electrostatic interactions.[37] For the strongly polar Stockmayer model used, the fluctuations of $\vec{M}$ were observed to be dependent on the relative size of the cutoff sphere and the computational periodic box. In other words, a dependence of $\langle \vec{M}^2 \rangle$ on the system size was reported. Therefore, we investigated the system size dependence of the values of $\varepsilon(0)$ and τ_D obtained with the external field method, using liquid water as the test system. The results are shown in Figure 5 and Table 3. While the variation of $P_z(t = \infty)$ decreases as the system size increases, because of better statistics, the average relaxation is independent of the system size (lowest panel in Figure 5), and so are the values obtained for τ_D and $\varepsilon(0)$.

One could think of avoiding periodicity artifacts by using a fixed nonperiodic spherical boundary in combination with the image charge representation of the reaction field.[38] However, such an approach introduces wall effects, which cannot easily be compensated via the use of special wall forces.[39−41]

Table 3. Calculated Values for the Relative Static Dielectric Permittivity $\varepsilon(0)$ and the Debye Dielectric Relaxation Time τ_D at 298 K and 1 atm for Water Using Three Different System Sizes and Two Different Electric Field Strengths[a]

number of H_2O	$\varepsilon(0)$	E_z^{ext}[e nm^{-2}]	τ_D[ps]	E_z^{ext}[e nm^{-2}]	τ_D[ps]
1024	63	0.03	6.1	0.05	6.0
5384	67	0.03	6.6	0.05	6.0
12 800	64	0.03	6.3	0.05	6.0

[a] The electric field strengths were chosen such that they were as large as possible while within the linear-response regime.

5. SUMMARY AND CONCLUSIONS

Expressions for the calculation of the static relative dielectric permittivity ($\varepsilon(0)$) and the Debye dielectric relaxation time (τ_D) of a molecular model from MD simulations of the liquid phase, where a constant external electric field is applied, were given for the case in which a spherical truncation with $\varepsilon_{cs} \geq 1$ inside the cutoff sphere combined with a Poisson−Boltzmann continuum reaction-field with $\kappa \geq 0$ outside the cutoff sphere is used. The external field method was applied to molecular models of four different molecular liquids (water (H_2O), dimethylsulfoxide (DMSO), methanol (MeOH), and chloroform ($CHCl_3$)), and the results were compared to values obtained through the dipole moment fluctuation methodology and from experiment. The dielectric permittivities and relaxation times calculated using the external field method agree with values resulting from the fluctuation method. The external field method is simple to implement and, compared to the dipole moment fluctuation methodology, it is computationally more efficient and can also be applied to uncharged, but polarizable molecular models or in coarse-grained simulations, where $\varepsilon_{cs} > 1$ is used.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: wfvgn@igc.phys.chem.ethz.ch.

■ ACKNOWLEDGMENT

We thank Phillippe Hünenberger for useful discussion. The financial support by the National Center of Competence in Research (NCCR) in Structural Biology, the Swiss National Science Foundation (Grant No. 200020-121913), and the European Research Council (Grant No. 228076) is gratefully acknowledged.

■ REFERENCES

(1) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Clarendon Press: Oxford, U.K., 1987.

(2) Neumann, M. *Mol. Phys.* **1983**, *50*, 841.

(3) Heinz, T. N.; van Gunsteren, W. F.; Hünenberger, P. H. *J. Chem. Phys.* **2001**, *115*, 1125.

(4) Adams, D. J.; Adams, E. M. *Mol. Phys.* **1981**, *42*, 907.

(5) Berendsen, H. J. C. Transport Properties Computed by Linear Response through Weak Coupling to a Bath. In *Computer Simulation in Materials Science*; Meyer, M., Pontikis, V., Eds.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1991; pp 139−155.

(6) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 1517.

(7) Luty, B. A.; van Gunsteren, W. F. *J. Phys. Chem.* **1996**, *100*, 2581.

(8) Hünenberger, P. H.; McCammon, J. A. *J. Chem. Phys.* **1999**, *110*, 1856.

(9) Hünenberger, P. H.; McCammon, J. A. *Biophys. Chem.* **1999**, *78*, 69.

(10) Weber, W.; Hünenberger, P. H.; McCammon, J. A. *J. Phys. Chem. B* **2000**, *104*, 3668.

(11) Nymand, T. M.; Linse, P. *J. Chem. Phys.* **2000**, *112*, 6386.

(12) Walser, R.; Hünenberger, P. H.; van Gunsteren, W. F. *Proteins* **2001**, *43*, 509.

(13) Tironi, I. G; Luty, B. A.; van Gunsteren, W. F. *J. Chem. Phys.* **1997**, *106*, 6068.

(14) Barker, J. A.; Watts, R. O. *Mol. Phys.* **1973**, *26*, 789.

(15) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1995**, *102*, 5451.

(16) Tironi, I. G.; van Gunsteren, W. F. *Mol. Phys.* **1994**, *83*, 381.

(17) Walser, R.; Mark, A. E.; van Gunsteren, W. F. *J. Chem. Phys.* **2000**, *112*, 10450.

(18) Geerke, D. P.; Oostenbrink, C.; van der Vegt, N. F. A.; van Gunsteren, W. F. *J. Phys. Chem. B* **2004**, *108*, 1436.

(19) Smith, P. E.; van Gunsteren, W. F. *J. Chem. Phys.* **1994**, *100*, 3169.

(20) Glättli, A.; Daura, X.; van Gunsteren, W. F. *J. Chem. Phys.* **2002**, *116*, 9811.

(21) van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*; vdf Hochschulverlag AG an der ETH Zürich: Zürich, Switzerland, 1996.

(22) Panofsky, W. K. H.; Phillips, M. *Classical Electricity and Magnetism*, 2nd ed.; Addison−Wesley: Reading, MA, 1964.

(23) Böttcher, C. J. F. *Dielectrics in Static Fields*; Theory of Electric Polarization, Vol. 1; Elsevier Science Publishers: Amsterdam, The Netherlands, 1973.

(24) Böttcher, C. J. F.; Bordewijk, P. *Dielectrics in Time-Dependent Fields*; Theory of Electric Polarization, Vol. 2; Elsevier Science Publishers: Amsterdam, The Netherlands, 1978.

(25) Christen, M.; Hünenberger, P. H.; Bakowies, D.; Baron, R.; Bürgi, R.; Geerke, D. P.; Heinz, T. N.; Kastenholz, M. A.; Kräutler, V.; Oostenbrink, C.; Peter, C.; Trzesniak, D.; van Gunsteren, W. F. *J. Comput. Chem.* **2005**, *26*, 1719.

(26) Neumann, M.; Steinhauser, O. *Chem. Phys. Lett.* **1983**, *102*, 508.

(27) Neumann, M. *J. Chem. Phys.* **1985**, *82*, 5663.

(28) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.

(29) Evans, W. A. B.; Powles, J. G. *Mol. Phys.* **1982**, *45*, 695.

(30) Kunz, A. P. E.; Eichenberger, A. P.; van Gunsteren, W. F. *Mol. Phys.* **2010**, *109*, 365.

(31) Lide, D. R. *Handbook of Chemistry and Physics*, 88th ed.; CRC Press/Taylor and Francis: Boca Raton, FL, 2007−2008.

(32) Riddick, J. A.; Bunger, W. B.; Sakand, T. K. *Organic Solvents: Physical Properties and Methods of Purification*; John Wiley and Sons: New York, 1986.

(33) Kaatze, U. *J. Chem. Eng. Data* **1989**, *34*, 371.

(34) Kaatze, U.; Pottel, R.; Schaefer, M. *J. Phys. Chem.* **1989**, *93*, 5623.

(35) Barbenza, G. H. *J. Chim. Phys. Phys.−Chim. Biol.* **1968**, *65*, 906.

(36) Antony, A. A.; Smyth, C. P. *J. Am. Chem. Soc.* **1964**, *86*, 152.

(37) Stenhammar, J.; Linse, P.; Karlström, G. *J. Chem. Phys.* **2009**, *131*, 164507.

(38) Friedman, H. L. *Mol. Phys.* **1975**, *29*, 1533.

(39) Juffer, A. H.; Botta, E. F. F.; van Keulen, B. A. M.; van der Ploeg, A.; Berendsen, H. J. C. *J. Comput. Phys.* **1991**, *97*, 144.

(40) Juffer, A. H.; Berendsen, H. J. C. *Mol. Phys.* **1993**, *79*, 623.

(41) Juffer, A. H. *On the Modelling of Solvent Mean Force Potentials*, Ph.D. Thesis, University of Groningen, The Netherlands, 1993.

Small Molecules in C_{60} and C_{70}: Which Complexes Could Be Stabilized?

Tatiana Korona*,[†] and Helena Dodziuk*,[§]

[†]Faculty of Chemistry, University of Warsaw, Pasteura 1, 02-093 Warsaw, Poland
[§]Institute of Physical Chemistry, Polish Academy of Sciences, Kasprzaka 44, 01-224 Warsaw, Poland

ABSTRACT: The recent syntheses of complexes involving some small molecules in opened fullerenes and those of hydrogen molecule(s) in C_{60} and C_{70} are accompanied in the literature by numerous computations for endohedral fullerene complexes which cope with the problem of the stability of these complexes. In this contribution, stabilization energies of endohedral complexes of C_{60} and C_{70} with H_2, N_2, CO, HCN, H_2O, H_2S, NH_3, CH_4, CO_2, C_2H_2, H_2CO, and CH_3OH guests have been estimated using symmetry-adapted perturbation theory, which, contrary to the standard DFT and some other approaches, correctly describes the dispersion contribution of the host−guest interactions. On the basis of these calculations, the endohedral complexes with all these guests were found stable in the larger fullerene, while the C_{60} cage was found too small to host the latter four molecules. Except for H_2 and H_2CO, a stabilization effect for most guests in the C_{60} cage is about 30 kJ/mol. For H_2 and H_2O guests, a typical supramolecular effect is observed; namely, the stabilization in the smaller cage is equal to or larger than that in the larger C_{70} host. Except for the water molecule where the induction interaction plays a non-negligible role, in all complexes the main stabilization effect comes from the dispersion interaction. The information on the stability of hypothetical endohedral fullerene complexes and physical factors contributing to it can be of importance in designing future experiments contributing to their applications.

INTRODUCTION

The captivating idea of an empty space inside the C_{60} cage that could be filled with atoms or molecules has been recognized from the early stages of fullerene study.[1] It was also soon realized that the confinement of a guest inside the fullerene cage would influence both host and guest[2] and could lead to various applications,[2,3] although most of them proved until now unrealizable.[4] Fullerenes can be used in medicine,[5−7] in photovoltaic devices,[8,9] and in electronics,[10−12] in particular, as single-molecule transistors for quantum computing.[13−16] Of special interest are so-called peapods formed by carbon nanotubes filled with endohedral fullerene complexes such as $^{15}N@C_{60}$ or $^{31}P@C_{60}$, which are considered as promising elements of quantum-processing architecture.[17] In view of exciting properties and the prospects of applications, a prediction of stability and properties of endohedral fullerene complexes on the basis of computations seems to be quite important. In particular, the question of how many hydrogen molecules can be hosted by C_{60} has drawn the attention of many researchers.[18,19] The published theoretical estimations vary from 1 to over 20, while only one H_2 molecule could be inserted in C_{60},[20] and in the synthesized mixture of $H_2@C_{70}$ and $2H_2@C_{70}$,[21] the latter complex was present as only a 4% admixture, which agrees well with the results obtained on the basis of a simplified model.[22] An often neglected fact that endohedral fullerene complexes are objects with distinct topological properties is worth mentioning.[19,23]

Fullerene endohedral complexes involving lanthanoide guests are probably the most studied, but recent syntheses of $H_2@C_{60}$[20] and of a mixture of $H_2@C_{70}$ and $2H_2@C_{70}$[21] using a "molecular surgery" approach consisting of the chemical opening of the cage, inserting the guest, and chemically closing the cage hole have paved the way to endohedral fullerenes with other guest molecules. Several open cage fullerenes with nonpolar or slightly polar molecular guests inserted have been reported. In addition to H_2 in opened C_{60},[20,24] they include $H_2@C_{70}$ and $2H_2@C_{70}$,[25] N_2,[26] H_2O,[27,28] NH_3,[29] CO,[30] and CH_4[31] in opened C_{60}, $CO@C_{60}$, and N_2 in C_{60} and C_{70}.[32] This vivid activity of synthetic chemists inspired us to analyze the stability of several endohedral fullerene complexes with small guest molecules.

It should be stressed that the calculations of endohedral fullerene complexes present a difficult task since, on the one hand, the systems under study are large and, on the other, weak dispersive interactions stabilizing the complexes with nonpolar or slightly polar guests are poorly described at low-level quantum chemical calculations tractable for such large systems. Earlier studies on endohedral fullerene complexes were summarized in two reviews.[18,19] A detailed discussion of published experimental and calculated data on endohedral fullerene complexes involving hydrogen guests has been published recently,[33] applying symmetry-adapted perturbation theory (SAPT)[34,35] on $H_2@C_{60}$, $2H_2@C_{60}$, and $2H_2@C_{70}$. Therefore, here, only new results for these complexes and those involving complexes with other nonpolar or slightly polar guests will be briefly discussed.

Endohedral complexes with C_{60}, C_{70}, two isomers of C_{76}, seven isomers of C_{80}, and nine isomers of C_{82} with a series of nonpolar or slightly polar guest molecules were analyzed by Dodziuk et al.[36] using molecular mechanics (MM) calculations using two different force fields. The most interesting conclusion from these calculations (the qualitative results did not depend on the force field applied) was that only H_2, H_2O, and, possibly, NH_3 were stabilized inside C_{60}. Thus, if endohedral fullerenes were to be applied (one of the potential applications envisaged by Stoddart[3] is their utilization as drug carriers), then the development of methods for production and purification of higher fullerenes was a must. Next, after learning about obtaining H_2

Received: February 16, 2011
Published: April 15, 2011

in C_{60} and two hydrogen molecules in opened C_{70},[37] Dodziuk et al. carried out MM calculations for 1−4 H_2 in C_{60}, C_{70}, two isomers of C_{76}, two isomers of C_{78}, and seven isomers of C_{80}, finding that one guest molecule is stabilized in both C_{60} and C_{70}.[36] Although two H_2's were also found stabilized in the latter host, the absolute value of stabilization energy was smaller by 1.8 kcal/mol than that of the H_2@C_{70} complex.[22] This effect is due to repulsion of two hydrogen molecules inside 2H_2@C_{70}.

The conclusions of these studies were that (a) it would be difficult to close the C_{70} cage with two hydrogen molecules inside and (b) C_{80} or larger fullerene would be necessary to host three H_2's. The former conclusion proved valid, and interestingly and somewhat accidentally, the experimental H_2@C_{70}:2H_2@C_{70} rate of 96:4[21] corresponded to the calculated energy difference. Extensive calculations for H_2 and N_2 in C_{60} at fixed geometries were carried out by Slanina et al., who utilized second-order Møller−Plesset (MP2), spin-component-scaled MP2 (SCS-MP2)[38] theories, and density functional theory (DFT) with the MPWB1K functional,[39] yielding the best estimates for the stabilization effect of 4 and 9 kcal/mol, respectively.[40] Ren and co-workers reported calculations for H_2 inside the C_{60} cage, but their choice of methods—a combined PM3 and DFT study—has not been appropriate for such types of complexes, which are mainly bound by dispersion interactions.[41] One and two H_2 molecules in C_{60} and two of them in C_{70} were calculated in our group[33] using SAPT(DFT) (SAPT with interacting molecules described on the DFT level), yielding one H_2 stabilized in C_{60} and two of them in C_{70}. On the other hand, on the basis of calculations for the structures optimized at the PBE-D/def2-TZVPP level using several DFT functionals and MP2 and SCS-MP2 methods, Kruse and Grimme claimed that 2H_2@C_{70} is more stable than H_2@C_{70}.[42] They even stated that we have obtained the same trend, although we have not considered the complex H_2@C_{70} explicitly.[33] In the interesting work of Sebastianelli et al.,[43] it was found that the translation−rotation zero-point energies constitute significant factors destabilizing the 2H_2@C_{60} and 3H_2@C_{70} complexes. The conclusions from their work should equally apply to other endohedral complexes, meaning, e.g., that even if the interaction energy for some fullerene complexes has been found negative, it does not automatically signify that the complex can be formed if too large zero-point translation and rotation motions are present. In a similar spirit, Yagi and Watanabe proposed to separate slow vibrational motions of a host within a new method called instantaneous vibrational analysis (IVA) and applied the new approach to study translational and rotational movement of water molecule inside the C_{60} cage.[44] They have found a considerable blue shift of vibrational levels of most configurations of water inside C_{60} and also noticed a significant dependence of the final results on the level of theory used to evaluate the potential for IVA. Recently, Min et al. studied the IR absorption of endohedral H_2 in C_{60} in a wide range of temperatures, from 6 to 300 K.[45]

Several studies for complexes involving molecular guests other than hydrogen have been carried out. Unfortunately, most of them employed the computational methods, like standard DFT or semiempirical models, which are known to give unreliable results when utilized for nonpolar or slightly polar guests. The lack of a proper description of dispersion interactions in DFT with standard functionals makes all conclusions about endothermal effects involved in the complexation questionable at best. Utilizing extremely small basis sets is another common drawback of many fullerene studies. Although it is often difficult to afford larger basis sets, an examination of rather subtle intermolecular interactions with inappropriate basis sets is of very limited significance. Therefore, such studies involving, e.g., geometry optimizations, should be at best accompanied with *a posteriori* single-point calculations for selected configurations, employing a more advanced method for describing electron correlation, like MP2 or SCS-MP2. In many cases, the stabilization only is achieved after the dispersion effect is taken into account. However, one should be aware that MP2 often overestimates the binding energy for the interacting molecules involving aromatic rings.[46] Some examples of such combined studies are given in works of Charkin et al., who performed a study on methane inside the C_{60} and C_{84} (T_d) cages[47] and carbon dioxide, ethyne, and several other small molecules in the C_{70} cage[48] with the B3LYP functional and the 6-31G and 6-31G* basis sets, followed by the MP2 computation used to obtain the dissociation energies. The same group calculated the properties of benzene and borazole inside the C_{84} fullerene.[49] Yet another application of this type has been reported by Slanina et al. for the nitrogen molecule in C_{60}.[40,50] The complexes of methane in C_{84} and C_{60} have also been calculated by Rehaman and co-workers.[51] However, only for the smaller host has the MP2 stabilization energy been reported, while for the CH_4@C_{84}, only the DFT value has been given. DFT calculations of Jin and co-workers on the complexes of C_2H_2@C_{60}, C_2H_4@C_{60}, and C_2H_6@C_{60}[52] also seem of too low accuracy to yield reliable results. Some GGA functionals have been applied by Gao et al.[53] to N_2@C_{60}, in which a considerable interaction of the lone pairs on nitrogen atoms with the π electrons of the cage can be expected. Therefore, the conclusion of Gao and co-workers on the small effect of the guest inclusion on the properties of the complex is questionable. Another example of an inappropriate choice of methodology is the work by Mazurek and Sadlej-Sosnowska,[54] who utilized very small basis sets for the Hartree−Fock (HF), DFT, and MP2 calculations of the complex of benzene enclathrated inside C_{60}. Therefore, their conclusion that such a complex would be stable cannot be trusted. Note that it can be easily seen that the distances between guest hydrogen atoms and host carbon atoms are considerably smaller than the sum of their van der Waals radii; therefore, large repulsion effects should be expected.

Ren and co-workers[55] performed B3LYP/6-31G(d) calculations for the hypothetical highly reactive tetrahedrane inside C_{60}. Although the effect of stabilization of a short-lived species in a fullerene cage has been already reported (see the discussion by Dodziuk)[56] and the cyclobutadiene synthesis inside a hemicarcerand has been executed by the Cram group,[57] for the case investigated by the Ren group, there is no way to insert a short-lived tetrahedrane into the fullerene cage to allow one to test its stabilization. Similarly, obtaining hypercoordinated cluster C_2 inside highly unstable C_{20} and C_{24}, for which Wang et al.[58] reported the DFT calculations, seems hardly possible, although the synthesis of the former cage has been achieved.[59]

In another study, Hu and Ruckenstein[60] reported HF calculations for H_2 and CO molecules inside the C_{58} cage with a seven-membered ring. It is, however, well-known from the early Cioslowski calculations[61] that the HF method is not capable of describing the stabilization of endohedral fullerene complexes with a nonpolar or slightly polar guest, due to the lack of dispersion interactions (note, however, that Cioslowski did not draw this conclusion in the paper).[61] The DFT calculations for the hypothetical saturated $C_{60}H_{60}$ cage with small guest molecules have been reported by Hu and Ruckenstein.[62] However, the cage has been shown by Saunders[63] and Dodziuk and Nowiński[64] to be highly strained if all CH bonds point outside. Therefore, the investigations by Hu and Ruckenstein have a very limited foreseeable value. The former authors also carried out HF and B3LYP calculations for the opened C_{60} cage which are indispensable for the manufacturing of endohedral fullerene complexes with nonpolar or slightly polar molecules.[60] Unfortunately, also here the selection of the computational methods put in doubt the reliability of the results obtained.

In this study, we want to present the calculations of the stabilization energies of endohedral complexes of the C_{60} and C_{70} fullerenes with several small guest molecules, performed using the accurate SAPT approach. This method, although more demanding in terms of the computational resources than the supermolecular HF and DFT methods, yields in return very reliable results, which for small molecules compare favorably to those obtained by the highly accurate supermolecular CCSD(T) method.[65,66] With increasing computer power, this method has become affordable and easy to use, even for the fullerene complexes, and can therefore allow one to evaluate the quality of other methods, utilized so far for systems of such sizes.[33,67−70] A separation of energy components into physically sound contributions allows us additionally to analyze the sources of stability or instability of a given endohedral complex under scrutiny. As mentioned above, several uses of the endohedral fullerene complexes have been proposed but not yet implemented. We believe that a better understanding of forces stabilizing the complexes can contribute to the development of their marketable applications.

■ METHOD

The interaction energy of two molecules A and B is conveniently defined as a difference of the energy of the inclusion complex and the sum of energies of constituent molecules:

$$E_{\mathrm{int}} = E(\mathrm{A@B}) - [E(\mathrm{A}) + E(\mathrm{B})] \qquad (1)$$

Note that in this definition the geometries of A and B do not change in the complex. The stabilization energy is then defined as the interaction energy calculated at a minimum. Two approaches are applied to calculate this energy: the supermolecular and the perturbational ones. In the first approach, the definition (eq 1) is directly used to calculate the interaction energy, where the energies of A@B, A, and B are obtained from the computational quantum chemistry method, like, e.g., CCSD(T), MP2, HF, or DFT. The supermolecular approach is easy to utilize, but one should be aware of the limitations of the methods used to obtain E(A@B), E(A), and E(B) so that no important components of the interaction energy are lost (like the dispersion energy, which is absent in the supermolecular HF interaction energy).

In the second approach, the interaction energy is derived from perturbation theory with the intermolecular interaction operator being the perturbation operator and with the sum of Hamiltonians of molecules A and B being the unperturbed operator. The most successful perturbational theory applied to the intermolecular interactions is the approach proposed and developed by Jeziorski et al., which bears the name symmetry-adapted perturbation theory (SAPT).[34,35] Since we do not know the exact wave functions for the unperturbed molecules A and B, which is usually a prerequisite of employing the perturbational method, several approaches have been developed to cope with this problem, among which the description of the interacting molecules on the DFT level is the only practical way to treat large complexes, especially if two-electron repulsion integrals are calculated with help of the density-fitting (DF) technique[71] (see, e.g., refs 65, 72−76).

The idea of applying DFT in SAPT is a successful example of "taking the best from both worlds", i.e., treating the intramolecular electron correlation by DFT and the intermolecular (i.e., the dispersion effect) by the *ab initio* perturbational method. The quality of the SAPT(DFT) (called alternatively DFT-SAPT) approach has been critically evaluated by comparing it with the supermolecular CCSD(T) results[65,66] and recently with accurate SAPT(CCSD) calculations (see, e.g., refs 77−79). The results of this comparison confirm that the SAPT(DFT) usually reproduces quite accurately the so-called intramonomer electron-correlation effects and that it is therefore the best possible method up-to-date capable of treating large van der Waals complexes.

In SAPT, the interaction energy is obtained as a sum of several physically sound energy contributions, i.e., the electrostatic, induction, and dispersion energies and their exchange counterparts, which arise from the imposition of the Pauli exclusion principle on the approximate wave function

$$E_{\mathrm{int}} = E^{(1)}_{\mathrm{elst}} + E^{(1)}_{\mathrm{exch}} + E^{(2)}_{\mathrm{ind}} + E^{(2)}_{\mathrm{exch\text{-}ind}} + E^{(2)}_{\mathrm{disp}} + E^{(2)}_{\mathrm{exch\text{-}disp}} \qquad (2)$$

Therefore, SAPT not only provides us the total interaction energy, like the supermolecular methods do, but additionally, it allows for an analysis of the importance of various contributions to the interaction energy. It should be noted parenthetically that in SAPT the intermolecular interaction operator is utilized in an exact, nonexpanded form; i.e., no multipole expansion is employed, thus avoiding possible complications arising from the divergent character of the inverse-distance series. It should be emphasized that for the case of the endohedral complex the usual expanded form of the interaction potential cannot be utilized at all. If for some reason the expanded form of the potential is necessary, the form presented in refs 80 and 81 should be applied.

In this work, DF-DFT-SAPT (i.e., DFT-SAPT with two-electron repulsion integrals calculated with help of the DF technique) implemented in the Molpro suite of programs[82] has been used to calculate the stabilization energies of the complexes. The interacting molecules in DFT-SAPT have been described by the PBE functional,[83] with the asymptotic correction as proposed in ref 84. For a calculation of this correction, we took the experimental ionization potentials from refs 85 and 86 for C_{60} and C_{70}, respectively, or from the Web site http://cccbdb.nist.gov/exp2.asp for other molecules, while the DFT HOMO energies have been calculated in the same basis as used in DFT-SAPT.

The geometries of the complexes were prepared using molecular mechanics with the MM3[87−89] parametrization. The minimized structures were then reoptimized using the ArgusLab package (http://www.arguslab.com/) with the PM3 method. Several input geometries have been tried in MM and (e.g., for the linear guests involving C_{60}) the guest has been orientated perpendicular to the six- or five-membered rings of the host. The steric energy of the system has been minimized, whereas for C_{70} complexes, three guest orientations for the starting geometry (along the long axis and perpendicular to it) have been minimized. Only the geometries corresponding to the lowest MM energies have been selected for the single-point DFT-SAPT calculations. We are aware that this method of selection of the configurations cannot locate the DFT-SAPT minimum exactly, but the full geometry optimization by a more advanced method,

Table 1. Calculated SAPT(DFT) Interaction Energies in the TZVP Basis Set for the Complexes of Linear, Planar, and Three-Dimensional Guests in C_{60} and C_{70} in kilojoules per mole[a]

guest	C_{60} host	C_{70} host
	linear guests	
H_2	−13.2[b]	−13.0
N_2	−15.9	−30.1
CO	−21.4	−31.2
HCN[c]	9.8	−40.2
HCN[d]	−6.3	not calculated
CO_2	117.9	−32.2
C_2H_2	73.0	−22.2
	planar guests	
H_2O	−30.8	−27.9
H_2S	−32.0	−33.1
H_2CO[c]	65.1	1.8
H_2CO[d]	not calculated	−8.6
	nonplanar guests	
NH_3	−22.6	−30.7
CH_4[c]	−6.2	−23.8
CH_4[d]	−16.5	not calculated
CH_3OH	199.1	−29.7

[a] For a few cases, the calculations were repeated in the larger def2-TZVPP basis. [b] In ref 35, where the better TZVPP basis has been used for this complex, the interaction energy was equal to (−19.3) − (−19.4) kJ/mol for three orientations studied, while for one orientation (the hydrogen molecule parallel to one hexagon), it was higher by 0.76 kJ/mol. [c] Calculated using TZVP basis set. [d] Calculated using def2-TZVPP basis set.

like DFT (with a possible addition of dispersion correction by the Grimme method[38]) or MP2 in a reasonable basis, is beyond our computer capacities. Our previous studies on just one complex ($H_2@C_{60}$)[78] indicate that the potential energy surface is almost flat in the minimum region and that, e.g., moving the H_2 molecule by 0.2 Å from the center of the cage changes the interaction energy by only 0.3 kJ/mol. We have also performed some test calculations for an example of larger water guests in C_{60} and changed the PM3 configuration by shifting H_2O by 0.03 to 0.10 Å off the fullerene center or by elongating the O−H bonds by 1%. Such configuration modifications resulted in the energy lowering or raising by at most 0.25 kJ/mol. Therefore, it seems that locating the minimum by a very approximate method, like MM or PM3, should not change the general conclusions about values of the DFT-SAPT stabilization energies by more than a few tenths of kilojoule per mole.

Sizes of the complexes prevent us from utilizing large basis sets. To counterbalance the computational time and accuracy, we decided to employ the TZVP basis set[90,91] with the corresponding auxiliary basis sets for density-fitting.[92,93] In our previous work on hydrogen molecules in C_{60}, we have checked that this basis set gives similar values of energy contributions to those of a larger TZVPP basis for all components, but for the dispersion energy, the TZVPP basis gave a 15% larger value. Since the main aim of the present calculations is to estimate the stability of various species, we can safely say that if the complex should be stable on the basis of the DFT-SAPT calculation in a smaller TZVP basis, it will certainly be stable in any more elaborate basis. In several cases, when the stabilization energy in the TZVP basis set turns out to be close to zero, the calculations in the def2-TZVPP[94] basis set have been performed as well in order to verify whether the better basis set yields stabilization of the complexes. However, since the calculation of a single point with DF-DFT-SAPT in the TZVP basis set takes 2−3 days with available computer facilities, depending on the complex size, and becomes 2.5−3 times longer in the def2-TZVPP basis, we renounce the idea of recalculating all complexes under study in the larger basis.

RESULTS AND DISCUSSION

The calculated interaction energies (see Table 1) obtained with the TZVP basis indicate that, except probably CH_2O, all molecules under scrutiny are stabilized in the C_{70} host, while HCN, CH_2O, C_2H_2, CO_2, and CH_3OH do not form stable complexes in C_{60}. Moreover, a rather small energy value for the $CH_4@C_{60}$ complex suggests that its stability can be jeopardized if, in addition, the zero-point energies are included, which can be equal to a few kilojoules per mole (see refs 43−47). As mentioned above, in order to verify the stability of this complex, an additional calculation in the larger def2-TZVPP basis has been performed. In the new basis, the energy becomes lower by as much as 10 kJ/mol. This finding is mainly attributed to the improved description of the dispersion component of the interaction energy (see Table 2). Similarly, a repeated calculation for the $CH_2O@C_{70}$ in the def2-TZVPP basis yields a small stabilization, which is mostly due to the improved description of the dispersion interactions. Therefore, for the case of a larger fullerene, the DF-DFT-SAPT method predicts that the endohedral complexes with all guests under study are stabilized.

For obvious reasons, for many pairs involving the same guest and different hosts, it can be observed that the stabilization is larger in the larger host. It is noteworthy that several complexes involving the latter cage (with N_2, CO, CO_2, H_2O, H_2S, NH_3, and CH_3OH guests) exhibit very close values of the stabilization energy (−30 ± 3 kJ/mol). Actually, only the complexes $H_2@C_{70}$ and $H_2CO@C_{70}$ have a very different stabilization energy of −13 kJ/mol and 1.8 (or −8.6 in the larger basis) kJ/mol, respectively, while the largest absolute value of 40 kJ/mol has been found for the complex $HCN@C_{70}$. However, the H_2 and, especially, H_2O complexes in both fullerenes under consideration exhibit an effect typical for supramolecular chemistry. Namely, the absolute value of the stabilization energy in the smaller, in size, cage is equal to or even larger than that in the larger C_{70} host. This effect can be rationalized by the fact that in a smaller cage more carbon atoms are close to the guest. Thus, and if the net interaction with the carbon atom is stabilizing, this results in a larger or equal stabilization of this guest in the smaller fullerene.

To analyze the factors influencing stabilization of the complexes, let us look at the SAPT components, which are shown in Table 2 for the fullerene C_{60} and in Table 3 for the larger fullerene. First, it should be noted that for endohedral complexes it is inappropriate to think in terms of the usual asymptotic description of the energy components (e.g., interaction of dipoles etc.), since the interacting molecules are too close to one another. Thus, in this case, the short-range part of the components, i.e., the one resulting from the overlapping of the electron clouds, comes into play.

Starting from the first-order corrections, one can say that the electrostatic energy is relatively large (i.e., it constitutes a large

Table 2. SAPT Interactions Energy Components (in kJ/mol) for the Complexes with C_{60}[a]

	$E^{(1)}_{\mathrm{elst}}$	$E^{(1)}_{\mathrm{exch}}$	$E^{(2)}_{\mathrm{ind}}$	$E^{(2)}_{\mathrm{exch\text{-}ind}}$	$E^{(2)}_{\mathrm{disp}}$	$E^{(2)}_{\mathrm{exch\text{-}disp}}$	E_1	E_2	E_{int}	E_{int}[d]
H_2	−7.4	21.5	−4.9	4.5	−30.9	4.1	14.1	−27.3	−13.2	−16.3
N_2	−18.1	51.4	−16.8	15.4	−55.1	7.4	33.3	−49.2	−15.9	−21.4
CO	−23.7	61.2	−27.8	24.1	−65.5	10.3	37.4	−58.8	−21.4	−28.0
HCN[b]	−91.1	233.0	−99.6	86.4	−151.2	32.5	141.9	−132.0	9.8	−5.3
HCN[c]	−90.4	234.2	−108.7	91.2	−165.7	33.1	143.8	−150.1	−6.3	
CO_2	−168.7	450.7	−189.3	181.2	−202.6	46.6	282.0	−164.1	117.9	97.7
C_2H_2	−166.2	443.8	−220.9	187.8	−224.1	52.6	277.6	−204.5	73.0	50.6
H_2O	−13.3	35.6	−18.3	10.7	−52.4	7.0	22.2	−53.0	−30.8	−36.0
H_2S	−69.8	159.9	−92.3	83.6	−143.3	29.9	90.1	−122.1	−32.0	−46.3
H_2CO	−147.9	369.2	−178.5	161.7	−181.7	42.3	221.3	−156.2	65.1	46.9
NH_3	−62.7	149.2	−85.8	74.8	−124.3	26.2	86.5	−109.1	−22.6	−35.1
CH_4[b]	−75.3	190.5	−83.0	78.4	−146.9	30.1	115.2	−121.4	−6.2	−20.9
CH_4[c]	−74.2	190.6	−81.1	76.4	−159.3	31.3	116.4	−132.9	−16.5	
CH_3OH	−269.1	712.0	−372.3	322.1	−262.5	68.9	442.9	−243.8	199.1	172.9

[a] E_1 and E_2 are the sums of the first-order and the second-order components, respectively. [b] Calculated using the TZVP basis set. [c] Calculated using the def2-TZVPP basis set. [d] Obtained by adding 10% of $E^{(2)}_{\mathrm{disp}}$ to E_{int} calculated in the TZVP basis

Table 3. SAPT Interaction Energy Components (in kJ/mol) for the Complexes with C_{70}[a]

	$E^{(1)}_{\mathrm{elst}}$	$E^{(1)}_{\mathrm{exch}}$	$E^{(2)}_{\mathrm{ind}}$	$E^{(2)}_{\mathrm{exch\text{-}ind}}$	$E^{(2)}_{\mathrm{disp}}$	$E^{(2)}_{\mathrm{exch\text{-}disp}}$	E_1	E_2	E_{int}	E_{int}[d]
H_2	−4.7	13.1	−2.8	2.5	−23.8	2.7	8.4	−21.4	−13.0	−15.4
N_2	−23.7	63.5	−19.7	18.8	−79.6	10.4	39.8	−70.0	−30.1	−38.1
CO	−28.9	73.4	−26.4	25.0	−86.8	12.5	44.5	−75.6	−31.2	−39.8
HCN	−52.3	131.1	−64.4	53.9	−132.4	24.0	78.8	−118.8	−40.1	−53.3
CO_2	−40.9	106.3	−37.0	34.2	−110.7	16.0	65.4	−97.6	−32.2	−43.2
C_2H_2	−63.8	167.1	−81.8	73.6	−145.5	28.1	103.3	−125.4	−22.2	−36.8
H_2O	−12.1	35.0	−18.6	11.8	−51.0	7.1	22.9	−50.8	−27.9	−33.0
H_2S	−66.3	151.4	−96.9	89.2	−139.7	29.3	85.0	−118.1	−33.1	−47.0
H_2CO[b]	−63.5	163.4	−95.1	91.4	−118.0	23.5	100.0	−98.2	1.8	−10.0
H_2CO[c]	−62.9	164.4	−72.9	67.6	−127.6	22.8	101.5	−110.1	−8.6	
NH_3	−37.1	91.0	−54.2	46.3	−94.2	17.4	−84.7	−70.0	−30.7	−40.1
CH_4	−43.8	114.9	−45.2	42.9	−112.0	19.4	71.1	−94.9	−23.8	−35.0
CH_3OH	−69.0	177.3	−75.9	65.5	−155.7	28.0	108.3	−138.0	−29.7	−45.3

[a] E_1 and E_2 are the sums of the first-order and the second-order components, respectively. [b] Calculated using TZVP basis set. [c] Calculated using def2-TZVPP basis set. [d] Obtained by adding 10% of $E^{(2)}_{\mathrm{disp}}$ to E_{int} calculated in the TZVP basis

part of the total interaction energy) in all cases, but it is completely quenched by the first-order exchange term (coming from the Pauli exclusion principle), and that in all cases the net contribution from the interaction of the nonperturbed electron clouds of the host and guest is strongly repulsive (see the column with the total first-order energy, i.e., the E_1 column). In addition, the amount of this repulsion is obviously dependent on the size of the guest. The examination of the pair of the second-order induction and exchange-induction energies reveals that in the majority of cases the positive exchange-induction term cancels the negative induction contribution, leaving a small negative stabilizing contribution. Such a behavior is characteristic for the interaction of highly symmetric molecules with one another and can be explained by the fact that the main component of the induction term comes in this case from the short-range effect, i.e., overlapping of the perturbed electron clouds. The notable exception is the water guest, in which case, almost a half of the induction contribution remains. For other polar molecules, the quenching is still big because even though they have sizable dipole moments, their electron clouds overlap strongly with the host's cloud, creating a short-range induction component. However, one can see that for polar molecules some small, but non-negligible, net contribution remains. Importantly, in some cases, the latter is crucial to ensuring the negative interaction energy or at least make it less positive (as for CH_3OH in C_{70} and C_{60}, respectively). Interestingly, for the methane molecule, the induction term contributes a lot to the negative interaction energy, although this molecule does not have a permanent dipole and quadrupole moment. The importance of the short-range induction contribution for the case of the methane molecule has also been observed in an investigation of the complex of Ar with CH_4 reported by Heijmen et al.[95]

The last pair of SAPT components, i.e., the second-order dispersion and exchange-dispersion energies, exhibits different behavior in comparison to the induction and exchange-induction pair. The exchange-dispersion energy never quenches the dispersion to a large extent. Quite on the contrary, it counterbalances at most 26% (usually 10−20%) of the dispersion effect. In this way, the dispersion is the major negative contribution and is mainly responsive for binding of the guest inside the host. For smaller guests, its magnitude is sufficient to counterbalance the destabilizing effect of the first-order repulsion term, making the guest −host system stable.

In Tables 2 and 3, the interaction energies of two complexes, HCN@C_{60} and H_2CO@C_{70}, calculated in the TZVP basis set assume small positive values. As mentioned above, to check whether this destabilization could be due to the small basis set applied, additional calculations with the def2-TZVPP basis set were performed. As expected, the dispersion energy turned out to be the most sensitive to the quality of the basis set, and an addition of the polarization Gaussian functions resulted in lowering of both the dispersion energy and total interaction energy, leading to the stabilization of the complexes. The resulting increase in the absolute value of the dispersion contribution was about 10% of its value in a smaller basis. Since other components (see the electrostatic energy, first-order exchange energy, and the sum of the second-order induction and exchange-induction energies) are not that sensitive to the basis set effect, we decided to estimate the increase in the stabilization caused by the saturation of the dispersion components by scaling of the $E^{(2)}_{\rm disp}$ term by 10%. The resulting estimated interaction energy is given in the last column in Tables 2 and 3. Such a rescaling procedure has been already utilized in SAPT calculations in similar cases, i.e., when the application of the large basis was not possible because of the hardware limitations, and has been shown to substantially improve the quality of the potential energy surface.[96] It can be seen that—as expected—the stabilization of the complexes becomes larger with the rescaled dispersion, or (for some larger guest in the smaller cage) they become less destabilized. The effect is especially pronounced for nonpolar guests, in which case the dispersion is the main attractive component of the interaction energy. It is also interesting to note that for two isoelectronic guests of approximately the same size, N_2 and CO, a difference of about 5 kJ/mol in the total energy appears. An inspection of data presented in Table 2 allows one to draw a conclusion that the nonsymmetric distribution of the electron density for CO causes a stronger interaction with the C_{60} cage. This effect for the unperturbed densities can be observed as larger absolute values of the electrostatic and first-order exchange energies. Similarly, the perturbed densities cause qualitatively the same effect for the induction and exchange-induction pair; i.e., charge redistribution in the cage caused by CO is larger than by N_2. Also, the dispersion and exchange-dispersion terms contribute to this effect, giving the net increase of the stabilization energy. Interestingly enough, the net stabilization energy for both species in the larger cage is virtually the same. However, the components are quite different (see Table 3).

Although the stabilization energy is the most popular and easy to obtain parameter of the stability of the endohedral complexes, an ultimate test of stability should be the thermodynamic functions, like enthalpies or Gibbs energies. However, in both cases, the calculation of the vibrational frequencies of the complex and the constituent molecules is necessary, which is quite expensive for a molecule of fullerenes' size. Therefore, such studies are much more rare. Slanina and Nagase[50] have obtained an estimate of the Gibbs energy for the N_2@C_{60} complex and have found that the entropic contribution is quite significant (the MP2 stabilization energy and the $T\Delta S$ value are equal to −9.8 kcal/mol and −5.9 kcal/mol, respectively), although the stabilization effect prevails. On the basis of this study, one can estimate that some of the complexes studied by us can be thermodynamically unstable at room temperature. However, due to the high energy of CC bond breaking, once formed, they can be, like cubane, kinetically stable. It should be also noted that the harmonic approximation, utilized almost exclusively to estimate the thermal contribution from the vibrations, could be insufficient for some modes of endohedral complexes.[45]

■ SUMMARY

We performed the calculations of the stabilization energies of several endohedral complexes with C_{60} and C_{70} fullerenes with the recently developed DF-DFT-SAPT approach. On the basis of these results, the stability of all guests under study in the latter, larger fullerene is predicted, while larger guests, i.e., CO_2, C_2H_2, H_2CO, and CH_3OH, are not stabilized in the C_{60} cage. However, a general conclusion of Dodziuk et al.[36] that mastering the manufacturing and purification of larger fullerenes is necessary for applications of their endohedral complexes with molecular guests remains in force. The analysis of the energy components reveals that the main stabilizing effect is always due to the dispersion energy, while the net contribution of the first-order terms is always repulsive. Only for the highly polar molecules does the induction effect of polarizing the cage by the electrostatic field of the guest contribute non-negligibly to the stabilization of the complexes. Taking into account that few small molecules have been found trapped in the opened C_{60}, we expect that the cages will be chemically closed in the near future, providing the corresponding endohedral complexes. At present, no devices on the basis of endohedral fullerene complexes are available on the market, in spite of numerous proposals of their applications. We believe that an understanding of forces stabilizing them can contribute to their development and practical use in the near future.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: tania@tiger.chem.uw.edu.pl, dodziuk@ichf.edu.pl.

■ ACKNOWLEDGMENT

We are grateful to Dr. K. S. Nowiński for the preparation of TOC graphics and to Mr. T. B. Demissie for the preparation of the input data.

■ REFERENCES

(1) Heath, J. R.; O'Brien, S. C.; Zhang, Q.; Liu, Y.; Curl, R. F.; Tittel, F. K.; Smalley, R. E. *J. Am. Chem. Soc.* **1985**, *107*, 7779–7780.

(2) Dresselhaus, M. S.; Dresselhaus, G.; Eklund, P. C. *Science of Fullerenes and Carbon Nanotubes: Their Properties and Applications*; Elsevier: Oxford, U. K., 1995.

(3) Stoddart, J. F. *Angew. Chem., Int. Ed. Engl.* **1991**, *30*, 70–71.

(4) Dodziuk, H. In *Strained Hydrocarbons. Beyond the van't Hoff and Le Bel hypothesis*; Dodziuk, H., Ed.; Wiley-VCH: Weinheim, Germany, 2009; pp 5−12.

(5) Cagle, D. W.; Kennel, S. J.; Mirzadeh, S.; Alford, J. M.; Wilson, L. J. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 5182–5187.

(6) Bolskar, R. D. *Nanomedicine—U.K.* **2008**, *3*, 201–213.

(7) Bakry, R.; Vallant, R. M.; Najam-ul-Haq, M.; Rainer, M.; Szabo, Z.; Huck, C. W.; Bonn, G. K. *Int. J. Nanomed.* **2007**, *2*, 639–649.

(8) Clarke, T. M.; Durrant, J. R. *Chem. Rev.* **2010**, *110*, 6736–6767.

(9) Ross, R. B.; Cardona, C. M.; Swain, F. B.; Guldi, D. M.; Sankaranarayanan, S. G.; Van Keuren, E.; Holloway, B. C.; Drees, M. *Adv. Funct. Mater.* **2009**, *19*, 2332–2337.

(10) Kobayashi, S.; Mori, S.; Iida, S.; Ando, H.; Takenobu, T.; Taguchi, Y.; Fujiwara, A.; Taninaka, A.; Shinohara, H.; Yoshihiro, I. *J. Am. Chem. Soc.* **2003**, *125*, 8116–8117.

(11) Shibata, K.; Kubozono, Y.; Kanbara, T.; Hosokawa, T.; Fujiwara, A.; Ito., Y.; Shinohara, H. *Appl. Phys. Lett.* **2004**, *84*, 2572–2574.

(12) Yasutake, Y.; Shi, Z. J.; Okazaki, T.; Shinohara, H.; Majima, Y. *Nano Lett.* **2005**, *5*, 1057–1060.

(13) Yu, C. Y. *Phys. Rev. A* **2007**, *75*, art. 012318.

(14) Garelli, M. S.; Kusmartsev, F. V. *Eur. Phys. J. B* **2005**, *48*, 199–206.

(15) Twamley, J. *Phys. Rev. A* **2003**, *67*, art. no. 052318.

(16) Meyer, C.; Harneit, W.; Weidinger, A.; Lips, K. *Phys. Status Solidi B* **2002**, *233*, 462–466.

(17) Yang, W. L.; Xu, Z. Y.; Wei, H.; Feng, M.; Suter, D. *Phys. Rev. A* **2010**, *81*, 023303.

(18) Dodziuk, H. *J. Nanosci. Nanotechnol.* **2007**, *7*, 1102–1110.

(19) Dodziuk, H. In *Mathematics and Topology of Fullerenes*; Graovac, A., Ori, O., Cataldo, F., Eds.; Springer: Hamburg, Germany, 2011; pp 117—151.

(20) Komatsu, K.; Murata, M.; Murata, Y. *Science* **2005**, *307*, 238–240.

(21) Murata, M.; Maeda, S.; Morinaka, Y.; Murata, Y.; Komatsu, K. *J. Am. Chem. Soc.* **2008**, *130*, 15800–15801.

(22) Dodziuk, H. *Chem. Phys. Lett.* **2005**, *410*, 39–41.

(23) Dodziuk, H.; Nowinski, K. S. *Tetrahedron* **1998**, *54*, 2917–2930.

(24) Rubin, Y.; Jarrosson, T.; Wang, W.; Bartberger, M. D.; Houk, K. N.; Schick, G.; Saunders, M.; Cross, R. J. *Angew. Chem., Int. Ed.* **2001**, *40*, 1543–1546.

(25) Murata, Y.; Maeda, S.; Murata, M.; Komatsu, K. *J. Am. Chem. Soc.* **2008**, *130*, 6702–6703.

(26) Ito, S.; Shimotani, H.; Takagi, H.; Dragoe, N. *Full. Nanotubes Carbon Nanostruct.* **2008**, *16*, 206–213.

(27) Iwamatsu, S.; Murata, S. *Tetrahedron Lett.* **2004**, *45*, 6391–6394.

(28) Xiao, Z.; Yao, J. Y.; Yang, D. Z.; Wang, F. D.; Huang, S. H.; Gan, L. B.; Jia, Z. S.; Jiang, Z. P.; Yang, X. B.; Zheng, B.; Yuan, G.; Zhang, S. W.; Wang, Z. M. *J. Am. Chem. Soc.* **2007**, *129*, 16149–16162.

(29) Whitener, K. E., Jr.; Frunzi, M.; Iwamatsu, S.-I.; Murata, S.; Cross, R. J.; Saunders, M. *J. Am. Chem. Soc.* **2008**, *130*, 13996–13999.

(30) Iwamatsu, S. I.; Stanisky, C. M.; Cross, R. J.; Saunders, M.; Mizorogi, N.; Nagase, S.; Murata, S. *Angew. Chem., Int. Ed.* **2006**, *45*, 5337–5340.

(31) Whitener, K. E., Jr.; Cross, R. J.; Saunders, M.; Iwamatsu, S.-I.; Murata, S.; Mizorogi, N.; Nagase, S. *J. Am. Chem. Soc.* **2009**, *131*, 6338–6339.

(32) Peres, T.; Cao, B. P.; Cui, W. D.; Lifshitz, C.; Khong, A.; Cross, R. J.; Saunders, M. *Int. J. Mass Spectrom.* **2001**, *210*, 241–247.

(33) Korona, T.; Hesselmann, M.; Dodziuk, H. *J. Chem. Theory Comput.* **2009**, *5*, 1585–1596.

(34) Jeziorski, B.; Moszynski, R.; Szalewicz, K. *Chem. Rev.* **1994**, *94*, 1887–1930.

(35) Szalewicz, K.; Patkowski, K.; Jeziorski, B. *Struct. Bonding (Berlin)* **2005**, *116*, 43–117.

(36) Dodziuk, H.; Dolgonos, G.; Lukin, O. *Carbon* **2001**, *39*, 1907–1911.

(37) Komatsu, K.; Murata, M.; Murata, Y. In *XIX International Winterschool on Electronic Properties of Novel Materials*, Kirchberg in Tirol, Austria, 2005.

(38) Grimme, S. *J. Chem. Phys.* **2003**, *118*, 9095–9102.

(39) Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2005**, *1*, 415–432.

(40) Slanina, Z.; Pulay, P.; Nagase, S. *J. Chem. Theory Comput.* **2006**, *2*, 782–785.

(41) Ren, Y. X.; Ng, T. Y.; Liew, K. M. *Carbon* **2006**, *44*, 397–406.

(42) Kruse, H.; Grimme, S. *J. Phys. Chem. C* **2009**, *113*, 17006–17010.

(43) Sebastianelli, F.; Xu, M.; Bacic, Z.; Lawler, R.; Turro, N. J. *J. Am. Chem. Soc.* **2010**, *132*, 9826–9832.

(44) Yagi, K.; Watanabe, D. *Int. J. Quantum Chem.* **2009**, *109*, 2080–2090.

(45) Min, G.; Nagel, U.; Hüvonen, D.; Rõõm, T.; Mamone, S.; Levitt, M. H.; Carravetta, M.; Murata, Y.; Komatsu, K.; Chen, J. Y.-C.; Turro, N. J. *J. Chem. Phys.* **2011**, *134*, art. 054507.

(46) Hobza, P.; Selzle, H. L.; Schlag, E. W. *J. Phys. Chem.* **1996**, *100*, 18790–18794.

(47) Charkin, O. P.; Klimenko, N. M.; Charkin, D. O.; Mebel, A. M. *Russ. J. Inorg. Chem.* **2004**, *49*, 723–733.

(48) Charkin, O. P.; Klimenko, N. M.; Charkin, D. O.; Mebel, A. M. *Russ. J. Inorg. Chem.* **2005**, *50*, 1903–1911.

(49) Charkin, O. P.; Klimenko, N. M.; Charkin, D. O.; Mebel, A. M. *Russ. J. Inorg. Chem.* **2005**, *50*, 1702–1709.

(50) Slanina, Z.; Nagase, S. *Mol. Phys.* **2006**, *104*, 3167–3171.

(51) Rehaman, A.; Gagliardi, L.; Pyykkö, P. *Int. J. Quantum Chem.* **2007**, *107*, 1162–1169.

(52) Jin, L. J.; Zhang, M.; Su, Z. M.; Shi, L. L. *J. Theor. Comput. Chem.* **2008**, *7*, 1–11.

(53) Gao, H.; Zhu, W. H.; Tang, C. M.; Geng, F. F.; Yao, C. D.; Xu, Y. L.; Deng, K. M. *Acta Phys. Sinica* **2010**, *59*, 1707–1711.

(54) Mazurek, A. P.; Sadlej-Sosnowska, N. *Int. J. Quantum Chem.* **2010**, *110*, 1354–1359.

(55) Ren, Y. X.; Jiang, C. Y.; Wang, J.; Liu, Z. Y. *J. Mol. Graphics Modell.* **2008**, *27*, 558–562.

(56) Dodziuk, H. In *Strained Hydrocarbons. Beyond the van't Hoff and Le Bel hypothesis*; Wiley-VCH: Weinheim, Germany, 2009; pp 449—458.

(57) Tanner, M. E.; Knobler, C. B.; Cram, D. J. *Angew. Chem., Int. Ed.* **1991**, *30*, 1924–1027.

(58) Wang, Y.; Huang, Y. H.; Liu, R. Z. *J. Mol. Struct. THEOCHEM* **2006**, *775*, 61–65.

(59) Prinzbach, H.; Weller, A.; Landenberger, P.; Wahl, F.; Worth, J.; Scott, L. T.; Gelmont, L.; Olevano, D.; von Issendorff, B. *Nature* **2000**, *407*, 60–63.

(60) Hu, Y. H.; Ruckenstein, E. *J. Chem. Phys.* **2003**, *119*, 10073–10081.

(61) Cioslowski, J. *J. Am. Chem. Soc.* **1991**, *113*, 4139–4141.

(62) Hu, Y. H.; Ruckenstein, E. *J. Chem. Phys.* **2005**, *123*, art. 144303.

(63) Saunders, M. *Science* **1991**, *253*, 330–331.

(64) Dodziuk, H.; Nowinski, K. S. *Chem. Phys. Lett.* **1996**, *249*, 406–412.

(65) Podeszwa, R.; Szalewicz, K. *Chem. Phys. Lett.* **2005**, *412*, 488–493.

(66) Hesselmann, A.; Jansen, G.; Schütz, M. *J. Chem. Phys.* **2005**, *122*, art. 054306.

(67) Hesselmann, A.; Korona, T. *Phys. Chem. Chem. Phys.* **2011**, *13*, 732–743.

(68) Podeszwa, R.; Bukowski, R.; Rice, B. M.; Szalewicz, K. *Phys. Chem. Chem. Phys.* **2007**, *9*, 5561–5569.

(69) Podeszwa, R.; Szalewicz, K. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2735–2746.

(70) Podeszwa, R. *J. Chem. Phys.* **2010**, *132*, art. 044704.

(71) Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R. *J. Chem. Phys.* **1979**, *71*, 4993–4999.

(72) Williams, H. L.; Chabalowski, C. F. *J. Phys. Chem. A* **2001**, *105*, 11158.

(73) Jansen, G.; Hesselmann, A. *J. Phys. Chem. A* **2001**, *105*, 11156–11157.

(74) Misquitta, A. J.; Szalewicz, K. *Chem. Phys. Lett.* **2005**, *357*, 301–306.

(75) Hesselmann, A.; Jansen, G. *Chem. Phys. Lett.* **2002**, *357*, 464–470.

(76) Hesselmann, A.; Jansen, G. *Chem. Phys. Lett.* **2002**, *362*, 319–325.

(77) Korona, T. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6509–6519.

(78) Korona, T. *J. Chem. Theory Comput.* **2009**, *5*, 2663–2678.

(79) Korona, T. In *Recent Progress in Coupled Cluster Methods*; Čársky, P., Paldus, J., Pittner, J., Eds.; Springer: Dordrecht, The Netherlands, 2010; pp 267−296.

(80) Pyykkö, P.; Wang, C.; Straka, M.; Vaara, J. *Phys. Chem. Chem. Phys.* **2007**, *9*, 2954–2958.

(81) Wang, C.; Straka, M.; Pyykkö, P. *Phys. Chem. Chem. Phys.* **2010**, *12*, 6187–6203.

(82) Werner, H. J.; Knowles, P. J.; Lindh, R.; Manby, F. R.; Schütz, M.; Celani, P.; Korona, T.; Mitrushenkov, A.; Rauhut, G.; Adler, T. B.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Goll, E.; Hampel, C.; Hetzer, G.; Hrenar, T.; Knizia, G.; Köppl, C.; Liu, Y.; Lloyd, A. W.; Mata, R. A.; May, A. J.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklass, A.; Palmieri, P.; Pflüger, K.; Pitzer, R.; Reiher, M.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T.; Wang, M.; Wolf, A. *Molpro*, version 2009.1; Cardiff University: Cardiff, U. K., 2009.

(83) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

(84) Grüning, M.; Gritsenko, O. V.; van Gisergen, S. J. A.; Baerends, E. J. *J. Chem. Phys.* **2001**, *114*, 652–660.

(85) Lichtenberger, D. L.; Nebesny, K. W.; Ray, C. D.; Huffman, D. R.; Lamb, L. D. *Chem. Phys. Lett.* **1991**, *176*, 203–208.

(86) Steger, H.; Holzapfel, J.; Hielscher, A.; Kamke, W.; Hertel, I. V. *Chem. Phys. Lett.* **1995**, *234*, 455–459.

(87) Lii, J.-H.; Allinger, N. L. *J. Am. Chem. Soc.* **1989**, *111*, 8566–8575.

(88) Lii, J.-H.; Allinger, N. L. *J. Am. Chem. Soc.* **1989**, *111*, 8576–8582.

(89) Allinger, N. L.; Kuohsiang, C.; Lii, J.-H. *J. Comput. Chem.* **1996**, *17*, 642–668.

(90) Godbout, N.; Salahub, D. R.; Andzelm, J.; Wimmer, E. *Can. J. Chem.* **1992**, *70*, 560–571.

(91) Schafer, A.; Huber, C.; Ahlrichs, R. *J. Chem. Phys.* **1994**, *100*, 5829–5835.

(92) Weigend, F.; Höser, M.; Patzelt, H.; Ahlrichs, R. *Chem. Phys. Lett.* **1998**, *294*, 143–152.

(93) Weigend, F.; Köhn, A.; Hättig, C. *Chem. Phys. Lett.* **2002**, *294*, 3175–3183.

(94) Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.

(95) Heijmen, T. G. A.; Korona, T.; Moszynski, R.; Wormer, P. E. S.; van der Avoird, A. *J. Chem. Phys.* **1997**, *107*, 902–913.

(96) Moszynski, R.; Wormer, P. E. S.; Jeziorski, B.; van der Avoird, A. *J. Chem. Phys.* **1994**, *101*, 2811–2824.

JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Free Energy Landscapes of Alanine Dipeptide in Explicit Water Reproduced by the Force-Switching Wolf Method

Yasushige Yonezawa,[*,†] Ikuo Fukuda,[‡] Narutoshi Kamiya,[†] Hiromitsu Shimoyama,[†] and Haruki Nakamura[†]

[†]Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan

[‡]Computational Science Research Program, RIKEN, 2-1, Hirosawa, Wako, Saitama 351-0198, Japan

ABSTRACT: Precise and rapid calculation of long-range interactions is of crucial importance for molecular dynamics (MD) and Monte Carlo simulations. Instead of the Ewald method or its high speed variant, PME, we applied our novel method, called the force-switching Wolf method, to computation of the free energy landscapes of a short peptide in explicit water. Wolf and co-workers showed that long-range electrostatic energy under a periodic boundary condition can be well reproduced even by truncating the contribution from the distant charges, when the charge neutrality is taken into account. We recently applied the procedure proposed by Wolf and co-workers to a mathematically consistent MD theory by means of a force-switching scheme, and we show that the total electrostatic energy for sodium chloride liquid was well conserved and stable during the MD simulation with the force-switching Wolf method. Our current results for an aqueous peptide solution with a series of canonical and multicanonical molecular dynamics simulations show that the force-switching Wolf method is not only in good accordance with the energies and forces calculated by the conventional PME method but also properly reproduces the solvation and the free energy landscapes of the peptide at 300 K.

INTRODUCTION

Molecular dynamics (MD) and Monte Carlo (MC) simulations are now indispensable and widely used for studies of condensed matter, such as nano- and biomolecular systems. In these simulations, most of the computational time is usually consumed by calculations of nonbonded long-range, electrostatic interactions. The long-range interactions determine thermodynamic, structural, and dynamical properties of the systems. Therefore, fast and accurate methods for computing the interactions are required for reliable MD and MC applications.

There are various algorithms proposed for computing the long-range electrostatic interactions. Truncation of interactions (the cutoff method) was formerly the most widely used, but it introduces serious errors and artifacts in the treatment of the energies and forces.[1] The reaction field method,[2] cell-multipole method,[3] fast-multipole method,[4] and smooth-cutoff potential method[5] have been developed to overcome these problems.

The Particle Mesh Ewald (PME) method,[6,7] a computationally efficient alternative to the Ewald method, takes advantage of the Fast Fourier Transform algorithm and is now widely accepted as the standard to calculate electrostatic interactions among charged particles of molecular systems under periodic boundary conditions. However, the Ewald method has intrinsic artifacts, as pointed out by many researchers.[8−10] Moreover, because of network communication problems of the Fast Fourier Transfer algorithm employed in the PME method, it is hard to accomplish good scalability with respect to large systems in highly parallel computations.[11,12]

Wolf et al. proposed an algorithm, in which charge neutrality and potential damping are applied to the cutoff method, and it achieves fairly good accuracy, comparable to the Ewald method, for simulations of NaCl and MgO in crystalline and liquid phases.[13] The algorithm, hereafter referred to as the Wolf method, is easy to parallelize and is computationally very efficient like the original cutoff method, because no reciprocal space calculations are necessary. In the past decade, the Wolf method attracted a great deal of attention and has been evaluated in many applications. Demontis et al. used this method in simulation studies of liquid water and in anhydrous and hydrated aluminosilicates. They showed that the method is computationally more efficient than the PME method when the damping parameter is carefully chosen.[14] Zahn et al. proposed a modification of the Wolf method to apply it to molecular liquids in MD simulations and obtained results comparable to those of the PME method in the TIP3P and SPC water systems.[15] Ma and Garofalini applied the Wolf method to β-SiC crystals with a short-range correction in MD and obtained results that were very close to experimental data.[16] Avendaño and Gil-Villegas used the Wolf method in MC to simulate properties of electrolyte solutions and molten salts and showed that the Wolf method reproduced the simulation data produced using the PME method.[17] Fennell and Gezelter proposed another modification to the original Wolf method, referred to as the Fennell method hereafter.[18] This method was also shown to reproduce the results of MD simulations of NaCl crystals using the PME method.

Kikugawa et al. applied the Fennell method both to globular proteins in explicit water and to membrane proteins immersed in lipid bilayers with explicit water.[19] They obtained results comparable to the PME method with respect to not only the energies and forces but also the dynamics of the radial distribution function for the solvent and conformational dynamics of the proteins. Moreover, they demonstrated that the Fennell method realizes fast MD calculations using the special purpose MD engine MDGrape-3 and is highly scalable in parallel on a PC cluster connected by high speed networks.

Received: June 28, 2010
Published: March 30, 2011

Recently, Fukuda et al. pointed out that there is an inconsistency between forces and energies in the formulation of the Wolf method. They then proposed an improved method, in which the inconsistency is completely removed using the force-switching scheme.[20] Hereafter, we call it the force-switching Wolf (FSw-Wolf) method. They confirmed that the electrostatic energy of the FSw-Wolf method for the sodium and chloride liquid system is comparable to that of the Ewald method and that it produces more exact energy conservation properties than those of the original Wolf method.

In this report, in order to investigate the applicability of the FSw-Wolf method to biological molecular systems, we evaluated the method for an alanine-dimer peptide in explicit water under periodic boundary conditions. We here focus our attention just on a comparison between the FSw-Wolf method and the PME method. This system would not produce any significant artifacts, due to the periodicity applied in the PME, if the configuration sampling is completed. In fact, in studies of similar systems, Villarreal and Montrich recently reported that incomplete sampling of explicit solvent of solvated biomolecular systems is likely to affect the results to a greater extent compared to the artifact induced by the Ewald method.[21] To effectively utilize the sampling data, the free energy is a good measure, since it is strongly affected by the quality of the sampling data, especially in biological systems. Consequently, we show that when the parameters are sufficiently optimized, the energies, forces, and radial distribution functions of the FSw-Wolf method are comparable to the PME method in canonical molecular dynamics simulations. Moreover, we show that the dipole−dipole interactions of water molecules calculated by the FSw-Wolf method are similar to those obtained by the PME method. In order to precisely evaluate the free energy landscape using the FSw-Wolf method, we carried out multicanonical molecular dynamics (McMD) simulations[22−27] for the peptide system, and we compared the results from the PME method with those from the FSw-Wolf method. The McMD simulation is one of the generalized-ensemble methods, and it allows us to investigate the free energy landscape involving rare events separated by the large barriers, which are not attained by canonical MD simulations. Such a reliable estimation of the free energy using the McMD simulation with the Wolf type method has not been ever performed for biomolecular systems.

MATERIAL AND METHODS

Wolf Method. The Wolf method has been described elsewhere.[13] We briefly introduce the algorithm. The electrostatic potential has a long-range nature as it slowly decreases as the inverse of the distance between the charged particles. The effective range is thought to be infinite. Therefore, some particular treatment is required for the integration of an infinite charge distribution. The form of the pure electrostatic potential E_{coloumb} is

$$E_{\text{coulomb}} = \frac{1}{2}\sum_{i=1}^{N}\sum_{j(\neq i)}^{N}\frac{q_i q_j}{|\mathbf{r}_{ij}|} \tag{1}$$

where the summation runs over all charged particles. q_i stands for the charge of the ith atom and $|r_{ij}|$ the distance between the ith and jth atoms.

The cutoff method calculates only interactions between the particles within radius R_c of each other, which creates an artificially charged sphere. This introduces large inaccuracies as indicated by Wolf et al.[13] They proposed a way to eliminate this artifact by introducing a charge neutrality condition within the cutoff sphere and potential damping using the erfc. The total electrostatic energy E_{wolf} can be written as

$$E_{\text{wolf}} = \sum_{i=1}^{N}\sum_{\substack{j(>i)\\|\mathbf{r}_{ij}|<R_C}}\left[\frac{q_i q_j \operatorname{erfc}(\alpha|\mathbf{r}_{ij}|)}{|\mathbf{r}_{ij}|} - \lim_{|\mathbf{r}_{ij}|\to R_c}\left\{\frac{q_i q_j \operatorname{erfc}(\alpha|\mathbf{r}_{ij}|)}{|\mathbf{r}_{ij}|}\right\}\right] - \left(\frac{\operatorname{erfc}(\alpha R_c)}{2R_c} + \frac{\alpha}{\sqrt{\pi}}\right)\sum_{i=1}^{N} q_i^2 \tag{2}$$

where R_c is the cutoff distance, α represents a damping parameter that determines the speed of the convergence of the summation, and N is the number of atoms. The second term on the right-hand of eq 2 is a representation of the charge neutrality regarding the atoms in the cutoff sphere other than the center atom i. The third term represents the contribution by atom i; the last is the self-energy. Using this approach, it is possible to precisely reproduce the Madelung energy.

THE FORCE-SWITCHING WOLF METHOD

Fukuda et al.[20] have pointed out that the negative derivative of the Wolf potential energy is not consistent with the force of the Wolf method, which was proposed in the original paper by Wolf et al.[13] This inconsistency should be a source of serious systematic error when applied to molecular dynamics simulations. Fukuda et al. then modified the Wolf method to fulfill the consistency using the force-switching scheme, and they demonstrated that the FSw-Wolf method satisfied energy conservation much better than the original Wolf method.

The total energy $E_{\text{FSw-Wolf}}$ of the FSw-Wolf method is presented by the following equations,

$$E_{\text{FSw-Wolf}} = \frac{1}{2}\sum_{i=1}^{N}\sum_{j(\neq i)} q_i q_j \tilde{V}(|\mathbf{r}_{ij}|) - \left[\frac{\operatorname{erfc}(\alpha r_1)}{2r_1} - \frac{V^*(r_1)}{2} + \frac{\alpha}{\sqrt{\pi}}\right]\sum_i q_i^2 \tag{1.1}$$

where

$$\tilde{V}(|\mathbf{r}_{ij}|) \equiv \begin{cases} \dfrac{\operatorname{erfc}(\alpha|\mathbf{r}_{ij}|)}{|\mathbf{r}_{ij}|} + V^*(r_1) - \operatorname{erfc}(\alpha r_1)/r_1 & \text{for } 0 < |\mathbf{r}_{ij}| < r \\ V^*(|\mathbf{r}_{ij}|) & \text{for } r_1 \leq |\mathbf{r}_{ij}| \leq r_c \\ 0 & \text{for } r_c < |\mathbf{r}_{ij}| < \infty \end{cases} \tag{1.2}$$

Here, r_1 and r_c stand for the switching length (described below) and the cutoff length, respectively. In the current study, according to a report by Fukuda et al.,[20] we employed $r_c - r_1 = 1$ Å. V^* is determined so that the force smoothness is satisfied in the entire range by the force-switching scheme as follows:

$$\mathbf{f}^{i}_{\text{FSw - Wolf}} = \sum_{j(\neq i)} q_i q_j \tilde{F}(|\mathbf{r}_{ij}|)\frac{\mathbf{r}_{ij}}{|\mathbf{r}_{ij}|} \tag{1.3}$$

where

$$\tilde{F}(|\mathbf{r}_{ij}|) = \begin{cases} F(|\mathbf{r}_{ij}|) & \text{for } 0 < |\mathbf{r}_{ij}| < r_1 \\ F^*(|\mathbf{r}_{ij}|) & \text{for } r_1 \leq |\mathbf{r}_{ij}| \leq r_c \\ 0 & \text{for } r_c < |\mathbf{r}_{ij}| < \infty \end{cases} \tag{1.4}$$

and

$$F(|\mathbf{r}_{ij}|) \equiv \frac{\operatorname{erfc}(\alpha|\mathbf{r}_{ij}|)}{|\mathbf{r}_{ij}|^2} + \frac{2\alpha}{\sqrt{\pi}} \frac{\exp(-\alpha^2|\mathbf{r}_{ij}|^2)}{|\mathbf{r}_{ij}|} \tag{1.5}$$

Here, we set $F^*(|\mathbf{r}_{ij}|)$ to have the following fourth polynomial equation form:[28]

$$F^*(|\mathbf{r}_{ij}|) \equiv \alpha' + \beta|\mathbf{r}_{ij}| + \gamma|\mathbf{r}_{ij}|^2 + \delta|\mathbf{r}_{ij}|^3 \tag{1.6}$$

Thereby, we obtained the values of the coefficients satisfying the smoothness condition at r_1 and r_c as

$$\begin{bmatrix} \alpha' \\ \beta \\ \gamma \\ \delta \end{bmatrix} = \frac{1}{(r_c - r_1)^3} \begin{bmatrix} (-r_c r_1 b + r_c a + r_1^2 b - 3ar_1)r_c^2 \\ (br_c^2 + r_c r_1 b - 2r_1^2 b + 6ar_1)r_c \\ -(2br_c^2 - r_c r_1 b + 3r_c a - r_1^2 b + 3ar_1) \\ br_c - br_1 + 2a \end{bmatrix} \tag{1.7}$$

with $a = F(r_1)$ and $b = \mathrm{d}F(r_1)/\mathrm{d}r$.

Thus, $V^*(|\mathbf{r}_{ij}|)$ is expressed as the following:

$$V^*(|\mathbf{r}_{ij}|) = -\left(\alpha'|\mathbf{r}_{ij}| + \frac{\beta|\mathbf{r}_{ij}|^2}{2} + \frac{\gamma|\mathbf{r}_{ij}|^3}{3} + \frac{\delta|\mathbf{r}_{ij}|^4}{4}\right) + \left(\alpha' r_c + \frac{\beta r_c{}^2}{2} + \frac{\gamma r_c{}^3}{3} + \frac{\delta r_c{}^4}{4}\right) \tag{1.8}$$

In the Appendix, we describe the details of implementation of the FSw-Wolf method for biomolecular system with the potential energies associated with the covalent bonds.

Canonical Molecular Dynamics Simulations. We employed alanine-dimer peptide for the current studies, because the alanine-dimer peptide (Ala−Ala) has been well studied as an ideal standard biological peptide. We capped the alanine-dimer peptide with an acethyl group (Ace) at the N terminus and with an N-methyl group (NMe) at the C terminus to eliminate large electrostatic interactions between the termini. The peptide, Ace−Ala−Ala−NMe, was immersed in a cubic box of water with an edge length of 36 Å and solvated with 1541 water molecules for a total of 4655 atoms in the system. We used the Amber96 force field for the peptide and the TIP3P model for water.[29]

Preparation and equilibration procedures for a production run of the system were as follows. The positions of the modeled hydrogen atoms were adjusted by energy minimization *in vacuo*. The peptide was then immersed in a water box, which had been pre-equilibrated under NPT conditions with a constant number of particles, temperature (300 K), and pressure (1 atm) using the Berendsen thermostat and barostat.[30] The atoms of the peptide were kept fixed while the solvent was allowed to equilibrate for 200 ps under NVT conditions at 300 K. After the solvent equilibration, we reduced the restraining force for the atoms of the peptide and simulated all atoms of the system under NPT conditions at 300 K and 1 atm of pressure. The covalent bonds and angles including polar hydrogens were constrained and treated as rigid bodies, thus allowing for a simulation time step of 1 fs. A production run of 10 ns was done under NVT conditions, using a Hoover−Evans thermostat at 300 K.[31]

In the PME method, the parameter α of the PME method was set to 0.35 Å^{-1}, and the real space cutoff distance was set to 10 Å for all runs. The mesh size was set to 36 × 36 × 36, thus ensuring a grid density of 1 Å for the system with sufficient accuracy from the Ewald method. Moreover, we employed the atom base cutoff scheme for the long-range interactions in both methods. The radius of the neighbor list is 1 Å larger than the cutoff, and it is updated every five steps with a time step of 1 fs for all simulations. The parameters for the FSw-Wolf method are described in detail in the Results and Discussion section.

To evaluate dipole−dipole interactions in a homogeneous polar molecular system, a cubic water system under periodic boundary conditions with an edge length of 36 Å including 1569 water molecules was prepared in addition to the peptide system. We performed canonical ensemble molecular dynamics simulations of the water system at 300 K, using the PME and FSw-Wolf methods. The dipole−dipole interactions were evaluated using the 1 ns trajectories of the water system.

Multicanonical Molecular Dynamics Simulations. We used the force-biased multicanonical molecular dynamics (FBMcMD) simulation method[24] to evaluate the free energy landscape of the peptide. The FBMcMD algorithm has been described elsewhere,[24,26] so here we briefly summarize it. We generated the FBMcMD ensemble by performing constant-temperature MD at an arbitrarily chosen temperature $T_0 = 1/k_B\beta_0$ with force scaling as

$$\frac{\mathrm{d}\mathbf{p}_i}{\mathrm{d}t} = \nu(E)\mathbf{f}_i \tag{3.1}$$

$$\nu(E) = \frac{\partial \alpha_{\mathrm{mc}}(E)}{\partial E} \tag{3.2}$$

where β_0 is the inverse temperature, k_B is the Boltzmann constant, $\mathbf{p}_i$ and $\mathbf{f}_i$ indicate the momentum and the force of the *i*th atom, respectively, $\alpha_{\mathrm{mc}}(E)$ is the multicanonical temperature, and $\nu(E)$ represents the force scaling factor. Since the $\nu(E)$ values have not been given *a priori*, they should be determined by the following iterative scheme:

$$\nu^{k+1}(E) = \nu^k(E) + \frac{1}{\beta_0} \frac{\partial \ln P^k(E)}{\partial E} \tag{4}$$

Here, $P^k(E)$ is the probability distribution of potential energy from *k*th iterative run of the FBMcMD. $\nu(E)$ relates to the density of states $\Omega(E)$ through the following equation:

$$\frac{1}{\Omega(E)} = \mathrm{e}^{-\beta_0 \alpha_{\mathrm{mc}}(E)} \tag{5}$$

Once the $\nu(E)$ has converged, the system can realize a random walk on the potential energy space. We set the temperature for the FBMcMD simulation to 300−700 K so that the peptide can sample various structures at this temperature.

We started FBMcMD simulations from the peptides equilibrated by canonical MD. The same system from the canonical MD was used. We set the reference temperature to 250 K. The FBMcMD simulations for the PME method and that with the modified Wolf method were done in the same manner. We note that only nonbonded electrostatic interactions are different between the two simulations. In both simulations, we then obtained flat energy distributions covering a temperature range from 300 to 700 K. We then executed production runs for each simulation for 2 × 10^7 steps and stored snapshots every 1 ps. All of the simulations were done using the *myPresto* molecular dynamics computing program.[32]

RESULTS AND DISCUSSION

Accuracy of Energies and Forces. In order to evaluate the accuracy of the FSw-Wolf method, we analyzed the trajectories of the energies and forces for the system from the canonical MD simulations as follows.

A total of 1000 snapshots of the atom coordinates, i.e., one every 1 ps, were extracted from the 1 ns trajectory as simulated using the PME method. Then, for each of these structures, we calculated the electrostatic energy using the FSw-Wolf method with four different α values (0.1, 0.12, 0.16, and 0.2 Å^{-1}) and 18 different r_c cutoff distances from 8.0 to 16.5 Å with an interval of 0.5 Å. Thus, the electrostatic energies with 72 different FSw-Wolf method parameters were compared to the corresponding energies from the PME method. The differences between the PME and the FSw-Wolf methods were calculated according to the following criteria:

$$E_{\mathrm{err}}(m) = \frac{|E_{\mathrm{FSw\text{-}Wolf}}(m) - E_{\mathrm{PME}}(m)|}{|E_{\mathrm{PME}}(m)|} \tag{6.1}$$

$$\langle E_{\mathrm{err}} \rangle = \frac{1}{M} \sum_{m=1}^{M} E_{\mathrm{err}}(m) \times 100 \tag{6.2}$$

Here, m is the index of each snapshot extracted from the trajectory and M is the total number of snapshots. A comparison of the force vectors of the two methods was made using eq 7:

$$F_{\mathrm{err}}(m) = \frac{1}{N} \sum_{i=1}^{N} \frac{|\mathbf{f}^{i}_{\mathrm{FSw\text{-}Wolf}}(m) - \mathbf{f}^{i}_{\mathrm{PME}}(m)|}{|\mathbf{f}^{i}_{\mathrm{PME}}(m)|} \tag{7.1}$$

$$\langle F_{\mathrm{err}} \rangle = \frac{1}{M} \sum_{m=1}^{M} F_{\mathrm{err}}(m) \times 100 \tag{7.2}$$

Here, $\mathbf{f}_{\mathrm{FSw\text{-}Wolf}}{}^{i}$ and $\mathbf{f}_{\mathrm{PME}}{}^{i}$ are the electrostatic force vectors acting on each atom using the FSw-Wolf method and the PME method, respectively. $F_{\mathrm{err}}(m)$ is the relative amplitude of the force vector difference at the mth snapshot structure, and $\langle F_{\mathrm{err}} \rangle$ is the average difference over all of the snapshots.

In Figure 1, we show the energy errors between the PME method and the FSw-Wolf method with several α parameters as a function of the cutoff length. We can see from the figure that the energy errors decrease with increasing cutoff length. The error with α = 0.05 Å^{-1} was large in the whole range of cutoff length. In the case where the cutoff length is greater than 12 Å, the error was significantly reduced when α was larger than 0.10 Å^{-1}. The error was as small as 0.1% with α = 0.12 Å^{-1} and a cutoff length of 16.5 Å.

For comparison with the other cutoff method, we estimated the error from the reaction field (RF) method, which has been useful in many extensive studies. The dashed line in Figure 1 shows the energy error obtained from the RF method[2] with $\varepsilon = \infty$. The error gradually decreases as the cutoff length increases. When we choose a value greater than 0.15 for the α parameter, the error of the FSw-Wolf method is less than that of the RF method for a wide range of cutoff lengths, in this peptide system.

In Figure 2, we show the force error between the PME method and the FSw-Wolf method with various α values as a function of the cutoff length. Although the force errors decreased with increasing cutoff length, the error ratios were 2 to 10 times larger than the energy error ratios. In Figure 2, the force error was about 4% at a cutoff length of 16.5 Å with α = 0.05 Å^{-1}. As α increased, the observed force error decreased, and it was as small as 0.82%, at a cutoff length of 16.5 Å with α = 0.12 Å^{-1}.

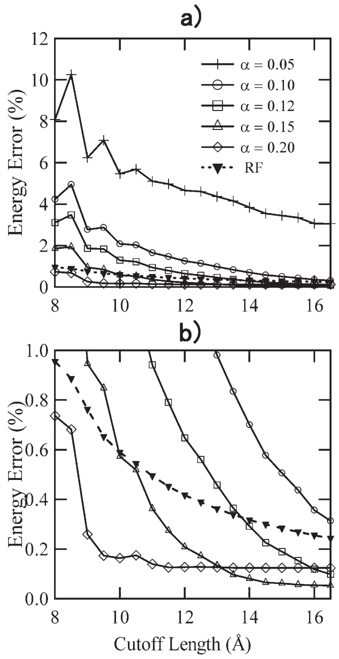


Figure 1. (a) Energy errors of the FSw-Wolf method with various α values from the PME method and the error of the reaction field (RF) method ($\varepsilon = \infty$). (b) Magnified view of the energy error at the small error region.

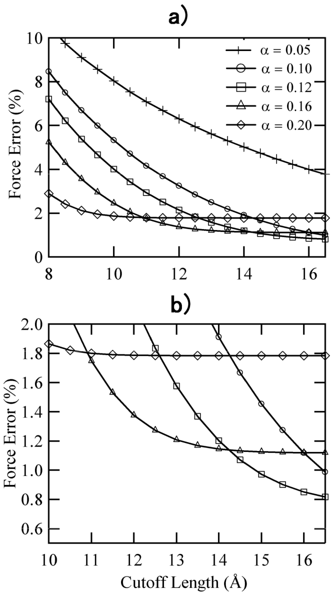


Figure 2. (a) Force errors of the FSw-Wolf method with various α values from the PME method. (b) Magnified view of the force error at the small error region.

The origin of the errors in energies and forces should be due to omission of the reciprocal space contribution of the PME method in the FSw-Wolf method. The reciprocal contributions are reduced by taking a larger cutoff length than the PME method.

As expected, the energy accuracy of the FSw-Wolf method is in good accordance with the PME method. On the basis of results from the canonical MD simulations, respecting the force accuracy, we found that a parameter set consisting of α equal to 0.12 Å^{-1} and a cutoff length of 16.5 Å is optimal for the FSw-Wolf method dealing

Table 1. Timings of the FSw-Wolf (FSWW) and PME Methods with Respect to the Number of CPU in Parallel Computations

	number of CPU				
method	8	16	32	64	128
	single step calculation time (sec)				
PME	0.87	0.61	0.54	0.48	0.51
FSWW (16.5 Å cutoff)	3.57	1.90	1.09	0.66	0.46
FSWW (12.0 Å cutoff)	1.60	0.93	0.59	0.42	0.34

with the peptide system. In addition to the optimal parameter set, we employed another parameter set consisting of α equal to 0.16 Å^{-1} and a cutoff length of 12 Å to examine the features of the smaller cutoff length relative to the optimized parameters. This parameter set yielded 0.2% and 1.37% errors for energy and force, respectively, from those by the PME method. Those parameters were used in the computations hereafter.

Parallel Timings. We investigated the parallel timing in the FSw-Wolf method, because good scalability is generally expected for the simple truncation method. To confirm that the FSw-Wolf method has this advantage, we estimated the timing of the methods for a water system, as a benchmark. We prepared a periodic boundary cubic system including 29 662 pure water molecules with an edge length of 98 Å. For the FSw-Wolf method, cutoff lengths of 16.5 Å and 12 Å were used, with the neighbor list 1 Å larger than the cutoff length. For the PME method, we set a 128 × 128 × 128 fine mesh for the reciprocal space and an 8 Å cutoff for the real space. The calculations were performed on a conventional PC cluster connected by a normal switching-hub with the program *myPresto*, which we previously developed.[32] We used MPI and slab decomposition for FFT parallelization in the PME method. The results are displayed in Table 1.

The table shows that the timing of the FSw-Wolf method with 16.5 Å and that with 12.0 Å cutoffs are faster than the timing of the PME method over 128 CPUs and that over 64 CPUs, respectively. Similar results for PME were also obtained in ref 33, in which a parallel simulation of dihydrofolate reductase, with 23 558 atoms, using a conventional PC cluster with 64 × 64 × 64 mesh and the NAMD program does not show a reduction in the calculation time over about 200 CPUs. Although special machines with optimized networks may show good PME scalability, the FSw-Wolf calculations are faster than the PME calculations with a fine mesh, performed on a conventional PC cluster with normal network connections, even if the system has 88 986 atoms.

The recent technology of the volumetric decomposition has shown to exhibit better scalability than that of the slab decomposition for more than 256 CPUs, when a special high-speed network is applied.[12] On the contrary, when an ordinary simple network is used, the slab decomposition shows good scalability for the smaller number of CPUs than the maximum number of the grid size.

Radial Distribution Function. We have calculated radial distributions between water and the peptide as a function of the radial distance. We calculated the radial distribution of density $g(r)$ using eq 8:

$$g(r) = \frac{\langle N(r)_{\delta r} \rangle}{\frac{4}{3}\pi[(r+\delta r)^3 - r^3]} \tag{8}$$

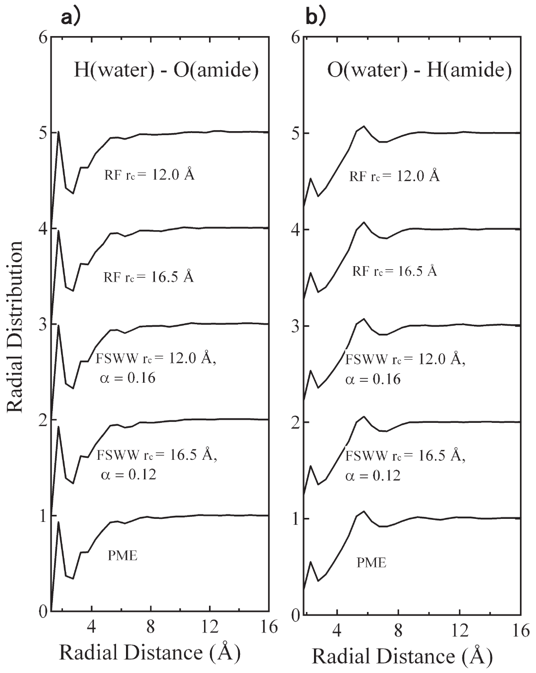


Figure 3. Radial distribution of (a) water hydrogen and peptide oxygen and (b) water oxygen and hydrogen in the peptide bond as a function of radial distance from the PME method and the FSw-Wolf methods (FSWW) and Reaction Field (RF) methods. Except for PME, the base lines are shifted for comparison. The distributions are normalized to unity at the bulk region of solvent.

where $N(r)_{\delta r}$ is the number of atoms of water between r and $r + \delta r$, and the angular brackets denote the ensemble average.

In Figure 3, we show the radial distribution function (RDF) of water oxygen and the amide hydrogen and that of water hydrogen and the peptide oxygen for the three methods: the PME method, the FSw-Wolf method using a cutoff of 16.5 Å with α = 0.12 Å^{-1} and that using a cutoff of 12 Å with α = 0.16 Å^{-1}, and RF methods with cutoffs 12.0 and 16.5 Å. All of the results are almost identical, indicating that the FSw-Wolf and RF methods reproduce the RDFs by PME, even if a cutoff value of 12 Å is used.

Dipole–Dipole Interactions. It has been pointed out that the simple truncation method suffers from either sensitivity or an artifact regarding the dielectric properties, especially for the solvent material.[34] To examine whether this disadvantage also appeared in the FSw-Wolf method, we evaluated the distance-dependent Kirkwood factor using both the PME and FSw-Wolf methods, for a cubic system with an edge length of 36 Å, including 1569 water molecules, at 300 K. The distance-dependent Kirkwood factor $G(r)$, as a function of distance r, describes the angular correlation of the molecules, which have permanent dipole moments, as

$$G(r) = \frac{1}{N}\left\langle \sum_i \left(\frac{\mu_i \sum\limits_{j,\, r_{ij} < r} \mu_j}{|\mu|^2} \right) \right\rangle \tag{9}$$

where r_{ij} is the distance between the ith and jth molecules, N is the number of molecules, and μ_i and μ_j are the dipole moments of the ith and jth molecules, respectively. <...> denotes the time ensemble average. In Figure 4, the distance-dependent Kirkwood

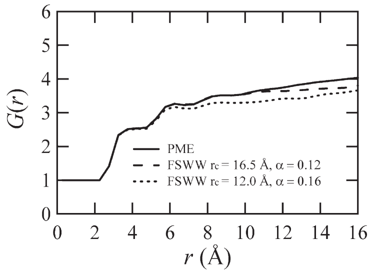


Figure 4. The distance-dependent Kirkwood factor from the PME method (solid line) and the FSw-Wolf (FSWW) methods with r_c = 16.5 Å (dotted line) and 12.0 Å (dashed line).

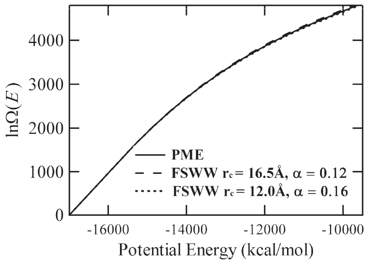


Figure 5. Logarithm of the densities of states $\Omega(E)$ of the peptide system from the PME method and the FSw-Wolf methods (FSWW).

factor, $G(r)$, of the FSw-Wolf method with r_c = 16.5 Å is very similar to that of the PME method, and that with r_c = 12.0 Å is slightly different at the larger distance. The dielectric constants calculated from the distance-independent Kirkwood factor[35,36] are 86.8 (PME), 82.2 (FSw-Wolf method with r_c = 16.5 Å), and 81.1 (FSw-Wolf method with r_c = 12.0 Å). These values are comparable with the results obtained in recent studies,[35,36] although $G(r)$ is very sensitive to the system size and the computation conditions. It is noted that the Kirkwood factor, by means of cutoff methods, is often very different from the PME method, even when much larger cutoff distances are employed.[35] The Kirkwood factors were also reported to deviate with the simple RF method from those with the PME method,[37,38] or to be similar to those with the improved RF method depending on the simulation conditions.[2] Thus, the current FSw-Wolf method with the appropriate parameters should give not only accurate absolute energy values but also reliable dynamic properties.

Free Energy Landscapes of the Alanine-Dimer Peptide. We carried out three different McMD simulations, the first one using the PME method, the second one using the FSw-Wolf methods employing the optimized parameters described above, and the third one employing the shorter cutoff length of 12 Å with α = 0.16 $Å^{-1}$. We obtained flat energy distributions and converged densities of states $\Omega(E)$ for all of the cases. The three densities of state are shown in Figure 5.

In this figure, the three densities of state coincide with each other, indicating that the three systems are thermodynamically very similar.

We have taken the trajectories from the McMD simulations. The length of each trajectory was 20 ns. We then plotted the free energy landscape of the peptide reweighted at 300 K with respect to the distance between the carbon atom of the C-terminal methyl and the carbon atom of the N-terminal methyl in the peptide, as shown in Figure 6.

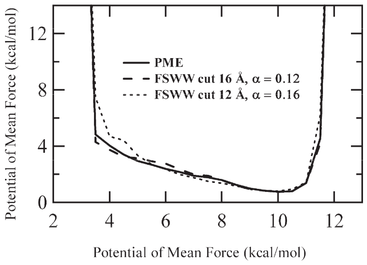


Figure 6. Potentials of mean force at 300 K as a function of the distance between the carbon atom of the C-terminal methyl and the carbon atom of the N-terminal methyl in the peptide from the PME method and the FSw-Wolf methods (FSWW).

From the figure, it is apparent that the peptides are distributed widely from the extended to the twisted conformations. Furthermore, we see that the free energy landscapes from the PME method and from the two FSw-Wolf methods are almost identical. It suggests that the conformational outlines of the peptide obtained from the two methods bear a very close resemblance.

The backbone dihedral angles of the peptide well describe the conformations. So far, there are experimental and theoretical conformational studies of short peptides, and they employed the dihedral angles to evaluate the conformational space.[39,40] To compare the conformational free energy landscapes from both methods, we used the potential of mean force with respect to dihedral angles of the peptide. We employed a pair of the dihedral angles, say, (ϕ_1, ψ_1) and (ϕ_2, ψ_2). Here, ϕ_1 and ψ_1 are first alanine backbone dihedral angles, and ϕ_2 and ψ_2 are the second ones. We show the free energy landscapes of the dihedral angles reweighted to 300 K in Figures 7−9.

In the lower panels of Figures 8 and 9, we depict the absolute free energy difference between the FSw-Wolf method and the PME method. The white regions of the figures indicate that the differences are smaller than 1 kcal/mol. As shown in the figures, the free energy landscapes have a strong resemblance to each other. We see that free energy local minimums, major basins, of C_5^{ext} in the vicinity of (ϕ = −150°, ψ = 150°), P_{II} in (−70°, 140°), and α_R' in (−60°, −60°) in the PME method are well reproduced in the FSw-Wolf method. On the contrary, slightly different distributions of the free energies in the FSw-Wolf method from the PME methods are found around the shallow basins, the free energies of which are much higher than those of the major basins. However, such differences in the free energy landscapes do not significantly contribute to the conformational distributions.

Consequently, from a comparison between the FSw-Wolf method and the PME in terms of the energies, forces, radial distribution functions, dipole−dipole interactions, and the free energy landscapes of the alanine-dimer peptide in explicit water, it is concluded that the FSw-Wolf method with the optimized parameters or even with the shorter cutoff length of 12 Å with α equal to 0.16 $Å^{-1}$ yields very similar results to the PME method. This fact shows that the criteria regarding the choice of the parameter values with respect to the force error in the FSw-Wolf method surely ensures a good reliability of the method. In addition, we can also get similar results along with the computational efficiency even if we loosen the criteria.

In order to avoid the discontinuity at the cutoff length r_c for the force function originally derived from the Wolf method, we used a switching force function instead of the shifted force approach as proposed in ref 18. Since these two approaches deform the original force function, leading to the deformation of the original

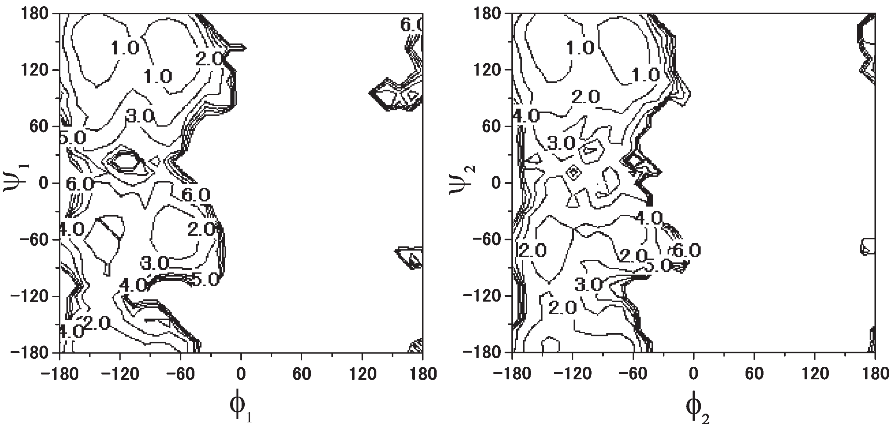


Figure 7. The free energy landscapes of the couples of dihedral angles of the peptide reweighted to 300 K using the PME method. The left-hand side and right-hand side display (ϕ_1, ψ_1) and (ϕ_2, ψ_2), respectively. Units are kcal/mol.

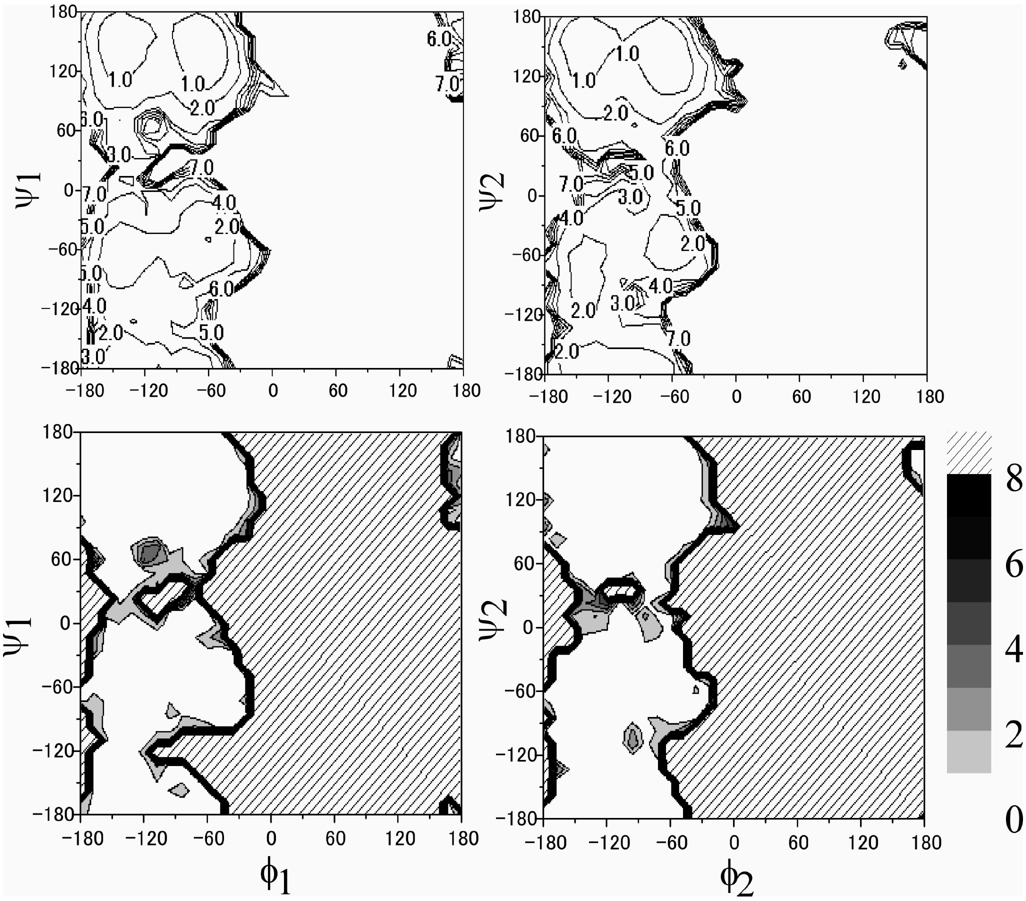


Figure 8. Upper panel: The free energy landscapes of the couples of the dihedral angles of the peptide reweighted to 300 K using the FSw-Wolf method with the optimized parameter set of cutoff length 16.5 Å with α = 0.12 $Å^{-1}$. Lower panel: The absolute value of the free energy difference from the PME method. Left-hand side and right-hand side display (ϕ_1, ψ_1) and (ϕ_2, ψ_2), respectively. The area illustrated by the oblique lines shows a difference larger than 8 kcal/mol. Units are kcal/mol.

potential function, they both need a correction of the total energy. As described in ref 20, this correction can be successfully done in the switching force approach, while it is nontrivial in the shifted force approach. In fact, in the latter, the correction depends on the potential parameters α and r_c, and also on $|r_{ij}| - r_c$, namely, the particle configuration. Thus, it should keep away from a simple correction.[41] In addition, in contrast to the shifted force approach, the current switching force approach can easily induce smoothness of the force function, which is required in the stable numerical integration in MD.[28]

A previous study in ref 20 showed that the electrostatic energy by means of the FSw-Wolf method of a highly charged system,

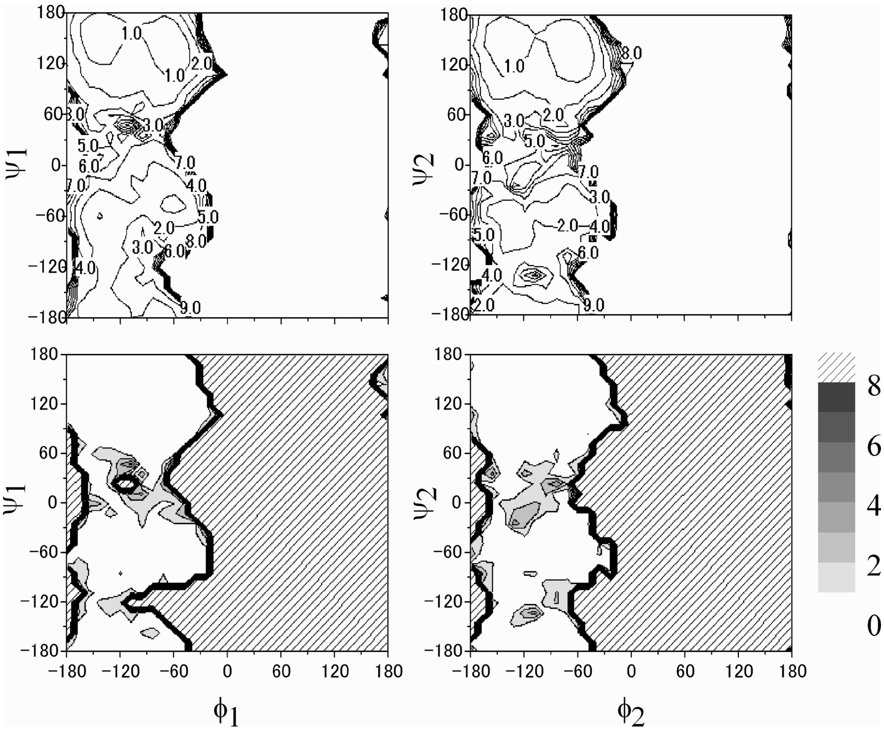


Figure 9. Upper panel: The free energy landscapes of the couples of dihedral angles of the peptide reweighted to 300 K using the FSw-Wolf method with the parameter set of cutoff length 12 Å with α = 0.16 Å^{-1}. Lower panel: The absolute value of free energy difference from the PME method. Left-hand side and right-hand side display (ϕ_1, ψ_1) and (ϕ_2, ψ_2), respectively. The area illustrated by the oblique lines shows a difference larger than 8 kcal/mol. Units are kcal/mol.

sodium chloride, was in accordance with that of the Ewald method. High statistical simulation studies on the charged protein system, using the FSw-Wolf method, are in progress.

CONCLUSION

Accurate and rapid computations for electrostatic interactions are of significant importance. Although the Ewald-like method is generally regarded as the standard for calculations of condensed matters' MD and MC simulations under periodic boundary conditions, many alternative methods have been proposed in order to overcome the problem of intrinsic artifacts and the lack of scalability on highly parallel computations. Utility of real space approximations for the Ewald summation is that they are free from the all-to-all networking procedure on the frequency-space calculation, which severely prevents a good scalability with respect to the number of particles, especially in a large system. In this respect, the original Wolf method[13] and the FSw-Wolf method proposed by Fukuda and co-workers[20] as well as other alternatives are expected to be promising alternatives to the Ewald method for treating large systems with highly parallel computations, which have been one of the major trends in computational science, such as in multiscale physics studies. In fact, as demonstrated in one of our previous works,[19] these Wolf type methods have significant advantages, regarding the scalability and the highly parallel calculations on massive PC clusters.

Among the variants of the Wolf method having these advantageous features regarding the computational cost, the FSw-Wolf method presents a consistent MD scheme with a successful energy correction and a sufficient smoothness of the energy function.[20] In this work, we have thus evaluated a physical reliability of the FSw-Wolf MD method in a biological system, using an alanine-dimer peptide in explicit water. Compared with the PME method using the canonical MD simulations, we first determined an optimized set of the parameters for the FSw-Wolf method. The optimized parameters and the other parameters employing a smaller cutoff length of 12 Å yielded comparable energies, forces, and radial distributions to those of the PME method. Then, we carried out force-biased multicanonical molecular dynamics simulations using the two parameter sets to evaluate the free energy landscape of the alanine-dimer peptide with respect to the conformational space. The results at 300 K obtained from the multicanonical molecular dynamics were in accordance with those of the PME method.

By using the canonical MD simulation and the McMD simulation, which is highly reliable, especially in view of the free energy estimation, we consequently show that the FSw-Wolf method is a promising alternative to the PME method with respect to physical accuracy. Advantageous features regarding the computational cost for the Wolf and the variant methods remain in the FSw-Wolf method. Thus, we believe that the FSw-Wolf method should be very useful for simulating large biological systems, in particular, for highly parallel computations.

APPENDIX

We here describe how the total electrostatic energy in the system that contains covalent bond interactions is evaluated in the FSw-Wolf method. The total energy of such a system is represented as

$$E^{\text{total}} = \frac{1}{2}\sum_{i=1}^{N}\sum_{j(\neq i)}^{N}\frac{q_iq_j}{r_{ij}} - \frac{1}{2}\sum_{i=1}^{N}\sum_{j\in N_{ii}^{\text{CB}}}\frac{q_iq_j}{r_{ij}} \qquad \text{(A1)}$$

where $j \in N_{ii}^{\mathrm{CB}}$ means that atoms i and $j(\neq i)$ are submitted to covalent bond interactions, e.g., bond, angle, torsion interactions for "1−2, 1−3, and 1−4 pairs". By using eq 1.1 for the first term in eq A1, we get

$$E^{\mathrm{total}} \sim \frac{1}{2}\sum_{i=1}^{N}\sum_{j(\neq i)}^{N} q_i q_j \tilde{V}(r_{ij}) - \left[\frac{1}{2}\frac{\mathrm{erfc}(\alpha r_1)}{r_1} - \frac{1}{2}V^*(r_1) + \frac{\alpha}{\sqrt{\pi}}\right]\sum_{i=1}^{N} q_i^2 - \frac{1}{2}\sum_{i=1}^{N}\sum_{j \in N_{ii}^{\mathrm{CB}}} \frac{q_i q_j}{r_{ij}} \tag{A2}$$

and thus

$$E^{\mathrm{total}} \sim \frac{1}{2}\sum_{i=1}^{N}\sum_{j \notin N_{ii}^{\mathrm{CB}} \cup \{i\}}^{N} q_i q_j \tilde{V}(r_{ij}) - \left[\frac{1}{2}\frac{\mathrm{erfc}(\alpha r_1)}{r_1} - \frac{1}{2}V^*(r_1) + \frac{\alpha}{\sqrt{\pi}}\right]\sum_{i=1}^{N} q_i^2 + \frac{1}{2}\sum_{i=1}^{N}\sum_{j \in N_{ii}^{\mathrm{CB}}} q_i q_j \left[\tilde{V}(r_{ij}) - \frac{1}{r_{ij}}\right] \tag{A3}$$

where $j \notin N_{ii}^{\mathrm{CB}} \cup \{i\}$ indicates that atoms i and j are submitted to only nonbonded interactions. It may be convenient for the implementation to use the following deformation in the last term of eq A3, viz.,

$$\tilde{V}(r_{ij}) - \frac{1}{r_{ij}} \sim -\frac{\mathrm{erf}(\alpha r_{ij})}{r_{ij}} - \frac{\mathrm{erfc}(\alpha r_1)}{r_1} + V^*(r_1) \tag{A4}$$

where we have used eq 1.2 and the assumption that r_1 is sufficiently large, i.e., $r_1 > r_{ij}$ for $j \in N_{ii}^{\mathrm{CB}}$. According to eq A4, we have

$$E^{\mathrm{total}} \sim \frac{1}{2}\sum_{i=1}^{N}\sum_{j \notin N_{ii}^{\mathrm{CB}} \cup \{i\}}^{N} q_i q_j \tilde{V}(r_{ij}) - \frac{1}{2}\sum_{i=1}^{N}\sum_{j \in N_{ii}^{\mathrm{CB}}}^{N} \frac{q_i q_j}{r_{ij}}\mathrm{erf}(\alpha r_{ij}) - \left[\frac{1}{2}\frac{\mathrm{erfc}(\alpha r_1)}{r_1} - \frac{1}{2}V^*(r_1)\right]\left\{\sum_{i=1}^{N} q_i\left(q_i + \sum_{j \in N_{ii}^{\mathrm{CB}}}^{N} q_j\right)\right\} - \frac{\alpha}{\sqrt{\pi}}\sum_{i=1}^{N} q_i^2 \tag{A5}$$

and so the force is

$$\mathbf{f}^{\mathrm{total}i} \sim \sum_{j \notin N_{ii}^{\mathrm{CB}} \cup \{i\}}^{N} q_i q_j \tilde{F}(r_{ij})\frac{\mathbf{r}_{ij}}{r_{ij}} - \sum_{j \in N_{ii}^{\mathrm{CB}}}^{N} \frac{q_i q_j}{{r_{ij}}^2}\left[\frac{2\alpha r_{ij}}{\sqrt{\pi}}e^{-(\alpha r_{ij})^2} - \mathrm{erf}(\alpha r_{ij})\right]\frac{\mathbf{r}_{ij}}{r_{ij}} \tag{A6}$$

AUTHOR INFORMATION

Corresponding Author

*E-mail: yasuyon33@protein.osaka-u.ac.jp.

ACKNOWLEDGMENT

We thank Dr. Daron M. Standley for the helpful comments and careful reading of the manuscript. This research was supported by Research and Development of the Next-Generation Integrated Simulation of Living Matter, a part of the Development and Use of the Next-Generation Supercomputer Project of the Ministry of Education, Culture, Sports, Science and Technology (MEXT).

REFERENCES

(1) Schreiber, H.; Steinhauser, O. *Biochemistry* **1992**, *31*, 5856–5860.

(2) Schulz, R.; Lindner, B.; Petridis, L.; Smith, J. C. *J. Chem. Theory Comput.* **2009**, *5*, 2798–2808.

(3) Ding, H.; Karasawa, N.; Goddard, W. A., III. *J. Chem. Phys.* **1992**, *97*, 4309.

(4) Board, J. A., Jr.; Causey, J. W.; Leathrum, J. F., Jr.; Windemuth, A.; Schulten, K. *Chem. Phys. Lett.* **1992**, *198*, 89.

(5) Steinbach, P. J.; Brooks, B. R., Jr. *J. Comput. Chem.* **1994**, *15*, 667.

(6) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577–8593.

(7) Darden, T.; York, D.; Pederson, L. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

(8) Hunenberger, P. H.; McCammon, J. A. *J. Chem. Phys.* **1999**, *110* (4), 1856–1872.

(9) Smith, P. E.; Pettitt, B. M. *J. Chem. Phys.* **1996**, *105* (10), 4289–4293.

(10) Figueirido, F.; Del Buono, G. S.; Levy, M. *J. Chem. Phys.* **1995**, *103* (14), 6133–6142.

(11) Kia, A.; Kim, D.; Darve, E. *J. Comput. Phys.* **2008**, *227*, 8551–8567.

(12) Oh, K. J.; Deng, Y. *Comput. Phys. Commun.* **2007**, *177*, 426–431.

(13) Wolf, D.; Keblinski, P.; Phillpot, S. R.; Eggebrecht, J. *J. Chem. Phys.* **1999**, *110*, 8254.

(14) Demontis, P.; Spanu, S.; Suffritti, G. B. *J. Chem. Phys.* **2001**, *114*, 7980.

(15) Zahn, D.; Schilling, B.; Kast, S. M. *J. Phys. Chem. B* **2002**, *106*, 10725.

(16) Ma, Y.; Garofalini, S. H. *J. Chem. Phys.* **2005**, *122*, 094508.

(17) Avendaño, C.; Gil-Villegas, A. *Mol. Phys.* **2006**, *104* (9), 1475.

(18) Fennell, C. J.; Gezelter, J. D. *J. Chem. Phys.* **2006**, *124*, 234104.

(19) Kikugawa, G.; Apostolov, R.; Kamiya, N.; Taiji, M.; Himeno, R.; Nakamura, H.; Yonezawa, Y. *J. Comput. Chem.* **2009**, *30* (1), 110–8.

(20) Fukuda, I.; Yonezawa, Y.; Nakamura, H. *J. Phys. Soc. Jpn.* **2008**, *77* (11), 114301–114305.

(21) Villarreal, M. A.; Montich, G. G. *J. Biomol. Struct. Dyn.* **2005**, *23*, 135–142.

(22) Hansmann, U. H. E.; Okamoto, Y.; Eisenmenger, F. *Chem. Phys. Lett.* **1996**, *259* (3−4, 6), 321−330.

(23) Nakajima, N.; Nakamura, H.; Kidera, A. *J. Phys.Chem. B* **1997**, *101*, 817–824.

(24) Kim, J. G.; Fukunishi, Y.; Kidera, A.; Nakamura, H. *Phys. Rev. E* **2003**, *68*, 21110.

(25) Nakajima, N.; Higo, J.; Kidera, A.; Nakamura, H. *J. Mol. Biol.* **2000**, *296*, 197–216.

(26) Watanabe, Y.; Kim, J. G.; Fukunishi, Y.; Nakamura, H. *Chem. Phys. Lett.* **2004**, *400*, 258–263.

(27) Kim, J. G.; Fukunishi, Y.; Kidera, A.; Nakamura, H. *Phys. Rev. E* **2004**, *70*, 57103.

(28) Queyroy, S.; Nakamura, H.; Fukuda, I. *J. Comput. Chem.* **2009**, *30*, 1799–1815.

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

(30) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.

(31) Hoover, W. G.; Ladd, A. J. C.; Moran, B. *Phys. Rev. Lett.* **1982**, *48*, 1818–1820. Evans, D. J. *J. Chem. Phys.* **1983**, *78*, 3297.

(32) Fukunishi, Y.; Mikami, Y.; Nakamura, H. *J. Phys. Chem. B* **2003**, *107*, 13201.

(33) Bowers, K. J.; Chow, E.; Xu, H.; Dror, R. O.; Eastwood, M. P.; Gregersen, B. A.; Klepeis, J. L.; Kolossváry, I.; Moraes, M. A.; Sacerdoti, F. D.; Salmon, J. K.; Shan, Y.; Shaw, D. E. Proceedings of the ACM/IEEE Conference on Supercomputing (SC06), Tampa, Florida, November 11−17, 2006.

(34) Alper, H. E.; Levy, R. M. *J. Chem. Phys.* **1989**, *91* (2), 1242–1251.

(35) Yonetani, Y. *J. Chem. Phys.* **2006**, *124*, 204501.

(36) Takahashi, K.; Narumi, T.; Yasuoka, K. *J. Chem. Phys.* **2010**, *133*, 014109.

(37) Belhadj, M.; Alper, H. E.; Levy, R. M. *Chem. Phys. Lett.* **1991**, *179*, 13–20.

(38) van der Spoel, D.; van Maaren, P. J. *J. Chem. Theory Compt.* **2006**, *2*, 1–11.

(39) Christoph, F. W.; Weisshaar, J. C. *J. Phys. Chem. B* **2003**, *107*, 3265–3277.

(40) Vargas, R.; Garza, J.; Hay, B. P.; Dixon, D. A. *J. Phys. Chem. A* **2002**, *106*, 3213–3218.

(41) Leach, A. R. *Molecular Modelling: Principles and Applications*, 2nd ed.; Prentice Hall: New York, 2001.

Quantum Chemical Modeling of Enzymatic Reactions: The Case of Decarboxylation

Rong-Zhen Liao,[†,‡] Jian-Guo Yu,[‡] and Fahmi Himo[*,†]

[†]Department of Organic Chemistry, Arrhenius Laboratory, Stockholm University, SE-10691 Stockholm, Sweden
[‡]College of Chemistry, Beijing Normal University, Beijing, 100875, People's Republic of China

Supporting Information

ABSTRACT: We present a systematic study of the decarboxylation step of the enzyme aspartate decarboxylase with the purpose of assessing the quantum chemical cluster approach for modeling this important class of decarboxylase enzymes. Active site models ranging in size from 27 to 220 atoms are designed, and the barrier and reaction energy of this step are evaluated. To model the enzyme surrounding, homogeneous polarizable medium techniques are used with several dielectric constants. The main conclusion is that when the active site model reaches a certain size, the solvation effects from the surroundings saturate. Similar results have previously been obtained from systematic studies of other classes of enzymes, suggesting that they are of a quite general nature.

I. INTRODUCTION

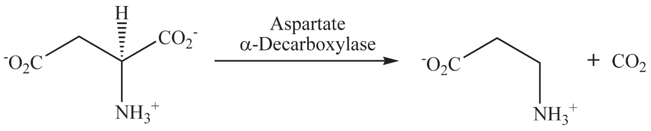
The use of quantum chemical models of enzyme active sites has proven very powerful in the study of both enzyme reaction mechanisms and various active site properties.[1] The philosophy of the approach, commonly called the cluster approach, is to cut out a rather limited part of the enzyme active site, a cluster, and use accurate electronic structure methods to calculate geometries, energies, and other properties. The electronic structure method of choice has been density functional theory (DFT), in particular, the B3LYP hybrid functional.[2]

In the cluster approach, two procedures are commonly used to compensate for the fact that a large part of the enzyme is not treated explicitly. To account for possible steric effects exerted by the enzyme surroundings on the cluster, a number of centers, typically where truncations are made, are kept fixed in the geometry optimizations. This procedure is necessary to prevent unrealistic movements of the various groups of the active site.

To account for electrostatic polarization effects, dielectric cavity techniques are usually used. The surrounding enzyme is assumed to be a homogeneous polarizable continuum with some dielectric constant ε. The choice of this dielectric constant is somewhat arbitrary and has been a matter of discussion, but usually $\varepsilon = 4$ is used.

The combination of these two approximations has been shown to be a quite robust protocol that indeed is sufficient to elucidate reaction mechanisms, distinguish between different mechanistic scenarios, and analyze the roles of various parts in the active site. Ten years ago, typical cluster models consisted of ca. 50 atoms, while today 150 atom models are quite common. Consequently, the scope of applications has been broadened considerably. Through the large number of applications in recent years, it has been demonstrated that the approach has a wide applicability, as essentially all classes of enzymes have been modeled quite successfully.[3]

Scheme 1. Reaction Catalyzed by AspDC



It is easy to realize that as the size of the model grows, a better description of the active site is achieved, and both the coordinate-locking scheme and the implicit solvation model will work better and better because the model will be more flexible and more of the polarization effects will be already explicitly included in the cluster model. The question has been how large a model one needs to use before the effects saturate. Saturation of solvation effects in this sense means that the addition of these does not influence the energy profile of the reaction under investigation; i.e., the relative energies are the same with and without the inclusion of implicit solvation. At that point, the exact choice of the dielectric constant becomes an irrelevant issue.

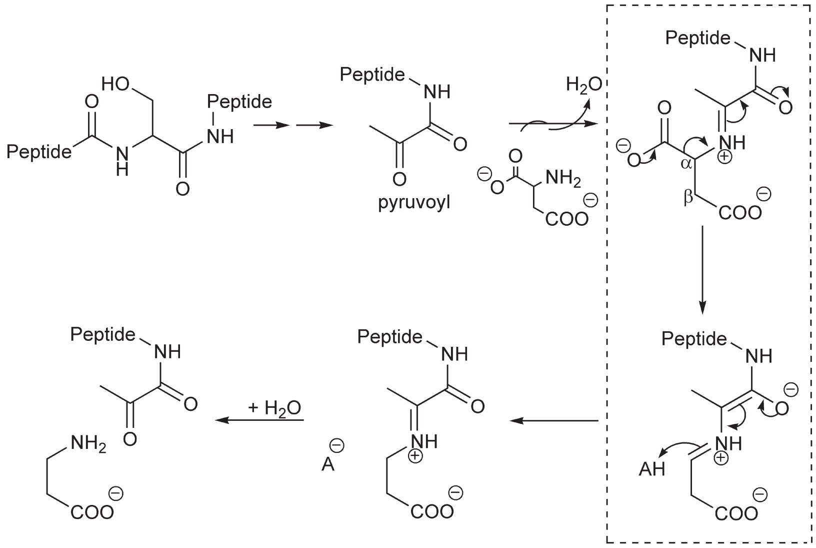
Recently, by performing systematic studies in which the size of the cluster model was gradually increased, we have shown that saturation of the solvation effect happens surprisingly fast, at a model size of less than 200 atoms. This has been demonstrated for three cases that potentially could be problematic for the cluster approach, namely, (a) the formation of an ion pair in the reaction of 4-oxalocrotonate tautomerase (4-OT),[4] (b) the release of a chloride ion in the reaction of haloalcohol dehalogenase (HheC),[5] and (c) the transfer of a methyl cation in the reaction of histone lysine methyltransferase (HKMT).[6]

In the present paper, we further examine the scope of this cluster methodology by studying another important class of enzymes, namely, decarboxylases. Here, from being an anionic moiety (R−COO), CO_2 is released as a neutral gas, which could provide additional challenges where the cluster approach needs to be evaluated.

Received: October 13, 2010
Published: March 29, 2011

dx.doi.org/10.1021/ct200031t | *J. Chem. Theory Comput.* 2011, 7, 1494–1501

Scheme 2. Suggested Reaction Mechanism of AspDC[a]



[a] The frame indicates the C−C cleavage step studied in the present work.

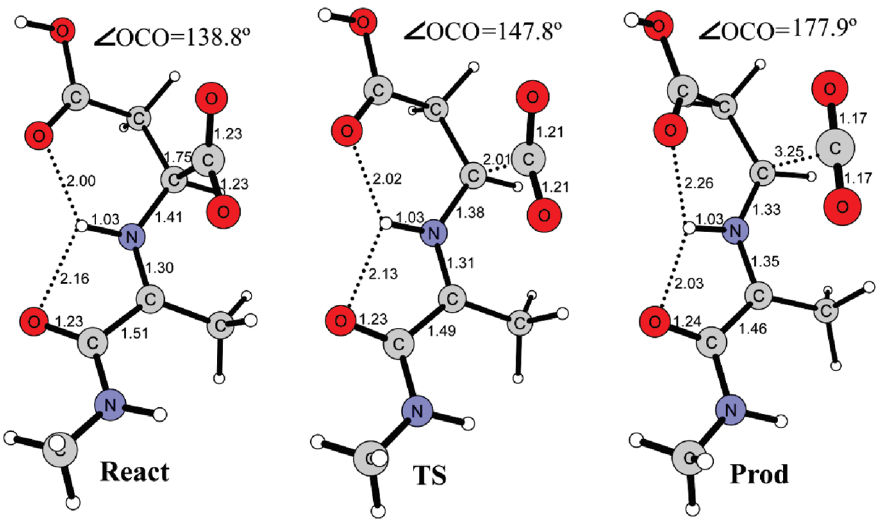


Figure 1. Optimized structures of the reactant, transition state, and product for model 0. Distances in Ångstroms. The carboxylate ∠OCO angle is indicated.

A number of enzymes catalyzing carboxylation/decarboxylation reactions have previously been studied both with the cluster approach[7] and with QM/MM methodologies.[8] The specific enzyme considered in the present work is L-aspartate α-decarboxylase (AspDC), which catalyzes the decarboxylation of L-aspartate to β-alanine (Scheme 1).[9] This reaction is essential for the biosynthesis of pantothenate (vitamin B5) and coenzyme A in bacteria.[10]

AspDC belongs to the class of decarboxylases that utilize a covalently bound pyruvoyl group that is generated through an autoproteolytic cleavage reaction.[9c,11] The suggested reaction mechanism for AspDC is given in Scheme 2.[12] It involves an initial iminium formation step, followed by a C−C bond cleavage step which releases the carbon dioxide. Protonation and hydrolysis steps complete the reaction to give the final product and regenerate the pyruvoyl cofactor.

Because we are, in the present paper, only interested in the methodological issues, we will assume that this mechanism is correct and will focus only on the key C−C cleavage step. Several models of the AspDC active site are systematically devised to investigate how the reaction energetics and solvation effects change with the model size.

II. COMPUTATIONAL DETAILS

All calculations presented herein were performed using the density functional theory method B3LYP as implemented in Gaussian 03.[13] For geometry optimizations, the 6-31G(d,p) basis set was used. In order to obtain more accurate energies, single-point calculations based on the optimized geometries were done using the 6-311+G(2d,2p) basis set. Solvation effects were calculated at the same level as the geometry optimizations by performing single-point calculations on the optimized structures using the conductor-like polarizable continuum model method (CPCM).[14] Five different dielectric constants were used, namely ε = 2, 4, 8, 16, and 80. For models 0, I, II, and III (see below), zero-point energy (ZPE) corrections were calculated at the same level as geometry optimizations. For models IV.1, IV.2, and V, the size of the models prohibited the frequency calculations. Thus, for these models, the ZPE correction was taken from model III.

As discussed in the Introduction, a number of atoms are kept fixed during the geometry optimizations to prevent unrealistic movements of the various groups in the models. This technique leads to a few small imaginary frequencies, in this case all below 40i cm^{-1}. These frequencies contribute insignificantly to the ZPE and can be ignored. However, they make the calculation of harmonic entropy effects inaccurate. Therefore, the entropy effects were not considered for models I−V, see discussion below.

It is important to point out here that when working with large models of enzyme active sites, like the ones used in the present work, multiple-minima problems can appear, which can lead to unreliable relative energies. We have, by careful visual inspection, confirmed that the parts that do not directly participate in the reaction are in the same local minima throughout the reaction.

Table 1. Summary of the Calculated Energetics (kcal/mol) for the Decarboxylation Step Using Various Models of AspDC

		ε = 1	ε = 2	ε = 4	ε = 8	ε = 16	ε = 80
model 0 (27 atoms)	$\Delta E^{\ddagger}$	0.1	2.3	3.7	4.5	4.9	5.2
	ΔE	−9.5	−5.7	−3.5	−2.2	−1.5	−1.0
model I (76 atoms)	$\Delta E^{\ddagger}$	8.3	12.0	14.2	15.5	16.1	16.7
	ΔE	+0.3	+2.9	+4.5	+5.5	+6.0	+6.5
model II (95 atoms)	$\Delta E^{\ddagger}$	8.8	11.9	13.6	14.7	15.3	15.7
	ΔE	−0.6	+2.6	+4.5	+5.6	+6.2	+6.7
model III (135 atoms)	$\Delta E^{\ddagger}$	9.0	11.4	12.9	13.5	13.9	14.2
	ΔE	+0.8	+2.5	+3.5	+4.1	+4.3	+4.6
model IV.1 (166 atoms)	$\Delta E^{\ddagger}$	13.9	13.1	12.7	12.5	12.4	12.3
	ΔE	+9.9	+8.7	+8.0	+7.6	+7.4	+7.2
model IV.2 (189 atoms)	$\Delta E^{\ddagger}$	13.0	15.6	17.0	17.8	18.2	18.5
	ΔE	+4.2	+7.7	+9.8	+9.9	+10.5	+10.9
model V (220 atoms)	$\Delta E^{\ddagger}$	13.5	13.5	13.4	13.3	13.3	13.3
	ΔE	+9.0	+9.6	+9.9	+10.0	+10.0	+10.0

The above-mentioned coordinate locking scheme facilitates this procedure to some extent.

III. RESULTS AND DISCUSSION

III.A. Pyruvoyl-Catalyzed Decarboxylation. We first consider the decarboxylation step of only the substrate covalently bound to the pyruvoyl, i.e., without any surrounding active site residues. In this model, which we call model 0, the cofactor is truncated at the α-carbon of Ile26 and the β-carboxylic group of the substrate aspartate is in the protonated form. In the active site, this group forms salt bridges to the Arg54′ residue, and when the latter is not explicitly included in the model, it is a better choice to protonate the group rather than using the anion model to avoid charge delocalization problems in the calculations.[15] The model consists of 27 atoms and has a total charge of 0. The optimized structures of the reactant, transition state (TS), and product species are shown in Figure 1.

In the catalytic cycle of AspDC, the formation of the Schiff base (iminium intermediate) leads to weakening of the C−C bond and hence the facilitation of the decarboxylation step. In the zwitterionic reactant structure of model 0, we see that the positive charge at NH is stabilized by two hydrogen bonds to the neighboring β-carboxylic acid group and the carbonyl oxygen. It is interesting to note how the scissile C−C bond is weakened, having a bond length of 1.75 Å, considerably longer than a normal C−C single bond length. Indeed, the barrier for decarboxylation is calculated to be very low for this model. In the gas phase, the step is practically barrierless (+0.1 kcal/mol) with an exothermicity of 9.5 kcal/mol. The addition of solvation effects, however, increases both the barrier and the reaction energy, because the zwitterionic reactant structure is stabilized more than the TS and product structures. The barrier increases to, e.g., 5.2 kcal/mol, and the reaction energy becomes only −1.0 kcal/mol when ε = 80 is used. All energies are reported in Table 1.

Before presenting the results concerning the active site models, one additional issue needs to be discussed here, namely, the entropic effects. The decarboxylation step results in the decomposition of the reactant molecule into two, an imine and a carbon dioxide. The entropy effects could potentially contribute in a non-negligible way to the energetics. We have calculated the harmonic entropy effects for model 0, and it turns out that at room temperature (298.15 K), the entropy effects increase the

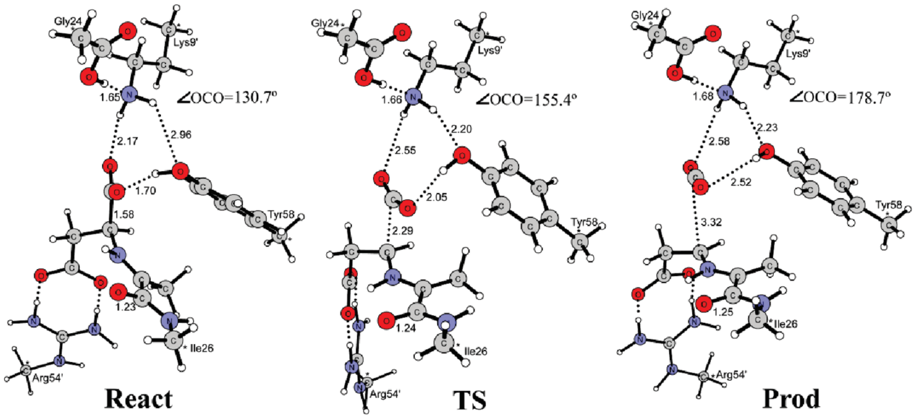


Figure 2. Optimized stationary points for model I. Centers indicated by asterisks are kept fixed during the geometry optimizations.

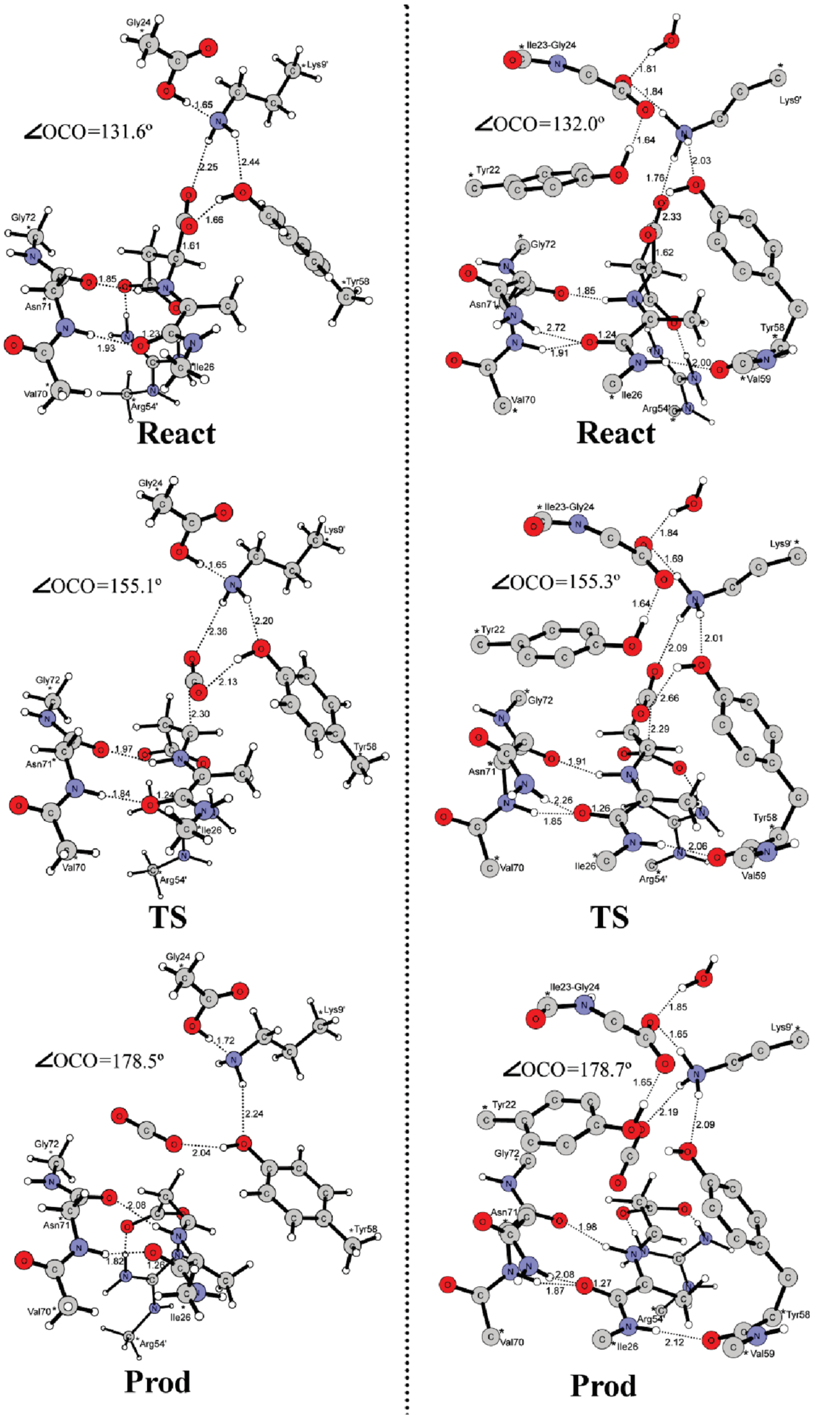


Figure 3. Optimized structures of the reactant, transition state, and product for model II (left) and model III (right). In this and the following figures, some hydrogen atoms are removed for clarity.

barrier by less than 0.1 kcal/mol. This is a very important result that justifies the omission of entropy effects for the barriers in the active site models. It is also consistent with results from QM/MM free energy calculations on the histone lysine methyltransferase enzyme, where it was found that the potential energy and the free energy barriers differed by only 1 kcal/mol.[16] Very similar conclusions were reached by Thiel and co-workers in their studies on *p*-hydroxybenzoate hydroxylase, 5′-fluoro-5′-deoxyadenosine synthase, P450cam, and chorismate mutase.[17]

For the products complex, on the other hand, the entropy effects become larger of course, since CO_2 has completely dissociated from the molecule. This leads to a lowering of the reaction energy by 4.2 kcal/mol.

III.B. Active Site Model I. In the following sections, we discuss how the inclusion of active site surrounding groups affects the energetics of this reaction. Six models, gradually increased from 76 to 220 atoms, were constructed on the basis of the high-resolution crystal structure of AspDC

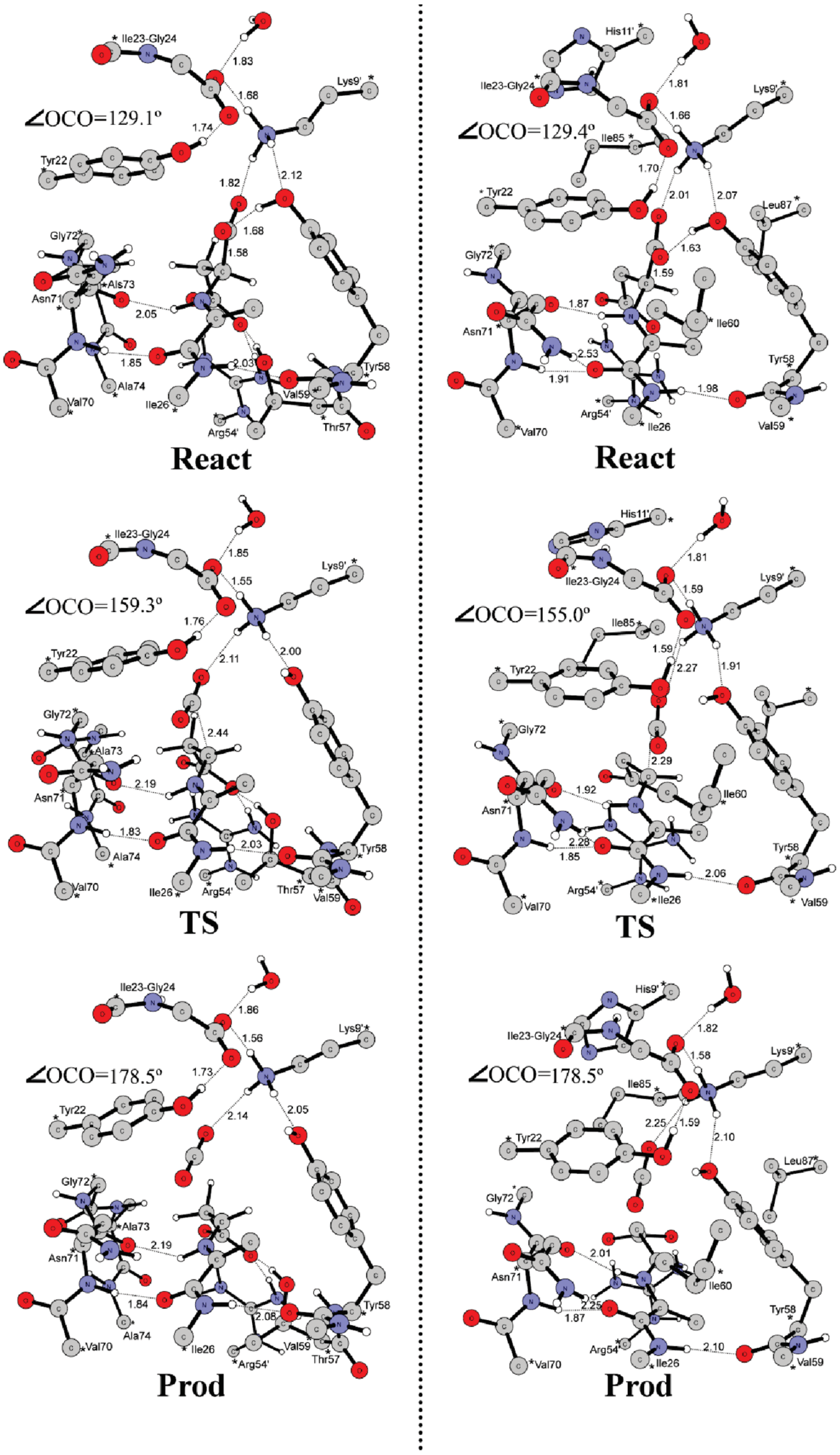


Figure 4. Optimized structures for models IV.1 and IV.2.

from *H. pylori* in complex with aspartate amide (PDB code: 1UHE).[18]

The first obvious places to add groups to model 0 are the two carboxylate groups of the substrate. To the α-carboxylate group that is going to be cleaved, two hydrogen bond donors were added, Tyr58 (modeled by a methylphenol) and Lys9′ (modeled by propylamine). In addition, the carboxylic moiety of the Gly24 (generated in the autocleavage step), which forms a hydrogen bond to Lys9′, is also included (modeled by an acetate molecule), as it will affect the hydrogen-bonding properties of the lysine. The β-carboxylate, which in model 0 was in the protonated from, is now in the ionized form but forming salt bridges to the cationic Arg54′ residue (modeled by a methyl-guanidinium). The resulting model, called model I and shown in Figure 2, consists thus of 76 atoms and has a total charge of 0.

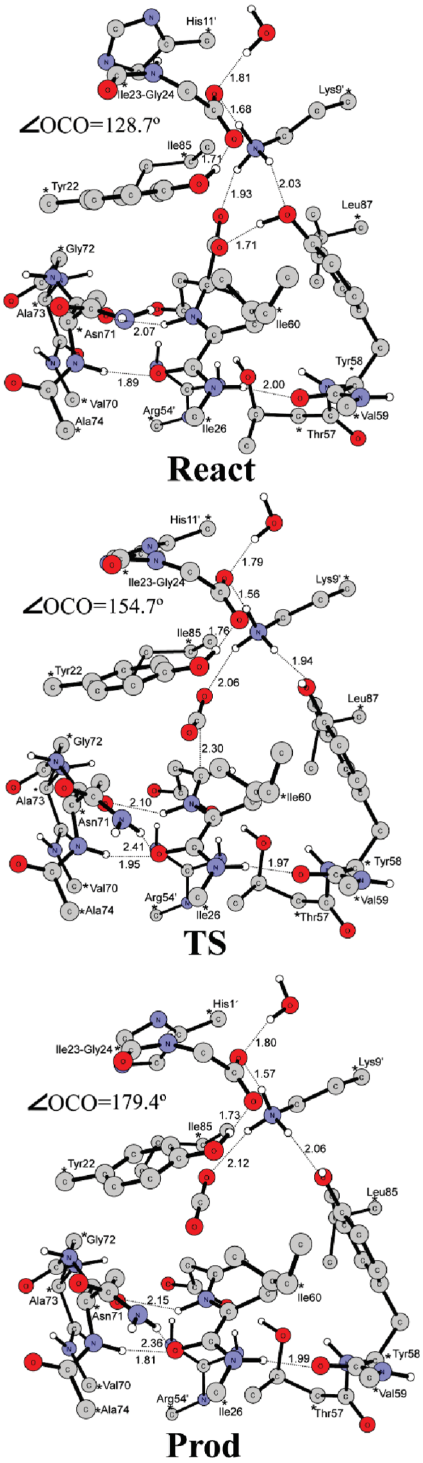


Figure 5. Optimized structures for model V.

Without solvation, the calculated barrier now is 8.3 kcal/mol and the reaction energy is +0.3 kcal/mol (Table 1). These values are 8−10 kcal/mol higher than the corresponding ones for model 0. This is mainly because the added hydrogen bonds provided by Lys9′ and Tyr58 are stronger to the anionic moiety of the reactant as compared to the TS and neutral carbon dioxide in the product, see Figure 2.

Upon inclusion of solvation effects, both the barrier and the reaction energy increase further (see Table 1). With $\varepsilon = 4$ and $\varepsilon = 80$, for example, the barrier increases to 14.2 and 16.7 kcal/mol, respectively, and the reaction energy increases to +4.5 and +6.5 kcal/mol, respectively. The solvation effects are quite large, and additional groups are clearly needed.

III.C. Active Site Model II.

Next, on the basis of model I, the peptide backbone chain between Val70, Asn71, and Gly72 was added (Figure 3). These groups form hydrogen bonds to the iminium NH and the amide carbonyl. With this addition, the active site model (model II) now consists of 95 atoms.

As seen from Figure 3, the critical geometric parameters in the transition state are quite similar to those obtained from model I, with a C−C distance of 2.30 Å and an ∠OCO angle of 155.1°. The calculated barrier for model II is 8.8 kcal/mol, and the reaction energy is −0.6 kcal/mol. Both of these values are also quite close to the ones calculated for model I. The solvation effects are still large. For example, both the barrier and the reaction energy increase by ca. 5 and 7 kcal/mol using $\varepsilon = 4$ and $\varepsilon = 80$, respectively.

We see thus that although two explicit hydrogen bonds to the substrate are added to the model, the energies and solvation effects are not changed significantly compared to model I, indicating that more groups need to be included in the active site model.

III.D. Active Site Model III.

On the basis of model II, a larger model consisting of 135 atoms was designed, called model III, see Figure 3. In this model, the Gly24 residue is extended to include the peptide bond to Ile23 to give the group more flexibility. Tyr22 and a crystallographically observed water molecule are included to stabilize the negative charge of Gly24 carboxylate. Furthermore, the Val59-Tyr58 peptide and the side chain of Asn71, which form two hydrogen bonds to the Ile26-pyruvoyl amide group, are also included.

Because of the newly added groups around the carboxylic moiety of Gly24, this group is now in the deprotonated anionic form (R−COO), and the Lys9′ side chain is in the protonated cationic form ($R{-}NH_3^+$) in the reactant complex of model III. The hydrogen bond (1.76 Å) between Lys9′ and the substrate carboxylate group becomes stronger compared to that in model II (2.25 Å). It is interesting to note that in the transition state, the key C−C distance is 2.29 Å and the ∠OCO is 155.2°, which are very close to those in model II. The barrier is 9.0 kcal/mol, and the product lies at +0.8 kcal/mol, also quite close to the values found for model II. The solvation effects are now somewhat smaller. For instance, the barrier increases to 12.9 and 14.2 kcal/mol, and the reaction energy increases to +3.5 and +4.6 kcal/mol using $\varepsilon = 4$ and $\varepsilon = 80$, respectively.

Although the solvation effects are now smaller than before, they are still of considerable size, and clearly more groups need to be added before saturation is reached.

III.E. Active Site Model IV.

Model III was increased in two different ways, and the resulting models are called models IV.1 and IV.2 and consist of 166 and 189 atoms, respectively. They differ in where the addition is made. In model IV.1, the region around the substrate β-carboxylate and the iminium part of the substrate is extended by Thr57 and the Gly72-Ala73-Ala74 peptide chain, while in model IV.2, the region around the α-carboxylate is extended by the Ile60, Ile85, and Leu87 residues (see Figure 4).

The two extensions lead to significant and different changes in both the energies and the solvation effects as compared to model III.

In model IV.1, the barrier is calculated to be 13.9 kcal/mol, and the reaction energy is +9.9 kcal/mol. In contrast to all previous models, the solvation now causes a lowering of both the barrier and the reaction energy. The solvation effects are also getting smaller compared to the previous models (less than 2 kcal/mol for the barrier and less than 3 kcal/mol for the reaction energy),

Model IV.2, on the other hand, has a barrier of 13.0 kcal/mol and a reaction energy of 4.2 kcal/mol, both of which are raised by up to 5−7 kcal/mol upon the addition of solvation.

These different results for models IV.1 and IV.2 show thus that the newly added groups influence the model energies in different ways. The next obvious model is to combine these two.

III.F. Active Site Model V. In model V, all of the groups added in models IV.1 and IV.2 are combined into a 220 atom model, the largest one used in this study (Figure 5).

The barrier now is 13.5 kcal/mol, and the reaction energy is +9.0 kcal/mol. As seen from Table 1, the addition of solvation effects, even with the largest dielectric constant ($\varepsilon = 80$), leads to a change of the vanishingly small 0.2 kcal/mol for the barrier and 1.0 kcal/mol for the reaction energy. Most of the polarization effects on the reactive parts are thus already explicitly included in the cluster model, and the solvation effects can be considered as saturated at this size.

It is also very interesting to note that the optimized transition state for this model has a very similar local geometry to those of the other models discussed above. For example, the dissociating C−C bond distance is 2.30 Å and the ∠OCO angle is 154.7°.

IV. CONCLUSIONS

In the present study, we have investigated how the quantum chemical cluster approach works for the case of enzymatic decarboxylation reactions, as exemplified by aspartate decarboxylase (AspDC). The size of the active site model is systematically increased, and the reaction barrier and energy are evaluated using several dielectric constants for the homogeneous surrounding.

The calculations show that once the model reaches a certain size (in this case 220 atoms) the solvation effects saturate; i.e., the relative energies are essentially the same whether the homogeneous surrounding is included or not (see Table 1). We have observed this quick convergence for several examples of different classes of enzymes, namely, 4-oxalocrotonate tautomerase,[4] in which an ion pair is formed during the reaction; haloalcohol dehalogenase HheC,[5] in which a chloride ion is released; and histone lysine methyltransferase,[6] in which a methyl cation is transferred. Taken together, these results suggest thus that this is a general feature of the cluster approach.

Of course, as pointed out by Ryde and co-workers,[19] convergence of the solvation effects is not equivalent to convergence of the energies (barriers and reaction energies), although they might be related. In this context, it is particularly interesting to note that the energies of all active site models (I−V) after application of some solvation corrections fall within a relatively narrow range of about 5 kcal/mol. This shows that the results are quite stable and are already using medium-sized models certainly in a sufficiently accurate manner to investigate mechanistic alternatives. Considering this, it is also unlikely that groups that are further away will affect the relative energies in any significant way.

Here, it should be remembered that geometry optimization of the structures is an essential requirement of the cluster approach, contributing to the quick convergence observed. By contrast, the QM/MM methodology exhibits a quite slow convergence behavior,[19−21] which in part could be due to the fact that geometries are not optimized.[19,20]

ASSOCIATED CONTENT

S Supporting Information. Cartesian coordinates for all structures. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author
*E-mail: himo@organ.su.se.

ACKNOWLEDGMENT

F.H. gratefully acknowledges financial support from The Swedish Research Council (Grants 621-2009-4736 and 622-2009-371) and computer time from the PDC Center for High Performance Computing. This work was also supported by grants from the National Natural Science Foundation of China (Grant No. 20873008).

REFERENCES

(1) For reviews, see: (a) Blomberg, M. R. A.; Siegbahn, P. E. M. *J. Phys. Chem. B* **2001**, *105*, 9375–9386. (b) Himo, F.; Siegbahn, P. E. M. *Chem. Rev.* **2003**, *103*, 2421–2456. (c) Noodleman, L.; Lovell, T.; Han, W.-G.; Li, J.; Himo, F. *Chem. Rev.* **2004**, *104*, 459–508. (d) Siegbahn, P. E. M.; Borowski, T. *Acc. Chem. Res.* **2006**, *39*, 729–738. (e) Himo, F. *Theor. Chem. Acc.* **2006**, *116*, 232–240. (f) Ramos, M. J.; Fernandes, P. A. *Acc. Chem. Res.* **2008**, *41*, 689–698. (g) Himo, F.; Siegbahn, P. E. M. *J. Biol. Inorg. Chem.* **2009**, *14*, 643–651. (h) Blomberg, M. R. A.; Siegbahn, P. E. M. *Biochim. Biophys. Acta* **2010**, *1797*, 129–142.

(2) (a) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652. (b) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.

(3) See for example the following applications: (a) Velichkova, P.; Himo, F. *J. Phys. Chem. B* **2005**, *109*, 8216–8219. (b) Hopmann, K. H.; Himo, F. *Chem.—Eur. J.* **2006**, *12*, 6898–6909. (c) Hopmann, K. H.; Guo, J.-D.; Himo., F. *Inorg. Chem.* **2007**, *46*, 4850–4856. (d) De Marothy, S. A.; Blomberg, M. R. A.; Siegbahn, P. E. M. *J. Comput. Chem.* **2007**, *28*, 528–539. (e) Leopoldini, M.; Chiodo, S. G.; Toscano, M.; Russo, N. *Chem.—Eur. J.* **2008**, *14*, 8647–8681. (f) Borowski, T.; Blomberg, M. R. A.; Siegbahn, P. E. M. *Chem.—Eur. J.* **2008**, *14*, 2264–2276. (g) Sevastik, R.; Whitman, C. P.; Himo, F. *Biochemistry* **2009**, *48*, 9641–9649. (h) Chen, S.-L.; Pelmenschikov, V.; Blomberg, M. R. A.; Siegbahn, P. E. M. *J. Am. Chem. Soc.* **2009**, *131*, 9912–9913. (i) Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. *Chem.—Eur. J.* **2009**, *15*, 4243–4247. (j) Roos, K.; Siegbahn, P. E. M. *Biochemistry* **2009**, *48*, 1878–1887. (k) Yang, L.; Liao, R.-Z.; Yu, J.-G.; Liu, R.-Z. *J. Phys. Chem. B* **2009**, *113*, 6505–6510. (l) Parks, J. M.; Guo, H.; Momany, C.; Liang, L. Y.; Miller, S. M.; Summers, A. O.; Smith, J. C. *J. Am. Chem. Soc.* **2009**, *131*, 13278–13285. (m) Liao, R.-Z.; Yu, J.-G.; Himo, F. *J. Phys. Chem. B* **2010**, *114*, 2533–2540.

(4) Sevastik, R.; Himo, F. *Bioorg. Chem.* **2007**, *35*, 444–457.

(5) Hopmann, K. H.; Himo, F. *J. Chem. Theory Comput.* **2008**, *4*, 1129–1137.

(6) Georgieva, P.; Himo, F. *J. Comput. Chem.* **2010**, *31*, 1707–1714.

(7) (a) Lee, J. K.; Houk, K. N. *Science* **1997**, *276*, 942–945. (b) Lundberg, M.; Blomberg, M. R. A.; Siegbahn, P. E. M. *J. Mol. Model* **2002**, *8*, 119–130. (c) Silva, P. J.; Ramos, M. J. *J. Phys. Chem. B* **2005**, *109*, 18195–18200. (d) Wang, J.; Hong, H.; Li, S.; He, H. *J. Phys. Chem. B* **2005**, *109*, 18664–18672. (e) Silva, P. J.; Ramos, M. J. *J. Phys. Chem. B* **2007**, *111*, 12883–12887. (f) Ito, Y.; Kondo, H.; Shiota, Y.; Yoshizawa, K. *J. Chem. Theory Comput.* **2008**, *4*, 366–374.

(8) (a) Wu, N.; Mo, Y.; Gao, J.; Pai, E. F. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 2017–2022. (b) Lee, T.-S.; Chong, L. T.; Chodera, J. D.; Kollman, P. A. *J. Am. Chem. Soc.* **2001**, *123*, 12837–12848. (c) Stanton, C. L.; Kuo, I. W.; Mundy, C. J.; Laino, T.; Houk, K. N. *J. Phys. Chem. B* **2007**, *111*, 12573–12581. (d) Moya-García, A. A.; Ruiz-Pernía, J.; Martí, S.; Sánchez-Jiménez, F.; Tuñón, I. *J. Biol. Chem.* **2008**, *283*, 12393–12401. (e) Hu, H.; Boone, A.; Yang, W. *J. Am. Chem. Soc.* **2008**, *130*, 14493–14503.(f) Lin, Y.-L.; Gao, J. L. *J. Am. Chem. Soc.* **2011** [Online] dx.doi.org/10.1021/ja108209w.

(9) (a) Williamson, J. M.; Brown, G. M. *J. Biol. Chem.* **1979**, *254* 8074–8082. (b) Cronan, J. E., Jr. *J. Bacteriol.* **1980**, *141*, 1291–1297. (c) van Poelje, P. D.; Snell, E. E. *Annu. Rev. Biochem.* **1990**, *59*, 29–59. (d) Ramjee, M. K.; Genschel, U.; Abell, C.; Smith, A. G. *Biochem. J.* **1997**, *323*, 661–669. (e) Chopra, S. C.; Pai, H.; Ranganathan, A. *Protein Expression Purif.* **2002**, *25*, 533–540.

(10) Brown, G. M.; Williamson, J. M. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1982**, *53*, 345–381.

(11) (a) Recsei, P. A.; Huynh, Q. K.; Snell, E. E. *Proc. Natl. Acda. Sci. U.S.A.* **1983**, *80*, 973–977. (b) Albert, A.; Dhanaraj, V.; Genschel, U.; Khan, G.; Ramjee, M. K.; Pulido, R. *Nat. Struct. Biol.* **1998**, *5*, 289–293. (c) Xiong, H.; Pegg, A. E. *J. Biol. Chem.* **1999**, *274*, 35059–35066. (d) Schmitzberger, F.; Kilkenny, M. L.; Lobley, C. M. C.; Webb, M. E.; Vinkovic, M.; Matak-Vinkovic, D.; Witty, M.; Chirgadze, D. Y.; Smith, A. G.; Abell, C.; Blundell, T. L. *EMBO J.* **2003**, *22*, 6193–6204.

(12) Saldanha, S. A.; Birch, L. M.; Webb, M. E.; Nabbs, B. K.; von Delft, F.; Smith, A. G.; Abell, C. *Chem. Commun.* **2001**, 1760–1761.

(13) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision D.01; Gaussian, Inc.: Wallingford, CT, 2004.

(14) (a) Klamt, A.; Schüürmann, G. *J. Chem. Soc., Perkin. Trans.* **1993**, *2*, 799–805. (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*, 669–691.

(15) (a) Himo, F.; Eriksson, L. A. *J. Am. Chem. Soc.* **1998**, *120*, 11449–11455. (b) Guo, J.-D.; Himo, F. *J. Phys. Chem. B* **2004**, *108*, 15347–15354. (c) Condic-Jurikic, K.; Perchyonok, V. T.; Zipse, H.; Smith, D. M. *J. Comput. Chem.* **2008**, *29*, 2425–2433.

(16) Hu, P.; Zhang, Y. *J. Am. Chem. Soc.* **2006**, *128*, 1272–1278.

(17) (a) Senn, H. M.; Thiel, S; Breidung, J.; Thiel, W. *J. Chem. Theory Comput.* **2005**, *1*, 494–505. (b) Senn, H. M.; Kästner, J.; Breidung, J.; Thiel, W. *Can. J. Chem.* **2009**, *87*, 1332–1337.

(18) Lee, B. I.; Suh, S. W. *J. Mol. Biol.* **2004**, *340*, 1–7.

(19) Hu, L. H.; Eliasson, J.; Heimdal, J.; Ryde, U. *J. Phys. Chem. A* **2009**, *113*, 11793–11800.

(20) Sumowski, C. V.; Ochsenfeld, C. *J. Phys. Chem. A* **2009**, *113*, 11734–11741.

(21) Solt, I.; Kulhánek, P.; Simon, I.; Winfield, S.; Payne, M. C.; Csányi, G.; Fuxreiter, M. *J. Phys. Chem. B* **2009**, *113*, 5728–5735.

JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Impact of Thermostats on Folding and Aggregation Properties of Peptides Using the Optimized Potential for Efficient Structure Prediction Coarse-Grained Model

Yannick G. Spill,[†] Samuela Pasquali, and Philippe Derreumaux*

Laboratoire de Biochimie Théorique, UPR 9080 CNRS et Université Paris Diderot (Paris 7), Institut de Biologie Physico Chimique, 13 rue Pierre et Marie Curie, 75005 Paris, France

ABSTRACT: The simulation of amyloid fibril formation is impossible if one takes into account all chemical details of the amino acids and their detailed interactions with the solvent. We investigate the folding and aggregation of two model peptides using the optimized potential for efficient structure prediction (OPEP) coarse-grained model and replica exchange molecular dynamics (REMD) simulations coupled with either the Langevin or the Berendsen thermostat. For both the monomer of blocked penta-alanine and the trimer of the 25−35 fragment of the Alzheimer's amyloid β protein, we find little variations in the equilibrium structures and heat capacity curves using the two thermostats. Despite this high similarity, we detect significant differences in the populations of the dominant conformations at low temperatures, whereas the configurational distributions remain the same in proximity of the melting temperature. Aβ_{25-35} trimers at 300 K have an averaged β-sheet content of 12% and are primarily characterized by fully disordered peptides or a small curved two-stranded β-sheet stabilized by a disordered peptide. In addition, OPEP molecular dynamics simulations of Aβ_{25-35} hexamers at 300 K with a small curved six-stranded antiparallel β-sheet do not show any extension of the β-sheet content. These data support the idea that the mechanism of Aβ_{25-35} amyloid formation does not result from a high fraction of extended β-sheet-rich trimers and hexamers.

1. INTRODUCTION

When an interacting particle is subject to Brownian motion in a solvent, under the assumption that the radius of the particle is significantly larger than the radius of the solvent constituents, its motion can be modeled by the Langevin equation:[1]

$$m\frac{\mathrm{d}^2\vec{X}(t)}{\mathrm{d}t^2} = -\vec{\nabla}V(t) - m\gamma\frac{\mathrm{d}\vec{X}(t)}{\mathrm{d}t} + \vec{R}(t) \qquad (1)$$
$$\langle R_i(t)R_j(t')\rangle = \delta(t-t')\delta_{ij}6\gamma k_B T_0 \text{ and } \langle\vec{R}(t)\rangle = 0$$

The action of the solvent at a temperature T_0 is represented implicitly through the action of a friction and a random force. The friction force corresponds to the macroscopic decay in momentum when a particle moves through a viscous fluid colliding with the small fluid particles; the random force represents the microscopic shocks with the solvent that have the effect of changing the particle trajectory. These shocks are disordered and tend to "thermalize" the particle, meaning that at equilibrium, the velocity distribution is Gaussian, with zero mean and variance $\sigma^2 = k_B T_0/m$. The parameter controlling the impact of viscosity on the dynamics is the friction constant γ, which can also be interpreted as the collision frequency between the particle and the solvent. Two limiting cases occur. When γ equals zero, the particle follows a Newtonian dynamic and evolves in the microcanonical ensemble. When γ tends to infinity, Newtonian forces are negligible compared to the Langevin forces, and the particle follows a purely diffusive Brownian dynamic. For all cases in between, the system is in the canonical ensemble, at constant temperature T_0.

Although originally postulated to describe the Brownian motion of a particle, this equation can be easily implemented to act as a heat bath in molecular dynamics (MD) simulations.[2] It can be shown that this formulation allows one to generate a canonical distribution of states.[3] In the context of molecular dynamics, a Langevin thermal bath has the disadvantage that the coupling of the system to the bath is not only global but also local due to the random shocks. For large values of γ, the local shocks between the particle and the solvent can become an important disturbance to the system's dynamics. Ideally one would want to control the temperature through a global coupling only in order to minimize the local disturbance. To overcome this limitation, a new thermostat was introduced by Berendsen et al.[4] The global Langevin coupling is maintained, while an integration over fast degrees of freedom smooths out local collisions. The modified equation of motion is

$$m\frac{\mathrm{d}^2\vec{X}(t)}{\mathrm{d}t^2} = -\vec{\nabla}V - m\frac{1}{2\tau}\left(\frac{T_0}{T} - 1\right)\frac{\mathrm{d}\vec{X}(t)}{\mathrm{d}t} \qquad (2)$$

where $\tau = 1/(2\gamma)$ and γ is the Langevin friction constant. In practice one sets the Berendsen time constant $\tau = 1/(2\gamma)$ rather than γ itself. Thermodynamics governed by a Berendsen thermostat depart from the canonical distribution for finite and nonzero values of τ. The statistical distribution followed by a system obeying eq 2 has been rigorously determined only recently.[5] It reduces to the canonical distribution for $\tau = 0$ and to the microcanonical distribution for $\tau = +\infty$.

Received: October 29, 2010
Published: April 11, 2011

Several computational studies have already discussed the effect of the Berendsen thermostat in MD and replica exchange molecular dynamics (REMD) simulations. The systems studied included a single butane molecule,[6] bulk water alone or in the presence of a penta-alanine,[7] simulated by use of a fully atomistic representation, and the 56-residue SH3 protein with tails of various lengths, simulated by use of a single C_α representation and the native structure-based Go potential.[8] These studies indicate that the Berendsen thermostat produces a noncanonical phase-space distribution, but the magnitude of the deviation is system-dependent: small for bulk water and larger for a peptide model using AMBER/TIP3P.[7] UNRES MD simulations of two model α-helical systems with the Berendsen and Langevin thermostats showed also differences in the folding dynamics, with the presence of explicit friction forces slowing down the folding.[9]

Unless a worldwide network of computers[10] or a specially built supercomputer[11] is used, atomic-level characterization of protein folding and aggregation in explicit solvent is limited to short time scales.[12] To go beyond this limitation, one solution is to reduce the number of degrees of freedon by resorting to simpler protein representations. The implicit solvent optimized potential for efficient structure prediction (OPEP) coarse-grained force field has been recently used to predict the 3D structures of peptides in their monomeric states via a greedy approach[13] and the early formed oligomers of various amyloid peptides by use of REMD and the Berendsen thermostat.[14,15]

Our aim is to investigate the influence of the Langevin and Berendsen thermostats on both folding and aggregation properties of peptides using the OPEP coarse-grained force field. To this end, we performed long REMD simulations on the penta-alanine peptide and the trimer of Alzheimer's $A\beta_{25-35}$ peptide. REMD is one generalized ensemble method that goes beyond conventional MD to accelerate convergence to equilibrium[16,17] and allows one to extract thermodynamic information. The penta-alanine model enables a comparison with the atomistic simulations in explicit solvent carried out by Rosta et al.[7] $A\beta_{25-35}$, due to its small size, is a convenient model system for studying the formation of $A\beta$ amyloids. It is known experimentally that this molecule can readily form β-sheet aggregates that are highly toxic to neurons.[18] Simulations addressing its dimerization in explicit water have been recently performed.[19,20] In contrast to other $A\beta$ fragments such as $A\beta_{16-22}$[21,22] or $A\beta_{29-42}$[23,24], considered as toy models for computational studies of amyloid formation, $A\beta_{25-35}$ has been identified as a physiological proxy with mutiple effects in neuronal intracellular components including plasma membranes, mitochondria, and cytosol, and its use has shown many of the same biochemical changes in animal models as those detected in Alzheimer's disease patients.[25] To enlarge the part of $A\beta_{25-35}$ amyloid dynamics and stability, we also carried out OPEP-MD simulations of hexamers built from structured trimer conformations.

2. MATERIALS AND METHODS

2.1. General Setup. REMD simulations[26] were performed with the coarse-grained OPEP force field.[27,28] In this model, the backbone N, H, C_α, C ,and O atoms are represented explicitly, while the side chains of all amino acids (except for the proline's heavy atoms) are represented by a unique bead. The interaction parameters include bond lengths, bond angles, improper torsions, dihedral angles, van der Waals interactions, and two- and four-body interactions for the backbone hydrogen bonds.[28] By removing numerous degrees of freedom, the OPEP force field allows significant acceleration with realistic sampling of conformational space.[29]

Table 1. Summary of Simulations Performed

starting config[a]	method	temp (K)	thermostat (τ or γ)	time (ns)
		Penta-alanine		
E	REMD	200−350	B (0.5 ps)	300 × 12
R	REMD	200−350	B (1 ps)	300 × 12
R	REMD	200−350	L (1 ps^{-1})	300 × 12
R	REMD	200−350	L (0.5 ps^{-1})	300 × 12
		$A\beta_{25-35}$ Trimers		
R	REMD	250−500	B (0.5 ps)	1000 × 24
R	REMD	250−500	L (1 ps^{-1})	1000 × 24
P	MD	300	B (0.5 ps)	100
P	MD	300	L (1 ps^{-1})	100
		$A\beta_{25-35}$ Hexamers		
AP_1	MD	300	B (0.5 ps)	100
AP_1	MD	300	L (1 ps^{-1})	100
AP_2	MD	300	B (0.5 ps)	100
AP_2	MD	300	L (1 ps^{-1})	100

[a] The starting conformation is either extended (E), random (R), parallel (P), or antiparallel (AP).

The Langevin thermostat was implemented in the OPEP-MD program by scaling each velocity:

$$v_{ij} \rightarrow e^{-\gamma\delta t} v_{ij} + \sqrt{(1 - e^{-\gamma\delta t})(1 + e^{-\gamma\delta t})}\sqrt{\frac{k_B T_0}{m_i}}\, G(t) \quad (3)$$

where v_{ij} is the jth component of the velocity of particle i and m_i is its mass; γ is the Langevin collision frequency constant; δt is the integration time step; k_B is Boltzmann's constant; T_0 is the target temperature; and $G(t)$ is a random Gaussian number with zero mean and unit variance. A typical value for γ is, for example, 50 ps^{-1}, thought to reproduce best the dynamics in water,[30] but values as low as 0.15 ps^{-1} have been used for amyloids with an implicit solvent model.[31]

Velocity rescaling for the Berendsen thermostat is[29]

$$v_{ij} \rightarrow \sqrt{\left(1 - \frac{\delta t}{\tau}\right)\left(1 - \frac{T_0}{T}\right) v_{ij}} \quad (4)$$

where T is the instantaneous temperature and τ is the Berendsen coupling constant. A typical value of τ suited to study aggregation is τ = 0.5 ps (see ref 15), as it simultaneously allows for a quick thermalization while not overconstraining the system.

2.2. Simulations. A summary of the setup details of all simulations is given in Table 1. For each system, we describe the method used (REMD or MD), the initial configuration, the temperatures and thermostats used (Berendsen or Langevin and the corresponding values of τ or γ), and the total simulation time.

The first system studied is the penta-alanine blocked by acetyl and N-methyl groups. An exponential temperature distribution of the replicas assured fast convergence. The temperature range varying between 200 and 350 K spans across the folded and the unfolded states: 200.0, 210.4, 221.4, 233, 245, 257.9, 271.4, 285.6, 300.5, 316.1, 332.6, and 350.0 K. For each thermostat, two 300 ns REMD simulations were launched with 12 replicas,

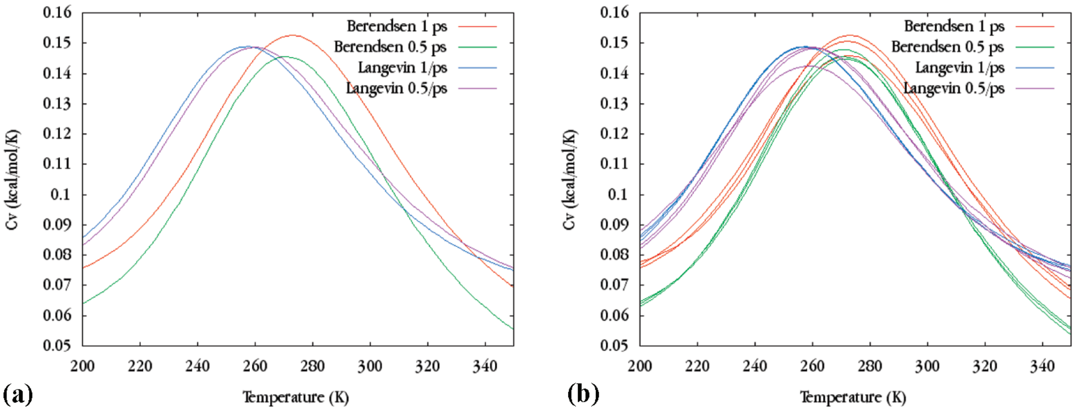


Figure 1. Plot of the heat capacity of the four penta-alanine simulations: (a) full 10−300 ns time interval; (b) 10−300, 10−150, and 150−300 ns intervals for each thermostat and coupling constant.

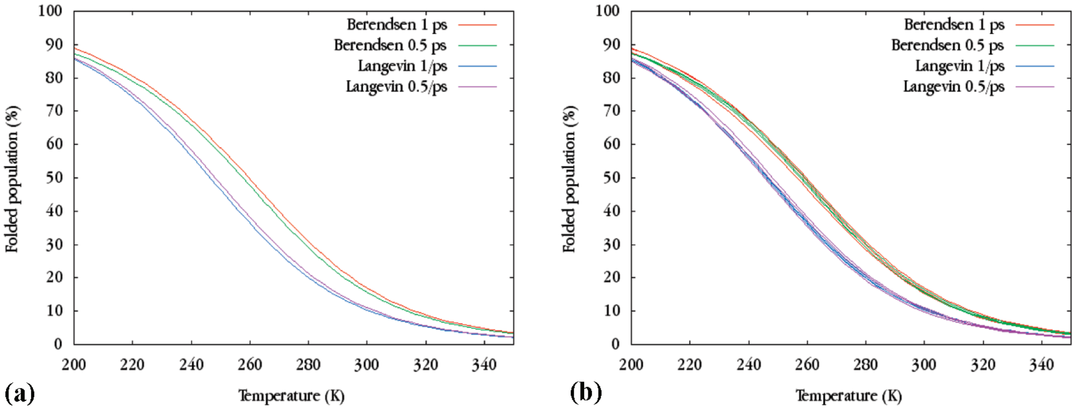


Figure 2. Populations of the folded state for four different penta-alanine simulations. The capped penta-alanine was considered to be folded if it had at least two *i*, *i* + 3 or *i*, *i* + 4 hydrogen bonds. (a) Full 10−300 ns time interval; (b) 10−300, 10−150, and 150−300 ns intervals for each thermostat and coupling constant.

with the following coupling constants: Berendsen τ, 0.5 ps and 1 ps; Langevin γ, 1 ps^{-1} and 0.5 ps^{-1}. This choice corresponds to two series of simulations with reciprocal time constants with respect to the equality $\tau = 1/(2\gamma)$. The initial configuration was an extended configuration.

The second system is a trimer of the hydrophilic−hydrophobic Aβ_{25-35} peptide, of sequence acetyl-Gly-Ser-Asn-Lys-Gly-Ala-Ile-Ile-Gly-Leu-Met-NH_2. Since it is a homotrimer, five different topologies are expected in principle:

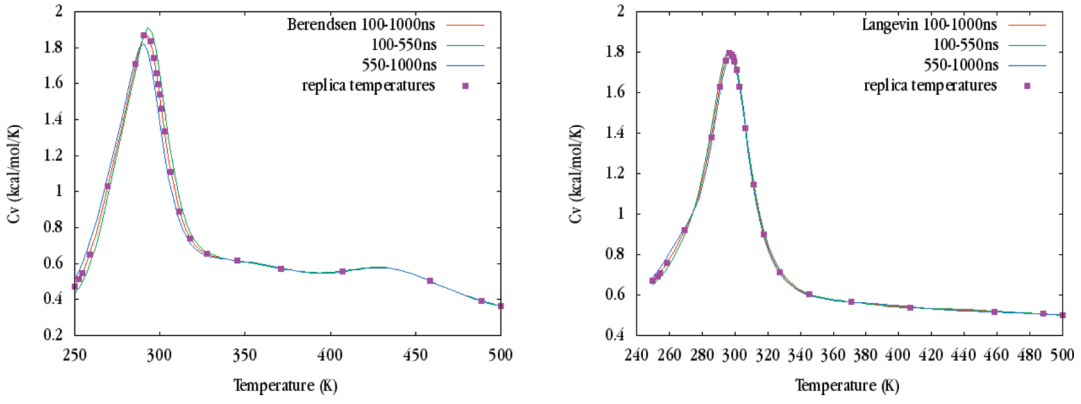
- antiparallel (AP): the three chains form a β-sheet where neighboring peptides are oriented opposite to each other
- mixed (Mix): the three chains form a β-sheet where two neighboring peptides are oriented parallel, while the third one is oriented in the opposite direction
- parallel (P): the three chains form a β-sheet with all peptides oriented in the same direction
- partially folded (2 + 1): two peptides form a β-sheet while the third one is random coil
- coil: all three peptides are random coil

Two REMD runs were performed, one with Berendsen $\tau = 0.5$ ps and the other with Langevin γ = 1 ps^{-1}. For each run, 24 replicas were used with the same set of temperatures: 250, 252.69, 254.48, 258.66, 269.33, 285.64, 290.74, 294.25, 296.55, 297.99, 298.96, 299.83, 300.96, 302.81 306.45, 311.48, 317.74, 327.46, 345.32, 370.95, 407.15, 458.41, 488.43, and 500.00 K. These temperatures were determined from the configurations obtained by a preliminary REMD simulation of 500 ns with 18 replicas starting from fully disordered peptides separated from each others by 15 Å via the procedure presented in ref 32. This choice corresponds to a replica flux-optimized temperature distribution for our system, as described in refs 17, 33, and 34. It is of interest to note that equilibration was not achieved within 500 ns in the preliminary REMD.

Each REMD simulation was performed for 1 μs per replica, starting with conformations belonging to the antiparallel, mixed, 2 + 1, and coil basins and obtained through the preliminary simulation. The integration time step was 1.5 fs, a sphere of 80 Å with reflecting boundary conditions was used, and energy and structure snapshots were taken every 1.5 ps. Replica exchanges were attempted every 6 ps, a time interval highly used and tested.[26,35] Two MD simulations of 100 ns at 300 K were also carried out on the trimer starting from a fully extended parallel β-sheet.

The third system is a hexamer of Aβ_{25-35} consisting of a single-layer antiparallel small β-sheet configuration. Four MD runs of 100 ns at 300 K were performed, two with Berendsen τ = 0.5 ps (B_1 and B_2) and two with Langevin γ = 1 ps^{-1} (L_1 and L_2). The runs B*i* and L*i* use the same structure, and the rmsd deviation between the two starting conformations is 10 Å.

2.3. 2D and 3D Structure Analysis. For each system, the secondary structure of all conformations was determined. For the



(a) Berendsen simulation

(b) Langevin simulation

Figure 3. Block analysis of the Aβ_{25-35} trimer heat capacity for each thermostat.

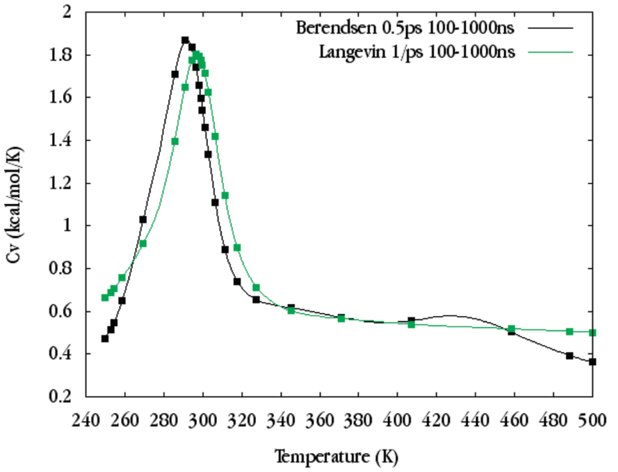


Figure 4. Heat capacities of the Aβ_{25-35} trimer with Langevin and Berendsen thermostats.

Aβ_{25-35} trimer, prior to cluster analysis based on C$_\alpha$ rmsd, we determined the topology of each configuration by using a combination of three metrics that includes vector cosines (where we compute the angles between the vectors representing the direction of each peptide), number of hydrogen bonds formed, and distance between the peptides. We also analyzed the transition times between the Mix and AP configurations and defined that a conformational change from Mix to AP (or vice versa) has occurred if, before and after the transition, the trimer explores both topologies for at least 100 ps. For the Aβ_{25-35} hexamer, we followed the C$_\alpha$ rmsd as a function of time using the core β-sheet region.

2.4. Thermodynamic Quantities. At the end of an REMD simulation, thermodynamic quantities can be extracted. Even though in principle these could be obtained from their definitions with respect to the independent variables of the system, in practice this direct evaluation leads, sometimes, to very large error bars in the estimates and alternative derivations turn out to be more convenient. This is the case for the specific heat C_V. With both thermostats, we could obtain C_V by computing the potential energy derivative with respect to the temperature, but different strategies give better results. For the Langevin thermostat, where the system obeys the canonical distribution, the heat capacity can be calculated from the potential energy as

$$C_V(T) = \frac{\langle (E - \langle E \rangle)^2 \rangle}{k_B T^2} \qquad (5)$$

For the weak-coupling Berendsen thermostat, where the distribution is noncanonical, the appropriate formula has been shown to be[5]

$$C_V(T) = \frac{k_B \langle (\delta\Phi)^2 \rangle}{(k_B T)^2 - \sqrt{\langle (\delta\Phi)^2 \rangle \langle (\delta K)^2 \rangle 2/(3N)}} \qquad (6)$$

where $\langle (\delta\Phi)^2 \rangle$ and $\langle (\delta K)^2 \rangle$ represent the variance of potential and kinetic energy, respectively, and broken brackets denote a time average.

Thermodynamic quantities are meaningful only if computed after the full convergence of the system has been established with respect to the particular quantity in question. Different quantities have different convergence times. We used a block analysis scheme to assess convergence over a given time frame. The quantity is evaluated once over the entire time interval, a second time over the first half of the interval, and a third time over the second half. The three curves are then superposed and the convergence is reached when the curves coincide within a preset error.

In practice, to extract thermodynamical quantities, histogram-based methods like WHAM are typically used.[36] We preferred implementing our own script based on the recently introduced MBAR method,[37] which has the conceptual advantage of producing curves accompanied by error bars and the technical advantage of reducing the computing time by avoiding some of the slow histogram calculations.

3. RESULTS

3.1. Penta-alanine. The first 10 ns of each REMD simulation was discarded, and the remaining 290 ns was analyzed. The acceptance rate was always higher than 30% in all four simulations, reaching its minimal values for the lowest and highest replicas.

When the heat capacity curves obtained from the two thermostats with reciprocal values of τ and γ are compared (Figure 1), the position of the peak (T_m) changes from 260 K for Langevin to 273 K for Berendsen. Compared to the Berendsen thermostat,

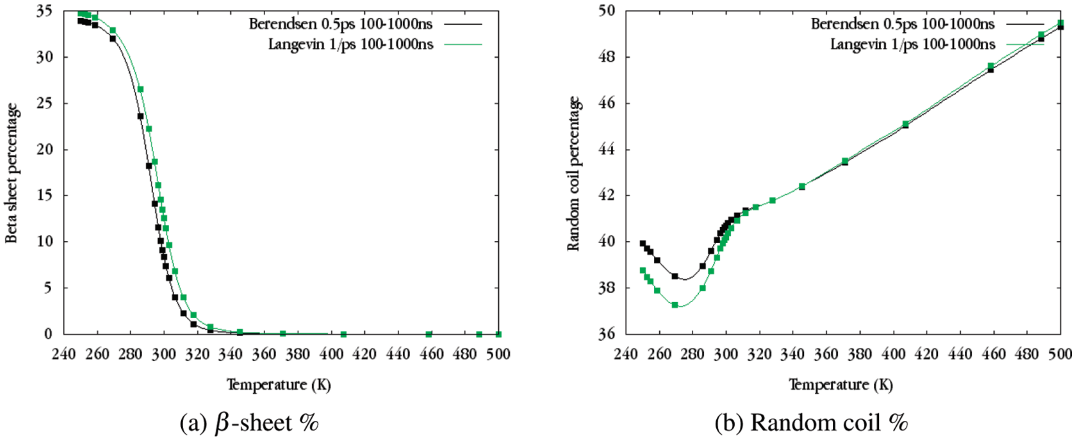


Figure 5. (a) β-Sheet and (b) random coil percentages in the two REMD simulations of Aβ_{25-35} trimer.

Table 2. Structural and Energetic Contents of the Two Aβ_{25-35} Trimer Simulations at the Lowest Temperature[a]

	Berendsen		Langevin	
class	%	energy (kcal/mol)	%	energy (kcal/mol)
AP	71.6	−134.1 ± 6.3	56.3	−133.8 ± 8.4
Mix	24.2	−130.6 ± 6.9	37.7	−131.4 ± 8.9
P	0.0		0.0	
2 + 1	4.3	−119.9 ± 6.7	5.9	−118.7 ± 8.9
coil	0.0		0.0	

[a] Obtained by classifying every tenth frame of the converged period of each simulation, totalling 60 000 structures. The population of coil conformations is too small to get accurate statistics at this temperature.

the Langevin thermostat shifts the melting temperature 13 K toward lower temperature.

If the coupling constant is modified, the position of T_m and the full width at half-maximum of the peak (fwhm) remain unchanged in the Langevin simulations, whereas in Berendsen simulations T_m remains constant while fwhm increases by 20% ± 5%. This variation in fwhm is larger than the error bars obtained from block analysis.

A recent all-atom REMD simulation on the same pentaalanine system in explicit solvent reported that, for a weak-coupling Berendsen thermostat, the folded state is overpopulated by about 10% at low temperatures and underpopulated at high temperatures.[7] To verify this hypothesis, we compared the population distribution as a function of temperature, obtained with both Langevin and Berendsen thermostats. In our simulations we define the folded state by a criterion in terms of the hydrogen-bond (H-bond) network. We considered a H-bond to be formed if the distance between donor and acceptor was smaller than 3 Å and the donor−H−acceptor angle was less than 60°. In all simulations, we found that H-bonds were established mainly between residues i and $i + 4$ (α-helix) and to a lesser extent between i and $i + 3$ (3_{10}-helix). We also observed a very small amount of $i + 5$ contacts (π-helix). We thus considered the peptide to be folded if it had two or more hydrogen bonds between residues three or four amino acids apart. The resulting folded populations are shown in Figure 2.

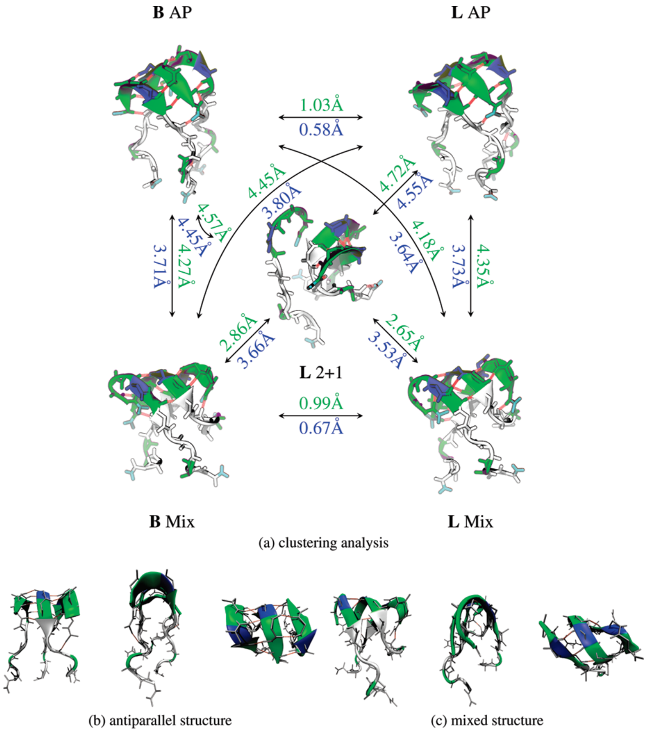
When we focus on the temperature range 290−340 K, our simulations lead to a small shift in the folded populations varying between 5% and 1% between the two thermostats, with an increased population for the Berendsen thermostat. At lower temperatures, the difference in the folded population slightly increases but never exceeds 10%. This overpopulation of the folded (helical) state at low temperatures with the weak Berendsen coupling constant is fully consistent with what was observed by Rosta et al.[7] using all-atom REMD simulations. In contrast to Rosta's results at 350 K, where the fraction of folded state is lowered by about 10% for Berendsen compared to Langevin but remains around 0.6 at 350 K, both thermostats with OPEP give a negligible population of folded state at 350 K.

3.2. Aβ_{25-35} Trimer.

Two REMD simulations were performed on the Aβ_{25-35} trimer, with reciprocal Langevin (γ = 1 ps^{-1}) and Berendsen (τ = 0.5 ps) thermostats starting with 24 conformations from different energy basins. From a block analysis of the potential energies and C_V values, the first 100 ns were discarded for each replica. Figure 3 shows the convergence of each simulation by comparing their heat capacity profiles over two time intervals, 100−550 and 550−1000 ns. The conformational properties are therefore based on 21.6 μs for each thermostat.

In Figures 4 and 5, we show the heat capacity and the percentages of β-sheet and random coil as determined by the STRIDE program[38] as a function of temperature. Overall, the characteristic transitions are shifted by approximately 5 K between the two thermostats, as assessed by the peak of heat capacity (297 K for Langevin vs 291 K for Berendsen) and the inflection point of β-sheet percentage (296 K for Langevin and 292 K for Berendsen). Random coil profiles also superpose well with a maximum difference of 1% at low temperatures. We note, however, that the excess heat capacity reaches slightly lower values in Berendsen simulation than in Langevin simulation at low (below 260 K) and high (>460 K) temperatures.

Of the five possible topologies of the Aβ_{25-35} trimer, all except the parallel (P) three-stranded β-sheet have been detected in each one of the REMD simulations. It should be noted that two MD simulations starting from the fully parallel geometry evolve toward the 2 + 1 topology within 100 ns. Table 2 reports the structural and energetic content of the REMD-generated topologies at the lowest temperature (250 K). As can be seen, the antiparallel conformation (AP) is the most populated, followed by the mixed (Mix) and 2 + 1 topologies in both simulations. Looking at the three populated AP, Mix, and 2 + 1 topologies representing 99.99% of the conformations, we find that the averaged



(a) clustering analysis

(b) antiparallel structure

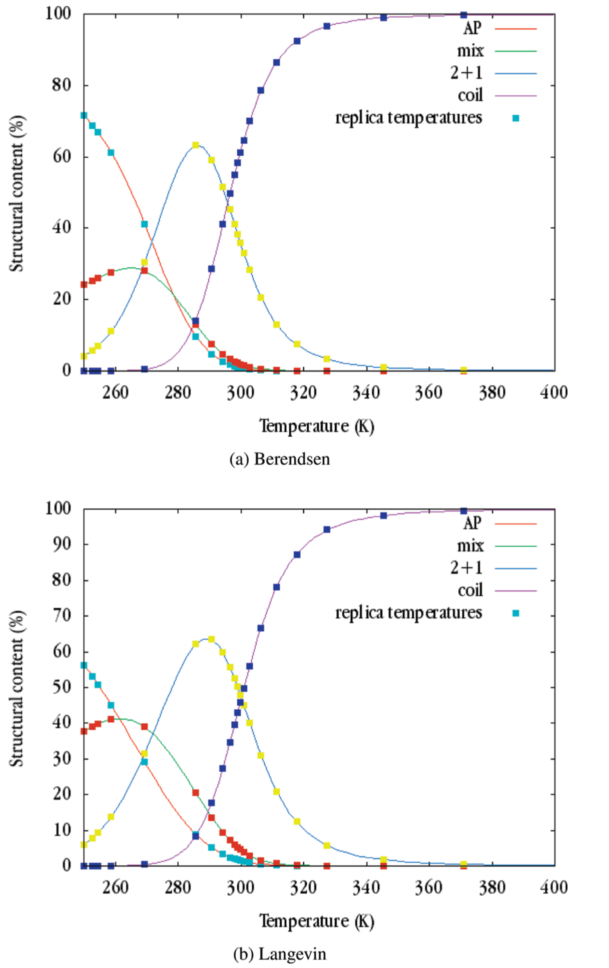
(c) mixed structure

Figure 6. Main structures encountered in the REMD simulations of $A\beta_{25-35}$ trimer at the lowest temperature. (a) Cluster analysis (images were generated by PyMOL).[39] **B** AP, first cluster of the Berendsen simulation (65% population); **L** AP, first cluster of the Langevin simulation (57%); **B** Mix, second cluster of the Berendsen simulation (16%); **L** Mix, second cluster of the Langevin simulation (16%); **L** 2 + 1, an example of a partially unfolded 2 + 1 topology (fifth cluster of Langevin simulation, 1% population). White, hydrophobic residues; green, hydrophilic residues; blue, polar residues; yellow, H-bonds. Numbers indicated are root-mean-square deviations (rmsds) in angstroms: numbers in green, rmsd on the whole backbone; numbers in blue, rmsds on the structured part only (amino acids 26−30). (b, c) Three different view angles of the (b) antiparallel and (c) mixed structures.

potential energies for both thermostats are very similar. For AP topology, we find −133.8 kcal/mol for Langevin versus −134.1 kcal/mol for Berendsen, and for 2 + 1 topology, we find −118.7 kcal/mol for Langevin versus −119.9 kcal/mol for Berendsen. However, the Berendsen simulation displays a narrower energy distribution than the Langevin simulation for each topology, with a difference in the root-mean-square fluctuation of the potential energy amounting to 2.0 kcal/mol.

The most populated clusters of each simulation at the lowest temperature are represented in Figure 6. Structural differences between the configurations sampled by the thermostats are very subtle, as reported by the cross rms deviations between the centers of the clusters. For instance, the mixed topologies or the antiparallel topologies in the two simulations deviate by less than 1 Å. Similar rms deviations are found at higher temperatures (data not shown).

We analyzed our trajectories in search of conformational events from the antiparallel to mixed topologies or vice versa. These events always involved a transient coil or partially folded (2 + 1) topology. Looking at each replica below the melting temperatures, we found the following statistics. For the Langevin simulation, we counted four events from Mix to AP, with an averaged $\tau_{\mathrm{Mix}\rightarrow\mathrm{AP}}$ of 268 ns, and five events from AP to Mix, with an averaged $\tau_{\mathrm{AP}\rightarrow\mathrm{Mix}}$ of 111 ns. For the Berendsen simulation we observed the same number of transitions as in Langevin, $\tau_{\mathrm{Mix}\rightarrow\mathrm{AP}}$ = 199 ns and $\tau_{\mathrm{AP}\rightarrow\mathrm{Mix}}$ = 89 ns. The two simulations thus behave very similarly. It should be noted that the interconversion between Mix and AP three-stranded β-sheets is slow despite the use of REMD and having optimized the replica temperatures distribution to ensure adequate exchange rates. This suggests the existence of significant potential and free energy barriers, as found also in other systems such as dimers of the $A\beta_{16-22}$[21]and GNNQQNY peptides.[40]



(a) Berendsen

(b) Langevin

Figure 7. Topological content percentage analyzed over the converged part of each $A\beta$ trimer simulation (60 000 structures): (a) Berendsen and (b) Langevin simulations.

A final comparison between the thermostats is reported in Figure 7, which compares the percentage of topological content generated by each simulation as a function of temperature. Although the distribution of each topology looks very similar for both thermostats, there is one striking difference. At low temperatures (<269 K), while the difference in the populations of the 2 + 1 topology is small (<1%), there are more antiparallel structures in the Berendsen simulation than in the Langevin simulation (see also Table 2). For instance, at 259 K (replica 4), 63% of the structures are antiparallel for Langevin and 46% for Berendsen. The reverse is true for the mixed state, where we observe a population of 26% for Berendsen against 41% for Langevin.

Around the melting temperatures, that is, between 280 and 340 K, the populations of the most ordered AP and mixed topologies

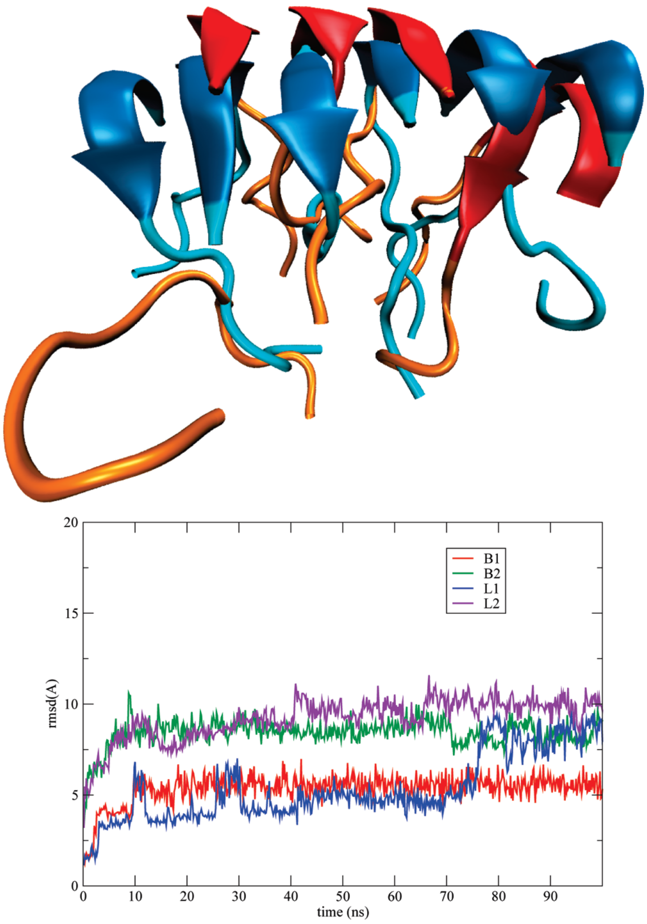


Figure 8. Aβ_{25-35} hexamer. (Top) Superposition of the initial structure on that obtained by Berendsen after 100 ns (image generated by VMD).[50]. The rmsd is 6.35 Å for the core. (Bottom) Time evolution of the Cα rmsd of all chains for residues 26−30 at 300 K, that is, a few degrees above the melting temperature predicted by Langevin and Berendsen REMD simulations.

vary little between the two thermostats. For instance, at 291 K (replica 9), the AP populations are 5.9% in one case and 6.2% in the other. The peak of the 2 + 1 topology is shifted 3 K toward higher temperatures for Langevin than Berendsen, and its height is the same in both simulations. However, its distibution is larger, with a full width at half-maximum of 36.4 K for Langevin versus 31.3 K for Berendsen. Similarly, the coil curve is shifted 3 K to the right for the Langevin simulation compared to Berendsen.

3.3. Aβ_{25-35} Hexamer. It is instructive to enlarge the part of the amyloid dynamics and stability by examining oligomers higher than the trimer. While changes in electrical fields associated with membranes and the presence of metals can play a role in Alzheimer's disease and are studied by standard or Car−Parrinello-type MD simulations,[41,42] we limit ourselves to mainly Aβ oligomers in aqueous solution.

Ma and Nussinov[43] studied the stability of Aβ_{25-35} octamers consisting of two β-sheets using short (5 ns) all-atom MD simulations at 330 K. Recent all-atom MD simulations by Shea and co-workers[19] showed that a V-shaped protofibril structure consisting of six Aβ_{25-35} peptides was stable at 310 K for 55 ns. Röhrig et al.[44] studied the stability of oligomers of Aβ_{16-22} from the dimer to the 32-mer. In all-atom MD simulations of ~30 ns at 300 and 348 K, a single-layer β-sheet of eight peptides was not stable in contrast to a two-layered octamer β-sheet, suggesting that the minimum nucleus size is on the order of 8−16 Aβ_{16-22} peptides.

It is well-established, however, that all-atom 100-ns MD simulations, while allowing one to study the stability of preformed oligomers, do not sample equilibrium structures, and one must resort to enhanced conformational technique and/or simplified protein−water representations. All-atom REMD simulations in explicit solvent showed that seven β2 m(83−89) peptides are in equilibrium between numerous topologies.[45] From different coarse-grained models, simulations pointed to the complexity of the free energy landscape of Aβ_{16-22} 6-mers and 7-mers.[46,47] Similarly, Masman et al.[48] found that pentamers of Aβ_{1-42} with fibril geometries remain stable by all-atom MD for 100 ns at 310 K, while Urbanc et al.[49] found disordered pentamers for the same system using very long discrete MD simulations.

Here, we constructed a hexamer of Aβ_{25-35} consisting of a single-layer antiparallel β-sheet configuration based on the antiparallel structure we found for the trimer, that is, with a β-sheet core spanning amino acids 26−30, and performed four MD runs of 100 ns at 300 K. Our goal is not to explore the full configurational space but rather to determine the differences between Langevin and Berendsen OPEP-MD simulations on a reasonable time scale. As can observed in Figure 8, the L_1 and B_1 runs lead to 5.2 ± 1.8 and 5.3 ± 0.8 Å rmsd, respectively, on the entire trajectory, while the L_2 and B_2 runs lead to 7.5 ± 1.2 and 6.7 ± 0.9 Å rmsd. This suggests (i) the same plasticity of the oligomers for both thermostats, characterized by the detachment of one or two external peptides from the β-sheet core, and (ii) the heterogeneity of the energy landscape, characterized by multiple isoenergetic conformations in dynamic equilibrium, consistent with many computational studies on other sequences.[31,46,47,49] As reported for the REMD simulations of the Aβ trimer, we observe similar averaged potential energies for the hexamer with both thermostats (−301 ± 15 kcal/mol for $B_1 + B_2$ vs −301 ± 20 kcal/mol for $L_1 + L_2$) and a narrower energy distribution for the Berendsen simulation.

4. CONCLUSIONS

In this study, we have determined the impact of Langevin and Berendsen thermostats on folding properties of penta-alanine and aggregation properties of the trimer of the 25−35 fragment of the Alzheimer's amyloid β protein with the OPEP coarse-grained model. Using long REMD simulations, we find small variations in the heat capacity curves for the two thermostats. There is, however, a small distortion in thermodynamic descriptions of the two systems with the Berendsen thermostat at low temperatures. While the structural contents of the folded state for the penta-alanine peptides and of the topologies for the Aβ_{25-35} trimer remain the same (rms deviations of less than 1 Å), their populations can vary by 15%. This finding is fully consistent with previous reports[7] and the physics behind this variation is the same: narrowed potential energy fluctuations modify the relative populations of the configurations. At higher temperatures, and precisely above and around the melting temperatures, REMD simulations with the Berendsen thermostat result in small effects. We can thus conclude that the aggregation properties of all the amyloids, with the OPEP force field and the Berendsen thermostat,[14,15] would be unaffected around their melting temperatures.

It is of interest to examine the results of the Aβ_{25-35} trimer simulations, since little is known about their structures experimentally, as they are transient. Contrary to other fragments of the full-length Aβ_{1-42} protein, the Aβ_{25-35} peptide has been subject to a small number of MD and REMD simulations. Wei et al.[19] and Kittner and Knecht[20] studied its dimerization in explicit water. Ma and Nussinov[43] studied the impact of N27Q substitution on the stability of preformed β-sheets, and Shea and coworkers[19] studied the stability of protofibrils with various topologies. Fu et al.[51] investigated the initial adsorption features and dynamics of Aβ_{25-35} on a single-walled carbon nanotube surface using MD in explicit solvent. Recently, Yu et al.[52] reported a hybrid computational approach to construct, search, optimize, and rank soluble micellelike Aβ_{25-35} structures with different side-chain packings at the atomic level.

Equilibration of Aβ_{25-35} trimers, as measured by the convergence of the heat capacity, is a very difficult task with the OPEP force field and REMD simulations. A replica flux-optimized temperature distribution and a larger number of replicas contribute to sampling efficiency. Our converged simulations, which total 48 μs in length, show that trimers at 300 K have an averaged β-sheet content of 12%. This content is very similar to the values found at 310 K in the monomer (13%) and the dimer (20%) by atomistic simulations.[19] The trimer configurational ensemble is primarily characterized by fully random coils or a small curved two-stranded β-sheet stabilized by a disordered peptide. Approximately 10% of the conformations consist, however, of curved small three-stranded β-sheets spanning the Ser$_{26}$-Asn-Lys-Gly-Ala$_{30}$ amino acids with mixed or fully antiparallel orientations of the chains. This impossibility to stabilize fully extended amyloid-like conformations has been also observed for dimers by Wei et al.[19] in explicit solvent. Whether the presence of straight-extended β-sheet-rich dimers and trimers in the assembly of Aβ_{25-35} peptides by use of an implicit water—atomistic protein model[53] results from finite-size effects[54] or is an artifact of the force field used remains to be determined, but the present OPEP-MD simulations of the hexamer do not reveal any extension of the small β-sheet core within 100 ns.

AUTHOR INFORMATION

Corresponding Author

*E-mail: philippe.derreumaux@ibpc.fr.

Present Addresses

†Unité de Bioinformatique Structurale, Institut Pasteur, 25 rue du Docteur Roux, 75015 Paris, France.

ACKNOWLEDGMENT

We thank Massimiliamo Bonomi and Alessandro Barducci (ETH Zurich) for useful discussions on the Langevin thermostat. Y.G.S. thanks Michael Nilges for his complementary computational support, John Chodera for his help in interfacing the MBAR library, and École Normale Supérieure for financial support. P.D. thanks CNRS and Institut Universitaire de France (IUF) for financial support.

REFERENCES

(1) Langevin, P. *C. R. Acad. Sci.* **1908**, *146*, 530–532.

(2) Schlick, T. *Molecular Modeling and Simulation: An Interdisciplinary Guide*; Interdisciplinary Applied Mathematics; Elsevier: Amsterdam, 2002; pp 435−440.

(3) McQuarrie, D. A. *Statistcal Mechanics*; Harper & Row: New York, 1976; pp 452−456.

(4) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Dinola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

(5) Morishita, T. *J. Chem. Phys.* **2000**, *113*, 2976–2982.

(6) D'Alessandro, M.; Tenenbaum, A.; Amadei, A. *J. Phys. Chem. B* **2002**, *106*, 5050–5057.

(7) Rosta, E.; Buchete, N.-V.; Hummer, G. *J. Chem. Theory Comput.* **2009**, *5*, 1393–1399.

(8) Mor, A.; Ziv, G.; Levy, Y. *J. Comput. Chem.* **2008**, *29*, 1992–1998.

(9) Khalili, M.; Liwo, A.; Jagielska, A.; Scheraga, H. A. *J. Phys. Chem. B* **2005**, *109*, 13798–13810.

(10) Snow, C. D.; Nguyen, H.; Pande, V. S.; Gruebele, M. *Nature* **2002**, *420*, 102–106.

(11) Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y.; Wriggers, W. *Science* **2010**, *330*, 341–346.

(12) Huet, A.; Derreumaux, P. *Biophys. J.* **2006**, *91*, 3829–3840.

(13) Maupetit, J.; Derreumaux, P.; Tuffery, P. *Nucleic Acids Res.* **2009**, *37*, W498–W503.

(14) Melquiond, A.; Dong, X.; Mousseau, N.; Derreumaux, P. *Curr. Alzheimer Res.* **2008**, *5*, 244–250(7).

(15) Chebaro, Y.; Mousseau, N.; Derreumaux, P. *J. Phys. Chem. B* **2009**, *113*, 7668–7675.

(16) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141–151.

(17) Trebst, S.; Troyer, M.; Hansmann, U. H. E. *J. Chem. Phys.* **2006**, *124*, 174903.

(18) Mattson, M. P.; Begley, J. G.; Mark, R. J.; Furukawa, K. *Brain Res.* **1997**, *771*, 147–153.

(19) Wei, G.; Jewett, A. I.; Shea, J.-E. *Phys. Chem. Chem. Phys.* **2010**, *12*, 3622–3629.

(20) Kittner, M.; Knecht, V. *J. Phys. Chem. B* **2010**, *114*, 15288–15295.

(21) Santini, S.; Wei, G.; Mousseau, N.; Derreumaux, P. *Structure* **2004**, *12*, 1245–1255.

(22) Nguyen, P. H.; Li, M. S.; Stock, G.; Straub, J. E.; Thirumalai, D. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 111–116.

(23) Itoh, S. G.; Okamoto, Y. *J. Phys. Chem. B* **2008**, *112*, 2767–2770.

(24) Lu, Y.; Wei, G.; Derreumaux, P. *J. Phys. Chem. B* **2011**, *115*, 1282–1288.

(25) Kaminsky, Y. G.; Marlatt, M. W.; Smith, M. A.; Kosenko, E. A. *Exp. Neurol.* **2010**, *221*, 26–37.

(26) Chebaro, Y.; Dong, X.; Laghaei, R.; Derreumaux, P.; Mousseau, N. *J. Phys. Chem. B* **2009**, *113*, 267–274.

(27) Forcellino, F.; Derreumaux, P. *Proteins: Struct., Funct., Bioinf.* **2001**, *45*, 159–166.

(28) Maupetit, J.; Tuffery, P.; Derreumaux, P. *Proteins: Struct., Funct., Bioinf.* **2007**, *69*, 394–408.

(29) Derreumaux, P.; Mousseau, N. *J. Chem. Phys.* **2007**, *126*, No. 025101.

(30) Pastor, R.; Brooks, B.; Szabo, A. *Mol. Phys.* **1988**, *65*, 1409–1419.

(31) Kim, S.; Takeda, T.; Klimov, D. K. *Biophys. J.* **2010**, *99*, 1949–1958.

(32) Nadler, W.; Meinke, J. H.; Hansmann, U. H. E. *Phys. Rev. E* **2008**, *78*, No. 061905.

(33) Katzgraber, H. G.; Trebst, S.; Huse, D. A.; Troyer, M. *J. Stat. Mech.* **2006**No. P03018.

(34) Nadler, W.; Hansmann, U. H. E. *Phys. Rev. E* **2007**, *75*, No. 026109.

(35) Sindhikara, D.; Meng, Y.; Roitberg, A. E. *J. Chem. Phys.* **2008**, *128*, No. 024103.

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

(37) Shirts, M. R.; Chodera, J. D. *J. Chem. Phys.* **2008**, *129*, No. 124105.

(38) Frishman, D.; Argos, P. *Proteins: Struct., Funct., Bioinf.* **1995**, *23*, 566–579.

(39) PyMOL, version 1.3; Schrödinger, LLC.

(40) Strodel, B.; Whittleston, C. S.; Wales, D. J. *J. Am. Chem. Soc.* **2007**, *129*, 16005–16014.

(41) Lugli, F.; Toschi, F.; Biscarini, F.; Zerbetto, F. *J. Chem. Theory Comput.* **2010**, *6*, 3516–3526.

(42) Morante, S. *Curr. Alzheimer Res.* **2008**, *5*, 508–524.

(43) Ma, B.; Nussinov, R. *Biophys. J.* **2006**, *90*, 3365–3374.

(44) Röhrig, U. F.; Laio, A.; Tantalo, N.; Parrinello, M.; Petronzio, R. *Biophys. J.* **2006**, *91*, 3217–3229.

(45) Simone, A. D.; Derreumaux, P. *J. Chem. Phys.* **2010**, *132*, No. 165103.

(46) Lu, Y.; Derreumaux, P.; Guo, Z.; Mousseau, N.; Wei, G. *Proteins: Struct., Funct., Bioinf.* **2009**, *75*, 954–963.

(47) Irbäck, A.; Mitternacht, S. *Proteins: Struct., Funct., Bioinf.* **2008**, *71*, 207–214.

(48) Masman, M. F.; Eisel, U. L. M.; Csizmadia, I. G.; Penke, B.; Enriz, R. D.; Marrink, S. J.; Luiten, P. G. M. *J. Phys. Chem. B* **2009**, *113*, 11710–11719.

(49) Urbanc, B.; Betnel, M.; Cruz, L.; Bitan, G.; Teplow, D. B. *J. Am. Chem. Soc.* **2010**, *132*, 4266–4280.

(50) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33–38.

(51) Fu, Z.; Luo, Y.; Derreumaux, P.; Wei, G. *Biophys. J.* **2009**, *97*, 1795–1803.

(52) Yu, X.; Wang, Q.; Zheng, J. *Biophys. J.* **2010**, *99*, 666–674.

(53) Cheon, M.; Chang, I.; Mohanty, S.; Luheshi, L. M.; Dobson, C. M.; Vendruscolo, M.; Favrin, G. *PLoS Comput. Biol.* **2007**, *3*, e173.

(54) Pawar, A.; Favrin, G. *PLoS ONE* **2008**, *3*, e2641.

JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Interplay of Correlation and Relativistic Effects in Correlated Calculations on Transition-Metal Complexes: The $(Cu_2O_2)^{2+}$ Core Revisited

Dimitrios G. Liakos and Frank Neese*

Lehrstuhl für Theoretische Chemie, Universität Bonn, Wegelerstrasse 12, D-53115 Bonn, Germany

S *Supporting Information*

ABSTRACT: Owing to the availability of large-scale computing facilities and the development of efficient new algorithms, wave function-based ab initio calculations are becoming more common in bioinorganic chemistry. In principle they offer a systematic route toward high accuracy. However, these calculations are by no means trivial. In this contribution we address some pertinent points through a systematic theoretical study for the equilibrium between the peroxo- and bis-(μ-oxo) isomers of the $[\{Cu(C_2H_8N_2)\}_2O_2]^{2+}$ complex. While this system is often regarded as a prototypical multireference case, we treat it with the single reference local-pair natural orbital coupled cluster method and reiterate that the multireference character in this system is very limited. A set of intermediate structures, for the interconversion between the two isomers, is calculated through a relaxed surface scan thus allowing the calculation of an energetic profile that cleanly connects the bis-(μ-oxo) and side-on peroxo minima on the ground-state potential energy surface. Only at the highest level of theory involving complete basis set extrapolation, triple excitation contributions as well as relativistic and solvent effects, the bis-(μ-oxo) isomer is found to be slightly more stable than the peroxo structure. This is in agreement with the experimental findings. The effects of basis set, triples excitation, relativity, and solvent contribution have all been analyzed in detail. Finally, the ab initio results are compared with density functional calculations using various functionals. It is demonstrated that the largest part of the discrepancies of the results reported in the literature are due to an inconsistent handling of relativistic effects, which are large in both ab initio and density functional theory calculations.

INTRODUCTION

The importance of copper enzymes hardly needs to be emphasized.[1−14] Within this class, enzymes featuring a binuclear copper active site have received significant attention. Prominent members include catechol oxidase[8] and tyrosinase[15,8−10,12] (both catalyzing the oxidation of catecholes to *o*-quinones) and hemocyanin,[13,16−22] an oxygen transportation protein. The common feature of these enzymes is the $Cu_2O_2{}^{2+}$ core in the active site. Up to six different isomers seem to be accessible for this core. Three of them (Figure 1) have been characterized spectroscopically[23−26] and crystallographycally.[23,24,26−28] Structure **A** is the μ-η^2:η^2-peroxo (side-on) isomer that will be referred to below as **P**. In this core, the Cu−Cu distance is close to 3.6 Å, and the O−O bond distance is ~1.4 Å. Structure **B**, the bis(μ-oxo) dicopper(III) isomer, is referred to as **O** below. Here the O−O distance has lengthened to 2.3 Å, which means that the O−O bond has effectively been broken, while the Cu−Cu distance is shortened to ~2.8 Å. Structure **C**, the trans μ-1,2-peroxo species, is less common but was the first motif to be observed crystallographycally.[27] All of these cores have singlet ground states. In structures **A** and **C** copper is in the formal oxidation state of 2+ (d^9 electronic configuration), while in structure **B** it is 3+. Thus, structures **A** and **C** are thought to represent magnetic coupling cases, in which case a closed shell determinant does not provide a good description of the electronic structure, and multireference approaches [or broken symmetry density functional theory, (DFT)] appear to be necessary to obtain reasonable results. Thus, the $(Cu_2O_2)^{2+}$ core has become a playground for theoreticians testing different theoretical approaches. A concise review of the literature up to 2009 has been published by Ghermann and Cramer.[7] Owing to the concept that **P** requires a multireference treatment, massive multireference calculations using the complete active space self-consistent field/complete active space second-order perturbation theory (CASSCF/CASPT2) and multireference configuration interaction (MRCI) methodologies have been undertaken.[29−34] However, as will be discussed below, structure **A** is so strongly coupled that it is outside the magnetic coupling regime, and treatments starting from a closed shell determinant are adequate. Probably the most accurate calculations to date have been performed by Cramer et al.[29,31] using the CR-CC-(2,3)[35−37] approach.

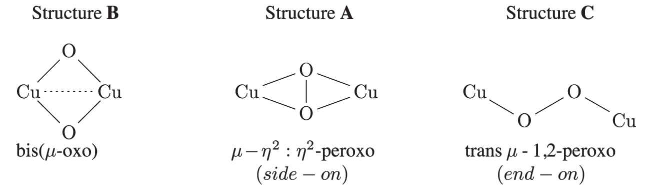


Figure 1. The three most common structures of the $CU_2O_2^{2+}$ core.

It has nevertheless become evident from the many theoretical studies performed on the system that proper theoretical

Received: December 2, 2010
Published: April 20, 2011

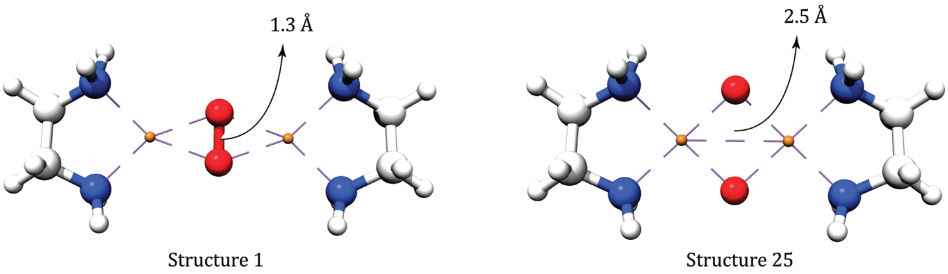


Figure 2. The structure of the first step (left) and the one after step 25 (right).

modeling of the $(Cu_2O_2)^{2+}$ core is challenging and that theoretical results with both DFT and wave function-based approaches scatter widely.[7,33,38−41] The available studies range from treatments of the bare core[29,34,40−43] up to complexes containing six histidines as terminal ligands.[44] It became evident that DFT results depend strongly on the specific form of the chosen functional.[29−32,38,39,44−52] The wave function-based calculations suffer from the fact that small models are not experimentally accessible, and hence comparison of theoretical results for the naked $(Cu_2O_2)^{2+}$ core to experiment are not possible. On the other hand, realistic size models are too large to be accurately treated with either single- or multireference wave function methods in conjunction with adequately large basis sets. At the multireference level, the CASSCF method has been used and found to be inadequate due to the lack of dynamic correlation contributions[53,54] and probably also due to excessive active space requirements. Adding part of the dynamical correlation through perturbation theory in the form of either the CASPT2[55] or the restricted active space second-order perturbation theory (RASPT2)[33] method can in principle improve the performance,[29−34] but satisfactory convergence with respect to the size of the reference space is difficult to achieve. In this respect the RASPT2 method, which allows for significantly more active space orbitals, is an important step. However, dictated by the high computational cost, the available correlated ab initio calculations that do include dynamic correlation contributions all featured relatively small double-ζ type basis sets that certainly fall short of coming close to the basis set limit. In addition, they were done on small models with ammonia model ligands. Hence, the theoretical results so far suffer from significant basis set incompleteness problems.

In this work we study the problem using the recently developed local pair natural orbital coupled cluster method.[56−58] This method can handle realistic models of actual dicopper cores while still employing quadruple-ζ size basis sets and reproducing the parent canonical correlation methods with an accuracy of 0.5 kcal/mol or better.[58] Therefore, these methods allow for reliably estimating the basis set limit of chemically relevant methods. In doing so, we have investigated a number of additional issues concerning such calculations, namely the interplay of correlation with scalar relativistic and solvent effects. As will be shown below, our most complete calculations are in full agreement with the available experiments. In order to provide a consistent set of calculations, we have also studied a number of DFT functionals and compared them to the ab initio results, with some surprising findings.

Another aspect that differs in our calculations from the literature is the use of a series of structures resulting from a relaxed surface scan along the O−O bond. Previous investigators have mainly used interpolated structures results (e.g., a rigid scan). As noted by Cramer et al.[31] such transit paths would be expected to overestimate the isomerization barrier between **P** and **O**.

The simplest type of dicopper core with saturated neutral amine ligands that avoids the complication associated with the use of ammonia (artificial hydrogen bonds, too much coordinative flexibility) contains simply ethylene diamine (*en*) as a capping ligand. Hence, we have chosen to study $[Cu_2(en)_2(O)_2]^{2+}$ as a representative model. This system is closely related to the one developed and studied by Stack and co-workers.[1,5,14,25,39,51,59,60] The experimental findings[59] demonstrate that for this system in dichloromethane solution, **O** should be the predominant species.

COMPUTATIONAL DETAILS

The ORCA suite of programs[61] was used for all calculations. A 25 point relaxed energy surface was constructed along the O−O bond distance between 1.3 and 2.5 Å using the Perdew−Burke−Ernzerhof (PBE)[62] functional together with Grimme's dispersion correction (Figure 2).[63] This way of describing the conversion between the two isomers has the advantage that the energy surface is smooth since the intermediate structures are optimized. Thus, the local maximum connecting the **P** and **O** should be a fairly good guess at the energy of the transition state describing the isomerization between the two forms.

For the wave function-based calculations the local-pair natural orbital coupled cluster (LPNO−CCSD)[57,58] method was used. The basis sets used were the def2-SVP,[64] def2-TZVP,[64,65] def2-TZVPP,[64,65] and def2-QZVP.[66] For all calculations, the def2-QZVP/C auxiliary basis set was used for the resolution of identity (RI)[67,68] approximation that was used throughout. Finally the complete basis set total energies were estimated based on a two-point extrapolation scheme that will be described in detail below.

In the absence of LPNO triples correction, canonical CCSD(T)[69,70] calculations were performed in order to estimate triple substation effects. For these calculations the def2-TZVP basis set was used for copper together with def2-SVP for the remaining atoms. This choice is dictated by computational cost.

For the DFT part of this study, the following functionals were used: B-LYP,[71,72] B3-LYP,[71−73] B1-LYP,[71,72,74] BHandHLYP,[72,73,75] and B2PLYP.[76] Motivated by the work of Siegbahn,[77] dispersion corrections were investigated according to semiempirical method developed by Grimme.[63] In these calculations, the def2-TZVP[65] basis set, together with the corresponding auxiliary basis set, was used.

For all DFT calculations, the restricted Kohn−Sham (RKS) formalism was used. The reason for this choice was that all unrestricted Kohn−Sham calculations gave identical results with

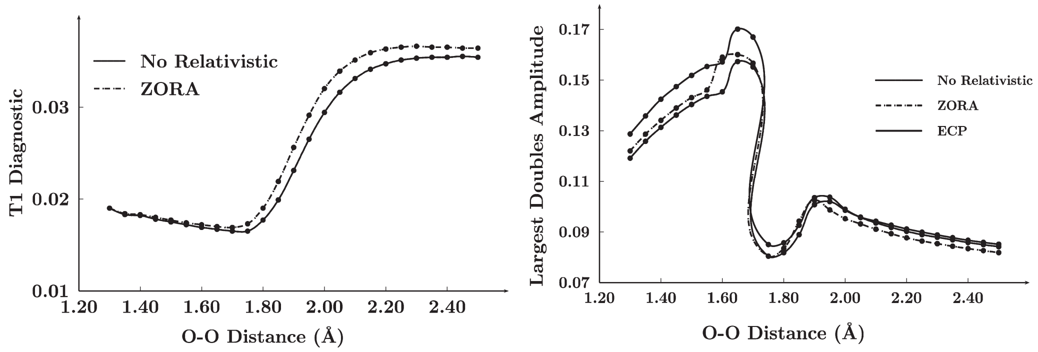


Figure 3. The T_1 diagnostic (left) and the largest doubles contribution in the wave function (right) across the isomerization coordinate, calculated with the LPNO−CCSD method.

the RKS ones (e.g., they maintained symmetry). Broken symmetry calculations were extensively examinded. The results showed that up to 20% exact exchange (EEX) (corresponding to the B3LYP functional), no broken symmetry solution was more stable than the RKS one, even in the region of the PES corresponding to **P**. Thus we feel that it is justified to focus on the results of the RKS calculations.

For both wave function- and DFT-based calculations, scalar relativistic effects were treated either explicitly [via the second-order Douglas−Kroll−Hess (DKH) transformation[78−82] or with the zeroth-order approximation for relativistic effects (ZORA)][83,84] or alternatively via effective core potentials (ECPs). In the latter case, the ECP10MWB[85,86] potential together with the corresponding basis set was used on copper. In the case of ZORA corrected calculations, the all-electron scalar relativistic basis sets described earlier were used.[87]

Solvent effects were treated with the conductor-like screening (COSMO)[88] approach,[89−91] as implemented in ORCA.[92]

WAVE FUNCTION RESULTS

Multireference Character of the $(Cu_2O_2)^{2+}$ Core. Since the LPNO−CCSD method chosen for the study is a single reference method, the question concerning the multireference nature of the wave function rises. In the literature, the T_1 diagnostic[93] is often employed to judge multireference character. The results in Figure 3 demonstrate that the T_1 diagnostic stays within reasonable bounds over the entire isomerization coordinate and slightly increases upon approaching the **O** isomer. If **P** would be a genuine multireference species (as might be expected from the formal d^9 electron configuration at the two copper(II) ions), the opposite trend would be expected. In fact, as discussed elsewhere,[94] one should distinguish between the terms 'multideterminantal' and 'muticonfigurational'. The open-shell singlet that dominates the wave function in the case of two weakly antiferromagnetically interacting d^9 sytems is multideterminantal but monoconfigurational because a single spatial configuration is involved in both determinants of the open-shell singlet. The term multiconfigurational should be reserved for cases in which different spatial configurations occur in the wave function with large weights.

In agreement with other researchers,[31] we do not think that the T_1 diagnostic is a good measure of multireference (or multideterminantal) character. In coupled cluster theory, the single excitation amplitudes essentially describe orbital relaxation, and hence, large single contributions are expected when the Hartree−Fock (HF) orbitals are poor. In the present case, the starting orbitals for **O** appear to be worse than those for **P**, which is a sensible result because the metal ligand covalency is certainly higher for the dicopper(III) species **O** compared to **P** and because HF theory is known to not provide bonds that are far too ionic.

In our opinion, a more valid criterion for multireference character is the largest double excitation amplitudes. For genuine diradicals, the largest amplitude should approach a value of unity in which case the single reference approach as such becomes invalid. Our results are shown in Figure 3. It is indeed observed that the largest doubles amplitudes occur on the **P** side of the isomerization surface. However, even there the largest double excitation amplitude does not exceed a value of 0.17.

Taken together these results imply that the single reference coupled cluster approach is very well suited for describing the $(Cu_2O_2)^{2+}$ core over the entire isomerization coordinate connecting the **P** and **O** minima. Obviously, this does not imply that other single reference methods are equally suitable for studying the $(Cu_2O_2)^{2+}$ core. Owing to the exponential Ansatz, coupled cluster theory (or its close variants) is certainly the most stable approach and tolerates much larger amounts of multireference character than, say, many body perturbation theory or configuration interaction approaches before breaking down.

Basis Set Limit Estimate. In order to approach the complete basis set (CBS) limit, large scale LPNO−CCSD calculations were performed. The usual practice to achieve this estimate is the use of a two-point extrapolation scheme. As described below, we have used various schemes for this extrapolation. The first combination of basis sets was based on the smallest possible basis set combination and involved the double-ζ/triple-ζ pair def2-SVP/def2-TZVP (referred ExtrapolationS). Second, the much more demanding triple-ζ/quadruple-ζ combination def2-TZVPP and def2-QZVP was used (referred to as ExtrapolationB). The results of this extrapolation are supposed to provide the most accurate results of this work.

The CBS extrapolated energy is estimated[95,96] according to the formula:

$$E^{(\mathrm{CBS})} + E_{\mathrm{HF}}^{(\mathrm{CBS})} + E_{\mathrm{LPNO-CCSD}}^{(\mathrm{CBS})} \qquad (1)$$

Here:

$$E_{\mathrm{HF}}^{(\mathrm{CBS})} = \frac{E_{\mathrm{HF}}^{X}\, e^{(-a\sqrt{Y})} - E_{\mathrm{HF}}^{Y}\, e^{(-a\sqrt{X})}}{e^{(-a\sqrt{Y})} - e^{(-a\sqrt{X})}} \qquad (2)$$

is the estimated CBS HF energy and

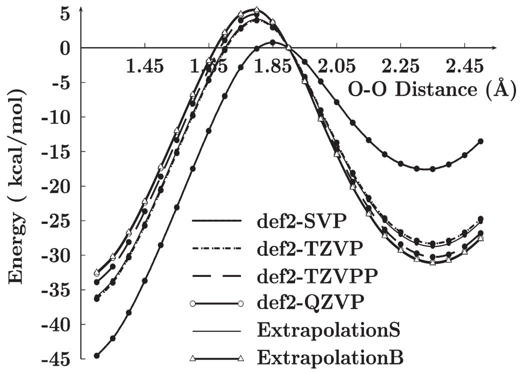


Figure 4. The HF energy calculated across the isomerization coordinate with different basis sets.

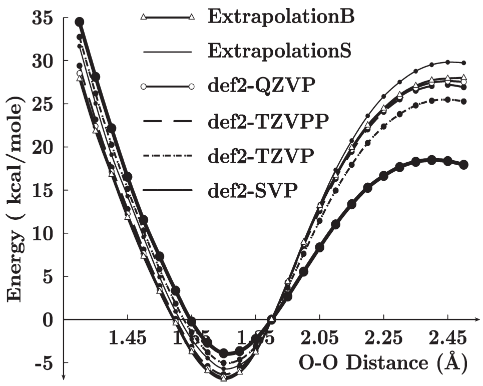


Figure 5. The LPNO−CCSD energy calculated across the interconversion coordinate with different basis sets.

$$E^{(\mathrm{CBS})}_{\mathrm{LPNO-CCSD}} = \frac{X^{\beta}E^{X}_{\mathrm{LPNO-CCSD}} - Y^{\beta}E^{Y}_{\mathrm{LPNO-CCSD}}}{X^{\beta} - Y^{\beta}} \qquad (3)$$

the estimated CBS LPNO−CCSD energy. In these equations X and Y are the smaller and larger cardinal numbers of the involved basis sets. We note in passing that the validity of extrapolating with the def2 basis sets has been investigated in ref 97 and was found to be excellent. The values of α and β used were 10.39 and 2.40 for ExtrapolationS and 7.88 and 2.971 for ExtrapolationB.[97]

In Figure 4 the SCF contribution to the total energy (without relativistic corrections and in the gas phase) is presented. It is obvious that the SCF energy converges smoothly to the CBS limit. The SCF energy has practically converged with the def2-QZVP basis set. The largest difference between the CBS energy calculated with ExtrapolationB and the energy calculated with def2-QZVP is 0.1 kcal/mol. The only result that deviates significantly from the CBS limit is, in fact, the uncorrected def2-SVP curve.

In Figure 5 the analogous results are shown for the LPNO−CCSD correlation energy. Fortunately, the correlation energy also appears to converge smoothly, and at the level of the def2-QZVP basis set, the LPNO−CCSD correlation energy has essentially converged to the CBS estimate. The largest deviation between the def2-QZVP and ExtrapolationB results is 0.65 kcal/mol, which is reasonable given that the correlation energy converges much more slowly to the basis set limit than the SCF energy.[95] However, these deviations are observed at the extreme points of the PES. In the more important regions of the PES, the deviations are again on the order of 0.2 kcal/mol.

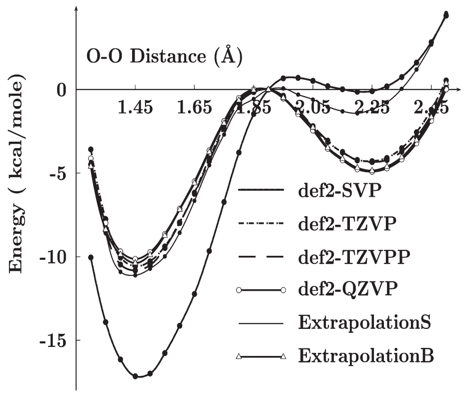


Figure 6. The total energy calculated across the interconversion coordinate with different basis sets.

Finally in Figure 6 the total PES is plotted as the sum of the previous two terms. The maximum deviation between the def2-QZVP and ExtrapolationB results is 0.48 kcal/mol. Again this occurs at the first and last points of the PES. In the more interesting relevant areas, the deviations are in the order of 0.1 kcal/mol. The deviations in the total energy are smaller than those obtained for the correlation energy alone, since SCF and correlation errors have opposite signs and tend to cancel. Nevertheless, it is concluded that with the def2-QZVP basis set, the PES has essentially converged to the CBS result.

Below, the LPNO−CCSD method is used in conjunction with ExtrapolationB in order to produce the most accurate results achievable with this methodology.

At this point a note concerning an alternative widely used extrapolation scheme is appropriate. In this scheme the CBS correlation energy is estimated on the basis of the MP2 CBS energy. The extrapolated correlation energy is obtained according to the formula:

$$E^{(\mathrm{CBS})}_{\mathrm{corr}} = E^{(\mathrm{CBS})}_{\mathrm{MP2}} + (E^{X}_{\mathrm{LPNO-CCSD}} - E^{X}_{\mathrm{MP2}}) \qquad (4)$$

Here $E^{(\mathrm{CBS})}_{\mathrm{corr}}$ is the CBS estimation for the correlation energy, calculated according to eq 3 only replacing the LPNO−CCSD energies with MP2 energies. $E^{X}_{\mathrm{LPNO-CCSD}}$ and E^{X}_{MP2} are the LPNO−CCSD and MP2 energies calculated with the small basis set. The HF energy is calculated in the same way as before. In the left part of Figure 7 the correlation energy calculated, as described above, using the cc-pVTZ/cc-pVQZ combination of basis sets (referred to as ExtrapolationC), is presented. For comparison in the same figure, the correlation energy calculated with the ExtrapolationB scheme is also plotted.

From the left part of Figure 7, ExtrapolationC strongly stabilizes **P** with respect to **O**. The stabilization of **P** over **O** is as much as 7.1 kcal/mol larger compared to what is obtained with the more rigorous ExtrapolationB scheme. It is obvious that a basis set limit estimate that introduces such a large error is useless for obtaining chemically meaningful results. The reason for this disappointing behavior can be found in the behavior of the MP2 correlation itself energy. In the right part of Figure 7, the correlation energies calculated with the canonical CCSD(T), LPNO−CCSD, and MP2 methods obtained with the same basis set are shown. While the LPNO−CCSD curve closely resembles

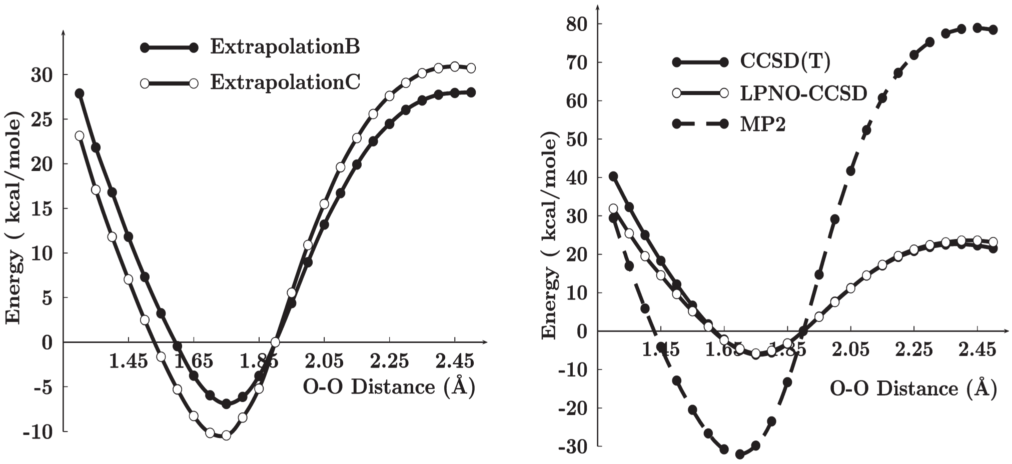


Figure 7. Complete basis set correlation energy estimates using extrapolation schemes ExtrapolationB and ExtrapolationC (left). Relative correlation energies calculated with the CCSD(T), LPNO−CCSD, and MP2 methods (right), (def2-TZVP basis set on for copper and def2-SVP for the remaining atoms).

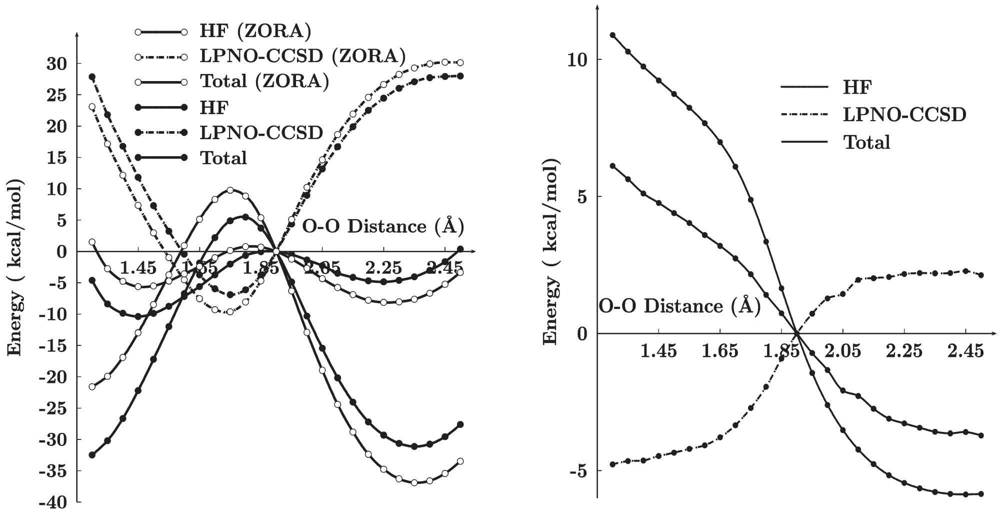


Figure 8. The PES and its components calculated with the LPNO−CCSD method and the ZORA scalar relativistic corrections together with the corresponding nonrelativistic curve (left). The effect of ZORA scalar relativistic corrections to the total energy across the interconversion coordinate (right).

the canonical CCSD(T) one, the MP2 estimate is miserable and dramatically overstabilizes **P**.

Interplay of Correlation and Relativity. Scalar relativistic effects are usually considered to be of lesser importance in the chemistry of the first transition row.[98] Nevertheless their importance has already been recognized for some time (e.g., Flock et al., ref 32). In Figure 8 the HF, LPNO−CCSD, and total energies calculated on the basis of ExtrapolationB and the inclusion of scalar relativistic ZORA corrections are shown (gas phase calculations).

The immediate conclusion from Figure 8 is that overall **O** is stabilized by about 8.4 kcal/mol relative to the **P**. The second important observation is that the net effect due to relativity arises from the interplay of two competing factors. The relativistic changes to the correlation energy work in favor of **P**, while the analogous effect on the SCF energy works in favor of **O**. The net outcome is the sum of these two contributions. Since the effect on the SCF energy is more pronounced, the latter dominates the overall relativistic correction, thus resulting in an overall stabilization of **O**. The 8.4 kcal/mol that **O** gains with respect to **P** is enough to even change the more stable minimum from **P** to **O**. The origin of the large scalar relativistic effects will be investigated below after the DFT results have been presented.

However, before proceeding to analyze the effects of relativity, the effects of perturbative triple excitations will be considered. The canonical CCSD(T) triple corrections obtained in the scalar relativistic and nonrelativistic cases are shown in Figure 9. It is evident that the relativistic corrections are fairly limited and that triple excitations slightly work in favor of **O** compared to **P**. Overall, the triples correction favors the **O** by 2.1 kcal/mol.

Solvent Effects. In this step of the investigation, solvation effects are added to the PES. They have been estimated at the level of the conductor-like screening (COSMO) model using CH_2Cl_2 as a solvent (def2-TZVPP basis set in this section). The difference between gas-phase and solvent results calculated at this level was added to the curve obtained with ExtrapolationB together with the triple substitution effects. The resulting curve is considered the most accurate result of this study and will serve as a reference for judging the DFT results. In Figure 10, the resulting PES is plotted together with the obtained solvent effect.

From the left part of the figure, the main conclusion is that **O** is more stable than **P** by 4.1 kcal/mol. The estimated transition-state energy is 8.6 kcal/mol higher above **O**. This result would imply facile interconversion of the two isomers with the thermodynamic equilibrium being significantly on the side of **O**. The true transition state must be slightly lower than the saddle point on the PES, as this point is obtained from a constraint optimization.

The additional stabilization of **O** relative to **P** is apparent from the right-hand side of Figure 10. This makes sense as **O** has the constituent atoms in higher formal oxidation states and is more compact than **P**. Both factors are thought to contribute to the extra stabilization of this dication.[7] The size of the effect is as large as 6 kcal/mol. Thus, the net result that **O** is more stable than **P** is caused by a combination of relativistic and solvent effects.

DENSITY FUNCTIONAL THEORY

Since the largest part of the literature concerning computational studies on dicopper complexes is done with DFT, a detailed study of the factors affecting the outcome of these calculations is presented below. The factors studied are the exact exchange contribution in the functional, dispersion forces, scalar relativity, and finally solvent effects.

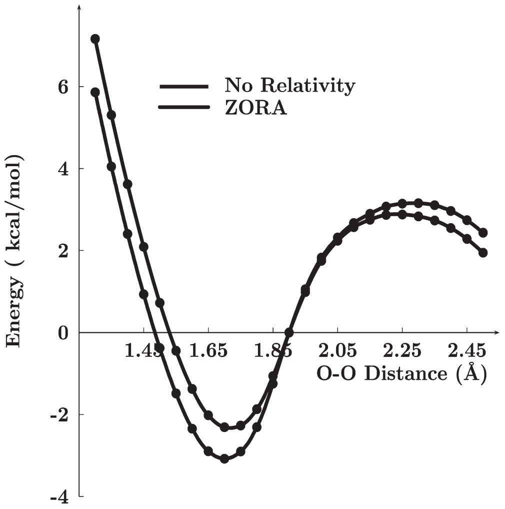


Figure 9. The correlation energy recovered by the perturbative triples correction across the isomerization coordinate calculated without relativistic corrections and after the ZORA correction.

Effect of Exact Exchange of the Functionals. In order to study the effect of exact exchange (EEX) in the calculated energies, a set of different functionals with varying EEX was used. The functionals were: BLYP[71,72] (0% EEX), B3LYP[71−73] (20% EEX), B1LYP[71,72,74] (25% EEX), BHandHLYP[72,73,75] (50% EEX), and finally also B2PLYP[76] (53% EEX). The choice of these functionals was made based on the fact that B3LYP is the most popular functional in current use and that the remaining functionals (except B2PLYP) use the same components and differ mainly in the fraction of EEX. The double hybrid B2PLYP functional also includes a MP2 correction that brings in semilocal correlation effects. The results of our calculations with these functionals in the gas phase and without corrections for relativistic or dispersion forces are presented in Figure 11.

It is obvious that the DFT results depend strongly on the fraction of EEX. In the right-hand part of Figure 11, the energy difference between the **O** and **P** minima is shown. Obviously, the **O** − **P** energy difference is almost linear to the fractional EEX. This was first noted by Rode et al.[99] and by Cramer et al.[29,31] and is extended here to the range of 0−50% EEX. A comparison reveals that B3LYP and B1LYP are in best agreement with LPNO−CCSD results (calculated with ExtrapolationB and no relativistic corrections), with B1LYP being slightly preferred. The BLYP functional erroneously predicts the wrong isomer. The B2PLYP functional does not seem to perform well in this application as it gives a much too high energy value for the **O** isomer. In fact it does not even predict a minimum for this species. This must be attributed to the badly failing MP2 component in the B2PLYP energy.

Weak Interactions — D. The correction due to Grimme[63] that has been shown to improve the results of DFT calculations[100,101] was investigated in this part. Parameters are only available for BLYP, B3LYP, and B2PLYP. Hence B1LYP and BHandHLYP were not investigated in this section.

In the left part of Figure 12, the PES calculated with the different functionals after the inclusion of dispersion forces correction is presented. In the right part the effect of these corrections in the functionals is plotted, and one can see there that the addition of dispersion forces correction does not have a significant effect on the relative stability of the two isomers. The

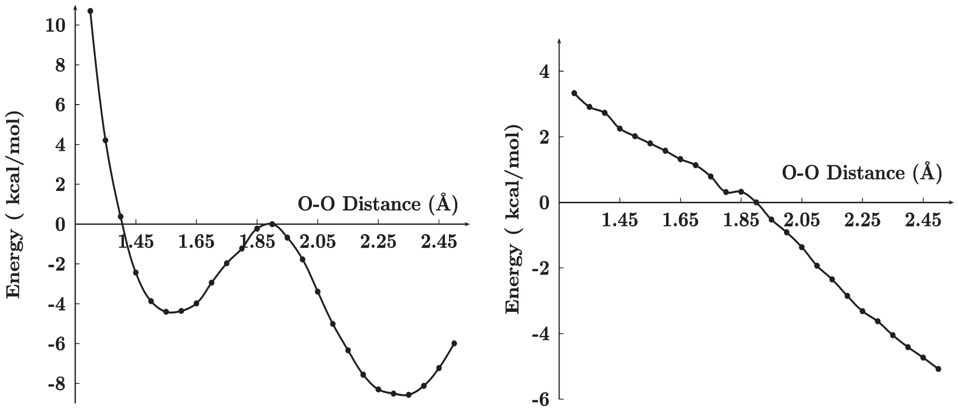


Figure 10. The total energy calculated across the isomerization coordinate estimated at the CBS limit including solvent effects, triples correlation energy, and ZORA relativistic corrections (left). The effect of COSMO on the energy calculated with the LPNO−CCSD method and the def2-TZVPP basis set, including ZORA corrections (right).

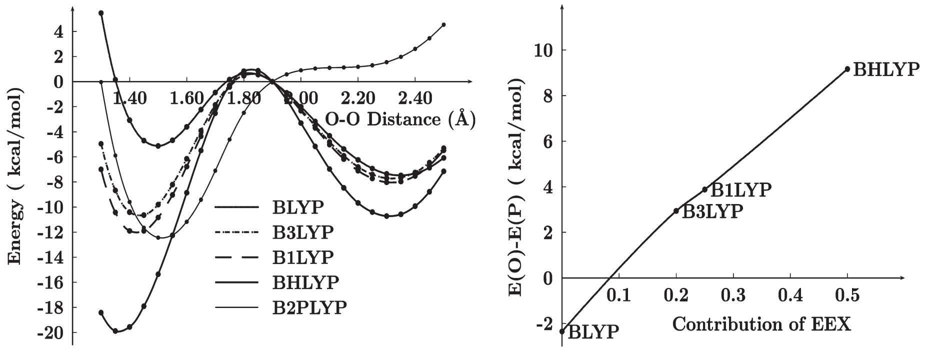


Figure 11. Energy path, following the isomerization coordinate, calculated with various functionals (left). Energy difference $E(\mathbf{O}) - E(\mathbf{P})$ plotted against the percentage of EEX contribution in the functional(right).

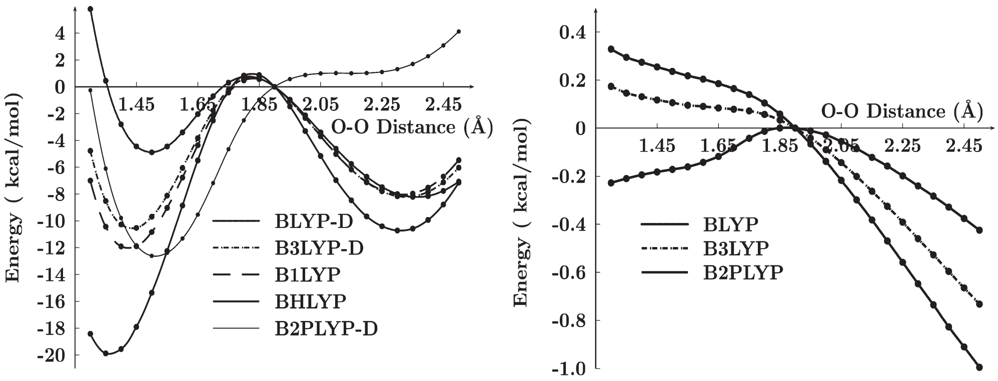


Figure 12. Single point energies calculated across the isomerization coordinate calculated with the correction for dispersion forces (left). Effect of dispersion forces correction (right).

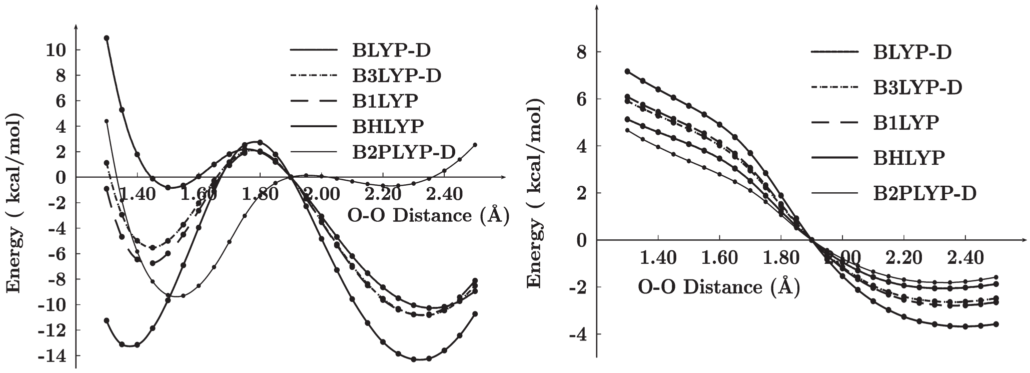


Figure 13. The potential energy using the density functional methods corrected for relativistic effects through ZORA (left part). The effect of ZORA (right part).

quantitative form of the PES remains unchanged, while the largest change in the relative energies is not larger than 1 kcal/mol. However, this result should not be overemphasized, as it has recently been shown that in $(Cu_2O_2)^{2+}$ complexes with large ligands, dispersion forces can be significant.[77]

Relativistic Effects. In most of the previous works, relativistic effects were included either implicitly through ECPs or have been ignored. Here an effort is made to systematically investigate the size of these effects. In order to accomplish this, three different approaches were used (the two scalar relativistic corrections, ZORA[83,84,102] and DKH[78−82] as well as the ECP ECP10MWB).[85,86] In Figure 13 the effect of the ZORA relativistic correction is presented, in Figure 14 the effect of the DKH correction, and finally in Figure 15 the effect of the use of an ECP for Cu.

It is obvious from these figures that the effect of relativity is as pronounced, as in the LPNO−CCSD case, and results in a net stabilization of the **O** isomer. As in the LPNO−CCSD case, relativity switches the order of stability of the two isomers. ZORA and DKH are almost indistinguishable. ECPs lead to a less pronounced relativistic effect. Despite the fact that again **O** is stabilized with respect to **P**, the stabilization is not large enough

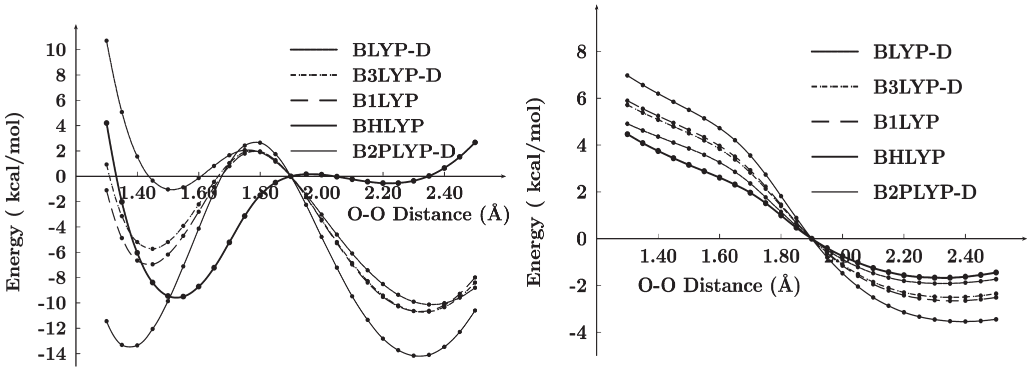


Figure 14. The potential energy using the density functional methods corrected for relativistic effects through DKH (left part). The effect of DKH (right part).

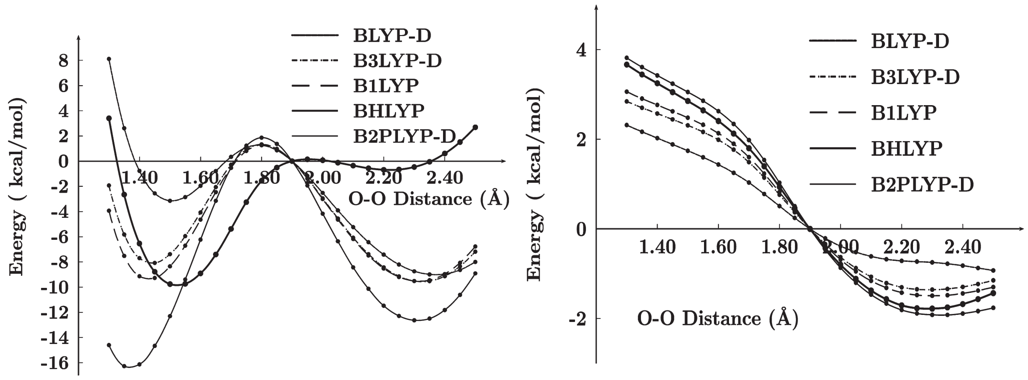


Figure 15. The potential energy using the density functional methods calculated using the quasirelativistic ECP ECP10MWB for copper.

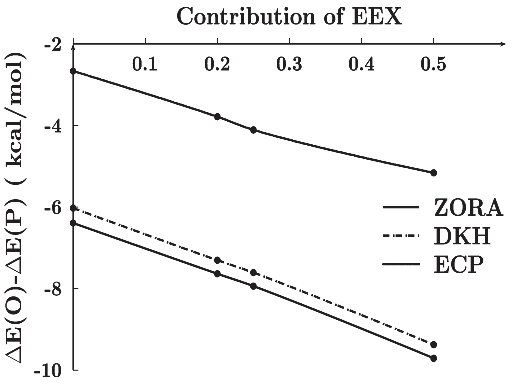


Figure 16. The relative stabilization of **O** with respect to **P** due to relativistic corrections calculated with BLYP, B3-LYP, B1-LYP, and BHandHLYP density functionals.

to change the position of the global minimum. It appears that a little less than half of the scalar relativistic effects are recovered in the ECP calculations (Figure 16).

Interestingly, the relativistic effect on the relative stability of the two isomers also strongly depends on the functional used. There appears to be a nearly linear correlation between the EEX contribution in the functional and the size of the relativistic corrections. This may well be related to the changes in metal−ligand covalency. With increasing EEX, the metal ligand bonds become more ionic. Thus, the d-electron count increases for both isomers with increasing EEX. As **O** has the lower formal d-electron count, it is expected to be increasingly stabilized relative to **P** as EEX increases.

Compared to the value of 6.8 kcal/mol calculated at the CBS limit with ZORA, LPNO−CCSD after triples correction, the best result is delivered by BLYP-D which predicts an energy difference of 6.4 kcal/mol. B3LYP-D is also excellent and predicts a value of 7.6 kcal/mol.

Solvent Effects. The resulting COSMO corrected PESs are shown in Figure 17.

The results follow the same pattern already found on the LPNO−CCSD calculations thus favoring **O**. In quantitative terms, BLYP-D provides the best result and predicts a stabilization of 7.4 kcal/mol relative to the reference value of 6.4 kcal/mol. B3LYP-D is slightly worse and gives a value of 8.2 kcal/mol.

■ ORIGIN OF RELATIVISTIC EFFECTS

The origin of the significant relativistic effects is investigated by looking at the changes that occur in the molecular orbitals due to relativity. This is reasonable as the net relativistic effect is dominated by the changes in the SCF energy. The effects of scalar relativity on the B3LYP orbitals 42 up to 77 [the highest occupied molecular orbital (HOMO) is orbital 69] are investigated in Figure 18, where the energy difference of the individual orbitals with and without the ZORA correction is plotted.

The plot can be divided into four regions depending on which fragments dominate the molecular orbitals within a given region. The deeper occupied valence orbitals are mainly ligand in character (see Table S1 of the Supporting Information). As seen in the left of Figure 18, the relativistic effects on these orbitals are

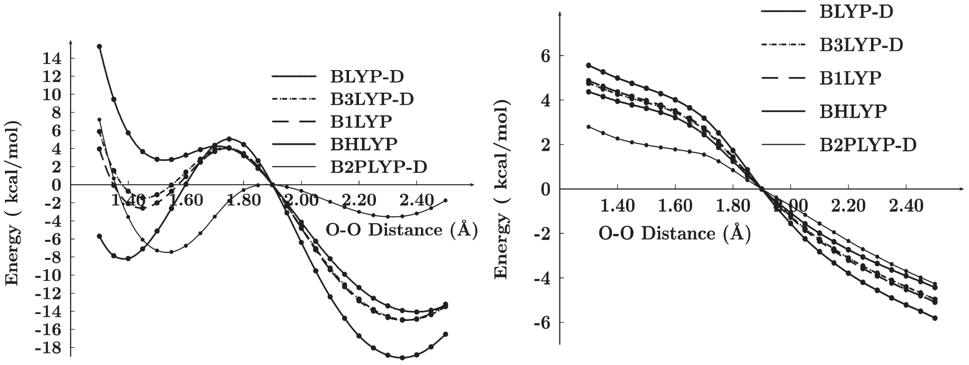


Figure 17. The potential energy using the density functional methods corrected for relativistic effects through ZORA, for dispersion forces through D and for solvent effects through COSMO, with CH_2Cl_2 solvent (left part). The effect of solvent (right part).

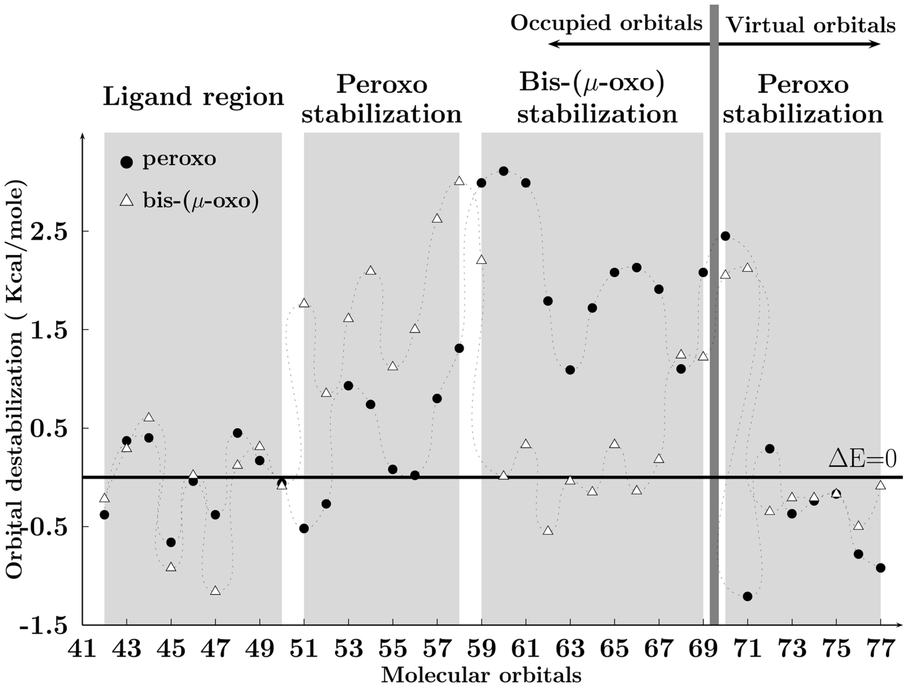


Figure 18. Changes in the B3LYP/def2-TZVP orbital energies due to ZORA for the two isomers. Orbitals are arbitrarily linked by wavy lines to facilitate visual comparison.

small and, more importantly, are almost identical for the two isomers. Hence these orbitals do not contribute to the preferential stabilization of **O** relative to **P**. Summed over all orbitals in that region, a net 0.9 kcal/mol stabilization of **O** results which is probably within the noise.

The second region is defined by orbitals 51−58. Here one finds that the **P** orbitals are preferentially stabilized. Summed over all contributions results in a value of 11.5 kcal/mol. From Table S2 of the Supporting Information it becomes evident that in this region the strongly stabilized copper d-orbitals of **O** are located, while for **P** these orbitals are mainly of ligand s and p character. Since d-orbitals are stabilized by relativity, this region provides relativistic corrections in favor of **P**.

The opposite case is met in region 3, the upper valence region. Here copper d-orbitals dominate for **P**, while for **O** the highest occupied orbitals are mainly of ligand s and p character. Since the latter are little affected by relativity, the sum over this region provides 18.4 kcal/mol relativistic effect in favor of **O**.

Thus the net relativistic effect can be traced back to the stronger destabilization of **P** relative to **O**. Summed over all molecular orbitals, a value of 7.8 kcal/mol in favor of **O** is obtained which is sufficiently close to the net relativistic stabilization energy of 5.3 kcal/mol. The preferential bias in favor of **O** is explained by the lower formal d-count (d^8 for **O** vs d^9 for **P**) and by the fact that **P** has less covalent metal ligand bonds and hence feels more of the d-orbital destabilization than **O**.

One final note concerns the effect in the lower lying unoccupied orbitals. In Figure 18 orbitals 70−77 have been included. It is apparent that **P** should be favored with respect to **O** as its orbitals are preferentially less destabilized than those of **O**. These decreased gaps between occupied and unoccupied orbitals will tend to increase the correlation energy. This may well explain the opposite trend for the correlation energies obtained in Figure 8.

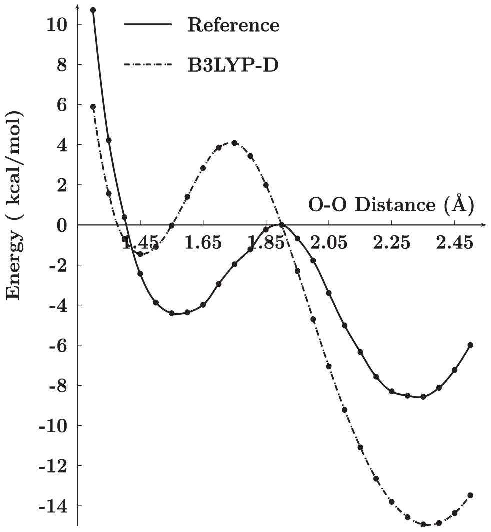


Figure 19. The most accurately calculated PES and the corresponding one calculated with B3LYP-D, ZORA and COSMO corrections using def2-TZVP basis set.

COMPARISON OF COUPLED CLUSTER AND DENSITY FUNCTIONAL CALCULATIONS

As reference curve, the scalar-relativistic CBS extrapolated LPNO−CCSD result together with solvent and triple-excitation corrections is used. Relative to this reference, the most accurate functional is B3LYP-D with a mean absolute error (MAE) of 4.4 kcal/mol, followed by B1LYP (4.5 kcal/mol) and BLYP-D (4.8 kcal/mol). In Figure 19 the reference PES is shown next to B3LYP-D. While the overall shape of the PES is correct with B3LYP-D, the quantitative performance is not quite satisfactory: The energy difference, between the two isomers, is overestimated with B3LYP by as much as 9.3 kcal/mol, and the position of the minimum for **P** is calculated to be 0.1 Å too short. Thus, none of the functionals can be considered as satisfactory if one gives preference to the wave function-based ab initio reference curve.

CONCLUSION

In this paper we have discussed some aspects of applying correlated wave function-based ab initio methods to transition-metal chemistry. To illustrate this subject, a fairly detailed study on the equilibrium between the bis-(μ-oxo) and peroxo isomers of $[Cu_2(en)_2(O)_2]^{2+}$ has been reported. A relaxed energy surface scan provided the reaction coordinate for the interconversion of the two isomers. Using the LPNO−CCSD method, the **O** isomer was found to be more stable than **P**, in agreement with the experimental results. The effect of relativity was found to be important, favoring the **O** isomer. This trend was consistent for the three different methods that were used to calculate the relativistic corrections, namely ZORA, DKH, and quasirelativistic ECP. The results for DKH and ZORA were almost identical. About half of the relativistic effect was obtained with ECPs. While this is somewhat displeasing, we note that ECPs also do not lead to dramatic computational savings over scalar relativistic all electron calculations. For this reason we prefer the latter. The effect of solvent (within the limits of the dielectric continuum model) was found to be significant as well and helps stabilizing **O** over **P** to an extent that mirrors the relativistic effect. Thus theoretical studies neglecting these two contributions to the **O**/**P** equilibrium are grossly wrong.

Flock et al.[32] in an early study on the **O**−**P** interconversion did include relativistic effects in their description but did not consider solvent effects. The same strategy was followed later by Rode and Werner[99] who included relativistic effects, though through the use of ECP's, but also left out solvent effects. Siegbahn in a very rigorous and detailed study[77] calculated the B3LYP energy difference between **O**/**P** and found it to be 8.3 kcal/mol. This result is in excellent agreement with the best ab initio results calculated in this work. However, the coincidentally perfect agreement is partly due to a cancellation of errors. Altough all necessary corrections were considered, Siegbahns calculations used the lacvp relativistic core potential. According to the calculations presented above, this implies that the stabilization of **O**, due to relativistic effects, is probably underestimated. In addition the lack of dispersion corrections also favors **P** relatively to **O**. Overall this would mean that some 4−5 kcal/mol of stabilization energy in favor of **O** should be added that would then lead to total stabilization close to 14 kcal/mol for **O**, which is in good agreement with our own B3LYP calculations. We note in passing that we agree with the comment of Cramer et al.,[31] that larger ligands (as mostly used in actual chemical studies) will greatly reduce the importance of solvent corrections.

Five different functionals were compared to the LPNO−CCSD results (BLYP, B3LYP, B2LYP, BHandHLYP, and B2PLYP). With the exception of B2PLYP, they mainly differ in the percentage of exact exchange that is incorporated. In agreement with the literature,[7,99] the addition of EEX systematically favors the **O** isomer with the stabilization being linearly proportional to the fractional EEX in the functional. The best agreement with the LPNO−CCSD results was obtained with the B3LYP functional once dispersion corrections were included. The double hybrid B2PLYP is not successful in this application because of the disastrous failure of MP2 for this system. However, none of the functionals can be considered to be in truly satisfactory agreement with the reference ab initio results. This clearly speaks in favor of continued efforts in the development of wave function-based ab initio results for applications in coordination and bioinorganic chemistry. The recent progress in the methodology is certainly encouraging in this respect.

ASSOCIATED CONTENT

S Supporting Information. The shape and composition of the molecular orbitals used for the analysis of the relativistic effects together with the Cartesian coordinates for all structures described. This material is available free of charge via the Internet at http://pubs.acs.org/.

AUTHOR INFORMATION

Corresponding Author

*E-mail: neese@thch.uni-bonn.de.

REFERENCES

(1) Mahadevan, V.; Gebbink, R. J. M. K.; Stack, T. D. P. Biomimetic modeling of copper oxidase reactivity. *Curr. Opin. Chem. Biol.* **2000**, *4*, 228–234.

(2) Sorrell, T. N. Synthetic models for binuclear copper proteins. *Tetrahedron* **1989**, *45*, 3–68.

(3) Solomon, E.; Lowery, M. Electronic structure contributions to function in bioinorganic chemistry. *Science* **1993**, *259*, 1575–1581.

(4) L. Holland, P.; B. Tolman, W. Dioxygen activation by copper sites: relative stability and reactivity of (μ-η^2:η^2-peroxo)- and bis(μ-oxo) dicopper cores. *Coord. Chem. Rev.* **1999**, *190−192*, 855–869.

(5) Mirica, L. M.; Ottenwaelder, X.; Stack, T. D. P. Structure and Spectroscopy of Copper−Dioxygen Complexes. *Chem. Rev.* **2004**, *104*, 1013–1046.

(6) Lewis, E. A.; Tolman, W. B. Reactivity of Dioxygen−Copper Systems. *Chem. Rev.* **2004**, *104*, 1047–1076.

(7) Gherman, B. F.; Cramer, C. J. Quantum chemical studies of molecules incorporating a $Cu_2O_2^{2+}$ core. *Coord. Chem. Rev.* **2009**, *253*, 723–753.

(8) Solomon, E. I.; Sundaram, U. M.; Machonkin, T. E. Multicopper Oxidases and Oxygenases. *Chem. Rev.* **1996**, *96*, 2563–2606.

(9) Solomon, E. I.; Baldwin, M. J.; Lowery, M. D. Electronic structures of active sites in copper proteins: contributions to reactivity. *Chem. Rev.* **1992**, *92*, 521–542.

(10) Klinman, J. P. Mechanisms Whereby Mononuclear Copper Proteins Functionalize Organic Substrates. *Chem. Rev.* **1996**, *96*, 2541–2562.

(11) Schindler, S. Reactivity of Copper(I) Complexes Towards Dioxygen. *Eur. J. Inorg. Chem.* **2000**, *2000*, 2311–2326.

(12) Kitajima, N.; Moro-oka, Y. Copper-Dioxygen Complexes. Inorganic and Bioinorganic Perspectives. *Chem. Rev.* **1994**, *94*, 737–757.

(13) Solomon, E. I.; Tuczek, F.; Root, D. E.; Brown, C. A. Spectroscopy of Binuclear Dioxygen Complexes. *Chem. Rev.* **1994**, *94*, 827–856.

(14) Cole, A. P.; Mahadevan, V.; Mirica, L. M.; Ottenwaelder, X.; Stack, T. D. P. Bis(μ-oxo)dicopper(III) Complexes of a Homologous Series of Simple Peralkylated 1,2-Diamines: Steric Modulation of Structure, Stability, and Reactivity. *Inorg. Chem.* **2005**, *44*, 7345–7364.

(15) Land, E. J.; Ramsden, C. A.; Riley, P. A. Tyrosinase Autoactivation and the Chemistry of ortho-Quinone Amines. *Acc. Chem. Res.* **2003**, *36*, 300–308.

(16) Karlin, K. Metalloenzymes, structural motifs, and inorganic models. *Science* **1993**, *261*, 701–708.

(17) Holm, R. H.; Kennepohl, P.; Solomon, E. I. Structural and Functional Aspects of Metal Sites in Biology. *Chem. Rev.* **1996**, *96*, 2239–2314.

(18) Kitajima, N.; Fujisawa, K.; Fujimoto, C.; Morooka, Y.; Hashimoto, S.; Kitagawa, T.; Toriumi, K.; Tatsumi, K.; Nakamura, A. A new model for dioxygen binding in hemocyanin. Synthesis, characterization, and molecular structure of the .mu.-.eta.2:.eta.2 peroxo dinuclear copper(II) complexes, [Cu(HB(3,5-R2pz)3)]2(O2) (R = isopropyl and Ph). *J. Am. Chem. Soc.* **1992**, *114*, 1277–1291.

(19) Kodera, M.; Katayama, K.; Tachi, Y.; Kano, K.; Hirota, S.; Fujinami, S.; Suzuki, M. Crystal Structure and Reversible O2-Binding of a Room Temperature Stable μ-η2:η2-Peroxodicopper(II) Complex of a Sterically Hindered Hexapyridine Dinucleating Ligand. *J. Am. Chem. Soc.* **1999**, *121*, 11006–11007.

(20) Magnus, K. A.; Ton-That, H.; Carpenter, J. E. Recent Structural Work on the Oxygen Transport Protein Hemocyanin. *Chem. Rev.* **1994**, *94*, 727–735.

(21) Gamez, P.; Koval, I. A.; Reedijk, J. Bio-mimicking galactose oxidase and hemocyanin, two dioxygen-processing copper proteins. *Dalton Trans.* **2004**, 4079–4088.

(22) Takano, Y.; Yamaguchi, K. Hybrid density functional study of ligand coordination effects on the magnetic couplings and the dioxygen binding of the models of hemocyanin. *Int. J. Quantum Chem.* **2007**, *107*, 3103–3119.

(23) Halfen, J. A.; Mahapatra, S.; Wilkinson, E. C.; Kaderli, S.; Young, V. G., Jr.; Que, L., Jr.; Zuberbuhler, A. D.; Tolman, W. B. Reversible Cleavage and Formation of the Dioxygen O—O Bond Within a Dicopper Complex. *Science* **1996**, *271*, 1397–1400.

(24) Mahapatra, S.; Halfen, J. A.; Wilkinson, E. C.; Pan, G.; Wang, X.; Young, V. G.; Cramer, C. J.; Que, L.; Tolman, W. B. Structural, Spectroscopic, and Theoretical Characterization of Bis(μ-oxo)dicopper Complexes, Novel Intermediates in Copper-Mediated Dioxygen Activation. *J. Am. Chem. Soc.* **1996**, *118*, 11555–11574.

(25) DuBois, J. L.; Mukherjee, P.; Collier, A. M.; Mayer, J. M.; Solomon, E. I.; Hedman, B.; Stack, T. D. P.; Hodgson, K. O. Cu K-Edge XAS Study of the [Cu2(μ-O)2] Core: Direct Experimental Evidence for the Presence of Cu(III). *J. Am. Chem. Soc.* **1997**, *119*, 8578–8579.

(26) Tyeklar, Z.; Jacobson, R. R.; Wei, N.; Murthy, N. N.; Zubieta, J.; Karlin, K. D. Reversible reaction of dioxygen (and carbon monoxide) with a copper(I) complex. X-ray structures of relevant mononuclear Cu(I) precursor adducts and the trans-(μ-1,2-peroxo)dicopper(II) product. *J. Am. Chem. Soc.* **1993**, *115*, 2677–2689.

(27) Jacobson, R. R.; Tyeklar, Z.; Farooq, A.; Karlin, K. D.; Liu, S.; Zubieta, J. A copper-oxygen (Cu2-O2) complex. Crystal structure and characterization of a reversible dioxygen binding system. *J. Am. Chem. Soc.* **1988**, *110*, 3690–3692.

(28) Kitajima, N.; Fujisawa, K.; Morooka, Y.; Toriumi, K. μ-η^2:η^2-Peroxo binuclear copper complex, [Cu(HB(3,5-(Me2CH)2pz)3)]2(O2). *J. Am. Chem. Soc.* **1989**, *111*, 8975–8976.

(29) Cramer, C. J.; Kinal, A.; Wloch, M.; Piecuch, P.; Gagliardi, L. Theoretical Characterization of End-On and Side-On Peroxide Coordination in Ligated Cu2O2Models. *J. Phys. Chem. A* **2006**, *110*, 11557–11568.

(30) Cramer, C. J.; Smith, B. A.; Tolman, W. B. Ab Initio Characterization of the Isomerism between the μ-η^2:η^2-Peroxo- and Bis(μ-oxo)dicopper Cores. *J. Am. Chem. Soc.* **1996**, *118*, 11283–11287.

(31) Cramer, C. J.; Wloch, M.; Piecuch, P.; Puzzarini, C.; Gagliardi, L. Theoretical Models on the Cu2O2 Torture Track: Mechanistic Implications for Oxytyrosinase and Small-Molecule Analogues. *J. Phys. Chem. A* **2006**, *110*, 1991–2004.

(32) Flock, M.; Pierloot, K. Theoretical Study of the Interconversion of O2-Binding Dicopper Complexes. *J. Phys. Chem. A* **1998**, *103*, 95–102.

(33) Malmqvist, P. A.; Pierloot, K.; Shahi, A. R. M.; Cramer, C. J.; Gagliardi, L. The restricted active space followed by second-order perturbation theory method: Theory and application to the study of CuO_2 and Cu_2O_2 systems. *J. Chem. Phys.* **2008**, *128*, 204109–10.

(34) Yanai, T.; Kurashige, Y.; Neuscamman, E.; Chan, G. K.-L. Multireference quantum chemistry through a joint density matrix renormalization group and canonical transformation theory. *J. Chem. Phys.* **2010**, *132*, 024105–9.

(35) Kowalski, K.; Piecuch, P. The method of moments of coupled-cluster equations and the renormalized CCSD[T], CCSD(T), CCSD(TQ), and CCSDT(Q) approaches. *J. Chem. Phys.* **2000**, *113*, 18–35.

(36) Piecuch, P.; Kowalski, K.; Pimienta, I. S. O.; Fan, P. D.; Lodriguito, M.; McGuire, M. J.; Kucharski, S. A.; Kuś, T.; Musiał, M. Method of moments of coupled-cluster equations: a new formalism for designing accurate electronic structure methods for ground and excited states. *Theor. Chem. Acc.* **2004**, *112*, 349–393.

(37) Piecuch, P.; Wloch, M.; Gour, J. R.; Kinal, A. Single-reference, size-extensive, non-iterative coupled-cluster approaches to bond breaking and biradicals. *Chem. Phys. Lett.* **2006**, *418*, 467–474.

(38) Saito, T.; Kataoka, Y.; Nakanishi, Y.; Matsui, T.; Kitagawa, Y.; Kawakami, T.; Okumura, M.; Yamaguchi, K. Theoretical studies on chemical bonding between Cu(II) and oxygen molecule in type 3 copper proteins. *Int. J. Quantum Chem.* **2009**, *109*, 3649–3658.

(39) Op't Holt, B. T.; Vance, M. A.; Mirica, L. M.; Heppner, D. E.; Stack, T. D. P.; Solomon, E. I. Reaction Coordinate of a Functional Model of Tyrosinase: Spectroscopic and Computational Characterization. *J. Am. Chem. Soc.* **2009**, *131*, 6421–6438.

(40) Kong, L.; Nooijen, M. Study of energetics of end-on and side-on peroxide coordination in ligated Cu2O2 models with State-Specific Equation of Motion Coupled Cluster Method. *Int. J. Quantum Chem.* **2008**, *108*, 2097–2107.

(41) Kurashige, Y.; Yanai, T. High-performance ab initio density matrix renormalization group method: Applicability to large-scale multireference problems for metal compounds. *J. Chem. Phys.* **2009**, *130*, 234114.

(42) Maddaluno, J.; Giessner-Prettre, C. Nonempirical calculations on dicopper(1+)-dioxygen: a possible model for oxyhemocyanin and oxytyrosinase active sites. *Inorg. Chem.* **1991**, *30*, 3439–3445.

(43) Bernardi, F.; Bottoni, A.; Casadio, R.; Fariselli, P.; Rigo, A. An ab initio study of the dioxygen binding site of hemocyanin: A comparison between CASSCF, CASPT2, and DFT approaches. *Int. J. Quantum Chem.* **1996**, *58*, 109–119.

(44) Siegbahn, P. E. M.; Wirstam, M. Is the Bis-μ-Oxo Cu2(III,III) State an Intermediate in Tyrosinase? *J. Am. Chem. Soc.* **2001**, *123*, 11819–11820.

(45) Aboelella, N. W.; Gherman, B. F.; Hill, L. M. R.; York, J. T.; Holm, N.; Young, V. G.; Cramer, C. J.; Tolman, W. B. Effects of Thioether Substituents on the O2 Reactivity of β-Diketiminate−Cu(I) Complexes: Probing the Role of the Methionine Ligand in Copper Monooxygenases. *J. Am. Chem. Soc.* **2006**, *128*, 3445–3458.

(46) Berces, A. Ligand Effects in the Models and Mimics of Oxyhemocyanin and Oxytyrosinase. A Density Functional Study of Reversible Dioxygen Binding and Reversible O—O Bond Cleavage. *Inorg. Chem.* **1997**, *36*, 4831–4837.

(47) Bérces, A. Density functional calculations of dioxygen binding in mono- and dinuclear copper complexes. *Int. J. Quantum Chem.* **1997**, *65*, 1077–1086.

(48) Lam, B. M. T.; Halfen, J. A.; Young, V. G.; Hagadorn, J. R.; Holland, P. L.; Lledos, A.; Cucurull-Sanchez, L.; Novoa, J. J.; Alvarez, S.; Tolman, W. B. Ligand Macrocycle Structural Effects on Copper-Dioxygen Reactivity. *Inorg. Chem.* **2000**, *39*, 4059–4072.

(49) Liu, X.-Y.; Palacios, A. A.; Novoa, J. J.; Alvarez, S. Framework Bonding and Coordination Sphere Rearrangement in the M2 × 2 Cores of Synthetic Analogues of Oxyhemocyanin and Related Cu and Pt Complexes. *Inorg. Chem.* **1998**, *37*, 1202–1212.

(50) Metz, M.; Solomon, E. I. Dioxygen Binding to Deoxyhemocyanin: Electronic Structure and Mechanism of the Spin-Forbidden Two-Electron Reduction of O2. *J. Am. Chem. Soc.* **2001**, *123*, 4938–4950.

(51) Mirica, L. M.; Rudd, D. J.; Vance, M. A.; Solomon, E. I.; Hodgson, K. O.; Hedman, B.; Stack, T. D. P. μ-η^2:η^2-Peroxodicopper(II) Complex with a Secondary Diamine Ligand: A Functional Model of Tyrosinase. *J. Am. Chem. Soc.* **2006**, *128*, 2654–2665.

(52) Siegbahn, P. E. M. Modeling aspects of mechanisms for reactions catalyzed by metalloenzymes. *J. Comput. Chem.* **2001**, *22*, 1634–1645.

(53) Roos, B. O. *The Complete Active Space Self-Consistent Field Method and its Applications in Electronic Structure Calculations*; John Wiley & Sons, Inc.: Hoboken, NJ, 2007; pp 399−445.

(54) Roos, B. O.; Taylor, P. R.; Siegbahn, P. E. M. A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach. *Chem. Phys.* **1980**, *48*, 157–173.

(55) Andersson, K.; Malmqvist, P. A.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. Second-order perturbation theory with a CASSCF reference function. *J. Phys. Chem.* **1990**, *94*, 5483–5488.

(56) Neese, F.; Wennmohs, F.; Hansen, A. Efficient and accurate local approximations to coupled-electron pair approaches: An attempt to revive the pair natural orbital method. *J. Chem. Phys.* **2009**, *130*, 114108–18.

(57) Neese, F.; Hansen, A.; Liakos, D. G. Efficient and accurate approximations to the local coupled cluster singles doubles method using a truncated pair natural orbital basis. *J. Chem. Phys.* **2009**, *131*, 064103–15.

(58) Liakos, D. G.; Hansen, A.; Neese, F. Weak Molecular Interactions Studied with Parallel Implementations of the Local Pair Natural Orbital Coupled Pair and Coupled Cluster Methods. *J. Chem. Theory Comput.* **2010**, *7*, 76–87.

(59) Mahadevan, V.; Henson, M. J.; Solomon, E. I.; Stack, T. D. P. Differential Reactivity between Interconvertible Side-On Peroxo and Bis-μ-oxodicopper Isomers Using Peralkylated Diamine Ligands. *J. Am. Chem. Soc.* **2000**, *122*, 10249–10250.

(60) Mirica, L. M.; Vance, M.; Rudd, D. J.; Hedman, B.; Hodgson, K. O.; Solomon, E. I.; Stack, T. D. P. Tyrosinase Reactivity in a Model Complex: An Alternative Hydroxylation Mechanism. *Science* **2005**, *308*, 1890–1892.

(61) Neese, F.; Becker, U.; Ganyushin, D.; Hansen, A.; Liakos, D. G.; Kollmar, C.; Kossmann, S.; Petrenko, T.; Reimann, C.; Riplinger, C.; Sivalingam, K.; Valeev, E.; Wezisla, B.; Wennmohs, F. *ORCA*; University of Bonn: Bonn, Germany, 2009.

(62) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865.

(63) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104–19.

(64) Schafer, A.; Horn, H.; Ahlrichs, R. Fully optimized contracted Gaussian basis sets for atoms Li to Kr. *J. Chem. Phys.* **1992**, *97*, 2571–2577.

(65) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.

(66) Weigend, F.; Furche, F.; Ahlrichs, R. Gaussian basis sets of quadruple zeta valence quality for atoms H−Kr. *J. Chem. Phys.* **2003**, *119*, 12753–12762.

(67) Feyereisen, M.; Fitzgerald, G.; Komornicki, A. Use of approximate integrals in ab initio theory. An application in MP2 energy calculations. *Chem. Phys. Lett.* **1993**, *208*, 359–363.

(68) Kendall, R. A.; Früchtl, H. A. The impact of the resolution of the identity approximate integral method on modern ab initio algorithm development. *Theor. Chem. Acc.* **1997**, *97*, 158–163.

(69) Lee, T. J.; Rendell, A. P.; Taylor, P. R. Comparison of the quadratic configuration interaction and coupled-cluster approaches to electron correlation including the effect of triple excitations. *J. Phys. Chem.* **1990**, *94*, 5463–5468.

(70) Pople, J. A.; Head-Gordon, M.; Raghavachari, K. Quadratic configuration interaction. A general technique for determining electron correlation energies. *J. Chem. Phys.* **1987**, *87*, 5968–5975.

(71) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic-Behavior. *Phys. Rev. A* **1988**, *38*, 3098–3100.

(72) Lee, C. T.; Yang, W. T.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron-Density. *Phys. Rev. B* **1988**, *37*, 785–789.

(73) Becke, A. D. Density-Functional Thermochemistry 0.3. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(74) Adamo, C.; Matteo, A. d.; Barone, V. From Classical Density Functionals to Adiabatic Connection Methods. the State of the Art. In *Advances in Quantum Chemistry*; P.-O. Löwdin, Sabin, J. R., Zerner, M. C., Brandas, E., Lami, A., Vincenzo, B., Eds.; Academic Press: San Diego, CA, 1999; Vol. Vol. 36, pp 45−75.

(75) Becke, A. D. A new mixing of Hartree−Fock and local density-functional theories. *J. Chem. Phys.* **1993**, *98*, 1372–1377.

(76) Grimme, S. Semiempirical hybrid density functional with perturbative second-order correlation. *J. Chem. Phys.* **2006**, *124*, 034108–16.

(77) Siegbahn, P. E. M. A comparison of the thermodynamics of O—O bond cleavage for dicopper complexes in enzymes and synthetic systems. *J. Biol. Inorg. Chem.* **2003**, *8*, 577–585.

(78) Hess, B. A. Relativistic electronic-structure calculations employing a two-component no-pair formalism with external-field projection operators. *Phys. Rev. A* **1986**, *33*, 3742.

(79) Hess, B. A.; Marian, C. M. *Computational Molecular Spectroscopy*; Jensen, P., Bunker, P. R., Eds.; Wiley: New York, 2000 pp 169.

(80) Jansen, G.; Hess, B. A. Revision of the Douglas-Kroll transformation. *Phys. Rev. A* **1989**, *39*, 6016.

(81) Wolf, A.; Reiher, M.; Hess, B. A. *Relativistic Quantum Chemistry, Theoretical and Computational Chemistry*; Schwerdtfeger, P., Ed.; Elsevier: Amsterdam, The Netherlands, 2002; Vol. 1, pp 622.

(82) Wolf, A.; Reiher, M.; Hess, B. A. *Recent Advances in Relativistic Molecular Theory*; Hirao, K.; Ishikawa, Y., Eds.; World Scientific: Singapore, 2004 pp 137.

(83) van Lenthe, E.; Baerends, E. J.; Snijders, J. G. Relativistic regular two-component Hamiltonians. *J. Chem. Phys.* **1993**, *99*, 4597–4610.

(84) van Lenthe, E.; Snijders, J. G.; Baerends, E. J. The zero-order regular approximation for relativistic effects: The effect of spin--orbit coupling in closed shell molecules. *J. Chem. Phys.* **1996**, *105*, 6505–6516.

(85) Dolg, M.; Wedig, U.; Stoll, H.; Preuss, H. Energy-adjusted ab initio pseudopotentials for the first row transition elements. *J. Chem. Phys.* **1987**, *86*, 866–872.

(86) Martin, J. M. L.; Sundermann, A. Correlation consistent valence basis sets for use with the Stuttgart−Dresden−Bonn relativistic effective core potentials: The atoms Ga−Kr and In−Xe. *J. Chem. Phys.* **2001**, *114*, 3408–3420.

(87) Pantazis, D. A.; Chen, X.-Y.; Landis, C. R.; Neese, F. All-Electron Scalar Relativistic Basis Sets for Third-Row Transition Metal Atoms. *J. Chem. Theory Comput.* **2008**, *4*, 908–919.

(88) Klamt, A.; Schuurmann, G. COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799–805.

(89) Tomasi, J. Thirty years of continuum solvation chemistry: a review, and prospects for the near future. *Theor. Chem. Acc.* **2004**, *112*, 184–203.

(90) Tomasi, J.; Persico, M. Molecular Interactions in Solution: An Overview of Methods Based on Continuous Distributions of the Solvent. *Chem. Rev.* **1994**, *94*, 2027–2094.

(91) Cramer, C. J.; Truhlar, D. G. Implicit Solvation Models: Equilibria, Structure, Spectra, and Dynamics. *Chem. Rev.* **1999**, *99*, 2161–2200.

(92) Sinnecker, S.; Rajendran, A.; Klamt, A.; Diedenhofen, M.; Neese, F. Calculation of Solvent Shifts on Electronic g-Tensors with the Conductor-Like Screening Model (COSMO) and Its Self-Consistent Generalization to Real Solvents (Direct COSMO-RS). *J. Phys. Chem. A* **2006**, *110*, 2235–2245.

(93) Lee, T. J.; Taylor, P. R. A diagnostic for determining the quality of single-reference electron correlation methods. *Int. J. Quantum Chem.* **1989**, *36*, 199–207.

(94) Neese, F. Prediction of molecular properties and molecular spectroscopy with density functional theory: From fundamental theory to exchange-coupling. *Coord. Chem. Rev.* **2009**, *253*, 526–563.

(95) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. Basis-set convergence of correlated calculations on water. *J. Chem. Phys.* **1997**, *106*, 9639–9646.

(96) Truhlar, D. G. Basis-set extrapolation. *Chem. Phys. Lett.* **1998**, *294*, 45–48.

(97) Neese, F.; Valeev, E. F. Revisiting the Atomic Natural Orbital Approach for Basis Sets: Robust Systematic Basis Sets for Explicitly Correlated and Conventional Correlated ab initio Methods? *J. Chem. Theory Comput.* **2010**, *7*, 33–43.

(98) Cramer, C. J. *Essentials of Computational Chemistry*, 2nd ed.; Wiley: New York, 2004; pp 179.

(99) Rode, M. F.; Werner, H.-J. Ab initio study of the O_2 binding in dicopper complexes. *Theor. Chem. Acc.* **2005**, *114*, 309–317.

(100) Huenerbein, R.; Schirmer, B.; Moellmann, J.; Grimme, S. Effects of London dispersion on the isomerization reactions of large organic molecules: a density functional benchmark study. *Phys. Chem. Chem. Phys.* **2010**, *12*, 6940–6948.

(101) Goerigk, L.; Grimme, S. A General Database for Main Group Thermochemistry, Kinetics, and Noncovalent Interactions, àí Assessment of Common and Reparameterized (meta-)GGA Density Functionals. *J. Chem. Theory Comput.* **2009**, *6*, 107–126.

(102) van Wullen, C. Molecular density functional calculations in the regular relativistic approximation: Method, application to coinage metal diatomics, hydrides, fluorides and chlorides, and comparison with first-order relativistic calculations. *J. Chem. Phys.* **1998**, *109*, 392–399.

JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

A Fast and Robust Poisson—Boltzmann Solver Based on Adaptive Cartesian Grids

Alexander H. Boschitsch*[†] and Marcia O. Fenley*[‡]

[†]Continuum Dynamics, Inc., 34 Lexington Ave., Ewing, New Jersey 08618, United States

[‡]Department of Physics and Institute of Molecular Biophysics, Florida State University, Tallahassee, Florida 32306-3408, United States

S *Supporting Information*

ABSTRACT: An adaptive Cartesian grid (ACG) concept is presented for the fast and robust numerical solution of the 3D Poisson—Boltzmann equation (PBE) governing the electrostatic interactions of large-scale biomolecules and highly charged biomolecular assemblies such as ribosomes and viruses. The ACG offers numerous advantages over competing grid topologies such as regular 3D lattices and unstructured grids. For very large biological molecules and their assemblies, the total number of grid points is several orders of magnitude less than that required in a conventional lattice grid used in the current PBE solvers, thus allowing the end user to obtain accurate and stable nonlinear PBE solutions on a desktop computer. Compared to tetrahedral-based unstructured grids, ACG offers a simpler hierarchical grid structure, which is naturally suited to multigrid, relieves indirect addressing requirements, and uses fewer neighboring nodes in the finite difference stencils. Construction of the ACG and determination of the dielectric/ionic maps are straightforward and fast and require minimal user intervention. Charge singularities are eliminated by reformulating the problem to produce the reaction field potential in the molecular interior and the total electrostatic potential in the exterior ionic solvent region. This approach minimizes grid dependency and alleviates the need for fine grid spacing near atomic charge sites. The technical portion of this paper contains three parts. First, the ACG and its construction for general biomolecular geometries are described. Next, a discrete approximation to the PBE upon this mesh is derived. Finally, the overall solution procedure and multigrid implementation are summarized. Results obtained with the ACG-based PBE solver are presented for (i) a low dielectric spherical cavity, containing interior point charges, embedded in a high dielectric ionic solvent—analytical solutions are available for this case, thus allowing rigorous assessment of the solution accuracy; (ii) a pair of low dielectric charged spheres embedded in an ionic solvent to compute electrostatic interaction free energies as a function of the distance between sphere centers; (iii) surface potentials of proteins, nucleic acids, and their larger-scale assemblies such as ribosomes; and (iv) electrostatic solvation free energies and their salt sensitivities—obtained with both linear and nonlinear Poisson—Boltzmann equations—for a large set of proteins. These latter results along with timings can serve as benchmarks for comparing the performance of different PBE solvers.

INTRODUCTION

The efficient and accurate implicit solvent-based electrostatic modeling of large complex and highly charged biomolecules in an aqueous electrolyte solution at finite ionic strengths remains an important and difficult challenge in computational molecular biophysics. Considerable success in modeling the long-range and nonspecific electrostatic interactions of biomolecules in ionic solution has been achieved on the basis of the Poisson—Boltzmann equation (PBE), which provides the electrostatic potential and other important derived quantities (e.g., electrostatic solvation free energies, electrostatic binding free energies, forces, and p*K* shifts) under varying ionic conditions.[1] Nevertheless, two challenges persist in the numerical calculation of such systems. First, for large molecules, the mesh topologies used to date—regular lattices and unstructured tetrahedral grids—are subject to various inefficiencies and/or mesh generation challenges that can be improved upon by considering an alternate mesh structure as well as selecting a representation of the solution that reduces the mesh resolution demands. In the current development, the PBE is solved upon a hierarchical mesh structure variously referred to as an adaptive Cartesian grid (ACG) or octree or simply a Cartesian mesh. The ACG terminology is adopted here to distinguish it from regular lattices, which are also commonly called Cartesian grids. The second challenge is achieving reliable and rapid solution convergence for highly charged biomolecular systems. The current article describes a methodology that addresses both challenges, resulting in a robust nonlinear PBE analysis capable of properly modeling salt-mediated and nonspecific electrostatic effects in nucleic acids and their associations with charged ligands such as cationic drugs, peptides, and larger proteins.

One goal of the ACG-based PBE solver is to facilitate computation of electrostatic properties for large-scale biomolecular systems at the atomic level of detail using readily accessible computational resources. For example, a recent experimental study suggests that the electrostatic interactions in the ribosomal exit tunnel can modulate the elongation rates of nascent peptides.[2] For such large-scale ribosomal systems, most Poisson—Boltzmann studies have necessarily been based on coarse-grained molecular

Received: December 4, 2010
Published: April 15, 2011

models due to memory constraints and convergence issues.[3] The PBE solver described here provides the variable mesh spacing necessary to efficiently accommodate such nanoscale biomolecular assemblies. Moreover, it contains robust iterative procedures that reliably converge the electrostatic solution at comparable rates for both the linear and nonlinear PBE of highly charged complex biomolecular systems such as ribosomes. This property is used to obtain a high resolution (0.3 Å) surface potential map of the highly charged large 50S ribosomal subunit. To the best of our knowledge, nonlinear PBE calculations for such a highly charged and large biomolecular system have not been previously performed on a serial platform—at least at such a fine grid resolution. Moreover, with the computational tools developed here, nonlinear PB calculations can be conducted in nearly the same amount of computer time as linear ones, as borne out in the Supporting Information, which provides such timings for a collection of proteins.

Solution Methods for the PBE. Numerical solutions to the PBE can be obtained using either finite difference (FD) techniques, which here include finite element (e.g., unstructured tetrahedral meshes) and finite volume-based discretizations, or boundary element methods (BEM). Each approach has inherent advantages, as reviewed in refs 1 and 4. Briefly, when solving the linear PBE using the BEM, (i) only a surface mesh is required, since the solution is expressed entirely in terms of surface distributions; (ii) far-field boundary conditions are automatically satisfied; (iii) the constraints upon the electrostatic potential and its normal gradient at the molecular surface are explicitly imposed; (iv) the potential fields associated with point charges are expressed analytically, thereby circumventing problems relating to representing singular solutions upon grids; and (v) the interactions between distant elements are evaluated using the exact expressions, thus conferring high accuracy. With the introduction of fast multipole methods, computational costs have been reduced from $O(N^2)$ to $O(N \log N)$ (N being the number of boundary elements), thus allowing much larger problems to be addressed. The first PBE solvers utilizing fast multipole-accelerated BEM were limited to zero salt conditions.[5−7] The extension to finite salt concentrations was first achieved by Boschitsch and co-workers[4,8] and subsequently by Lu and co-workers using a different form of the fast multipole expansion.[9,10]

A major limitation of BEM-based approaches is the forfeiture of a pure surface-based solution representation and the attendant increase in computational effort when solving the nonlinear form of the PBE. Our experience[11] has consistently shown that even for very simple cases, computation times can easily increase by O(10)−(100) when using the BEM for the linear part and nonlinear terms expressed as source distributions, where the latter appear as volume integrals over the entire computational domain. Hybrid schemes offer one venue for retaining the advantages of a BEM while allowing the nonlinear PBE to be addressed.[11]

In the FD method, the differential form of the equations is solved on a volume mesh that fills the region of interest. The discrete equations can be derived according to variational principles which underlie the finite element (FE) method or by classical finite difference schemes based on the Taylor series expansions about a given mesh point. The FE method provides a general and systematic approach for developing the discrete model upon a variety of meshes including unstructured (tetrahedral) and curvilinear grids. In some instances however, the FD method offers a more efficient approximation. For example, when adopting a regular lattice mesh, the FD approximation to the Laplacian operator at a mesh point both is second-order accurate and involves only six neighboring mesh points, whereas a FE model is only first-order accurate (at least in the most common implementations[12,13] using linear order tetrahedral elements) and involves all 26 neighboring mesh points, thus increasing computational requirements.

The mesh structure employed in a FD method directly influences the performance and the quality of the results obtained. Historically, FD-based PBE modeling has employed two basic grid structures:

Regular 3D Lattice. This is the grid arrangement adopted in the popular PBE solvers such as APBS,[14] UHBD,[15] PBEQ,[16] MEAD,[17] ZAP,[18] DelPhi,[19] and PBSA-Amber[20−22] and consists of a uniformly spaced rectangular grid superimposed over the biomolecule of interest. While no attempt is made to align the mesh with the molecular surface, good estimates of the electrostatic potential solution are nevertheless obtained because this solution is continuous across the surface. Regular lattices allow one to readily develop a simple and efficient discretization of the differential operators and to implement effective multigrid procedures. However, the lack of a variable or adaptive grid spacing capability leads to a restrictive tradeoff between accuracy and storage constraints as larger biomolecules are considered. Furthermore, to minimize errors generated at the outer boundary of the grid (such errors introduce biases in computed electrostatic potential and energies), the grid must be extended sufficiently far from the molecule so that the potential at the outer boundaries is negligible. Nonlinear PBE calculations of highly charged biomolecular systems are especially challenging in this regard since consistent outer-boundary treatments for the nonlinear PBE have only recently become available.[23] To reduce calculation effort and maintain good accuracy, the focusing[24] procedure is invoked where the solution obtained on a global mesh with large mesh spacing is interpolated onto a collection of finer, localized grids. This approach improves local accuracy but entails multiple PBE calculations for a given molecular configuration.

Unstructured Grids. To address the shortcomings of regular lattice grids, efforts have been directed at the use of unstructured tetrahedral grids for biomolecules (e.g., ref 25). Such grids can achieve good resolution over a wide range of length scales and also offer the opportunity for solution-dependent mesh adaptation. Unstructured grids have been used in the finite element solution of the PBE to produce accurate predictions of biomolecular electrostatic properties.[12,13,25] A useful feature of unstructured grids is the ability to conform to the molecular surface so that no edges or elements intersect the surface. This allows for inherently more accurate estimates of surface properties, particularly the electrostatic field, which is essential for reliable prediction of electrostatic PBE forces. On the other hand, unstructured meshes are subject to several limitations: (i) The generation of good quality meshes is complex and time-consuming, especially for grids that conform to the molecular boundary; however active research in this area is expected to reduce the associated computation times.[26] (ii) Neighboring nodes must be explicitly identified, thus increasing storage costs (each node has approximately 14 neighbors, compared to six on a regular lattice grid). (iii) Mesh adaptation procedures are complex and expensive due to the large number of refinement possibilities in 3D. (iv) Multigrid implementation is challenging because defining coarser level meshes and linear-order accurate (the minimum order needed for second-order PDEs) interpolation procedures between multigrid levels is nontrivial. (v) The discrete approximation to the PBE equation generally has first-order errors (errors are $O(h)$ where h is the local mesh spacing) compared to the approximation on a regular lattice, which is second-order accurate (errors are $O(h^2)$), so that slower convergence with mesh spacing is obtained.

Herein, an alternate grid structure is proposed that combines the adaptation and variable resolution features of unstructured grids with the simple cube geometry and multigrid capabilities enjoyed by regular lattice methods. This mesh, referred to here as an adaptive Cartesian grid, derives from the hierarchical decomposition of the computational domain known as an octree,[27] which is obtained by recursive and selective subdivision of a cube into smaller nested cubes (see for example Figure 2). It is noted that an article utilizing the ACG concept to solve the nonlinear PBE has recently appeared[28] to model supercapacitor behavior of porous electrodes. Their approach embodies several of the same methodology details described below, including the derivation of the finite difference formulas. Their applications do not appear to call for a decomposition of the solution to eliminate singular behavior at charge sites, and applications were limited to comparatively simple geometries. ACGs have been widely used in fluid mechanics applications to model flows about complex geometries.[29,30] Often, the most time-consuming and challenging task in such applications is constructing a good quality mesh (for a complex geometry, this can require several man-months), and ACGs were developed in response to the need for a fast and fully automated grid-generation capability.[31−33] Like unstructured grids, the ACG allows the analysis to "zoom" into regions where the solution is varying rapidly—e.g., near the molecular surface. Elsewhere, where variations are more gradual, fewer, larger cells may be used for optimal computational efficiency. Outer boundaries can be placed far from the molecular boundary to minimize the influence of boundary errors without incurring appreciable computational cost. Compared to unstructured grids, the ACG generation and adaptation procedures are both simpler and less expensive computationally (for example, a mesh containing a million nodes is easily generated in under a minute using standard nonoptimized code on a readily accessible PC hardware). Finally, ACG facilitates implementation of multigrid schemes since the underlying octree data structure already prescribes a complete hierarchy of coarser level meshes and linearly accurate interpolation between levels is readily achieved.

In addition to using an ACG, the PBE solution methodology presented here adopts a decomposition of the electrostatic potential field similar to that in ref 34 to eliminate the singularities at fixed atomic charge sites. In the exterior regions, the usual total electrostatic potential is computed. Inside the molecule, one develops the reaction field potential which contains no singularities and so is accurately resolved on a mesh. The interior and exterior solutions are connected by calculating the Coulombic potential for nodes near the molecular boundary using fast multipole acceleration methods.[8] By eliminating the singularities, this decomposition (i) increases overall accuracy and reduces sensitivity to grid translations/rotations, (ii) alleviates mesh spacing requirements (no refinement near charge sites is required), and (iii) allows one to directly and accurately compute *total* electrostatic free energies (the grid-dependent self-energies[35] are completely absent) and forces. An interesting consequence of i is that the regions where the computed solution varies most rapidly are at the molecular surface rather than at atomic charge sites. This implies that the finest mesh spacing is warranted at the surface, and coarser elements can be employed away from the surface.

The sections below describe the generation of the ACG mesh; the discrete approximation of the Poisson−Boltzmann equation on this grid, including the decomposition of the solution into the full and reaction field potentials and the imposition of outer boundary conditions; the solution procedure using Gauss−Seidel iteration and multigrid; and postprocessing operations. Results are obtained using the ACG-based PBE solver for classical idealized problems involving one and two low dielectric spheres, containing interior charges to affirm the overall accuracy of the method, high resolution calculations of the electrostatic potential and other important derived electrostatic properties for medium-sized biomolecules, and demonstration calculations for a selected large-scale and highly charged ribosome. In the Supporting Information, ACG-PB predictions of electrostatic solvation free energies are provided for a variety of proteins with varying size, shape, and charge density along with timing information for both linear and nonlinear PB solutions.

METHODOLOGY

Generation of the ACG. Generation of the ACG grid for a given molecular structure presumes availability of the atomic coordinates (ρ_k), radii (σ_k), and partial charges (Q_k). The atomic coordinates can be obtained from structural biology databases such as the RCSB Protein Data Bank (PDB files) or Nucleic Acid Database (NDB files). The atomic radii can be assigned using one of many available atomic radii sets (e.g., Bondi[36]). Assigning atomic charges is more involved, especially when proper protonation state assignment is required, but typically either a formal charge set is adopted or partial atomic charges derived from molecular mechanics force fields, such as AMBER[37] or CHARMM,[38] are used. In addition to this structural description, a molecular surface definition must also be specified. Common surface definitions available in the ACG generation software include the van der Waals (vdW) surface, which is the exposed surface of the collection of overlapping spheres, and the solvent-excluded (SE) surface (also commonly referred to as the molecular or Connolly surface) obtained by rolling a probe sphere of radius, r_{probe} (usually, r_{probe} = 1.4 Å for water), over the van der Waals surface and identifying the points which can be reached by the probe (exterior points) and which ones cannot (interior). Other surface definitions, such as various Gaussian function-based descriptions (e.g., ref 39), can also be used and are available in the ACG software. Developing the ACG and assigning the dielectric map to the resulting mesh nodes requires the ability to determine whether a given point lies within the molecular surface. For the vdW surface, this determination is straightforward using an inside-sphere test. For the SE surface, the test is somewhat more involved—here, the procedures described by Chan and Purisima[40] are employed.

The ACG generation process begins by placing an initial cube over the entire molecule. This initial cube is sized to be several times larger than the maximum dimension of the molecule so that the boundary condition at the outer boundary can be accurately imposed (see below). The cube is then uniformly subdivided a fixed number of times, L, to produce a uniform lattice starting mesh containing $(2^L + 1)^3$ nodes (or 8^L cube-shaped cells).

Recursive adaptation of this initial mesh then proceeds by identifying which individual mesh cells intersect the molecular surface. Each intersected cell is tested to determine whether one of the following refinement criteria is satisfied:

(i) the user-specified finest mesh spacing, $\Delta_{\min}$, is reached or
(ii) the intersected cell lies more than a prescribed distance from the nearest atomic charge site.

Each intersected cell that does not meet either of these criteria is uniformly subdivided into eight smaller cells. The resulting ACG is then again subjected to these mesh intersection and refinement tests and the grid generation process continued. The refinement process naturally terminates since eventually all intersected cells meet the refinement criteria i or ii.

To prevent excessive cell size variation that can be detrimental to solutions accuracy, the ACG is smoothed by requiring that no terminal cell (a cell that has not been refined into smaller ones) be larger than twice any of its neighbors. This requirement also facilitates development of the finite difference procedures and implementation of multigrid. If requested, a Stern or ion exclusion layer of specified thickness, t, is defined by appropriately marking all nodes that are outside the molecule and less than a distance, t, away from the nearest interior node.

When conducting electrostatic interaction or binding energy calculations where the electrostatic energy of, say, a charged ligand−nucleic acid complex is subtracted from the electrostatic energies of the charged ligand and nucleic acid considered in isolation, all three calculations (charged ligand, nucleic acid, and charged ligand−nucleic acid pair) are conducted on the same mesh. This is because the electrostatic interaction or binding energy is often several orders of magnitude smaller than the individual electrostatic energy contributions so that small errors (e.g., due to finite mesh size) in the individual electrostatic energies appear large relative to the electrostatic interaction energy. In such calculations, the ACG is generated with respect to all three geometries as if the molecule actually consisted of the superposition of all three molecular surfaces. The same mesh is then employed for all three energy calculations using the respective dielectric maps.

Finite Differencing on the ACG. The ACG contains "hanging" nodes, which are nodes that neighbor an element but are not a vertex of that element (for example, a node that lies on a midedge or the face of an element). The presence of hanging nodes complicates the application of a variational or finite element framework for deriving the governing equations (specifically, compatibility between different sized elements is not easily enforced). For this reason, a finite difference approach is adopted to obtain a discrete expression of the PBE on the ACG. In developing a FD approximation to the weighted Laplacian, $\nabla\cdot(\varepsilon\nabla\Phi)$, it is desirable to simultaneously achieve the following properties: (i) compactness, to ensure robust convergence and numerical stability, the formula should be compact (i.e., only involve immediately neighboring nodes); (ii) consistency, as mesh spacing is reduced, the difference formula should converge to the exact analytical result; (iii) positive weights, the final expression relates the potential at a point, i, to the weighted sum of the neighboring node potentials; ensuring that the associated weights are positive is important for stable convergence and conformance with maximum principles for elliptic PDEs.[41] An additional consideration for continuum electrostatic modeling is that the dielectric "constant" changes discontinuously at the molecular surface (MS). This makes it difficult to develop formally consistent FD rules. However, the errors committed in applying the FD formulas near the MS can be viewed as perturbations of the surface geometry. Also, the success of FD applied upon regular lattices indicates that good PBE predictions can be obtained with simple interpolation of the dielectric/ionic map (e.g., as currently done in any of the lattice code such as APBS). Here, we will adopt such interpolation schemes and confirm their effectiveness by subsequent numerical studies. Current work is being directed at addressing accurate interpolation at the surface.

The FD method begins by distinguishing between various types of nodes. Denoting the collection of terminal octree cells, i_b, that touch a node, i, by $\{N_i\} = \{i_b\colon i_b$ incident to node, $i\}$ (this implies that node i lies on the surface of i_b), then three types of mesh nodes can be distinguished.

Type 0 Node i is a *vertex* of all terminal cells, $i_b \in \{N_i\}$, which implies that it is *not* a hanging node. Type 0 nodes are further distinguished into two subtypes:
Type 0A All $i_b \in \{N_i\}$ are of equal size.
Type 0B The members $i_b \in \{N_i\}$ differ in size.
Type 1 The node lies on the *midedge* of at least one terminal cell, $i_b \in \{N_i\}$.
Type 2 The node lies on the *face* center of exactly one terminal cell, $i_b \in \{N_i\}$.

Examples of these nodes are shown in Figure 1. Note that type 1 and 2 nodes are necessarily adjacent to cells of differing size. Also, in all cases, the members of $\{N_i\}$ differ by no more than a factor of 2 in size. Finally, each node can only be of one type (e.g., it cannot simultaneously lie on a midedge and a face center). This is a result of the size constraint between neighboring cells. Under these constraints, the finite difference expressions for the differential operator, $\partial/(\partial x)(\varepsilon(\partial\Phi)/(\partial x))$, are now developed for each of the node types.

The finite differencing expression for type 0A nodes is the same as that used on a regular lattice. Along the x direction, the contribution to $\nabla\cdot(\varepsilon\nabla\phi)$ is

$$\frac{\partial}{\partial x}\left(\varepsilon\frac{\partial\Phi}{\partial x}\right)_i = \bar{\varepsilon}_{i,i+1}\frac{(\Phi_{i+1}-\Phi_i)}{\Delta x^2} - \bar{\varepsilon}_{i,i-1}\frac{(\Phi_i-\Phi_{i-1})}{\Delta x^2} + O(\Delta x^2) \qquad (1)$$

where $\mathbf{R}_i$ is the position of node i, $\phi_i = \phi(\mathbf{R}_i)$, $\mathbf{R}_{i\pm1}$=$\mathbf{R}_i \pm \mathbf{i}(\Delta x)$, $\mathbf{i}$ is the unit vector along x, and Δx is the size of the surrounding cells. The second-order error estimate, $O(\Delta x^2)$, formally only applies when the dielectric constant is not changing, which is the case away from the molecular surface. The dielectric constant, $\bar{\varepsilon}_{i,i+1}$, is evaluated at the connecting edge midpoint.

Referring to Figure 1, the unique consistent finite difference formula at a type 0B node involving the triplet of collinear nodes, $\{0, 1, 2\}$, is

$$\frac{\partial}{\partial x}\left(\varepsilon\frac{\partial\Phi}{\partial x}\right)_1 = \bar{\varepsilon}_{1,2}\frac{\Phi_2-\Phi_1}{3\Delta x^2} - 2\bar{\varepsilon}_{0,1}\frac{\Phi_1-\Phi_0}{3\Delta x^2} + O(\Delta x) \qquad (2)$$

where Φ_i is the value of Φ at the indicated vertex. Note that the formula is only first-order accurate.

The FD formulas for type 1 and type 2 nodes are developed by identifying the neighboring nodes, developing Taylor series expansions for these nodes, and then considering how to combine these series so that only the desired second-order derivatives remain. For type 1 nodes such as node M in Figure 1, this process

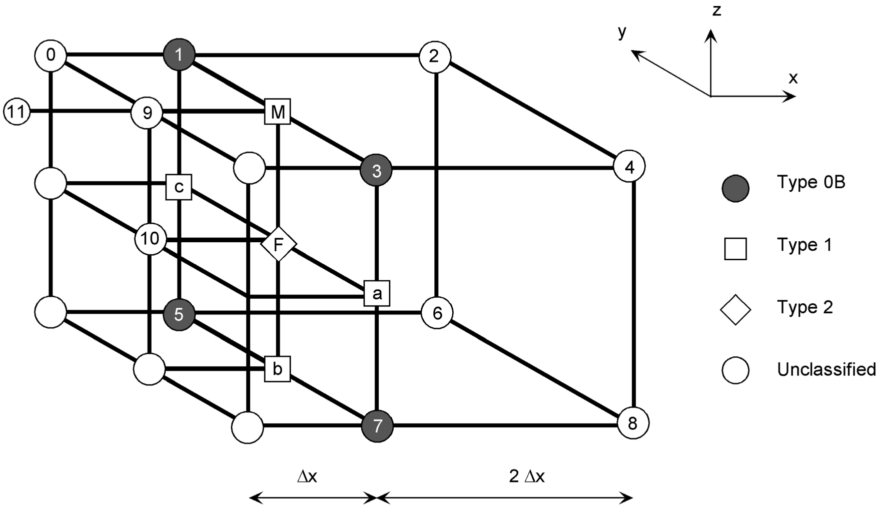


Figure 1. ACG nodes lying on the interface between two different sized cells.

leads to the first order formula:

$$\frac{\partial}{\partial x}\left(\varepsilon\frac{\partial\Phi}{\partial x}\right)_{\mathrm{M}} = \varepsilon(\mathbf{R}_{\mathrm{c1}})\frac{1}{3\Delta x^2}\left(\frac{\Phi_2+\Phi_4}{2}-\frac{\Phi_1+\Phi_3}{2}\right) - 2\bar{\varepsilon}_{\mathrm{M},9}\left(\frac{\varphi_{\mathrm{M}}-\varphi_9}{3\Delta x^2}\right)+O(\Delta x) \quad (3)$$

where the center of the face formed from nodes 1−2−3−4,

$$\mathbf{R}_{\mathrm{c1}} = \frac{\mathbf{R}_1+\mathbf{R}_2+\mathbf{R}_3+\mathbf{R}_4}{4} \quad (4)$$

For a type 2 node such as node F in Figure 1, one obtains the first order accurate formula:

$$\frac{\partial}{\partial x}\left(\varepsilon\frac{\partial\Phi}{\partial x}\right)_{\mathrm{F}} = \varepsilon(\mathbf{R}_{\mathrm{c2}})\frac{1}{3\Delta x^2}(\Delta^+\Phi) - 2\bar{\varepsilon}_{\mathrm{F},10}\left(\frac{\varphi_{\mathrm{F}}-\varphi_{10}}{3\Delta x^2}\right)+O(\Delta x) \quad (5)$$

where the cell center,

$$\mathbf{R}_{\mathrm{c2}} = \frac{1}{8}\sum_{k=1}^{8}\mathbf{R}_k \quad (6)$$

and the central difference approximation to $\partial\Phi/\partial x$ at $\mathbf{R}_{\mathrm{c2}}$ is

$$\Delta^+\Phi = \frac{1}{4}(\Phi_2+\Phi_4+\Phi_6+\Phi_8) - \frac{1}{2}(\Phi_{\mathrm{M}}+\Phi_{\mathrm{a}}+\Phi_{\mathrm{b}}+\Phi_{\mathrm{c}}-2\Phi_{\mathrm{F}}) \quad (7)$$

This form is preferred over other options since it promotes positive weights in the final assembled Laplacian approximation, eq 8.

Summary of the FD Formulas. The FD formulas are first-order accurate (errors are of $O(\Delta x)$) for other than type 0A nodes. It is possible to extend them to higher order by including additional nearby points, but this invites other problems such as nonpositive weights and numerical instability associated with the stronger influence from the more distant neighbors. The finite element method applied using linear elements is also first-order accurate (this is easily demonstrated in 1D when computing $\partial^2\Phi/\partial x^2$ upon an unevenly spaced grid), whereas the regular lattice methods are second-order accurate. The ACG-based FD method offers intermediate accuracy since, depending upon the degree of smoothing, the grid is populated mostly with type 0A nodes. Hence, the discretization is second-order accurate over most of the mesh and thus approaches the order of accuracy of a regular lattice.

The FD approximations to the second-order derivatives in the x direction extend naturally to the y and z derivatives which, when assembled, yield the discrete approximation to $\nabla\cdot(\varepsilon\nabla\Phi)$. At a node, i, this approximation can be cast in the form of a weighted sum:

$$(\nabla\cdot\varepsilon\nabla\Phi)_i = \sum_{\text{neighbors}, j}\omega_{ij}(\Phi_j-\Phi_i) \quad (8)$$

where ω_{ij} are the weights. For regions where the dielectric is constant, $\omega_{ij} > 0$. Across the surface where dielectric changes however, some weights for type 1 and type 2 nodes at the molecular surface may become negative, but no convergence issues have occurred in our PBE calculations to date. It is easy to show that this discrete approximation of $\nabla\cdot(\varepsilon\nabla\Phi)$ upon the ACG is compact and consistent, satisfies a discrete maximum principle (all weights, ω_{ij}, are positive[42]), and reverts to the classical finite difference expressions when implemented upon a regular lattice.

Application to the Poisson−Boltzmann Equation. The PBE is expressed over three distinct regions: (i) the molecular interior or solute region, Ω_1, which contains the atomic point charges, has a low dielectric constant, ε_1, and is enclosed by the molecular surface defined previously; (ii) the exterior or ionic solvent region, Ω_2, which has a high dielectric constant, ε_2, and contains the dissolved ions; and (iii) a charge-free Stern layer (or ion exclusion region where no mobile ions are present), Ω_3, of specified thickness about the molecular surface with dielectric constant $\varepsilon_3 = \varepsilon_2$. The Stern layer can be used to account for the ion size, and its thickness corresponds roughly to the hydrated radius of the ion. In all PB calculations below, the Stern layer thickness is set to zero.

The reduced (or dimensionless) electrostatic potential of any arbitrary 3D complex-shaped biopolyelectrolyte, Φ, at location $\mathbf{R}$ in the computational domain, is governed by

$$\nabla\cdot[\varepsilon(\mathbf{R})\nabla\Phi(\mathbf{R})]+\frac{4\pi e}{k_{\mathrm{B}}T}\rho(\mathbf{R}) = 0 \quad (9)$$

where the volume charge density in the different regions is given by

$$\rho(\mathbf{R}) = \rho^{\mathrm{f}}(\mathbf{R}) = \sum_k Q_k \delta(\mathbf{R} - \rho_k), \mathbf{R} \in \Omega_1 \tag{10a}$$

$$\rho(\mathbf{R}) = \rho^{\mathrm{m}}(\mathbf{R}) = -2eI_{1:1} \sinh(\Phi), \mathbf{R} \in \Omega_2 \tag{10b}$$

$$\rho(\mathbf{R}) = 0, \mathbf{R} \in \Omega_3 \tag{10c}$$

Note that the expression for $\rho^{\mathrm{m}}(\mathbf{R})$ in eq 10b pertains to a 1:1 electrolyte solvent (e.g., NaCl), which is assumed here for ease of presentation. The extension to more general salt environments is straightforward, and the ACG-based PBE solver currently accommodates mixtures of 1:1 and 2:1 salts[11,43] and asymmetric salts.

Introducing the Debye–Hückel screening parameter, κ, as

$$\kappa^2 = \frac{8\pi e^2(I_{1:1})}{\varepsilon_2 k_{\mathrm{B}} T} \tag{11}$$

allows one to rewrite the PBE in the exterior domain, Ω_2, as

$$\nabla\cdot(\varepsilon\nabla\Phi) = \varepsilon_2\kappa^2 \sinh(\Phi) \equiv f(\Phi) \tag{12}$$

The linearized form of eq 12, valid for small electrostatic potentials, $\Phi \ll 1$, is obtained by setting $f(\Phi) \approx f_L(\Phi) = \varepsilon_2\kappa^2\Phi$.

In the ACG-based FD implementation, a discrete approximation of eq 9 is solved at every mesh node. The discretization of the dielectric-weighted Laplacian, $\nabla\cdot(\varepsilon\nabla\Phi)$, is given by eq 8. For nodes outside the molecule (in Ω_2 and Ω_3), evaluation of the charge density in eq 9 according to eq 10b or 10c is straightforward. However, evaluation of the charge density within the molecular interior presents numerical difficulties because the potential becomes singular at the fixed solute charge sites (i.e., atomic centers), ρ_k. To eliminate this singular behavior, an alternate representation of the interior potential field is adopted.

Representation of the Interior Electrostatic Potential. The interior total electrostatic potential can be expressed as the sum $\Phi = \Phi^{\mathrm{rf}} + \Phi^{\mathrm{c}}$, where Φ^{rf} is the reaction field potential satisfying

$$\nabla\cdot[\varepsilon(\mathbf{R})\nabla\Phi^{\mathrm{rf}}(\mathbf{R})] = 0 \quad (\underline{R} \in \Omega_1) \tag{13}$$

and Φ^{c} is the singular Coulombic potential given by

$$\Phi^{\mathrm{c}}(\mathbf{R}) = \frac{1}{4\pi} \sum_{\mathrm{charges},\, k} \frac{q_k}{|\mathbf{R} - \rho_k|} \tag{14}$$

Here, the reduced charge centered at position ρ_k is $q_k = (4\pi e/\varepsilon_1 k_{\mathrm{B}} T)Q_k$. The reaction field potential contains no singularities and therefore can be accurately resolved on the ACG. Thus, at all interior points, the analysis solves for Φ^{rf} governed by eq 13 rather than Φ. This approach closely resembles the one implemented by Zhou et al.[35] upon a regular lattice. In the exterior region, $\Omega_2 \cup \Omega_3$, the full electrostatic potential, Φ, is retained, and eq 9 is solved.

To connect these two representations, Φ^{rf} and Φ, at the dielectric interface, first distinguish between the following four possible arrangements for a grid point, i, and its neighbors, j:

a. Point i, and all of its neighbors, j, lie inside the molecular interior, Ω_1.

b. Point i and all of its neighbors, j, lie inside the molecular exterior, $\Omega_2 \cup \Omega_3$.

c. Point i lies in the exterior region, $\Omega_2 \cup \Omega_3$, but at least one of its neighbors lies inside the molecule in Ω_1.

d. Point i lies inside Ω_1, but at least one of its neighbors lies outside the molecule in $\Omega_2 \cup \Omega_3$.

Cases a and b pose no difficulty since the discrete approximation eq 8 can be directly applied without modification. In case c, one solves eq 9 and thus seeks to evaluate $\nabla\cdot(\varepsilon\nabla\Phi)$. Here, the total electrostatic potential is available at node i and all neighbors lying in the exterior domain. However, for those neighbors located inside the interior region, Ω_1, one has only the reaction field potential. Hence, the Coulombic potential, evaluated according to 14, must be added to these interior grid points before evaluating the weighted Laplacian. Thus, eq 8 is modified to

$$(\nabla\cdot\varepsilon\nabla\Phi)_i = \sum_{\substack{\text{neighbors},\\ j \in \Omega_2 \cup \Omega_3}} \omega_{ij}(\Phi_j - \Phi_i) + \sum_{\substack{\text{neighbors},\\ j \in \Omega_1}} \omega_{ij}(\Phi_j^{\mathrm{rf}} + \Phi_j^{\mathrm{c}} - \Phi_i) \tag{15}$$

Case d is treated similarly. One can solve eq 13 and subtract Φ^{c} from all exterior neighbors, j. This option requires evaluating the Coulombic potential at exterior grid points. Alternatively, if both i and its neighbors, j, are sufficiently distant from the nearest charge, then one can instead solve the equation for the full potential, $\nabla\cdot(\varepsilon\nabla\Phi) = 0$. Then, as for case c, the Coulombic potential must be added to each of the interior nodes (including node i) before evaluating the weighted Laplacian.

All cases can be expressed in terms of the generalized potential,

$$\Phi^{\mathrm{g}}(\underline{R}) = \begin{cases} \Phi^{\mathrm{rf}}(\mathbf{R}), & \mathbf{R} \in \Omega_1 \\ \Phi(\mathbf{R}), & \mathbf{R} \in \Omega_2 \cup \Omega_3 \end{cases} \tag{16}$$

which is the discontinuous quantity actually represented upon the ACG mesh. The evaluation of the Laplacian can then be expressed as a weighted summation over all neighbors (without distinction as to whether they lie inside or outside the molecule):

$$(\nabla\cdot\varepsilon\nabla\Phi)_i = \sum_{\text{neighbors},\, j} \omega_{ij}(\Phi_j^{\mathrm{g}} - \Phi_i^{\mathrm{g}}) + \sigma_i \tag{17}$$

where σ_i represents the source terms originating from the Coulombic potentials at neighboring *interior* points such as those appearing in 15. Note that the source term is only nonzero for points having one neighbor across the molecular surface. Thus the Coulombic potential need only be evaluated at interior points lying adjacent to the molecular surface, thereby minimizing the number of Coulombic potential evaluations. To further expedite the computation, Φ^{c} is evaluated using the fast multipole acceleration method.[4,8,44]

After including the ionic source contributions from the PBE in the exterior region, the final discrete form of the PBE can be written:

$$\lambda_i = \sum_{\text{neighbors},\, j} \omega_{ij}(\Phi_j^{\mathrm{g}} - \Phi_i^{\mathrm{g}}) + \sigma_i - f(\Phi_i^{\mathrm{g}}) = 0 \tag{18}$$

where $f_i = f(\Phi_i^{\mathrm{g}}) = 0$ at interior points since one is solving for the reaction field potential there.

Outer Boundary Conditions. The governing equations are closed by specifying the potential at the outer boundary. One option is to set $\Phi = 0$ at the outer boundary and place the outer boundary sufficiently far away to minimize the effects of outer boundary errors—this can be accomplished more readily with the variable mesh spacing features of the ACG. Another option is to evaluate the outer boundary potential using the Debye–Hückel approximation:

$$\Phi^{\mathrm{b}} \cong \frac{1}{4\pi(\varepsilon_2/\varepsilon_1)} \sum_{\text{charges},\, k} \frac{q_k}{|\mathbf{R} - \rho_k|} \exp\{-\kappa|\mathbf{R} - \rho_k|\} \qquad (19)$$

which is useful when solving the linear PBE. Solutions to the nonlinear PBE, however, generally decay more quickly away from the surface (where, $|\Phi| > 1$) than their linear counterparts. Thus, when considering the nonlinear PBE, eq 19 tends to overestimate the boundary potentials, which introduces a bias into the computed solution.

The approach[23] adopted here is to approximate the electrostatic potential outside the computational domain by the approximate monopole formula, $\Phi^{\mathrm{b}} = Be^{-\kappa(r-h)}/r$ where the constant, B, is determined from electroneutrality conditions and $2h$ is the side length of the overall grid. This approach is equally valid for both the linear and nonlinear forms of the PBE and requires only that the magnitude of the potential at the outer boundaries $|\Phi| \ll 1$. An explicit expression for B is given elsewhere.[23]

Iterative Solution Scheme. The discrete system, eq 18, comprises a sparse algebraic set of coupled equations to be solved for the potentials, Φ_i^{g}. For large numbers of nodes, direct inversion of the equation system is not feasible, and an iterative inversion method must be used. The choice of iteration method has direct bearing upon the robustness and rate of solution convergence. Here, a standard Gauss−Seidel iteration method and multigrid are combined to achieve the good convergence. Gauss−Seidel iteration usually results in an initially rapid, but then slowed convergence rate. This slackening in convergence is due to the persistence of long wavelength errors. Like most simple iteration schemes, Gauss−Seidel updating effectively eliminates short wavelength errors that fluctuate most rapidly between grid points but is less efficient at removing long wavelength components. To promote faster convergence at moderate computational expense (storage and CPU), multigrid acceleration is employed in the ACG-based PBE solver.

Multigrid methods exploit the error smoothing properties of the Gauss−Seidel iteration process. After several Gauss−Seidel iterations, the short wavelength errors are mostly eliminated, and only long wavelength errors remain. These errors can therefore be accurately resolved upon a coarser grid. Moreover, because the mesh spacing is larger, the errors fluctuate more rapidly between grid points on the coarser mesh. Therefore, Gauss−Seidel applied on the coarser level is more effective at eliminating those errors. This basic insight motivates the multigrid concept, which attempts to eliminate errors over all wavelengths by projecting the solution onto a hierarchy of increasingly coarser meshes. Descriptions of the multigrid method are available elsewhere[45] (including applications to the PBE[46,47]). Therefore, only a brief description of the overall method is presented here, with emphasis reserved for those implementation details that are specific to the use of an ACG.

The multigrid algorithm begins by defining a sequence of nested meshes, $\{M^{\ell} : \ell = 0,\ \text{nlev}\}$ where M^0 is the finest level mesh. Next, interpolation procedures for transferring solutions and errors between successive levels are defined. In multigrid terminology, these are referred to as "restriction" (transferring a solutions from the finer grid, $M^{\ell-1}$, to the coarser level mesh, M^{ℓ}), and "prolongation" (transferring from M^{ℓ} to $M^{\ell-1}$) operators. Here bilinear interpolation is employed for the prolongation step and its adjoint operator (full weighting[45]) used for restriction.

In a two-level multigrid implementation, the solution process begins by conducting a series of single-level Gauss−Seidel iterations on the finest level, M^0. The errors (or residuals), λ_i, from eq 18 are then evaluated and restricted to the next coarser level, $\ell = 1$, using full weighted averaging. A discrete approximation to the PBE is then developed on this coarser level. However, the Coulombic source terms, σ_i, on this coarser level are set to zero and replaced everywhere by the restricted errors which now "drive" the coarser level solution. Gauss−Seidel iteration is then conducted on this coarser level to obtain a correction potential on this level, $\{\Phi^1\}$. The final step is to linearly interpolate the corrections to the finer level, $\ell = 0$, and add them to the existing solution, $\{\Phi^0\} = \{\Phi^{\mathrm{g}}\}$. The extension to multiple levels is straightforward and explained elsewhere.[45−47]

Post-Processing. The total electrostatic free energy expression is taken from eq 8 of ref 48. After integration by parts of the last term and substituting using the governing equation, eq 9, one obtains the total electrostatic free energy, G^{el}, in $k_{\mathrm{B}}T$ units:[11]

$$G^{\mathrm{el}} = G_{\mathrm{f}} + G_{\mathrm{m}} - \Delta\Pi \qquad (20)$$

where

$$G_{\mathrm{f}} = C\int_{\Omega_1} \frac{4\pi e}{k_{\mathrm{B}}T\varepsilon_1}\rho_{\mathrm{f}}\Phi \,\mathrm{d}V = C\sum_{\text{charges},\, k} q_k\Phi(\rho_k) \qquad (21\mathrm{a})$$

$$G_{\mathrm{m}} = C\tilde{\varepsilon}\kappa^2\int_{\Omega_2} \Phi \sinh(\Phi)\,\mathrm{d}V \qquad (21\mathrm{b})$$

$$\Delta\Pi = C\tilde{\varepsilon}\kappa^2\int_{\Omega_2} (2\cosh\Phi - 2)\,\mathrm{d}V \qquad (21\mathrm{c})$$

and the conversion factor to express the energies in $k_{\mathrm{B}}T$ units is

$$C = \frac{(k_{\mathrm{B}}T)^2\varepsilon_1}{8\pi e^2} \qquad (22)$$

Here, G_{f} is the energy due to fixed charges, and G_{m} and $\Delta\Pi$ are the electrostatic stress and excess osmotic pressure terms, respectively. The excess osmotic pressure contribution has special significance when assessing salt dependencies of the electrostatic free energies, since one can show[49,50]

$$\frac{\mathrm{d}G^{\mathrm{el}}}{\mathrm{d}\kappa} = -\frac{2}{\kappa}\Delta\Pi \qquad (23)$$

It is also useful to define the reaction field energy,

$$G_{\mathrm{rf}} = \int_{\Omega_1} \frac{4\pi e}{k_{\mathrm{B}}T\varepsilon_1}\rho_{\mathrm{f}}(\Phi - \Phi^{\mathrm{c}})\,\mathrm{d}V = \sum_{\text{charges},\, k} q_k\Phi^{\mathrm{g}}(\rho_k) \qquad (24)$$

which is the difference between G_{f} and the Coulombic energy.

For interior points in Ω_1 (e.g., the charge sites, ρ_k), $\Phi = \Phi^{\mathrm{c}} + \Phi^{\mathrm{g}}$, is obtained by interpolating the reaction field potential from the ACG and adding the Coulombic potential, Φ^{c}, from eq 14. The electrostatic energy contribution from the interior region, G_{f}, is computed by summing the product of charge times the electrostatic potential at the fixed charge sites. The volume integrals over the exterior region, Ω_2, are evaluated by looping over the cubic cells, i_{b}, of the ACG and approximating the volume integral of any function, $g(\Phi)$, by

$$\int_{V_{i_{\mathrm{b}}} \cap\, \Omega_2} g(\Phi)\,\mathrm{d}V \cong \frac{\Delta s^3}{8} \sum_{\text{external vertices},\, k} g(\Phi_k) \qquad (25)$$

where the sum is taken over the exterior forming nodes, k, of cell i_b, and Φ_k is the potential at forming node k. In addition, a correction term is added to the volume integrals to account for the contribution outside the computational domain. This correction term is based on the same monopole approximation used for the outer boundary treatment and is developed fully in ref 23.

■ RESULTS

The results presented here serve two main objectives. The first is to assess error by comparing ACG PBE predictions against analytical solutions or results obtained by alternate highly accurate means.[11] In these studies, simple model geometries involving one or two low dielectric spherical cavities, containing charges, embedded in a high dielectric ionic solvent medium are considered. The second goal is to demonstrate the effectiveness of the ACG-based PBE solver as a practical tool for modeling high resolution medium- and large-scale biomolecules ranging from proteins to more highly charged nucleic acids and its large and complex biomolecular assemblies for which numerous X-ray crystal structures are now available. This second goal is accomplished by computing the electrostatic potential maps. Calculations were performed on a Dell Precision M2300 laptop (3 GHz with 4 GB of installed memory) or a dual-processor Intel Xeon Linux workstation (3 GHz with 1 GB of memory).

In addition to the results below, the Supporting Information contains the computed electrostatic (solvation) energies, timings, and scaling with system size for proteins with varying charge densities, shapes, and sizes and modeled both with the linear and the nonlinear PBE. Those results show that for the same system the computation time to solve nonlinear PBE is, on average, 6% more than that for the linear PBE.

Linear PBE Solved for a Low Dielectric Spherical Cavity with a Unit Charge Embedded in a High Dielectric Ionic Solvent. The first model configuration studied solves the linear PBE for a unit radius low dielectric spherical cavity, containing a single interior charge, embedded in 0, 0.1, and 5 M salt solutions. The dielectric constants are set to $\varepsilon_1 = 2$ and $\varepsilon_2 = 80$, and the temperature is set to $T = 298.15$ K. No ion exclusion or Stern layer is modeled in this or subsequent PBE calculations presented here. Since analytical expressions, developed by Kirkwood,[51] for the solution of the linear Poisson–Boltzmann equation are available for all $\kappa > 0$,[51,52] this case constitutes a useful benchmark for establishing the overall accuracy of the ACG-PB solver. The computational domain extends over four radii, and the mesh is generated by requiring that any surface-intersected cell whose size is larger than 0.125 times the distance to the nearest charge is subdivided. As the charge is displaced toward the surface, this subdivision criterion produces an increasingly finer mesh about the surface point closest to the unit charge (see Figure 2). The time to complete the calculation for all 15 charge locations was 105 s on the PC laptop machine.

The mesh and contours of constant electrostatic potential for the case where the charge is closest to the surface ($1 - \rho = 3.125 \times 10^{-3}$ Å so that the distance from the surface is 0.3% of the atom radius) are shown in Figure 2. With the variable mesh spacing capability, the full mesh involves only approximately 115 000 mesh points (a comparable resolution calculation on a regular 3D lattice would entail over a trillion points). Note that the finest resolution provided by the ACG PB solver is at the surface nearest the charge, not at the charge itself. This is where the exterior full electrostatic potential and the interior reaction field electrostatic potential vary most rapidly. The rapid variation in the solution about the charge reflected in the contours is due to the analytically evaluated Coulombic potential contribution. Both outside and inside the molecule the potential maps are smooth and well-behaved, including near the unit charge placed inside the low dielectric sphere.

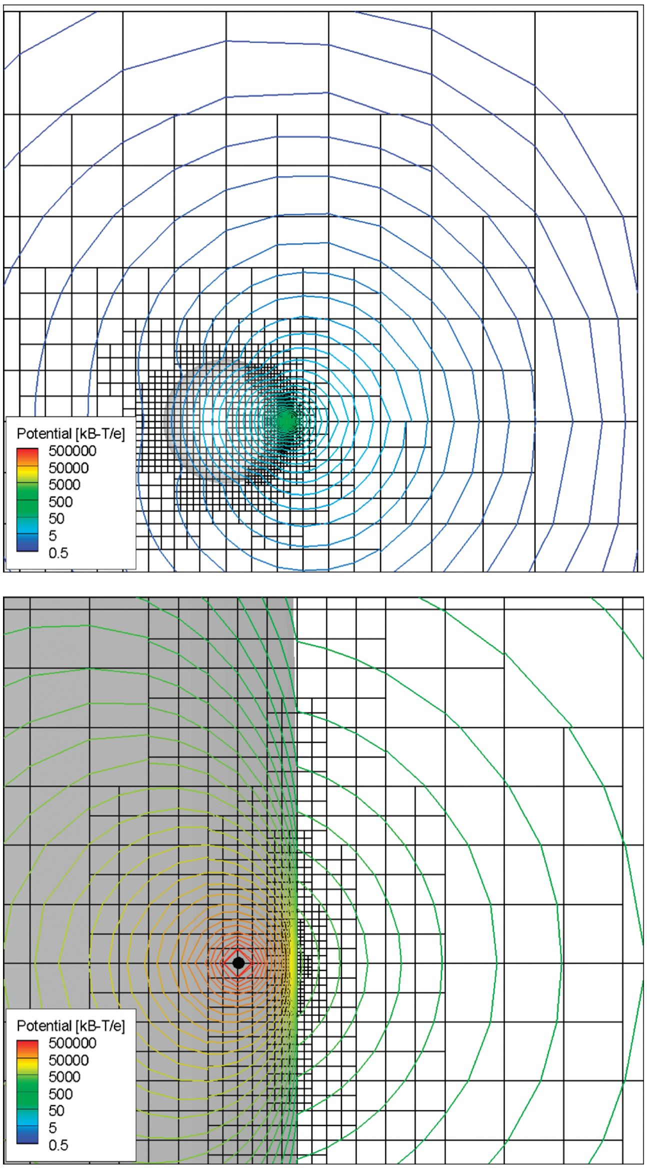


Figure 2. Cut through ACG for unit charge placed at $\rho = 0.9969\mathbf{i}$ inside a low dielectric spherical cavity. The lower graph provides a closeup of the mesh and solution near the charge site inside the low dielectric spherical cavity.

The numerical error defined as given by $E_{\rm rf} = G_{\rm rf}^{\rm comp}/G_{\rm rf}^{\rm exact} - 1$ is plotted as a function of distance below the surface in Figure 3 and shows that the error remains small even when the charge comes very close to the surface. For charges located within 99% of the spherical cavity radius, errors remain 1% or less.

Nonlinear PBE for a Low Dielectric Spherical Cavity of Varying Central Charge in a Salt Solution. The nonlinear behavior of a spherical cavity containing a centrally located charge is considered to verify accurate recovery of nonlinear solutions and demonstrate stable convergence at high net charge values. The governing PBE in this case reduces to a second order ordinary

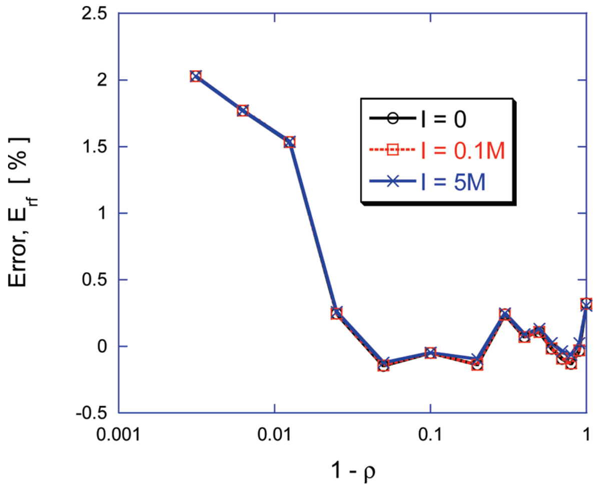


Figure 3. Error, $E_{rf} = G_{rf}^{comp}/G_{rf}^{exact} - 1$, plotted as a function of the distance of the interior unit charge, from the surface of the sphere. The linear PBE is solved over a unit radius spherical cavity with $\varepsilon_1 = 2$ and $\varepsilon_2 = 80$ under three different salt conditions: $I_{1:1} = 0$, 0.1, and 5.0 M corresponding to Debye−Hückel screening parameters, $\kappa = 0$, 0.103, and 0.728 Å^{-1}, respectively.

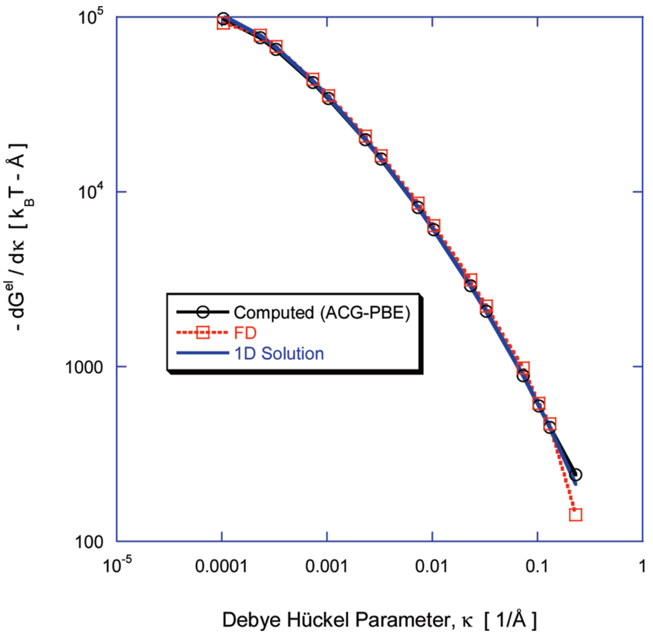


Figure 4. Comparison of salt sensitivity, $\partial G^{el}/\partial\kappa$, obtained from (i) the osmotic pressure, $\Delta\Pi$, computed with the ACG-PBE solver and relation 23; (ii) finite differencing of the G^{el} vs κ curve obtained with the ACG-PBE solver; and (iii) a 1D high resolution calculation. Nonlinear energy contributions as a function of the Debye−Hückel parameter, κ, for the single centrally located charge of 50e inside a 20 Å radius sphere. The dielectric constants are $\varepsilon_1 = 4$ and $\varepsilon_2 = 78.5$.

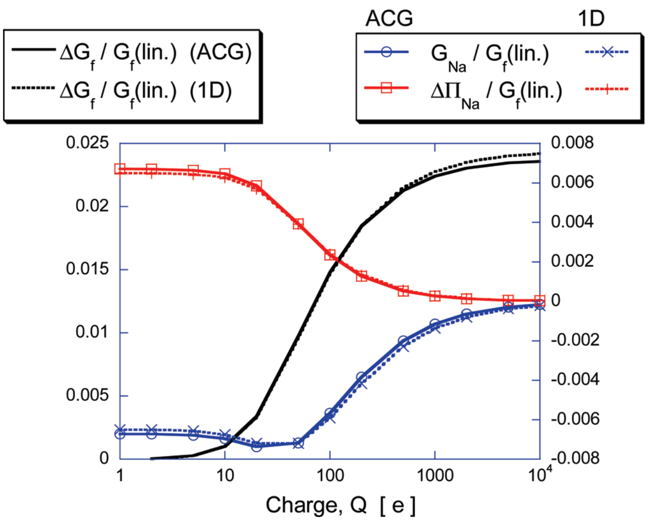


Figure 5. Charge dependence of electrostatic free energy ratios for I = 0.03 M. Results are obtained using the ACG scheme and the 1D finite element solution for the case of a spherical cavity with centrally located charge. The plotted electrostatic free energies, $\Delta G_f = G_f - G_f$(lin), G_{Na}, and $\Delta\Pi_{Na}$, are normalized by the fixed charge energy obtained from the linear PBE, G_f(lin). In this case, this equals the reaction field energy, G_{rf}, since the Coulombic energy is zero.

differential equation (ODE) that can be solved by alternative means (e.g., Appendix A of ref 11). Two cases are considered in this study. In the first, a centrally located 50e charge is placed inside a 20 Å radius sphere and the 1:1 salt concentration varied. The dielectric constant inside and outside the sphere are 4 and 78.5, respectively, and the temperature of the salt solution is T = 300 K. This case was examined by Zhou,[53] and his results closely agree with the ones obtained here. The variation of the total electrostatic free energy, G^{el}, as a function of a 1:1 salt concentration is also considered in Figure 4. According to eq 23, the slope of the G^{el} vs κ curve is related to the excess osmotic pressure energy contribution. This relationship thus constitutes an internal consistency check valid for general molecular geometries. The plot compares three different predictions of this electrostatic energy slope: (i) the right-hand side of eq 23, where the excess osmotic pressure, $\Delta\Pi$, is obtained using ACG-PBE; (ii) differentiation of a piecewise quadratic fit to the $G^{el} \sim \kappa$ curve where G^{el} is obtained from ACG-PBE; and (iii) the excess osmotic pressure predicted using the 1D analysis.[11] Close agreement is established over the entire 1:1 salt concentration range. The minor departure at the highest salt concentration appears to be due to the finite differencing algorithm, the excess osmotic pressure energy contributions obtained with the 1D solver, and ACG-PBE remaining in close agreement.

Next, the net charge is increased from 1e, where the PBE solution is essentially linear, to 10 000e, where nonlinear behavior dominates and the ability of the ACG solver to converge the solution in a robust manner is put to the test. In all cases, the number of multigrid cycles required to converge the solution ranged between 40 and 60. Figure 5 records the electrostatic free energy contributions, normalized by G_f for the linear problem, G_f(lin). Note that G_f(lin) can be expressed analytically, and the resulting values are in close agreement with the numerical predictions. Normalizing the electrostatic free energy contributions this way highlights the relative importance of the various nonlinear contributions. Again, good agreement between the 1D and ACG PB results is obtained. The change in normalized fixed charge energy, $\Delta G_f/G_f$(lin) (here, $\Delta G_f = G_f - G_f$(lin) and $G_f = G_{rf}$), is seen to be negligible at small charge values but to dominate the nonlinear contributions at higher charge values. Also, $\Delta G_f/G_f$(lin) seems to asymptote to a constant value at very high net charges. The opposite trend holds for the other two normalized electrostatic free energy contributions, G_{Na}/G_f(lin) and $\Delta\Pi_{Na}/G_f$(lin), which individually contribute a fixed fraction of total electrostatic energy at the low charge range. As net charge increases, however, their relative contributions diminish to zero. Also, since these two electrostatic free energy terms have opposite sign, their combined contribution is quite small over the entire charge range.

Electrostatic Interaction Free Energies for Two Low Dielectric Charged Spherical Cavities Embedded in an Aqueous Salt Medium. The long-range and nonspecific electrostatic interactions can modulate the kinetic rates of association of protein−protein and protein−nucleic acid association processes.[54] For instance, changes in the ionic conditions and charge distribution of the binding partners have a significant impact on the kinetic association rates of various biomolecular complexes.[55,56] The quantity of interest here is the electrostatic interaction free energy, which is the difference between the total electrostatic free energy of the complex and the summed total electrostatic energies of the individual molecules considered in isolation. Simple model two low dielectric spherical cavity systems have been studied previously using semianalytical treatments and are useful for validation purposes.[57] Electrostatic interaction or binding free energies can be difficult to calculate because they are usually much smaller in magnitude than the quantities being differenced. Therefore, the effects of truncation and other discretization errors upon the electrostatic interaction energy may be much more pronounced than for the total electrostatic energies of the two interacting partners.

Accurate electrostatic binding free energies for realistic and large-scale biomolecular systems are given below and elsewhere using the ACG PBE solver.[58−62] Here, the electrostatic interaction between a pair of low dielectric spherical cavities, containing interior charges, is considered as a model problem for verifying the ability to accurately calculate these interaction energies. The first sphere has a radius of 14 Å and contains three interior charges, $\{Q_i\} = \{-2.29, +8, +2.29\}e$, distributed along the x axis at locations $x_i = \{-7.8, 0, +7.8\}$ Å, relative to the center. The second sphere has a radius of 21 Å and also contains three interior charges, $\{Q_i\} = \{-2.21, -12, +2.21\}e$, distributed along the x axis at locations $x_i = \{X_2 - 11.7, X_2, X_2 + 11.7\}$ Å, where the separation, X_2, is the x location of the second sphere center. The dielectric constants chosen for this example are $\varepsilon_{\text{in}} = 4$ and $\varepsilon_{\text{out}} = 78.5$. The Debye−Hückel screening parameter $\kappa = 0.1316$ Å^{-1}.

Figure 6 compares the electrostatic interaction free energy obtained using the ACG PBE solver with semianalytical predictions[57] demonstrating excellent agreement when the same (van der Waals) surface is used to define the solute boundary that separates the interior and exterior dielectric regions. The total electrostatic free energies of the isolated 14 Å and 21 Å low dielectric charged spheres are −199.3 and −283.2 kcal/mol, respectively. Hence, the electrostatic interaction free energy is 2 orders of magnitude smaller than the individual electrostatic free energies. As one would expect, the choice of molecular surface affects the computed electrostatic interaction energy when spheres are closer than the solvent probe diameter (2.8 Å). Figure 6 also compares the electrostatic interaction energy obtained using the solvent excluded molecular surface. The resulting curve deviates significantly from the one using the van der Waals surface with a factor of 4 difference being obtained at the 35 Å separation.

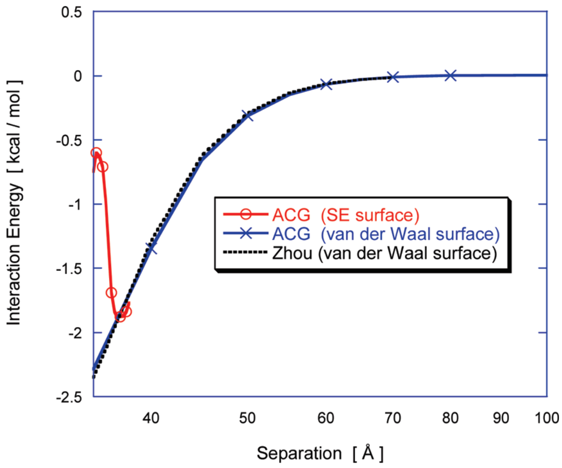


Figure 6. Electrostatic interaction energy for two low dielectric spherical cavities with interior charges embedded in a high dielectric ionic solvent, as a function of separation distance between the centers of the charged spherical cavities.

High Resolution Surface Potential Maps of Nucleic Acids and Their Binding Partners. Surface potential maps are now routinely used to identify potential binding or recognition sites on biomolecules at atomic resolution. For example, unique recognition or ion binding sites in irregular RNA structures, that contain noncanonical base pairs (e.g., GU wobble base pairs) and/or extruded (non)canonical bases, have been identified using the hybrid boundary element and finite difference nonlinear PB solver[11] and confirmed using the ACG-PB technology.[63,64] Obtaining such quantitative or high resolution electrostatic potential maps of large-scale biomolecules—especially highly charged ones like nucleic acids and their complexes with various charged binding partners—is very challenging for any PB solver. Surface potential maps are generated for three configurations: a small, low-charge RNA binding protein along with a single-stranded (ss) RNA binding partner, and a more highly charged noncanonical DNA structure using the nonlinear form of the PBE. The first case examines the binding of the cationic Fox-1 protein (net charge = +3e) to the RNA element UGCAUG (PDB id: 2err, model 1), where the latter is a simple single-stranded RNA structure. The solute boundary is modeled using the solvent excluded surface with atomic radii and charges specified using the CHARMM27 force field parameters.[38] As previously, the ion exclusion region is omitted; also $T = 298$ K, $\varepsilon_1 = 2$, and $\varepsilon_2 = 80$. The surface mesh spacing resolution is set to 0.3 Å and the outer boundary set to approximately 3 times the largest molecule dimension. In this study, the first model of the NMR ensemble was employed to assess the error incurred in electrostatic potential calculations. The histidine residues were considered unprotonated, while other charged residues were assigned protonation states based on a physiological pH value of 7. Thus, the Asp and Glu residues had a charge of −1e, whereas a charge of +1e was assigned to the Lys and Arg residues. The 1:1 (i.e., NaCl) salt concentration was fixed at 0.1 M.

The surface maps are obtained by first identifying the mesh edges intersected by the surface (i.e., those edges with an end point in the interior and exterior domains) and then calculating the intersection points. A triangulation of the intersection points is then developed and the potentials at the intersection points developed by extrapolating the ACG solution to the surface using the nearest exterior mesh nodes. All surface potential maps are produced using the commercial program, TecPlot.

As evident from Figure 7a,b, the single RNA element lies in a distinct pocket of very positive electrostatic potential on the RNA

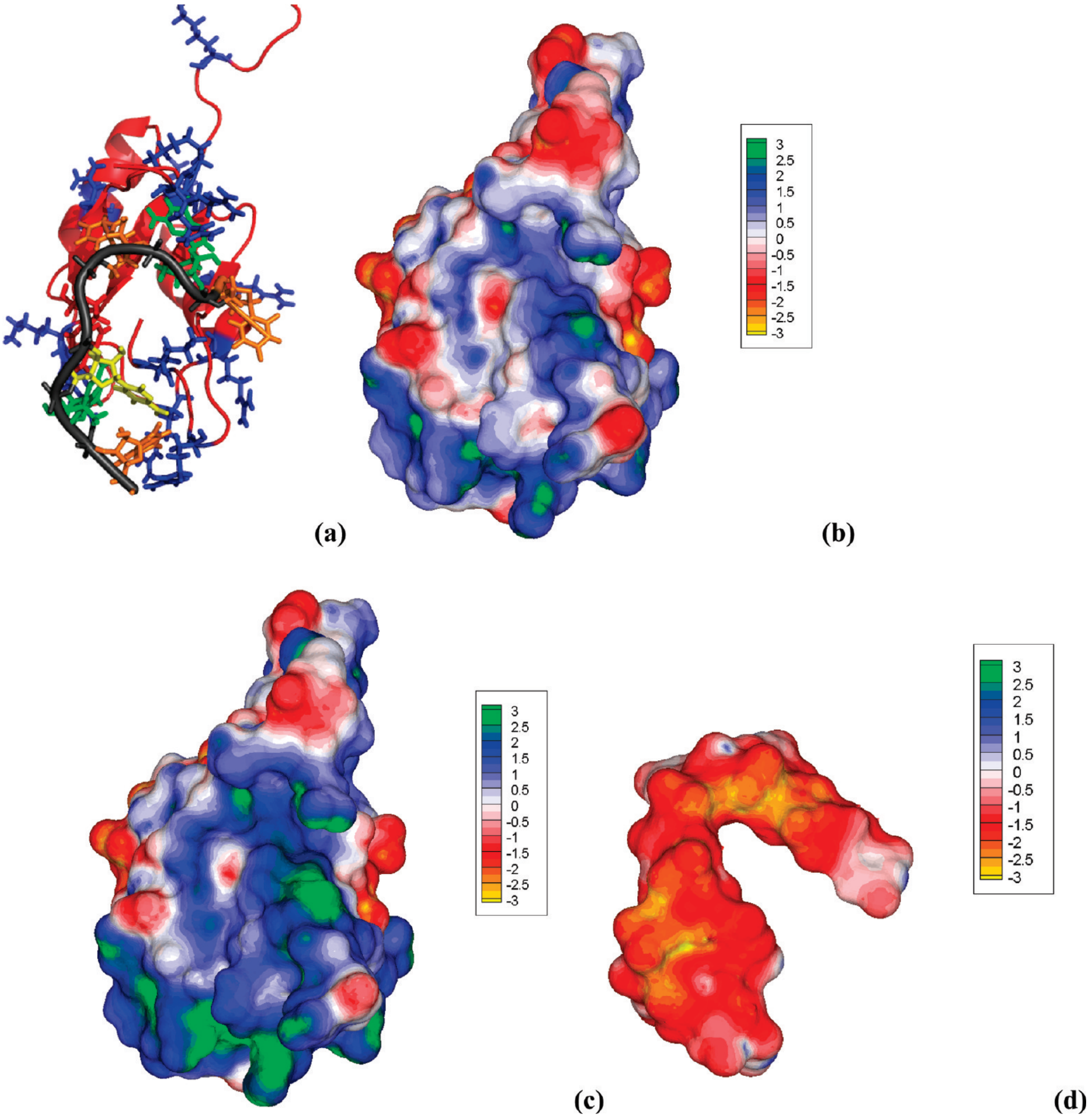


Figure 7. Different depictions of the NMR structure of RNA binding domain (RBD) of Fox-1 in complex with the single-stranded UGCAUGU RNA element (PDB id: 2err, model 1). (a) RNA phosphate backbone depicted by the dark gray ribbon and the different bases colored as adenine = red, uracil = orange, guanine = green, and cytosine = yellow. The protein peptide backbone adopts an orange ribbon representation with positively and negatively charged side chains shown as blue and red sticks, respectively. This view emphasizes the clustering of various cationic protein residues at the RNA binding interface. (b) Surface electrostatic potential (in kcal/mol/e) and overall shape of the RBD of Fox-1. A well-defined and concave region of positive potential, generated by the cationic protein residues and traced by the bent RNA structure, is clearly shown when the nonlinear PBE solution is employed. This cationic protein (net charge = $+3e$) follows the electrostatic pattern of other RNA binding proteins that have a distinct positive potential patch on their binding interface.[78] Thus, the RNA fills most of the concave blue/green protein surface to which it is complementary in both shape and electrostatic potential. (c) Same view and color map as b but using the linear PBE. The positive electrostatic potential is now overestimated and the positive region much broader than in b. (d) Electrostatic potential of the single-stranded bent and overall negatively charged RNA structure. As expected, an overall negative potential covers most of the RNA surface due to the presence of the anionic phosphate groups.

binding domain of the Fox-1 protein. It has been previously reported[65,66]—on the basis of linear PBE evaluations—that the surface potential in the RNA binding site of this protein is neutral. However, this is not borne out by the results obtained here with the nonlinear PBE (Figure 7b), thus highlighting the drawbacks of forming conclusions on the basis of linear PBE calculations. Predictions using the linear ACG-PB solver also show a significantly different surface potential distribution for the protein RNA binding domain of Fox-1 relative to the nonlinear one (see Figure 7b,c). The electrostatic potential of the ssRNA is negative over most of its surface (Figure 7d), whereas the negative potential regions of the RNA are attenuated with the presence of the cationic protein (results not shown).

The second example is the deformed and nonlinear DNA structure in association with the Tc3 transposase protein (PDB id: 1tc3). Charges and radii are assigned using the AMBER force

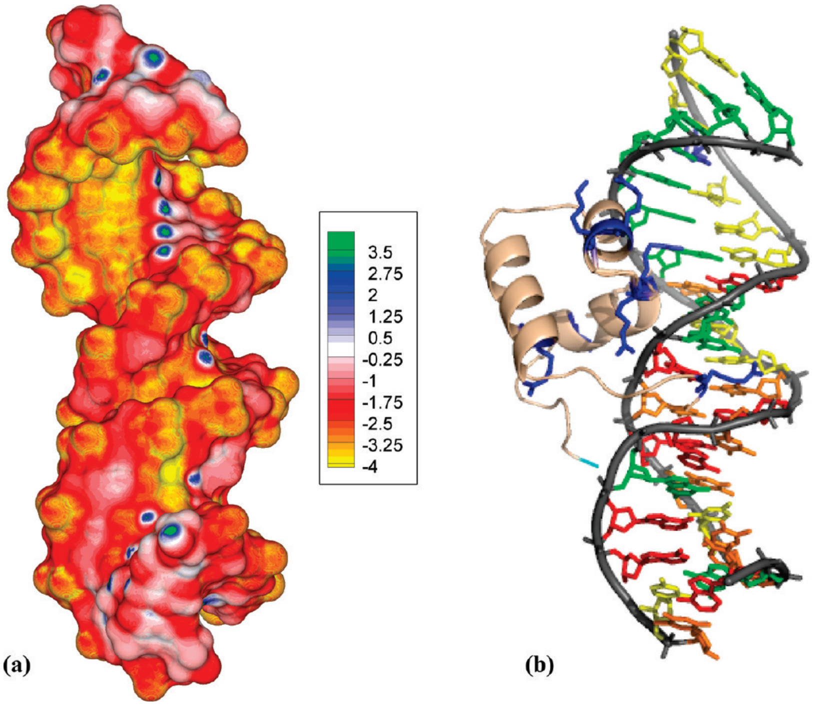


Figure 8. (a) The surface potential of the deformed and nonlinear DNA. Radii and atomic charges are assigned using the Amber force field. This unique A/B junction DNA structure generates a surface potential map with characteristics of an A-DNA major groove and B-DNA minor groove. A continuous high negative potential band along the G stretch along with electropositive spots due to amino groups of cytosine is observed for this nonlinear DNA structure. The electrostatic potential is given in units of kcal/mol/*e*. (b) The N-terminal DNA-binding domain of Tc3 transposase bound to the DNA (PDB id: 1TC3). The ribbon or tube-like representation of the DNA phosphate backbone is shown in light gray and that of the peptide backbone in gold. The bases are colored as adenine = red, thymine = orange, guanine = green and cytosine = yellow. The cationic residues that penetrate in the narrow minor groove or face the G-stretch side of the major groove are shown in blue stick representation.

field,[37] and the solute boundary is represented using the solvent excluded molecular surface. Also, $T = 298$ K, $\varepsilon_1 = 2$, and $\varepsilon_2 = 80$. The 1:1 salt concentration is set to 0.1 M. It is desirable that the electrostatic potential maps for nucleic acids capture unique local sequence dependent features and the intricate phosphate charge distribution (e.g., close clustering of phosphate groups that occurs at helical junctions in RNA and DNA structures). Here, high resolution surface potential maps capturing these features are produced. As portrayed in Figure 8a, the G stretch of the major groove side of the DNA structure and its locally narrowed minor groove have a deep negative potential.[67] The extensive region of negative potential along the G stretch of the major groove is mostly due to the deformation of the DNA structure. Figure 8b shows how the protein Tc3 *transposase* positions several positively charged side chains along one side of the major groove and in the narrow minor groove, forming numerous hydrogen bonds and salt bridges in these grooves. The linear PB solution produces much larger and more negative potential patches on the grooves (results not shown).

High Resolution Surface Electrostatic Potential Maps of Large-Scale Biomolecular Assemblies: Ribosomes. Due to the large-scale and highly charged nature of biomolecular assemblies such as ribosomes, which can contain more than a million atoms, it is very challenging to obtain stable and accurate electrostatic properties with standard 3D lattice nonlinear PB solvers. To date, these calculations require access to supercomputers and special techniques such as parallel focusing.[14,68] Moreover, these calculations often encounter convergence issues when using the nonlinear PBE necessary to properly model these highly charged systems at the all-atom level and for resolutions finer than 0.6 Å.[69−72]

The computation of the surface potential can be computationally demanding for the large-scale biomolecular assemblies here considered, making such computations inaccessible to desktop computers and even large clusters. Thus, to the best of our knowledge, the results here shown represent the first nonlinear PB calculation done on a serial platform for such a large-scale biomolecular system at a level of fine grid resolution of 0.3 Å using an all-atom model of a ribosomal subunit. All other reported surface potential maps of large-scale biomolecular assemblies, such as the small 30S ribosomal subunit or viruses, that were done using serial computers were obtained with the linear PBE solution, coarser grid resolutions, or more approximate generalized Born-based approaches.[3,69−76]

Here, a high resolution electrostatic potential map of the large 50S ribosomal subunit structure from *H. marismortui* (PDB id: 3cc4) was computed using the nonlinear PBE. This large ribosomal subunit consists of 5S and 23S RNA and numerous proteins with 150 970 atoms and has a net charge of −2949*e* (see Figure 9a). The cocrystal structure of anisomycin bound to the 50S ribosomal subunit was taken from the RCSB PDB Databank. All cofactors including metals and drugs were removed from the structure and only the protein and RNA chains retained. The CHARMM[38] force field atomic radii and charges were used for these PB calculations after the missing hydrogen atoms were added to the structure using the pdb2pqr server.[77] The 1:1 salt concentration was 0.1 M ($\kappa = 0.1030$ Å^{-1}), and the dielectric constants ε_1 and ε_2 were 2 and 80, respectively.

Using a finest mesh spacing of 0.3 Å results in an ACG mesh with a total of 52.5 million nodes. Meshing the minimum enclosing

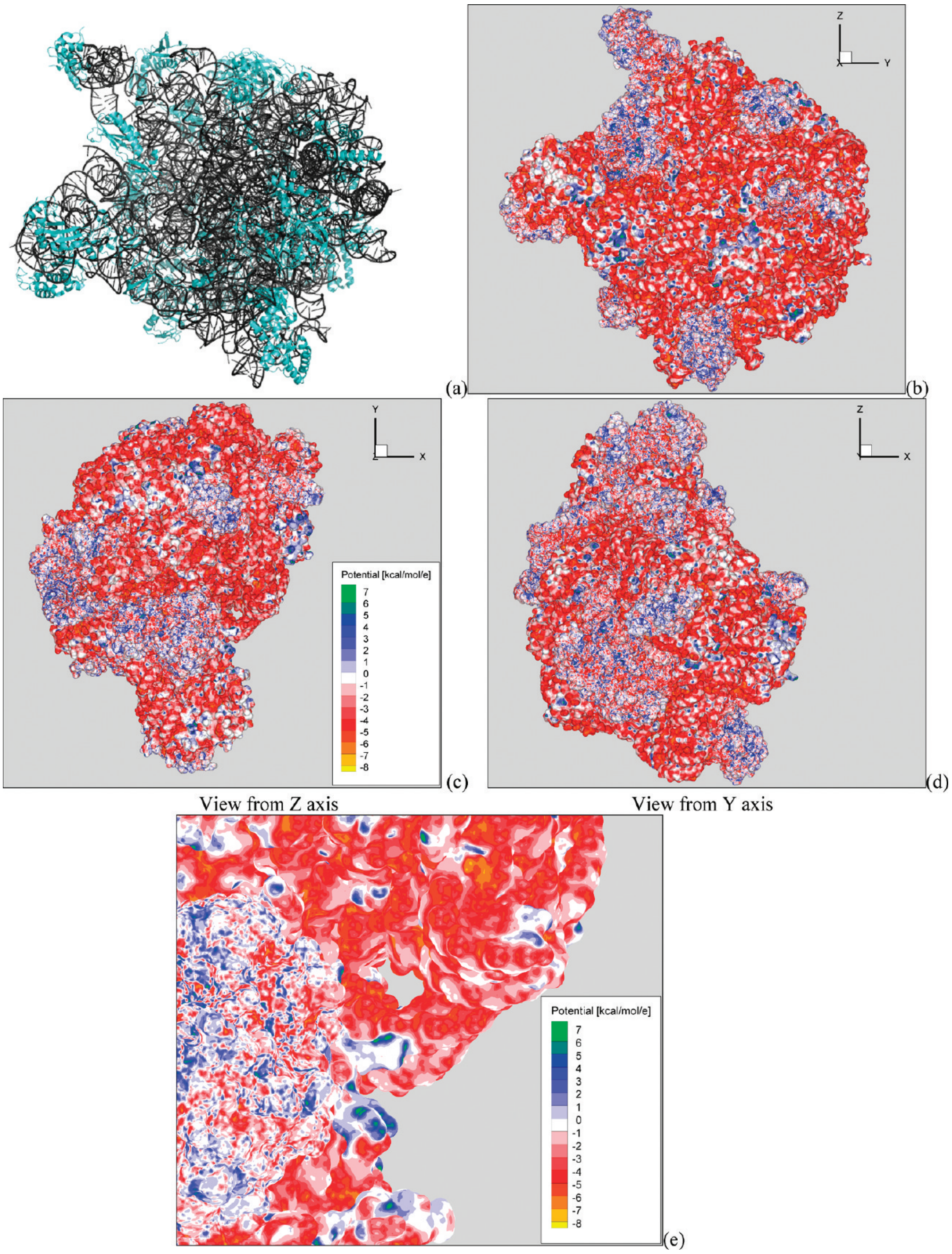


Figure 9. (a) Ribbon representation of the 50S ribosomal subunit (PDB id: 3cc4; net charge: −2949e; 150 970 atoms). The protein and rRNA molecules are shown in cyan and dark gray, respectively. (b−d) Different views of the surface potential (in kcal/mol/e) of the whole 50S ribosomal subunit. Note that the red and blue patches correspond to regions where the RNA and protein lie, respectively. (e) Closeup view of a particular intricate region of the complex 50S subunit showing the high quality of the generated surface potential map using the ACG nonlinear PB solver at the required mesh spacing to resolve the surface geometry.

box using a regular lattice grid with the same finest spacing would require 354 million nodes. Here, the outer boundary side length is 3 times larger than the longest molecular dimension, and thus the complete mesh spans 1228 Å. Solving the nonlinear PBE for this configuration produces stable and converged results within 170 iterations. With the ACG-based PBE solver, the nonlinear PBE solution for this large-scale and highly charged biomolecular assembly took 13.5 h using a 10-node 64-bit SGI Altix workstation. Machines of this caliber are widespread in university research departments and small businesses conducting computational biophysics research.

Figure 9b−d shows the electrostatic potential maps for the 50S ribosomal subunit (PDB id: 1CC4) and viewed from z, y, and x axes. The regions of positive and negative potential on the molecular surface correspond to the locations of the proteins and RNA, respectively. The presence of the proteins is essential in order to neutralize the close repulsive phosphate−phosphate interactions and thus help stabilize this intricate large-scale protein−RNA complex. A closeup of a portion of the surface potential in Figure 9e reveals the resolution and attendant quality of the surface potential map. The linear PB solution provides a different potential distribution on the surface of this highly charged biomolecular entity (results not shown). Thus, the nonlinear Poisson−Boltzmann solution should be used when modeling nonspecific electrostatic interactions of the ribosome, its assembly process, and associations with charged drug ligands.

CONCLUDING REMARKS AND FUTURE DIRECTIONS

A finite difference method to solve the linear/nonlinear Poisson−Boltzmann equation has been formulated and implemented on the grid structure known as an adaptive Cartesian mesh or octree. The generation of the mesh about a biomolecular structure, the construction of the finite difference operators upon the mesh, the representation of the electrostatic potential inside and outside the molecular surface, and preliminary computational results have been presented in this paper. Properties and advantages of the ACG-based PBE approach include the following:

- fast mesh generation due to the simple fundamental shape of the ACG cells
- optimized grid spacing where fine cells are used where the potential gradients are changing most rapidly (i.e., at the surface) and coarser elements used elsewhere
- use of compact finite difference formulas to evaluate the dielectric-weighted Laplacian and tailored for implementation on the ACG
- a representation of the potential (total potential outside the molecule and reaction field potential inside), which completely eliminates charge singularities and numerically induced self-charging energies
- a robust multigrid-accelerated convergence scheme
- the incorporation of a recently developed outer boundary treatment to estimate the boundary potential and provide first-order (i.e., based on a monopole approximation) corrections to computed energies.

Application of the method to idealized configurations involving charged and low dielectric spheres embedded in a high dielectric ionic solvent has confirmed that the method successfully maintains high accuracy as a charge is placed near the surface, properly predicts the electrostatic interaction energies for a pair of charged spheres, and reliably converges solutions for very highly charged systems. Comparisons with semianalytical solutions to the nonlinear PBE have verified that the ACG-based method accurately reproduces the salt-dependent behavior of highly charged spheres immersed in 1:1 salt solutions. PB calculations involving very complex biomolecular systems involving highly charged nucleic acid assemblies including the 50S ribosomal subunit have also been carried out successfully. The ACG-PB solver in conjuction with molecular dynamics or Brownian dynamics techniques should allow more careful and systematic studies of the role of nonspecific electrostatic interactions on the binding of various antibacterial drugs to the ribosome and ribosome and virus assembly processes at atomic resolution. An assessment of the performance of the nonlinear and linear PB predictions of electrostatic solvation free energies for a test set of 55 proteins—that vary in size, shape, and charge distribution—is also provided in order to establish benchmark test cases for comparisons with other PB solvers.

Ongoing activity in the development of the ACG-based PBE solver includes improved treatment of the solution near the dielectric interface to obtain more accurate predictions of the surface potential and normal gradients and, hence, forces; incorporation of nonuniform ion size effects; the calculation of electrostatic interaction energies between the two molecules where the bound and unbound states differ; and validation/testing of all of the above new ACG-PB features for a variety of biomolecular systems.

ASSOCIATED CONTENT

Supporting Information. Electrostatic free energies and its salt sensitivities for a set of 55 proteins based on the linear and nonlinear ACG-PB solution. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: alex@continuum-dynamics.com; mfenley@sb.fsu.edu.

ACKNOWLEDGMENT

This work has been supported by the NIH under Grant numbers 5 R44 GM57764-03 and 5R44GM073391-03 with technical monitors, Dr. Janna Wehrle and Jeni Smits. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health. One of the authors (M.O.F.) acknowledges the support from NIH-GM07838-01 (PI: Dr. Michael S. Chapman, OHSU, Co-PI: Dr. Marcia O. Fenley). One of the authors (M.O.F.) would like to thank Dr. Cristina Russo, Mr. Travis Mackoy, and Mr. Robert Harris for help in preparing structures, figures, and some PB calculations.

NOMENCLATURE

C = energy conversion factor, as defined in eq 22
e = protonic charge
$f(\Phi)$ = mobile ion charge function, as defined in eq 12
G^{el} = total electrostatic free energy, as defined in eq 20
G_f = fixed charged energy contribution, as defined in eq 21a
G_m = dielectric stress energy contribution, as defined in eq 21b
G_{rf} = reaction field energy, as defined in eq 24
$\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ = unit vectors along x, y, and z, respectively
$I_{1:1}$ = ionic strength of the 1:1 (monovalent) salt
k_B = Boltzmann constant
M^{ℓ} = ℓth level mesh in multigrid scheme (M^0 is the finest level mesh)
q_k = normalized charge, $q_k = (4\pi e/\varepsilon_1 k_B T)Q_k$, where ε_1 is the interior dielectric constant
Q_k = value of the kth charge, in units of e (electron charge)
$\mathbf{R}_i$ = position of the ith node in the ACG mesh
t = Stern layer or ion-exclusion thickness, in Å
T = absolute temperature of the aqueous salt solution, in K
$\delta(\mathbf{r})$ = 3D Dirac δ function centered at $\mathbf{r} = \mathbf{0}$
Δ_i = size (side length) of the ith mesh cell
$\Delta\Pi$ = excess osmotic pressure energy, as defined in eq 21c
ε = dielectric constant
ε_1, ε_2 = dielectric constant in the interior (Ω_1) and exterior (Ω_2 and Ω_3) regions, respectively
Φ = reduced (or dimensionless) total electrostatic potential
Φ^c = Coulombic potential, as defined in eq 14

Φ^g = potential field computed on the ACG grid and defined in eq 16

Φ^{rf} = reaction field potential, $\Phi^{rf} = \Phi - \Phi^c$

κ = Debye—Hückel screening parameter, as defined in eq 11, in $Å^{-1}$

λ_i = residual, as defined in eq 18

ρ = charge density, in e/$Å^3$

ρ^f = contribution of fixed solute charges to the total charge density, in $e/Å^3$

ρ^m = contribution of mobile ions to the total charge density, in $e/Å^3$

ρ_k = position of kth charge

σ_i = Coulombic source term

ω_{ij} = weights in the discrete approximation to the weighted Laplacian, e.g., eq 17

Ω_i = volume domains corresponding to the molecule interior (Ω_1), the Stern layer (Ω_3), and the remaining region, $\Omega_2 = R^3 - \Omega_1 - \Omega_3$

ACG = adaptive Cartesian grid

ACG-PBE = ACG-based PBE solver

BEM = boundary element method

FD = finite difference

FE = finite element

PBE = Poisson—Boltzmann equation

PDB = Protein Data Bank

PDE = partial differential equation

SE = solvent-excluded (surface)

vdW = van der Waals (surface)

REFERENCES

(1) Lu, B.; Zhou, Y. C.; Holst, M. J.; McCammon, J. A. Recent progress in numerical methods for the Poisson-Boltzmann equation in biophysical applications. *Commun. Comput. Phys.* **2008**, *3*, 973–1009.

(2) Lu, J.; Deutsch, C. Electrostatics in the Ribosomal Tunnel Modulate Chain Elongation Rates. *J. Mol. Biol.* **2008**, *384*, 73–86.

(3) Zhang, D.; Konecny, R.; Baker, N. A.; McCammon, J. A. Electrostatic interaction between RNA and protein capsid in cowpea chlorotic mottle virus simulated by a coarse-grain RNA model and a Monte Carlo approach. *Biopolymers* **2004**, *75*, 325–337.

(4) Boschitsch, A. H.; Fenley, M. O.; Zhou, H.-X. Fast Boundary Element Method for the Linear Poisson-Boltzmann Equation. *J. Phys. Chem. B* **2002**, *106*, 2741–2754.

(5) Bharadwaj, R.; Windemuth, A.; Sridharan, S.; Honig, B.; Nicholls, A. The Fast Multipole Boundary Element Method for Molecular Electrostatics: An Optimal Approach for Large Systems. *J. Comput. Chem.* **1995**, *16*, 898–913.

(6) Purisima, E. O. Fast Summation Boundary Element Method for Calculating Solvation Free Energies of Macromolecules. *J. Comput. Chem.* **1998**, *19*, 1494–1504.

(7) Zauhar, R. J.; Varnek, A. A Fast and Space Efficient Boundary Element Method for Computing Electrostatic and Hydration Effects in Large Molecules. *J. Comput. Chem.* **1996**, *17*, 864–877.

(8) Boschitsch, A. H.; Fenley, M. O.; Olson, W. K. A Fast Adaptive Multipole Algorithm for Calculating Screened Coulomb (Yukawa) Interactions. *J. Comput. Phys.* **1999**, *151*, 212–241.

(9) Lu, B.; Cheng, X.; Huang, J.; McCammon, J. A. Order N algorithm for computation of electrostatic interactions in biomolecular systems. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 19314–19319.

(10) Lu, B.; Cheng, X.; McCammon, J. A. "New-version-fast-multipole-method" accelerated electrostatic calculations in biomolecular systems. *J. Comput. Phys.* **2007**, *226*, 1348–1366.

(11) Boschitsch, A. H.; Fenley, M. O. Hybrid Boundary Element and Finite Difference Method for Solving the Nonlinear Poisson—Boltzmann Equation. *J. Comput. Chem.* **2004**, *25*, 935–955.

(12) Baker, N. A.; Holst, M. J.; Wang, F. Adaptive Multilevel Finite Element Solution of the Poisson-Boltzmann Equation II. Refinement at Solvent-Accessible surfaces in Biomolecular Systems. *J. Comput. Chem.* **2000**, *21*, 1343–1352.

(13) Holst, M.; Baker, N.; Wang, F. Adaptive Multilevel Finite Element Solution of the Poisson-Boltzmann Equation I: Algorithms and Examples. *J. Comput. Chem.* **2000**, *21*, 1319–1342.

(14) 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.

(15) 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.

(16) Jo, S.; Vargyas, M.; Vasko-Szedlar, M.; Roux, B.; Im, B. PBEQ-Solver for online visualization of electrostatic potential of biomolecules. *Nucleic Acids Res.* **2008**, *36*, W270–W275.

(17) Bashford, D.; Gerwert, K. Electrostatic Calculations of the pKa Values of Ionizable Groups in Bacteriorhodopsin. *J. Mol. Biol.* **1992**, *224*, 473–486.

(18) Grant, J. A.; Pickup, B. T.; Nicholls, A. A smooth permittivity function for Poisson-Boltzmann solvation methods. *J. Comput. Chem.* **2001**, *22*, 608–640.

(19) Gilson, M. K.; Sharp, K.; Honig, B. Calculating the electrostatic potential of molecules in solution: Method and error assessment. *J. Comput. Chem.* **1987**, *9*, 327–335.

(20) Cai, Q.; Hsieh, M.-J.; Wang, J.; Luo, R. Performance of Nonlinear Finite-Difference Poisson-Boltzmann Solvers. *J. Chem. Theory Comput.* **2010**, *6*, 203–211.

(21) Luo, R.; David, L.; Gilson, M. K. Accelerated Poisson-Boltzmann calculations for static and dynamic systems. *J. Comput. Chem.* **2002**, *23*, 1244–1253.

(22) Wang, J.; Luo, R. Assessment of Linear Finite-Difference Poisson-Boltzmann Solvers. *J. Comput. Chem.* **2010**, *31*, 1689–1698.

(23) Boschtisch, A. H.; Fenley, M. O. A New Outer Boundary Formulation and Energy Corrections for the Nonlinear Poisson—Boltzmann Equation. *J. Comput. Chem.* **2007**, *28*, 909–921.

(24) Gilson, M. K.; Sharp, K. A.; Honig, B. Calculating electrostatic interactions in biomolecules: method and error assessment. *J. Comput. Chem.* **1988**, *9*, 327–335.

(25) Cortis, C. M.; Friesner, R. A. An Automatic Three-Dimensional Finite Element Mesh Generation System for the Poisson-Boltzmann Equation. *J. Comput. Chem.* **1997**, *18*, 1570–1590.

(26) Bajaj, C. L.; Xu, G.; Zhang, Q. A Fast Variational Method for the Construction of Resolution Adaptive C^2-Smooth Molecular Surfaces. *Comput. Methods Appl. Mech. Eng.* **2009**, *198*, 1684–1690.

(27) Samet, H. *The Design and Analysis of Spatial Structures*; Addison-Wesley Publishing Company, Inc.: Reading, MA, 1990; p 510.

(28) Mirzadeh, M.; Theillard, M.; Gibou, F. A second-order discretization of the nonlinear Poisson—Boltzmann equation over irregular geometries using non-graded adaptive Cartesian grids. *J. Comput. Phys.* **2011**, *230*, 2125–2140.

(29) Aftosmis, M. J.; Berger, M. J.; Melton, J. E. Robust and Efficient Cartesian Mesh Generation for Component-Based Geometry. *AIAA J.* **1998**, *36*, 952–960.

(30) Murman, S. M.; Aftosmis, M. J.; Berger, M. J., Simulations of Store Separation from an F/A-18 with a Cartesian Method. *J. Aircraft* **2004**, *41*, (4).

(31) Aftosmis, M. J. *Solution Adaptive Cartesian Grid Methods for Aerodynamic Flows with Complex Geometries*; Lecture Notes 1997—02; Von Karman Institute for Fluid Dynamics: Rhode-St-Genèse, Belgium, 1997.

(32) Aftosmis, M. J.; Berger, M. J.; Melton, J. E. Robust and Efficient Cartesian Mesh Generation for Component-Based Geometry. In *35th AIAA Aerospace Sciences Meeting & Exhibit, AIAA-97-0196*, Reno, NV, 1997; AIAA: Reston, VA, 1997.

(33) Berger, M. J.; LeVeque, R. J. An Adaptive Cartesian Mesh Algorithm for the Euler Equations in Arbitrary Geometries. In *89-1930-CP*; AIAA: Reston, VA, 1989.

(34) Wang, J.; Cai, Q.; Li, Z.-L.; Zhao, H.-K.; Luo, R. Achieving energy conservation in Poisson−Boltzmann molecular dynamics: Accuracy and precision with finite-difference algorithms. *Chem. Phys. Lett.* **2009**, *468*, 112–118.

(35) Zhou, Z.; Payne, P.; Vasquez, M.; Kuhn, N.; Levitt, M. Finite-Difference Solution of the Poisson-Boltzmann Equation: Complete Elimination of Self-Energy. *J. Comput. Chem.* **1996**, *11*, 1344–1351.

(36) Bondi, A. Van der Waals Volumes and Radii. *J. Phys. Chem.* **1964**, *68*, 441–451.

(37) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A 2nd generation force-field for simulation of proteins, nucleic-acids, and organic-molecules. *J. Am. Chem. Soc.* **1995**, *117*, 11946–11975.

(38) Foloppe, N.; MacKerell, A. D. J. All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *J. Comput. Chem.* **2000**, *21*, 86–104.

(39) Friedrichs, M.; Zhou, R.; Edinger, S. R.; Friesner, R. A. Poisson-Boltzmann Analytical Gradients for Molecular Modeling Calculations. *J. Phys. Chem. B* **1999**, *103*, 3057–3061.

(40) Chan, S. L.; Purisima, E. O. Molecular surface generation using marching tetrahedra. *J. Comput. Chem.* **1998**, *19*, 1268–1277.

(41) Protter, M.; Weinberger, H. F. *Maximum Principles in Differential Equations*; Prentice Hall: Englewood Cliffs, NJ, 1967; reprint by Springer-Verlag: New York, 1984.

(42) Barth, T. J. Numerical Aspects of Computing Viscous High Reynolds Number Flows on Unstructured Meshes. In *29th Aerospace Sciences Meeting*, AIAA-91-0721, Reno, NV, 1991; AIAA: Reston, VA, 1991.

(43) Chen, S.-W. W.; Honig, B. Monovalent and Divalent Salt Effects on Electrostatic Free Energies Defined by the Nonlinear Poisson-Boltzmann Equation: Application to DNA Binding Reactions. *J. Phys. Chem. B* **1997**, *101*, 9113–9118.

(44) Greengard, L.; Huang, J. *A New Version of the Fast Multipole Method for Screened Coulomb Interactions in Three Dimensions*, 01-002; Courant Mathematics and Computing Laboratory, Courant Institute: New York, 2001; p 18.

(45) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes*, 2nd ed.; Cambridge University Press: Cambridge, U. K., 1992; p 960.

(46) Holst, M.; Saied, F. Multigrid Solution of the Poisson-Boltzmann Equation. *J. Comput. Chem.* **1993**, *14*, 105–113.

(47) Oberoi, H.; Allewell, N. M. Multigrid Solution of the Nonlinear Poisson-Boltzmann Equation and Calculation of Titration Curves. *Biophys. J.* **1993**, *65*, 48–55.

(48) Sharp, K.; Honig, B. Calculating total electrostatic energies with the nonlinear Poisson-Boltzmann equation. *J. Phys. Chem.* **1990**, *94*, 7684–7692.

(49) Sharp, K. A. Polyelectrolyte Electrostatics: Salt Dependence, Entropic, and Enthalpic Contributions to Free Energy in the Nonlinear Poisson-Boltzmann Model. *Biopolymers* **1995**, *36*, 227–243.

(50) Sharp, K. A.; Friedman, R. A.; Misra, V.; Hecht, J.; Honig, B. Salt Effects on Polyelectrolyte-Ligand Binding: Comparison of Poisson-Boltzmann, and Limiting Law/Counterion Binding Models. *Biopolymers* **1995**, *36*, 245–262.

(51) Kirkwood, J. G. Theory of Solutions of Molecules Containing Widely Separated Charges with Special Applications to Zwitterions. *J. Chem. Phys.* **1934**, *2*, 351–361.

(52) Sader, J. E.; Lenhoff, A. M. Electrical Double-Layer Interaction between Heterogeneously Charged Colloidal Particles: A Superposition Formulation. *J. Colloid Interface Sci.* **1998**, *201*, 233–243.

(53) Zhou, H.-X. Macromolecular Electrostatic Energy Within the Nonlinear Poisson-Boltzmann Equation. *J. Chem. Phys.* **1994**, *100*, 3152–3162.

(54) Schreiber, G.; Haran, G.; Zhou, H.-X. Fundamental aspects of protein-protein association kinetics. *Chem. Rev.* **2009**, *109*, 839–860.

(55) Getzoff, E. D.; Cabelli, D. E.; Fisher, C. L.; Parge, H. E.; Viezzoli, M. S.; Banci, L.; Hallewell, R. A. Faster superoxide dismutase mutants designed by enhancing electrostatic guidance. *Nature* **1992**, *358*, 347–351.

(56) Lago, H.; Parrott, A. M.; Moss, T.; Stonehouse, N. J.; Stockley, P. G. Probing the kinetics of formation of the bacteriophage MS2 translational operator complex: identification of a protein conformer unable to bind RNA. *J. Mol. Biol.* **2001**, *305*, 1131–1144.

(57) Zhou, H.-X. Boundary Element Solution to Macromolecular Electrostatics: Interaction energy between two proteins. *Biophys. J.* **1993**, *65*, 955–963.

(58) Bredenberg, J.; Boschitsch, A. H.; Fenley, M. O. The role of anionic protein residues on the salt dependence of the binding of aminoacyl-tRNA synthetases to tRNA: A Poisson-Boltzmann analysis. *Commun. Comput. Phys.* **2008**, *3*, 1051–1070.

(59) Bredenberg, J. H.; Russo, C.; Fenley, M. O. Salt-mediated electrostatics in the association of TATA binding proteins to DNA: A combined molecular mechanics/Poisson-Boltzmann study. *Biophys. J.* **2008**, *94*, 4634–4645.

(60) Bredenberg, J. H.; Fenley, M. O. Salt dependent association of novel mutants of TATA-binding proteins to DNA: Predictions from theory and experiments. *Commun. Comput. Phys.* **2008**, *3*, 1132–1153.

(61) Fenley, M. O.; Harris, R. C.; Jayaram, B.; Boschitsch, A. H. Revisiting the association of cationic groove-binding drugs to DNA using a Poisson-Boltzmann approach. *Biophys. J.* **2010**, *99*, 879–886.

(62) Harris, R. C.; Bredenberg, J. H.; Silalahi, A. R. J.; Boschitsch, A. H.; Fenley, M. O. Understanding the physical basis of the salt dependencd of the electrostatic binding free energy of mutated charged ligand-nucleic acid complexes. *Biophys. Chem.* **2011**.

(63) Xu, D.; Greenbaum, N. L.; Fenley, M. O. Recognition of the spliceosomal branch site RNA helix on the basis of surface and electrostatic features. *Nucleic Acids Res.* **2005**, *33*, 1154–1161.

(64) Xu, D.; Landon, T.; Greenbaum, N. L.; Fenley, M. O. The electrostatic characteristics of G·U wobble base pairs. *Nucleic Acids Res.* **2007**, *35*, 3836–3847.

(65) Auweter, S. D.; Fasan, R.; Reymond, L.; Underwood, J. G.; Black, D. L.; Pitsch, S.; Allain, F., H.-T. Molecular Basis of RNA Recognition by the Human Alternative Splicing Factor Fox-1. *EMBO J.* **2006**, *25*, 163–173.

(66) Auweter, S. D.; Oberstrass, F. C.; Allain, F., H.-T. Sequence-specific Binding of Single-stranded RNA: Is there a Code for Recognition? *Nucleic Acids Res.* **2006**, *34*, 4943–4959.

(67) Xu, D. *Electrostatics of Nucleic Acids and Hydration Properties of the Pseudouridine Dependent Spliceosomal Branch Site Helix*; Florida State University: Tallahassee, FL, 2007.

(68) Sayyed-Ahmad, A.; Miao, Y.; Ortoleva, P. Poisson-Boltzmann theory of bionanosystems. *Commun. Comput. Phys.* **2008**, *3*, 1100–1116.

(69) Trylska, J.; Konecny, R.; Tama, F.; Brooks, C. L. I.; McCammon, J. A. Ribosome motions modulate electrostatic properties. *Biopolymers* **2004**, *74*, 423–431.

(70) Trylska, J.; McCammon, J. A.; Brooks, C. L. I. Exploring Assembly Energetics of the 30S Ribosomal Subunit Using an Implicit Solvent Approach. *J. Am. Chem. Soc.* **2005**, *127*, 11125–11133.

(71) Devkota, B.; Petrov, A.; Lemieux, S.; Bpz, M. B.; Tang, L.; Schneemann, A.; Jonhson, J. E.; Harvey, S. C. Structural and Electrostatic Characterization of Pariacoto Virus: Implications for Viral Assembly. *Biopolymers* **2009**, *91*, 530–538.

(72) Dlugosz, M.; Trylska, J. Aminoglycoside association pathways with the 30S ribosomal subunit. *J. Phys. Chem. B* **2009**, *113*, 7322–7330.

(73) Qin, S.; Zhou, H.-X. Dissection of the High Rate Constant for the Binding of a Ribotoxin to the Ribosome. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 6974–6979.

(74) Lu, B.; Cheng, X.; Huang, J.; McCammon, J. A. An Adaptive Fast Multipole Boundary Element Method for Poisson-Boltzmann Electrostatics. *J. Chem. Theory Comput.* **2009**, *5*, 1692–1699.

(75) Gordon, J. C.; Fenley, A. T.; Onufriev, A. An Analytical Approach to Computing Biomolecular Electrostatic Potential. II. Validation and Applications. *J. Chem. Phys.* **2008**, *129*, 075102−1–075102−11.

(76) Pacios, L. F.; Garcia-Arenal, F. Comparison of Properties of particles of Cucumber mosaic virus and Tomato aspermy virus based on the analysis of molecular surfaces of capsids. *J. Gen. Virol.* **2006**, *87*, 2073–2083.

(77) Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A. PDB2PQR: an automated pipeline for the setup, execution, and analysis of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* **2004**, *32*, W665–W667.

(78) Bahadur, R. P.; Zacharias, M.; Janin, J. Dissecting Protein-RNA Recognition Sites. *Nucleic Acids Res.* **2008**, *36*, 2705–2716.

JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Role of the Axial Base in the Modulation of the Cob(I)alamin Electronic Properties: Insight from QM/MM, DFT, and CASSCF Calculations

Neeraj Kumar,[†] Mercedes Alfonso-Prieto,[‡,§] Carme Rovira,[‡,‖] Piotr Lodowski,[⊥] Maria Jaworska,[⊥] and Pawel M. Kozlowski*[,†]

[†]Department of Chemistry, University of Louisville, Louisville, Kentucky 40292, United States

[‡]Institut de Química Teòrica i Computacional (IQTCUB) and Computer Simulation and Modeling Laboratory (CoSMoLab), Parc Científic de Barcelona, Baldiri Reixac 10-12, 08028 Barcelona, Spain

[§]Institute for Computational Molecular Science, Temple University, 1900 North 12th Street, Philadelphia, Pennsylvania 19122, United States

[‖]Institució Catalana de Recerca i Estudis Avançats (ICREA)

[⊥]Department of Theoretical Chemistry, Institute of Chemistry, University of Silesia, Szkolna 9, PL-40 006 Katowice, Poland

S *Supporting Information*

ABSTRACT: Quantum chemical computations are used to study the electronic and structural properties of the cob(I)alamin intermediate of the cobalamin-dependent methionine synthase (MetH). QM(DFT)/MM calculations on the methylcobalamin (MeCbl) binding domain of MetH reveal that the transfer of the methyl group to the substrate is associated with the displacement of the histidine axial base (His759). The axial base oscillates between a His-on form in the Me-cob(III)lamin:MetH resting state, where the Co−N(His759) distance is 2.27 Å, and a His-off form in the cob(I)alamin:MetH intermediate (2.78 Å). Furthermore, QM/MM and gas phase DFT calculations based on an unrestricted formalism show that the cob(I)alamin intermediate exhibits a complex electronic structure, intermediate between the Co(I) and Co(II)-radical corrin states. To understand this complexity, the electronic structure of Im···[Cob(I)alamin] is investigated using multireference CASSCF/QDPT2 calculations on gas phase models where the axial histidine is modeled by imidazole (Im). It is found that the correlated ground state wave function consists of a closed-shell Co^{I} (d^8) configuration and a diradical contribution, which can be described as a Co^{II} (d^7)-radical corrin $(\pi^*)^1$ configuration. Moreover, the contribution of these two configurations depends on the $Co-N_{Im}$ distance. At short $Co-N_{Im}$ distances (<2.5 Å), the dominant electronic configuration is the diradical state, while for longer distances it is the closed-shell state. The implications of this finding are discussed in the context of the methyl transfer reaction between the Me-H_4folate substrate and cob(I)alamin.

1. INTRODUCTION

The key step in the catalytic cycle of the cobalamin-dependent methionine synthase (MetH) enzyme is the transfer of a methyl group from the methylcobalamin (MeCbl) cofactor to the homocysteine (Hcy) substrate.[1−14] The resulting cob(I)alamin intermediate is remethylated by methyl-tetrahydrofolate (Me-H_4folate) to generate back methyl-cob(III)alamin and tetrahydrofolate (H_4folate; Scheme 1).[15−19] In other words, during the catalytic cycle, the cobalt center oscillates between methyl-cob(III)alamin and cob(I)alamin.

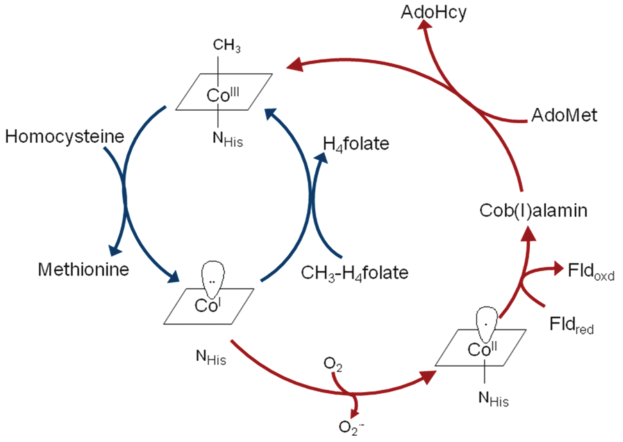
The MetH enzyme has a modular architecture, and the catalytic methyl transfer involves the interaction among different domains.[20] The crystal structure of the whole enzyme has not yet been resolved, mainly owing to the very high degree of conformational flexibility, but the X-ray crystal structure of individual domains,[18,21−23] including the one that binds MeCbl,[15] has been well characterized. During the course of catalytic reaction (Scheme 1), the two domains binding the MeCbl cofactor and the Hcy substrate form a reaction complex in which the substrate interacts with the MeCbl from the upper face of the cofactor. However, there is no structural information available with regard to the reaction complex, and therefore, the details of the reaction mechanism involving methylation of the Hcy substrate remain a subject of debate. It is generally believed that the enzyme operates via an S_N2-type nucleophilic displacement,[7,8] though an alternative mechanism in which the one-electron reduction of the MeCbl cofactor takes place has also been proposed.[24,25] This pathway does not impose specific geometrical and distance constraints with respect to the substrate and cofactor as does the S_N2 mechanism, which may be advantageous from the enzymatic point of view.

The transfer of the methyl group from the MeCbl cofactor to the Hcy substrate results in the formation of the cob(I)alamin intermediate.[26,27] In solution, this complex is tetra-coordinated because the change of the Co oxidation state induces the detachment of the axial 5,6-dimethylbenzimidazole (DBI) base (Figure 1a).[28] However, the question is how such displacement of the axial His base takes place inside the enzyme. Due to its high reactivity, there are no structural data available for the enzyme-bound cob(I)alamin.

Several theoretical studies have investigated the electronic and structural properties of the tetra-coordinated cob(I)alamin complex (i.e., without the axial base) using density functional theory

Received: January 27, 2011
Published: April 14, 2011

Scheme 1. The Catalytic Cycle (Shown in Blue) and the Reactivation Cycle (Shown in Red) for the Cobalamin-Dependent Methionine Synthase (MetH)[a]



[a] H_4folate, tetrahydrofolate; CH_3-H_4folate, methyl-tetrahydrofolate; Fld_{oxd}, oxidized form of flavodoxin; Fld_{red}, reduced form of flavodoxin; AdoHcy, adenosyl-homocysteine; AdoMet, adenosyl-methionine.

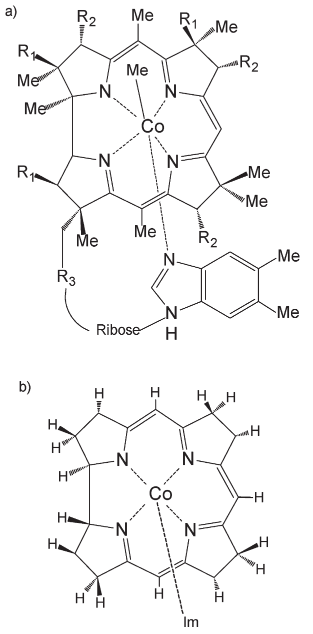


Figure 1. (a) Molecular structure of free methylcobalamin, where R = CH_3, $R_1 = CH_2CONH_2$, $R_2 = CH_2CH_2CONH_2$, and $R_3 = (CH_2)_2CONHCH_2$-$CH(CH_3)OPO_3^-$. (b) Gas phase model of the Im···[Cob(I)alamin] enzymatic intermediate employed in the present work.

(DFT) as well as multireference CASSCF calculations.[29−32] Jensen and Ryde[30] first noticed that the B3LYP-based wave function of the ground-state has a singlet instability, indicative of a complex wave function involving other configurations than $Co^I(d^8)$. Later, Jensen[31] demonstrated that the ground state of cob(I)alamin is multiconfigurational using CASSCF/CASPT2 calculations. The correlated wave function was found to consist of a closed-shell $Co^I(d^8)$ configuration (67%) and a diradical $Co^{II}(d^7)$-radical corrin $(\pi^*)^1$ configuration (23%). The formation of such unusual electronic configuration was explained by the overlap of low-lying metal d

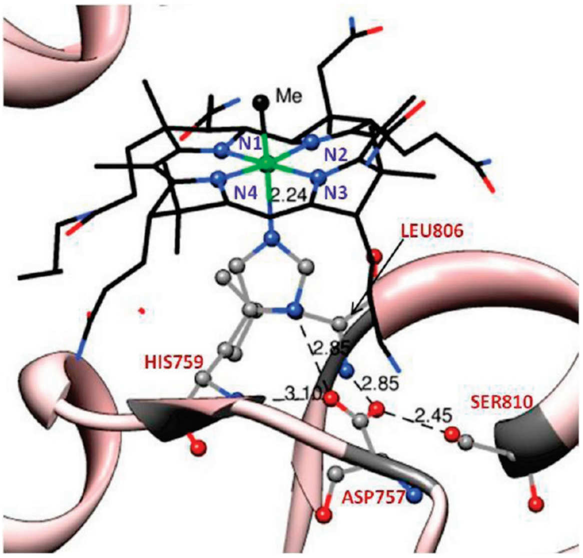


Figure 2. Close view of active site of MeCbl in the enzyme showing the interaction of lower axial ligand His759 with the other residues of the triad (Asp759, Ser810) as well as the nearby Leu806 based on the 1BMT crystal structure.[15]

orbitals with ligand orbitals at the singlet ground state, allowing the transfer of an electron from the Co to the corrin ligand. In this regard, the electronic structure of cob(I)alamin resembles the heme-based compound I intermediates, as pointed out by Ryde and Jensen.[33] Both cofactors have a noninnocent macrocycle that can exchange electrons with the metal. In compound I, one electron is transferred from the porphyrin to the Fe metal atom, whereas in the case of cob(I)alamin, electron transfer takes place in the opposite direction, i.e., from Co to the corrin ligand. Indeed, this proposed electron transfer is in agreement with a previous TD-DFT(B3LYP) study by Jaworska and Lodowski,[29] who found that the lowest energy bands (at 700 and 554 nm) in the electronic absorption spectrum of cob(I)alamin have $d \rightarrow \pi^*$ metal-to-ligand charge transfer (MLCT) character. However, a subsequent TD-DFT(PBE) study by Liptak and Brunold[32] questioned the open-shell antiferromagnetic contribution to the ground state. As will be shown later, the use of a nonhybrid DFT functional is probably the cause of this discrepancy.

Nevertheless, all of those studies were carried out for isolated models of cob(I)alamin without the presence of the axial base or inclusion of the enzymatic environment. Therefore, the actual coordination number and electronic structure of the cob(I)alamin intermediate in MetH remain an open subject. During the catalytic methyl transfer reaction (Scheme 1), different enzyme domains interact with each other in order to either cleave or generate the Co−C bond. Simultaneously, the His759 residue is expected to move with respect to the Co metal center to break/form the Co−N_{His759} bond by analogy with the solution chemistry of the B_{12} cofactor.[28] However, the crystal structure of the Me-cob(III)lamin resting state[15] (PDB code: 1BMT, 3 Å resolution) shows that His759 interacts with Asp757 and Ser810 through a network of hydrogen bonds, suggesting that His759 would be fixed with respect to the Co center (Figure 2). On the other hand, a structural study of the reactivation complex[22] (Scheme 1) formed by the adenosyl−methionine (AdoMet) and cob(I)alamin binding domains shows that the axial His759 moves away from the cobalt center and makes specific contacts with the AdoMet domain. Thus, it is still ambiguous how His759 interacts with the cob(I)alamin cofactor during the catalytic cycle.

The purpose of this study is two-fold: First, the coordination of the axial His759 and the cob(I)alamin cofactor inside the enzyme is explored by employing the QM/MM approach. Second, the complex electronic structure of the cob(I)alamin intermediate suggested by the QM/MM results is analyzed using DFT and CASSCF/QDPT2 calculations on gas phase models. Finally, the implications of this finding are discussed in the context of the methyl transfer reaction between the Me-H_4folate and cob(I)alamin.

2. COMPUTATIONAL DETAILS

Three different theoretical methods were used to study the cob(I)alamin intermediate. First, the formation of the Co(I) state was investigated inside the cobalamin-dependent methionine synthase (MetH) enzyme employing QM(DFT)/MM. Second, the complex electronic properties of the cob(I)alamin and the influence of the axial base were investigated using gas phase DFT calculations. Finally, the multiconfigurational character of the Im···[Cob(I)alamin] was further analyzed using CASSCF/MC-XQDPT2 calculations.

2.1. QM(DFT)/MM Calculations. The crystal structure of the MeCbI binding module of MetH (PDB code: 1BMT, at 3 Å resolution)[15] was used to model the cob(I)alamin intermediate, by removing the methyl ligand and adjusting the number of electrons of the QM system, consistent with the Co(I) oxidation state. Initially, the His759 residue is bound to the Co center, since the starting crystal structure corresponds to the Me-cob(III)lamin resting state. However, upon QM/MM geometry optimization, the axial base is observed to detach from the Co(I) center. The hybrid QM(DFT)/MM calculations were performed using the method developed by Laio et al.,[34] which combines the first-principles MD method of Car and Parrinello[35] with a force-field MD methodology (i.e., QM/MM CPMD). All details regarding the QM/MM calculations can be found in ref 25. In short, the geometry of the cob(I)alamin intermediate was optimized as a closed-shell singlet using the BP86 functional,[36,37] a plane wave basis set with a 70−90 Ry kinetic energy cutoff, and Martins−Troullier pseudopotentials[38] to describe the interaction between the ionic cores and the valence electrons. Two pseudopotentials were tested for the cobalt atom: one with nine valence electrons supplemented with nonlinear core corrections[39] and another with 17 valence electrons. To explore the possibility of spin polarization between the cobalt and the corrin, a single point calculation within the local spin density approximation (LSD) was also performed. Finally, since the GGA BP86 functional is suspected to have problems describing a possible open-shell singlet state in cob(I)alamin,[30,31] the B3LYP functional was also tested.

2.2. Gas Phase DFT Calculations. To get further insight into the complex electronic properties of the cob(I)alamin intermediate suggested by the QM/MM calculations, we carried out gas phase calculations. First, we studied the His-off form of cob(I)alamin for comparison with previous studies.[29−32] Two different gas phase models were used: a big model containing the full cofactor with all of the side chains and a truncated model with C2 symmetry where the side chains have been replaced by hydrogen atoms. Second, we investigated the influence of the axial His759 in the electronic properties of cob(I)alamin. An imidazole (Im) molecule modeling the axial His759 was placed under the Co center in accordance with the orientation obtained from the QM/MM calculations, and a series of displacements along the Co(I)−N_{Im} coordinate was generated. Geometry optimization of all models was carried out employing the Becke−Perdew (BP86)[36,37] functional and the 6-31G(d) (5d components) basis set, as implemented in Gaussian 03.[40] This level of theory constitutes an appropriate platform for describing the structural and electronic properties analysis of alkyl-cobalt(III) complexes, as documented in the literature.[41−43] However, pure GGA functionals such as PBE[44] or BP86 are known to be unable to converge to a spin polarized solution for Co(I) complexes.[30,31] This is not the case for hybrid functionals such as B3LYP, despite the fact that this functional underestimates the strength of the Co−C bond.[45,46] Therefore, we used B3LYP to test the possibility of a spin-polarized solution, which may be indicative of a more complex electronic wave function.[47] In particular, B3LYP calculations were initially performed on all models having an even number of electrons, assuming a singlet closed-shell wave function. Then, the unrestricted Kohn−Sham formalism using UB3LYP was applied by mixing HOMO and LUMO orbitals to obtain the corresponding open-shell singlet. For each broken-symmetry solution, we examined the extent of spin polarization between the cobalt and the corrin by analyzing the spin density distributions.

2.3. CASSCF/MC-XQDPT2 Calculations. Since the Kohn−Sham formalism (which is the base for DFT-based computations) is restricted to a single Slater determinant description, it cannot describe the multiconfigurational character of the cob(I)alamin system. Thus, we carried out CASSCF multireference calculations, followed by quasi-degenerate perturbation theory (QDPT2)[48] calculations with a multiconfigurational self-consistent-field reference function (MC-XQDPT2) to include the dynamical correlation, as implemented in the PC GAMESS/Firefly QC package.[49] All of the CASSCF calculations were performed on the DFT optimized structures of the cob(I)alamin models (without and with the Im base) using the 6-31G(d) basis set. The details regarding the active space chosen for these multireference calculations are described in section 3.3.

3. RESULTS AND DISCUSSION

3.1. Structure of Cob(I)alamin Inside MetH. MetH is a modular enzyme composed of four functional domains.[20] The one considered here is the MeCbl binding module,[15] in which the His759 side chain serves as the lower axial ligand (Figure 2), instead of the DBI base of the free MeCbl cofactor (Figure 1a). During the course of enzymatic methyl transfer, the cobalt center oscillates between the Co(III) and the Co(I) oxidation states, and the axial His is expected to dissociate from the cobalt center by similarity with the behavior of the B_{12} cofactor in solution. However, this assumption remains to be proved since the crystal structure of the cob(I)alamin intermediate is not available. Here, we have investigated the coordination of the axial His759 and the cob(I)alamin in the MetH enzyme by means of QM/MM calculations.

The QM/MM optimized structure of the cob(I)alamin:MetH intermediate is shown in Figure 3, and the main structural parameters are listed in Table 1. The optimized structure of the Me-cob(III)lamin:MetH resting state[25] is also included for comparison. The fact that the X-ray Co−N_{ax} bond length for Me-cob(III)alamin:MetH is very well reproduced at the QM/MM level strengthens our confidence that the demethylated cob(I)alamin:MetH form would also be well characterized from

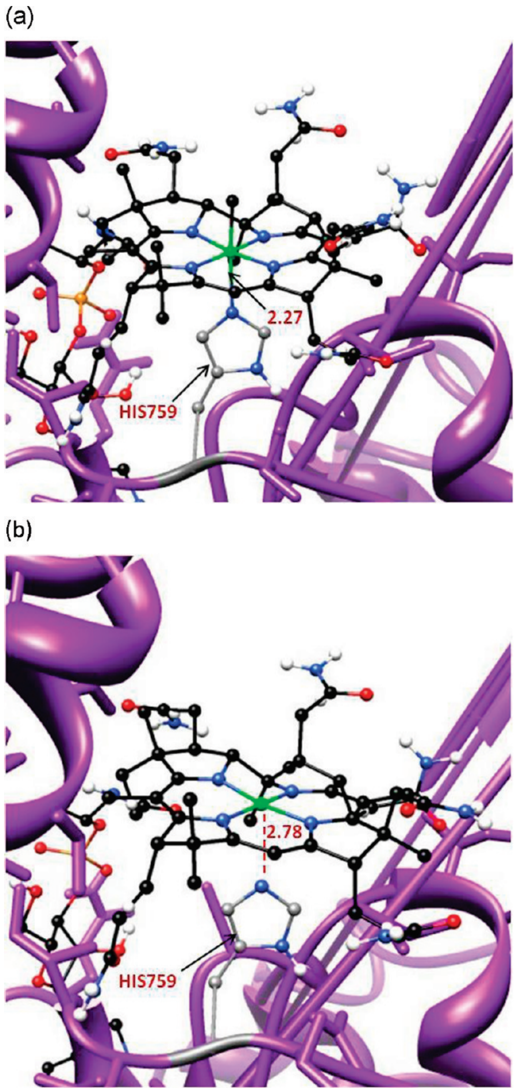


Figure 3. QM(DFT)/MM optimized structures of MeCbl binding domain of the MetH (a) hexa-coordinated Me-Cob(III)alamin resting state and (b) Cob(I)alamin intermediate.

Table 1. Key Structural Parameters of the Cofactor Binding Domain of MetH in the Me-Cob(III)alamin Resting State (MeCbl:MetH) and the Cob(I)alamin Intermediate (Co(I):MetH)[a]

	MeCbl:MetH		Co(I):MetH
parameter	QM/MM[25]	X-ray[15]	QM/MM[b]
Co−C	1.99	1.96	
Co−N_{Im}	2.27	2.24	2.78
Co−N1	1.88	1.91	1.84
Co−N2	1.86	1.93	1.81
Co−N3	1.93	2.02	1.91
Co−N4	1.92	2.02	1.87

[a] All distances are given in Å. [b] This work.

a structural point of view using the same level of theory (i.e., the BP86 functional). The calculations show that the Co−N_{ax}(His759) distance increases from 2.27 Å in the Co(III) oxidation state to 2.78 Å for Co(I). In other words, the axial ligand is displaced due to the change of the Co oxidation state. This is in agreement with the X-ray absorption spectroscopic studies showing that cob(I)alamin in solution is not axially coordinated.[28] Nevertheless, it should be noted that, differently from the cofactor in solution, the Co atom inside the enzyme remains weakly coordinated to the His. Most likely the axial His in MetH cannot move away further from the Co because it is hydrogen-bonded to Asp757 (Figure 2). Both His759 and Asp757 residues are in a loop, conferring to them a certain degree of flexibility that allows the axial base to be displaced during the catalytic cycle without breaking the hydrogen bond between them. However, this flexibility is not unlimited, because Asp757 is also interacting with a serine of an α helix (Ser810) and a leucine of a β sheet (Leu806), and these secondary structures are not as flexible as the loop. As a consequence, the Co···N(His) cannot increase beyond ∼2.8 Å without disrupting this hydrogen bond network, something that probably has a high energy cost. Maintaining a weak Co(I)−His coordination may be advantageous from the enzymatic point of view, since it would allow the axial His to recoordinate easily to the cobalt center when the cofactor is remethylated. The His−Asp−Ser triad is expected to play a key role in the Co(I) remethylation reaction by modulating the interaction between the Co and the axial base. Interestingly, the His−Asp−Ser triad is conserved in all corrinoid-based proteins catalyzing methyl transfer reactions (except in an iron−sulfur corrinoid protein), and thus it is tempting to suggest that the triad plays a similar role in all of the members of this protein family. Indeed, Hegemeier et al.[50] draw a similar conclusion for another cobalamin-dependent enzyme, i.e., the methanol−cobalamin methyltransferase (MetABC), on the basis of the long Co−N distance (2.51 Å) observed in the crystal structure of the cob(I)amide intermediate and the methylation rate of Co(I) being completely dependent on the presence of the axial base.

The structure of the cob(I)alamin:MetH intermediate was initially optimized considering a closed-shell singlet electronic configuration. However, since previous DFT as well as CASSCF calculations[29−32] have shown that free cob(I)alamin has a more complex electronic structure, we checked the possibility of having a spin polarizatied solution inside the enzyme by using the local spin density approximation. Despite several attempts, we were unable to converge the open-shell singlet either with BP86 or B3LYP. The obtained electronic state was found to be indeed an intermediate state between two electronic configurations: the closed-shell $Co^{I}(d^{8})$ singlet and the open-shell $Co^{II}(d^{7})$-corrin radical $(\pi^{*})^{1}$ singlet. The spin density distribution showed some unpaired electronic density on the cobalt and the corrin ring with opposite signs, i.e., some diradical character. Moreover, the total integrated absolute value of the spin density was $1e^{-}$ (intermediate between the 0 unpaired electrons expected for a pure closed-shell singlet and the two unpaired electrons for an open-shell singlet). This suggests that the closed-shell and the open-shell singlet states are very close in energy, so unless we force the system to be a closed-shell singlet, the single-determinantal DFT calculations give an intermediate state between Co(I) and Co(II)-corrin radical.

In summary, our QM/MM calculations confirm the early proposal of Wirt et al.[28] that the axial His ligand oscillates between the His-on form in the Me-cob(III)alamin:MetH resting state (Co−N(His) distance = 2.27 Å) and the His-off form in the cob(I)alamin:MetH intermediate (2.78 Å). Such movement of His759 in the enzymatic environment is due to the changes in the electronic structure at the metal center, i.e., from Co(III) to an intermediate Co(I)/Co(II) state. Hereafter, the complex electronic structure of the cob(I)alamin intermediate as well as the changes associated with the displacement of the axial base are further analyzed using DFT and CASSCF/QDPT2.

3.2. Structural and Electronic Analysis Based on DFT Calculations.

3.2.1. His-off Cob(I)alamin Intermediate. Table 2 shows the main structural parameters obtained for the gas phase models of the His-off form of the cob(I)alamin intermediate employing both the BP86 and B3LYP functionals. The geometries do not differ significantly from the one inside the MetH enzyme (Table 1), regardless of the model employed for the cofactor (full model with all of the side chains or truncated model with C2 symmetry) or the functional used. In particular, the calculated Co−N_{eq} distances (1.82−1.85 Å) are in agreement with the average experimental value obtained from EXAFS studies (1.86−1.88 Å).[28]

Table 2. Main Structural Parameters of the Full and Truncated Cob(I)alamin Gas Phase Models[a]

model	full			truncated			
parameter	B3LYP[b]	BP86[b]	PBE[32]	B3LYP[b]	BP86[b]	PBE[32]	CASSCF[31]
fold angle[61]				6.2	6.9	7.3	6.1
Co−N1	1.84	1.82	1.83	1.84	1.83	1.84	1.85
Co−N2	1.85	1.83	1.85	1.84	1.83	1.84	1.85
Co−N3	1.91	1.90	1.92	1.91	1.89	1.91	1.91
Co−N4	1.90	1.89	1.90	1.91	1.89	1.91	1.91

[a] All distances are given in Å, and angles are in deg. N1−N4 refers to the corrin nitrogen atoms. [b] This work.

As for the electronic structure, the main difference between the two functionals is that only B3LYP gives a spin polarized solution. This may reflect that the nonhybrid functionals such as BP86 do not describe correctly the spin polarization,[51,52] even though they give correct Co−C bond dissociation energies (BDE).[45] Figure 4 shows the (B3LYP) spin density distribution along with the number of unpaired electrons. There is one unpaired electron on the cobalt atom coupled antiferromagnetically with an unpaired electron on the corrin ring, consistent with a $Co^{II}(d^7)$-corrin radical $(\pi^*)^1$ diradical state. The spin density distributions of the full (Figure 4a) and the truncated (Figure 4b) models are almost identical. The side chains do not show any spin density, and the density exhibits 2-fold symmetry. Consequently, the truncated model with C2 symmetry was used for further analysis (Figure 4b).

From the energetic point of view, the open-shell singlet was found to be ~4 kcal/mol (B3LYP) lower in energy than the corresponding closed-shell configuration. Moreover, the ferromagnetic counterpart of the open-shell singlet, a triplet $Co^{II}(d^7)$-radical corrin $(\pi^*)^1$ state, was also found to be 3 kcal/mol lower in energy than the closed-shell singlet, further validating the diradical contribution to the cob(I)alamin intermediate. Although these small energy differences are

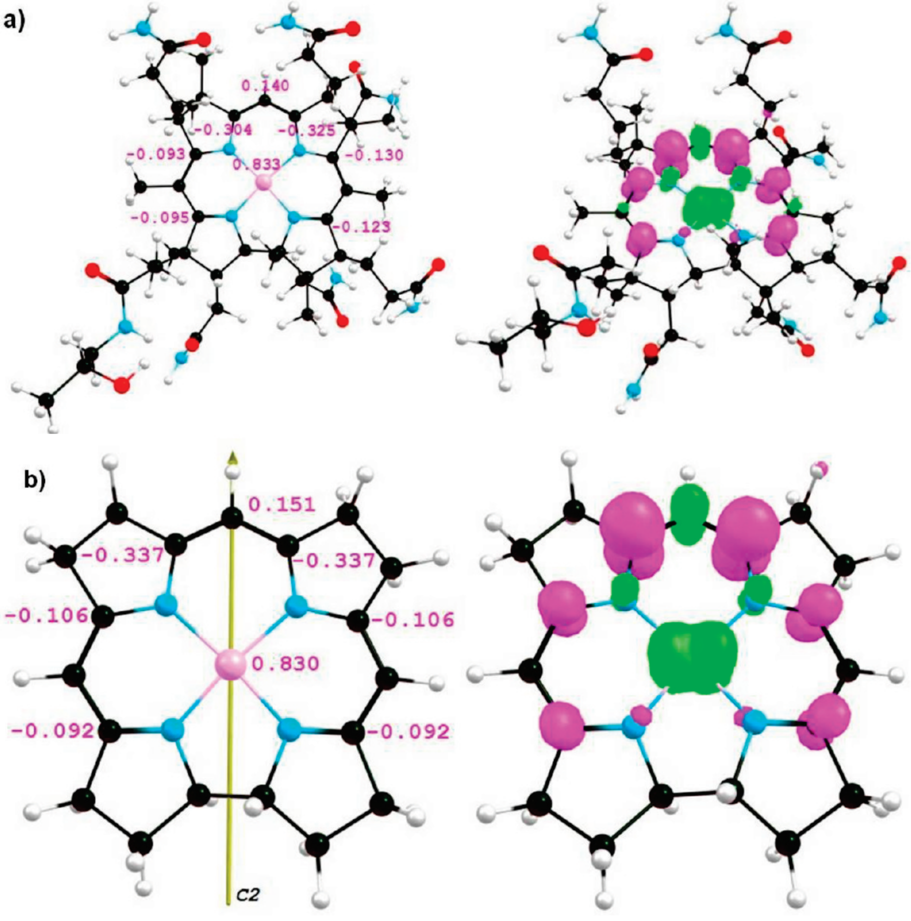


Figure 4. Spin density of (a) the full cob(I)alamin model and (b) the truncated cob(I)alamin model with C2 symmetry, computed at the B3LYP/6-31G(d) 5d level of theory. Left, spin populations; right, α and β spin density distributions colored as green and magenta, respectively.

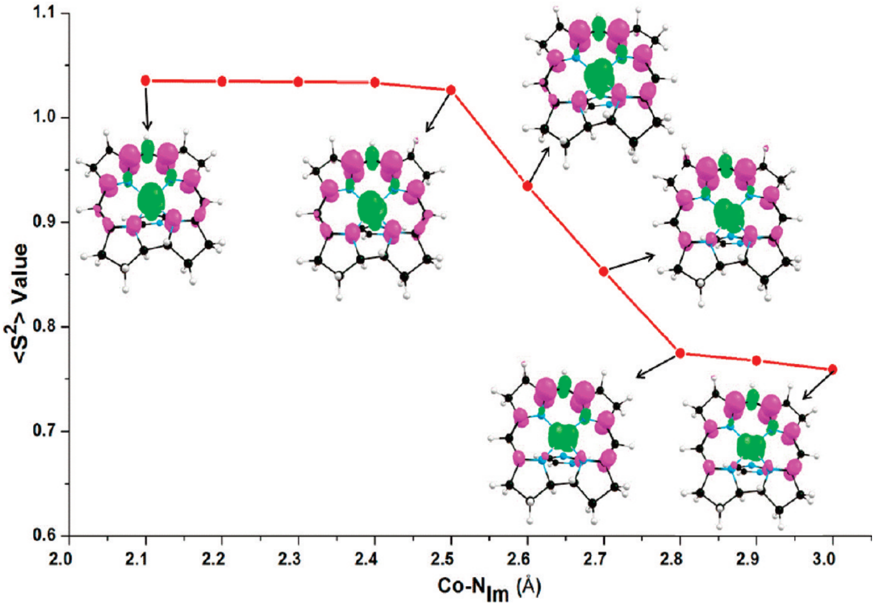


Figure 5. Evolution of the expectation value of the total spin, $\langle S^2 \rangle$, along with the spin density distributions during the elongation of the Co−N$_{Im}$ distance in the Im···[Cob(I)alamin] model system.

within the error of the B3LYP calculation, they indicate that the closed-shell and the diradical states are close in energy, thus suggesting that the electron transfer from the Co^I(d^8) to the corrin(π) to originate the Co^{II}(d^7)-corrin radical(π*)^1 may be feasible (see section 3.3.).

3.2.2. Influence of the Axial Base His. The electronic properties of cob(I)alamin were also evaluated in the presence of an imidazole (Im) as a model of the His axial base. Since the His759 moves between 2.27 and 2.78 Å with respect to the Co center inside the enzyme, the Co(I)−N$_{Im}$ distance was systematically varied between 2.1 and 2.8 Å in the gas phase DFT calculations. For each Co−N$_{Im}$ distance, the structure was optimized with both the BP86 and the B3LYP functionals. As for the His-off model (section 3.2.1), the structure does not differ significantly between the two functionals, but a spin polarized solution was only obtained in the case of the B3LYP functional

The spin density profiles (Figure 5 and Figure S1, Supporting Information) show unpaired spin density on the Co and the corrin ring, with opposite signs, consistent with a Co^{II}(d^7)-corrin radical (π*)^1 diradical state. However, it should be noted that the expectation value of the total spin $\langle S^2 \rangle$ is much larger (1.05−0.75) than the value expected for a singlet, $S(S + 1) = 0$, indicating that the open-shell singlet is significantly contaminated by the triplet state. Therefore, the single-determinantal DFT results need to be interpreted with caution. Nevertheless, we believe that the analysis of $\langle S^2 \rangle$ with respect to the Co−N$_{Im}$ distance (Figure 5) could help to assess the extent of the diradical character. The initial $\langle S^2 \rangle \sim$ 1.05 value remains constant until the distance reaches ~2.5 Å and decreases up to $\langle S^2 \rangle \sim 0.75$ at longer distances. This could indicate a decrease in the mixing between the open-shell singlet and triplet states due to an increase in the energy gap between them (see below).

Interestingly, the change in $\langle S^2 \rangle$ is accompanied by a change in the symmetry of spin density distribution on the cobalt atom. This switch in the d orbital of the cobalt atom bearing the unpaired electron can be further explained by analysis of the energies of relevant molecular orbitals near the HOMO/LUMO gap as a function of the Co−N$_{Im}$ distance (Figure 6). The antiferromagnetic coupling occurs between the upper occupied orbital of the corrin (π*$_{corr}$) and the singly occupied d orbital of the cobalt (d$_{Co}$). This d$_{Co}$ orbital is the d$_{z^2}$ orbital at short Co−N$_{Im}$ distances, but it is the d$_{yz}$ orbital at long distances, in agreement with the d$_{yz}$ → π* MLCT band observed in a previous study of the cob(I)alamin intermediate without the axial base.[31] Moreover, the energy gap between the π*$_{corr}$ and the d$_{Co}$ orbitals increases with the Co−N$_{Im}$ distance, further supporting our previous suggestion that the energy gap between the open-shell singlet and the triplet states increases with the Co−N$_{Im}$ distance.

In summary, the DFT calculations suggest that the cob(I)alamin intermediate has significant diradical character. They also indicate that the axial base modulates the nature of the coupling associated with Co^{II}(d^7)-corrin radical (π*)^1 configuration by changing the energy of the d orbitals of the cobalt atom. However, the multiconfigurational character of the cob(I)alamin can only be assessed correctly using a multireference method.

3.3. CASSCF/MC-XQDPT2 Analysis. The multiconfigurational nature of the wave function of cob(I)alamin was further investigated using multireference methods (CASSCF and CASSCF/MC-QDPT2). The selection of active orbitals for complex systems such as cob(I)alamin always represents a challenging problem. Using previously reported CASSCF calculations on corrin[31,53−55] or corrol-based[56] complexes as a guide, we have examined several kinds of active spaces and finally chosen to use a space of 10 active electrons distributed in 11 active orbitals.

Initially, CASSCF(10,11) calculations were carried out for the His-off form of cob(I)alamin (truncated model with C2 symmetry, see section 3.2.1.) in order to reproduce the weight of the diradical contribution reported by Jensen.[31] The active space, as shown in Figure S2 (Supporting Information), is very similar to that reported in ref 31. Second, we probed the effect exerted by the axial ligand on the electronic structure of the Im···[Co^I(corrin)] complex by performing a series of different CASSCF calculations with the same active space (Figures 7 and Supporting Information, S3−S9) but varying the Co−N$_{Im}$ distance

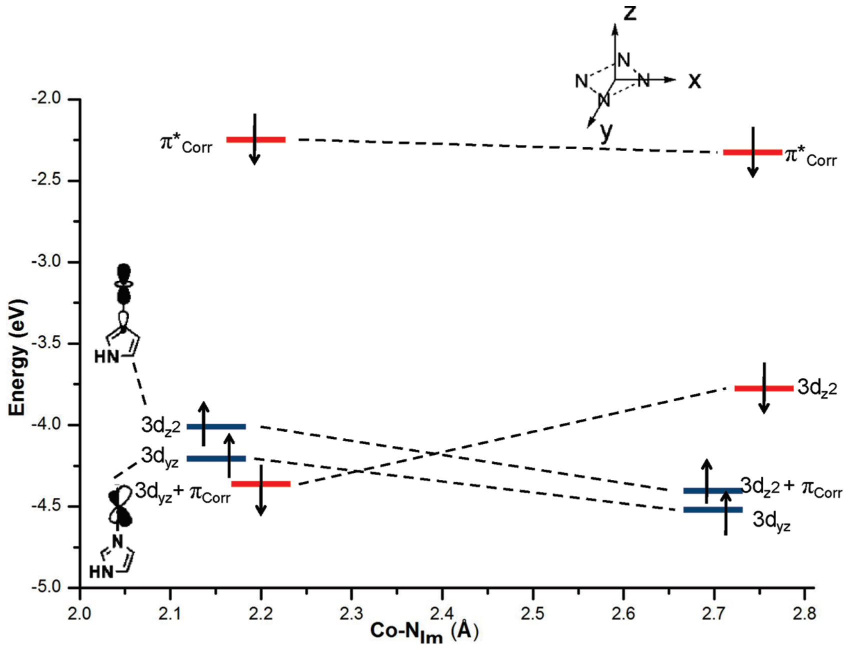


Figure 6. Variation of HOMO and HOMO−1 based orbital energies with the Co−N_{Im} distance calculated at the UB3LYP level of theory.

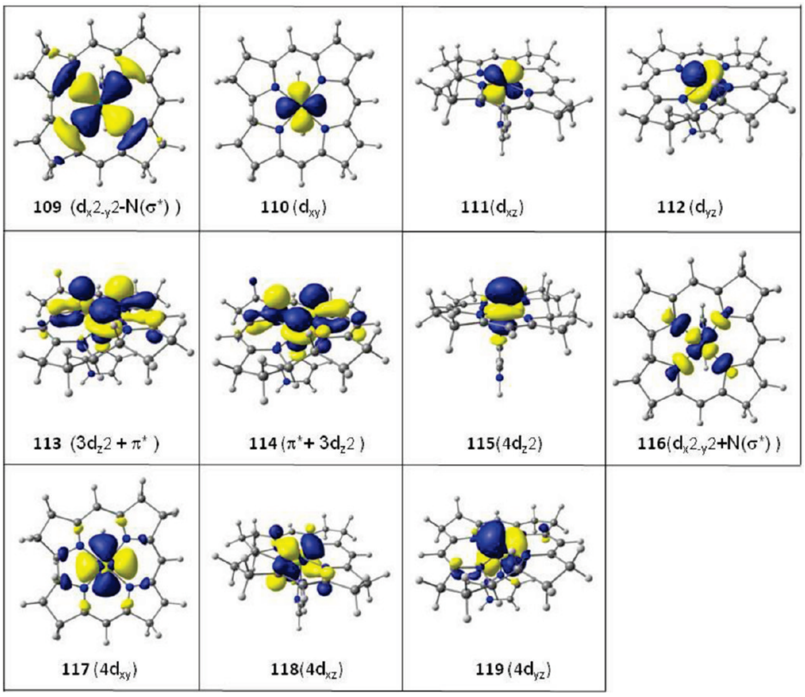


Figure 7. CASSCF active space orbitals used in the calculations of the Im···[Cob(I)alamin] model system at a Co−N_{Im} distance = 2.2 Å.

from 2.1 Å to 2.8 Å. In particular, the active space for Im···[Co^I(corrin)] is comprised of the $d_{x^2-y^2}$−N(σ*), d_{xy}, d_{xz}, d_{yz}, and $3d_{z^2}$ + π* orbitals, and the respective correlating orbitals are π* + $3d_{z^2}$, $4d_{z^2}$, $d_{x^2-y^2}$ + N(σ*), $4d_{xy}$, $4d_{xz}$, and $4d_{yz}$ (Figure 7). Note that this active space includes not only cobalt orbitals but also the corrin orbitals that may be important for the charge transfer between the corrin ring and the metal. The pair of correlating orbitals numbered 109 and 116 describe a σ* donation from the lone electron pairs of the equatorial nitrogen atoms to the $d_{x^2-y^2}$ orbital of cobalt. In addition, we have included the lowest unoccupied π* orbital of corrin, because it has been shown to be necessary to describe the electron configuration of the His-off form of cob(I)alamin.[31] This orbital mixes with the $3d_{z^2}$ cobalt orbital at a short Co−N_{Im} distance (CASSCF orbitals numbered 113 and 114, see Figure 7 and Figure S3−S5, Supporting Information), but with the $3d_{yz}$ orbital at long distances (CASSCF orbitals numbered 112 and 114, see Figures S6−S9, Supporting Information). Finally, an extra orbital has been added

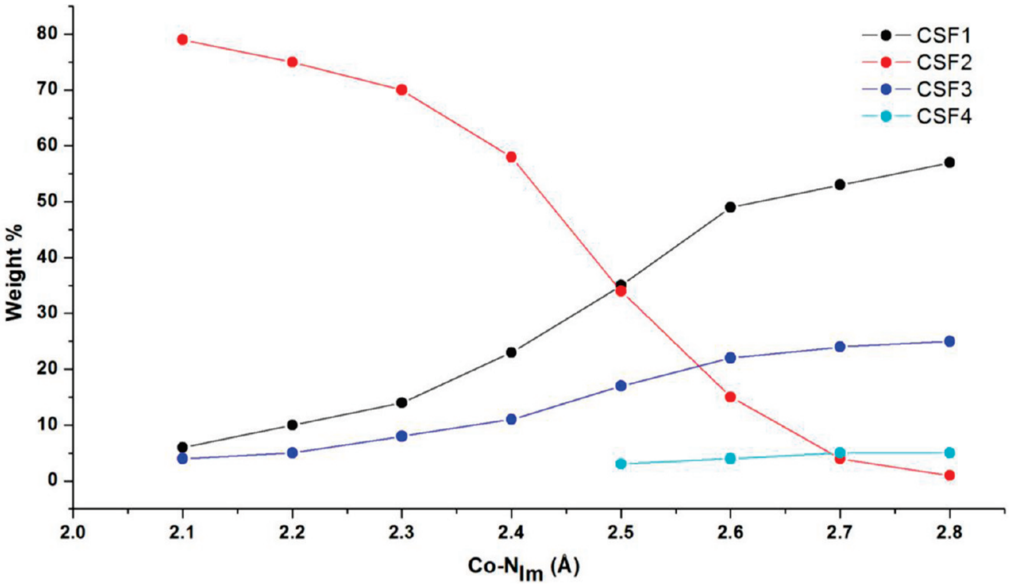


Figure 8. Weight (%) of the major configurations (CSFs) contributing to the ground state CASSCF wave function as a function of the Co−N$_{Im}$ distance for the Im···[Cob(I)alamin] model system. The description of the CSFs is given in Table 3.

Table 3. Composition of the CASSCF Wave Function of Cob(I)alamin for Each Co−N$_{Im}$ Distance, in Terms of the Mulliken Occupation Numbers of the Three Natural Orbitals Involved in the Open-Shell Singlet Description and the Weights of the Major Configurations State Functions (CSF1, CSF2, CSF3, and CSF4)

Co−N$_{Im}$ (Å)	orbital occupation[a]	CSF1 $(d_{z^2})^2(\pi^*)^0$ (wt %)	CSF2 $(d_{z^2})^1(\pi^*)^1$ (wt %)	CSF3 $(d_{yz})^1(\pi^*)^1$ (wt %)	CSF4 $(d_{yz})^2(\pi^*)^0$ (wt %)
2.1	...$(d_{yz})^{1.97}(d_{z^2})^{1.19}(\pi^*)^{0.81}$...	6	80	3	
2.2	...$(d_{yz})^{1.97}(d_{z^2})^{1.21}(\pi^*)^{0.79}$...	10	76	6	
2.3	...$(d_{yz})^{1.97}(d_{z^2})^{1.23}(\pi^*)^{0.76}$...	11	72	5	
2.4	...$(d_{yz})^{1.96}(d_{z^2})^{1.34}(\pi^*)^{0.66}$...	19	60	8	
2.5	...$(d_{yz})^{1.90}(d_{z^2})^{1.50}(\pi^*)^{0.57}$...	35	36	15	3
2.6	...$(d_{yz})^{1.50}(d_{z^2})^{1.94}(\pi^*)^{0.51}$...	51	13	21	4
2.7	...$(d_{yz})^{1.59}(d_{z^2})^{1.95}(\pi^*)^{0.42}$...	57	4	24	5
2.8	...$(d_{yz})^{1.59}(d_{z^2})^{1.95}(\pi^*)^{0.42}$...	59	1	24	5

[a] The Mulliken population analysis is shown for three natural orbitals involved in the open-shell singlet state. The rest of the active space orbitals are represented by dots.

to the correlating orbitals, in order to include the double shell effect[57] to all d orbitals.

The calculations reveal that the overall wave function of cob(I)alamin has multireference character, comprising five major configuration state functions (CSFs) that include single and double excitations (Figure S10, Supporting Information). By means of a unitary transformation to localized orbitals, this complex wave function can be simplified to four dominant configuration state functions (CSFs), as plotted in Figure 8. CSF1 is the closed-shell singlet, corresponding to the Co^{I}(d^8)-corrin $(\pi^*)^0$ configuration. The open-shell singlet, representing the Co^{II}(d^7)-corrin radical $(\pi^*)^1$ configuration, consists of two major configurations (CSF2 and CSF3), both corresponding to charge transfer states where one electron from a d orbital of Co (d_{z^2} or d_{yz}, respectively) has been shifted to the corrin π^* orbital. In addition, there is a small contribution of the double metal-to-ligand excitation, i.e., Co^{III}(d^6)-corrin anion$(\pi^*)^2$ (CSF4). Interestingly, the weight of these different CSFs varies with the Co−N$_{Im}$ distance. The weight of the closed-shell Co^{I}(d^8) configuration increases with the Co−N$_{Im}$ distance, as the overall diradical character decreases. However, the behavior of the two-diradical contributions is different. The weight of CSF2 decreases abruptly for Co−N$_{Im}$ distances > 2.5 Å, whereas the weight of CSF3 increases. The crossing between the two diradical configurations takes place at ∼2.6 Å and can be explored by analyzing the Mulliken populations (Table 3) and the pure fragment orbitals obtained after localization (Figure 9). When the Co−N$_{Im}$ distance is shorter than 2.5 Å, the electron is transferring from the d_{z^2} cobalt orbital to the corrin π^* orbital, whereas when it is longer than 2.5 Å the electron is shifted from d_{yz} to π^*. Indeed, the localized orbitals are consistent with the orbitals resulting from adding and subtracting the original HOMO (or HOMO−1 for longer Co−N$_{Im}$ distances) and LUMO orbitals for each Co−N$_{Im}$ distance:

$$\mathrm{Co}-\mathrm{N}_{\mathrm{Im}} < 2.5\ \text{Å}\quad \phi_{113}+\phi_{114} \rightarrow d_{z^2}\ \text{and}\ \phi_{113}-\phi_{114} \rightarrow \pi^*$$

$$\mathrm{Co}-\mathrm{N}_{\mathrm{Im}} > 2.5\ \text{Å}\quad \phi_{112}+\phi_{114} \rightarrow d_{yz}\ \text{and}\ \phi_{112}-\phi_{114} \rightarrow \pi^*$$

This switch in the cobalt d orbitals with the change in the axial bond length is consistent with the spin-polarized results obtained from DFT(UB3LYP) calculations. The availability of different cobalt d orbitals in the active space chosen for the calculations allows the system to change the d_{Co} orbital that participates in the diradical state during the elongation of the Co−N_{Im} distance. It should also be noted that the overlap between the π^* corrin orbital and the singly occupied cobalt d orbital is larger for d_{yz} than for d_{z^2}, resulting in a higher weight for the CSF3 configuration than for CSF2, although the sum of both contributions reduces significantly with increasing Co−N_{Im} distance. This is in line with the increasing closed-shell character as well as the appearance of the small contribution from the double excitation (CSF4).

In addition, state average CASSCF/MC-XQDPT2 calculations also reveal that each excited state has a significant diradical character, showing metal-to-ligand charge transfer (Table S1, Supporting Information). Thus, at Co−N_{Im} = 2.1 Å, the correlated wave function of the ground state is mainly composed of the open-shell singlet, $Co^{II}(d^7)$-corrin radical $(\pi^*)^1$ configuration. This diradical weight decreases as the axial bond is elongated, such that at Co−N_{Im} = 2.8 Å the diradical weight is only 24%, similar to the result obtained by Jensen[31] for the cob(I)alamin model without an axial base. In other words, the dominant contribution to the ground state wave function is the diradical configuration when the His is bound to the cobalt atom, but the closed-shell singlet, $Co^{I}(d^8)$ configuration when the axial His is weakly coordinated.

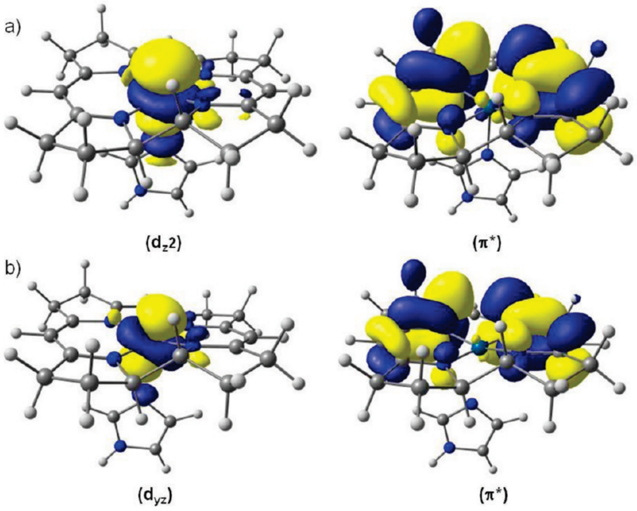


Figure 9. Pure fragment localized orbitals, describing the electron transfer between the cobalt and the corrin. (a) Co−N_{Im} distance < 2.5 Å; the electron is transferred from Co(d_{z^2}) to corrin(π^*). (b) Co−N_{Im} distance > 2.5 Å; the electron is shifted from Co(d_{yz}) to corrin(π^*).

3.4. Energy Changes as Function of Co−N_{Im} Distance. Finally, we have calculated the energy of the Im···[Co^I(corrin)] complex as a function of the distance between Co and N_{Im} atoms in order to evaluate the energy cost of displacing the axial ligand (Figure 10). We have employed single-determinant (DFT with either the BP86 or the B3LYP functional) and multireference (CASSCF and CASCF/MC-XQDPT2) methods and different basis sets (6-31G(d) and 6-311G(d,p)). Regardless of the computational method used, the curve is clearly repulsive. However, it should be noted that the (CASSCF/MC-XQDPT2) energy at a Co−N_{Im} distance ~ 2.3 Å (i.e., in the Me-Cob(III)alamin resting state of the MetH enzyme) is only 4 kcal/mol higher than at ~2.8 Å (i.e., the calculated value for the cob(I)alamin intermediate, see section 3.1). Therefore, although the change of distance causes noticeable changes in terms of electronic properties of the Co center (from Co^{II}-corrin radical to Co^{I}), it is energetically not costly.

3.5. Implications for the Remethylation of the Cob(I)alamin Intermediate in MetH. The methyl transfer reaction from Me-H_4folate to cob(I)alamin to generate the Me-cob(III)alamin resting state (Scheme 1) is generally assumed to occur through a S_N2-type displacement. However, in view of the present work, it can be suggested that (i) a radical mechanism, consisting of an electron transfer (ET) followed by a methyl radical transfer, is also possible and that (ii) the axial base would have a decisive role regarding the enhancement of the ET-based mechanism by modulating the cob(I)alamin electronic properties.

At shorter Co−N_{Im} distances (<2.5 Å), the dominant electronic configuration is diradical, with a Co(II) and an electron located on the corrin ring. This would favor the radical mechanism, in which an electron is initially transferred from the corrin to

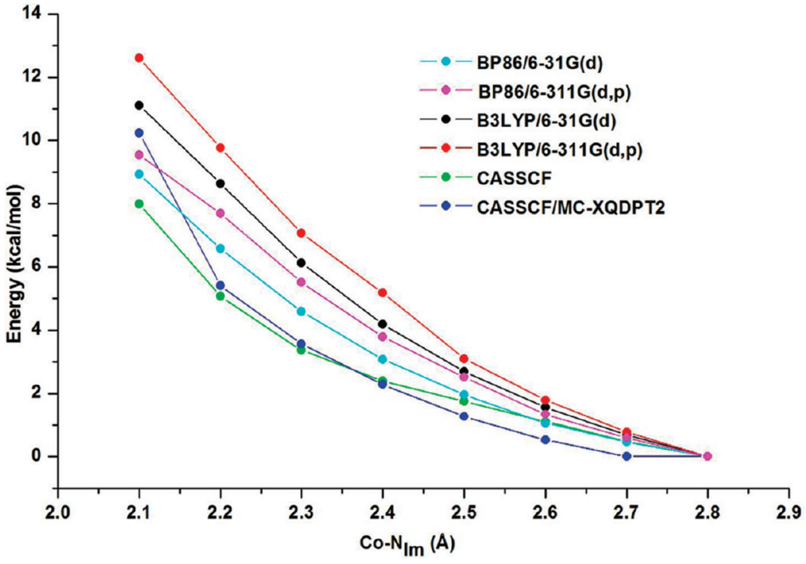


Figure 10. Change in the ground state energy with the Co−N_{Im} distance, computed with different DFT functionals and *ab initio* CASSCF/MC-XQDPT2 methods.

the Me-H_4folate in order to generate a Co(II) and [Me-H_4folate]$^{\bullet-}$ state. Subsequently, a methyl radical would be transferred from [Me-H_4folate]$^{\bullet-}$ to Co(II), forming Me-Cob(III)lamin. This methyl radical transfer would be in line with the reductive cleavage mechanism proposed for the methyl transfer from MeCbl to the Hcy substrate (Scheme 1), where the electron is transferred from the Hcy to the MeCbl, followed by a methyl radical transfer from the one electron reduced form of MeCbl to the Hcy substrate.[24,25] In other words, the reaction mechanism of MetH would not only involve the Co metal but also the corrin ligand, which can activate both substrates (Me-H_4folate and Hcy) of the MetH enzyme by ET. On the other hand, at longer Co$-$N$_{Im}$ distances (>2.5 Å), the dominant configuration is closed-shell, with an electron pair located on the metal center consistent with a Co(I) oxidation state. This would favor an S$_N$2-type mechanism in which the Co(I) nucleophile directly abstracts the methyl group of Me-H_4folate.

4. SUMMARY AND CONCLUSION

In the present theoretical study, the electronic and structural properties of the cob(I)alamin intermediate have been analyzed using QM(DFT)/MM, gas phase DFT, and CASSCF/QDPT2 calculations. Because previous studies[29−32] did not take into account the influence of the protein environment, we initially performed QM/MM calculations to study the formation of the Co(I) state inside the MeCbl domain of the methionine synthase (MetH) enzyme. The observed displacement of the axial His759 from the cobalt center is in agreement with the model proposed by Wirt et al.[28] based on the tetracoordinated state of free cob(I)alamin. However, His759 is weakly coordinated inside the enzyme (QM/MM optimized Co$-$N(His) distance = 2.78 Å), where the presence of the catalytic triad (His759$-$Asp757$-$Ser810) and the hydrogen bonds with other residues reduce its conformational freedom. This implies that the remethylation of the cob(I)alamin cofactor by Me-H_4folate (Scheme 1) occurs in the presence of the axial ligand, and thus His759 may have an influence in this methyl transfer reaction.

Consequently, we have analyzed the influence of the axial ligand in the electronic structure of cob(I)alamin cofactor using gas phase DFT followed by CASSCF/QDPT2 calculations. Irrespective of the theoretical method used, our results show that the ground state of cob(I)alamin is multiconfigurational, in agreement with a previous study of the cob(I)alamin cofactor without the axial ligand.[31] In addition to the closed-shell Co(I), a diradical Co(II)-corrin radical configuration (formed by electron transfer from the cobalt to the corrin ring) contributes to the electronic structure of the cob(I)alamin intermediate, revealing the noninnocent behavior of the corrin ring.[33] The weight of these two configurations depends on the distance of the axial base His from the Co center. The main contribution to the ground state wave function at short Co$-$N$_{Im}$ distances is the diradical configuration, whereas at long distances, it is the closed-shell. Therefore, our results suggest that (i) the standard description of the Co(I) nucleophile is not appropriate for cob(I)alamin, due to the noninnocent character of the corrin ring, and (ii) the distance between cob(I)alamin and the axial His plays an important role in modulating the nucleophilicity of Co(I).

In view of our results, we have proposed that the remethylation reaction in MetH (Scheme 1) could involve not only the metal (i.e., the closed-shell Co(I) configuration) but also the corrin ring (i.e., the diradical Co(II)-corrin radical configuration). In other words, in addition to the traditionally assumed S$_N$2 mechanism, our findings suggest the possibility of an alternative radical mechanism, in which an electron is transferred from the corrin to the Me-H_4folate in order to generate Co(II) and [Me-H_4folate]$^{\bullet-}$. It should be noted that this ET does not require the presence of any strong reducing agent near the $CH_3$$-$$H_4$folate substrate but rather the cofactor-induced formation of anion-radical-like species within the cob(I)alamin:[Me-H_4folate] reactant complex. Indeed, earlier studies by Marcus,[58] Shaik et al.,[59] and Zipse[60] already discussed that such ET bond-breakage mechanisms significantly enhance the reaction rates in comparison to S$_N$2 mechanisms. The energetics and dichotomy of the S$_N$2 and radical mechanisms in the methyl transfer reaction from Me-H_4folate to cob(I)alamin are currently being investigated in our group.

ASSOCIATED CONTENT

S Supporting Information. Spin density distributions along with the spin populations of the Im···[Cob(I)alamin] gas phase models, CASSCF active space orbitals of cob(I)alamin, CASSCF active space orbitals for the Im···[Cob(I)alamin] model at Co$-$N$_{Im}$ distances = 2.1$-$2.8 Å, major configurations contributing to the ground state CASSCF wave function before localization, relative energy of ground state and low lying excited states for each Co$-$N$_{Im}$ distance, and Cartesian coordinates of the QM region of the QM/MM optimized structures of the cofactor binding domain of methionine synthase (MetH) in the Me-Cob(III)alamin and Cob(I)alamin intermediate. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Phone: (502) 852-6609. Fax: (502) 852-8149. E-mail: pawel@louisville.edu.

ACKNOWLEDGMENT

N.K. and P.M.K. acknowledge the computer support and assistance provided by Cardinal Research Cluster Supercomputer Center at the University of Louisville. This work was also supported by the Ministry of Science and Higher Education (Poland) under grant no. N204 028336.

C.R. thanks the Spanish Ministry of Science and Innovation (MICINN, Grant FIS2008-03845) and the Generalitat de Catalunya (GENCAT, Grant 2009SGR-1309) for its financial assistance. M.A.-P. acknowledges an FI fellowship from GENCAT. C.R. and M.A.-P. also acknowledge the computer support, technical expertise, and assistance provided by the Barcelona Supercomputing Center—Centro Nacional de Supercomputación (BSC$-$CNS).

REFERENCES

(1) Dolphin, D. *B12*; Wiley-Interscience: New York, 1982.

(2) Banerjee, R. *Chem. Biol.* **1997**, *4*, 175–186.

(3) Ludwig, M. L.; Matthews, R. G. *Annu. Rev. Biochem.* **1997**, *66*, 269–313.

(4) *Vitamin B_{12} and B_{12} Proteins*; Kräutler, B., Arigoni, B., Golding, B. T., Eds.; Wiley-VCH: New York, 1998 (Lectures Presented at the 4th European Symposium on Vitamin B12 and B12 Proteins).

(5) Marzilli, L. G. In *Bioinorganic Catalysis*; Reedijk, J., Bouwman, E., Eds.; Marcel Dekker: New York, 1999; pp 423−468.

(6) Banerjee, R. *Chemistry and Biochemistry of B12*; Wiley: New York, 1999.
(7) Matthews, R. G. *Acc. Chem. Res.* **2001**, *34*, 681–689.
(8) Banerjee, R.; Ragsdale, S. W. *Annu. Rev. Biochem.* **2003**, *72*, 209–247.
(9) Brown, K. L. *Chem. Rev.* **2005**, *105*, 2075–2149.
(10) Randaccio, L.; Geremia, S.; Nardin, G.; Wuerges, J. *Coord. Chem. Rev.* **2006**, *250*, 1332–1350.
(11) Matthews, R. G.; Koutmos, M.; Datta, S. *Curr. Opin. Struct. Biol.* **2008**, *18*, 658–666.
(12) Randaccio, L.; Geremia, S.; Demitri, N.; Wuerges, J. *Trends Inorg. Chem.* **2009**, *11*, 1–19.
(13) Matthews, R. G. In *Metal Ions in Life Sciences*; Sigel, A., Sigel, H., Sigel, R. K. O., Eds.; Royal Society of Chemistry: Cambridge, U. K., 2009; Vol. 6, pp 53−114.
(14) Randaccio, L.; Geremia, S.; Demitri, N.; Wuerges, J. *Molecules* **2010**, *15*, 3228–3259.
(15) Drennan, C. L.; Huang, S.; Drummond, J. T.; Matthews, R. G.; Ludwig, M. L. *Science* **1994**, *266*, 1669–1674.
(16) Goulding, C. W.; Matthews, R. G. *Biochemistry* **1997**, *36*, 15749–15757.
(17) Pearisco, K.; Goulding, C. W.; Huang, S.; Matthews, R. G.; Panner-Hahn, J. E. *J. Am. Chem. Soc.* **1998**, *120*, 8410–8416.
(18) Evans, J. C.; Huddler, D. P.; Hilgers, M. T.; Romanchuk, G.; Matthews, R. G.; Ludwig, M. L. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 3729–3736.
(19) Koutmos, M.; Pejchal, R.; Bomer, T. M.; Matthews, R. G.; Smith, J. L.; Ludwig, M. L. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, -3286–3291.
(20) Goulding, C. W.; Postigo, D.; Matthews, R. G. *Biochemistry* **1997**, *36*, 8082–8091.
(21) Dixon, M. M.; Huang, S.; Matthews, R. G.; Ludwig, M. *Structure* **1996**, *4*, 1263–1275.
(22) Datta, S.; Koutmos, M.; Pattridge, K. A.; Ludwig, M. L.; Matthews, R. G. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 4115–4120.
(23) Bandarian, V.; Pattridge, K. A.; Lennon, B. W.; Huddler, D. P.; Matthews, R. G.; Ludwig, M. L. *Nat. Struct. Biol.* **2002**, *9*, 53–56.
(24) Kozlowski, P. M.; Kuta, J.; Galezowski, W. *J. Phys. Chem. B* **2007**, *111*, 7638–7645.
(25) Alfonso-Prieto, M.; Biarnés, X.; Kumar, M.; Rovira, C.; Kozlowski, P. M. J. *Phys. Chem. B* **2010**, *114*, 12965–12971.
(26) Lexa, D.; Saveant, J. M. *Acc. Chem. Res.* **1983**, *16*, 235–243.
(27) Banerjee, R. V.; Frasca, V.; Ballou, D. P.; Matthews, R. G. *Biochemistry* **1990**, *29*, 11101.
(28) Wirt, M. D.; Sagi, I.; Chance, M. R. *Biophys. J.* **1992**, *63*, 412–417.
(29) Jaworska, M.; Lodowski, P. *THEOCHEM* **2003**, *631*, 209–223.
(30) Jensen, K. P.; Ryde, U. *ChemBioChem* **2003**, *4*, 413–424.
(31) Jensen, K. P. *J. Phys. Chem. B* **2005**, *109*, 10505–10512.
(32) Liptak, M. D.; Brunold, T. C. *J. Am. Chem. Soc.* **2006**, *128*, 9144–9156.
(33) Jensen, K. P.; Ryde, U. *Coord. Chem. Rev.* **2009**, *253*, 769–778.
(34) Laio, A.; VandeVondele, J.; Rothlisberger, U. *J. Chem. Phys.* **2002**, *116*, 6941–6947.
(35) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471–2474.
(36) Becke, A. D. *J. Chem. Phys.* **1986**, *84*, 4524–4529.
(37) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822–8824.
(38) Troullier, N.; Martins, J. L. *Phys. Rev. B* **1991**, *43*, 1993–2006.
(39) Louie, S. G.; Froyen, S.; Cohen, M. L. *Phys. Rev. B* **1982**, *26*, 1738–1742.
(40) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision C.02; Gaussian, Inc.: Wallingford, CT, 2004.
(41) Rovira, C.; Kozlowski, P. M. *J. Phys. Chem. B* **2007**, *111*, 3251–3257.
(42) Kuta, J.; Patchkovskii, S.; Zgierski, M. Z.; Kozlowski, P. M. *J. Comput. Chem.* **2006**, *27*, 1429–1437.
(43) Kozlowski, P. M.; Kamachi, T.; Toraya, T.; Yoshizawa, T. *Angew. Chem., Int. Ed.* **2007**, *46*, 980–983.
(44) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
(45) Jensen, K.; Ryde, U. *J. Phys. Chem A* **2003**, *107*, 7539–7545.
(46) Ruiz, E.; Cirera, J.; Alvarez, S. *Coord. Chem. Rev.* **2005**, *249*, 2649–2660.
(47) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
(48) Nakano, H. *J. Chem. Phys.* **1993**, *99*, 7983–7992.
(49) Granovsky, A. A. PC GAMESS/Firefly version 7.1.G. www. http://classic.chem.msu.su/gran/games/index.html.
(50) Hagemeier, C. H.; Kruer, M.; Thauer, R. K.; Warkentin, E; Ermler, U. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 18917–18922.
(51) Goel, S.; Masunov, A. E. J. *Chem. Phys.* **2008**, *129*, 214302–214316.
(52) Jensen, K. P.; Roos, B. O.; Ryde, U. J. *Chem. Phys.* **2007**, *126*, 014103–014117.
(53) Kozlowski, P. M.; Kamachi, T.; Kumar, M.; Nakayama, T.; Yoshizawa, K. *J. Phys. Chem. B* **2010**, *114*, 5928–5939.
(54) Kornobis, K.; Kumar, N.; Wong, B. M.; Jaworska, M.; Lodowski P.; Andruniow, T.; Ruud, K.; Kozlowski, P. M. *J. Phys. Chem. A* **2011**, *115*, 1280–1292.
(55) Kumar, N.; Jaworska, M.; Lodowski, P.; Kumar, M.; Kozlowski, P. M. *J. Phys. Chem. B* **2011**, doi:10.1021/jp200945a.
(56) Roos, B. O.; Veryazov, V.; Conradie, J.; Taylor, P. R.; Ghosh, A. *J. Phys. Chem. B* **2008**, *112*, 14099–14102.
(57) Roos, B. O.; Lindh, R.; Malmqvist, P.-Å.; Veryazov, V.; Widmark, P.-O. *J. Phys. Chem.* **2005**, *109*, 6575–6579.
(58) Marcus, R. A. *J. Phys. Chem. A* **1997**, *101*, 4072–4087.
(59) Shaik, S. S; Schlegel, H. B.; Wolfe, S. *Theoretical Aspects of Physical Organic Chemistry. The SN2Mechanism*; Wiley-Interscience: New York, 1992.
(60) Zipse, H. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 1697–1700.
(61) Rovira, C.; Biarnés, X. *Inorg. Chem.* **2004**, *43*, 6628–6632.

JCTC Journal of Chemical Theory and Computation

ARTICLE

pubs.acs.org/JCTC

Secondary Structure Assignment of Amyloid-β Peptide Using Chemical Shifts

Geoffrey P. F. Wood* and Ursula Rothlisberger*

Laboratory of Computational Chemistry and Biochemistry, BCH 4107 EPF Lausanne, CH-1015 Lausanne, Switzerland

S Supporting Information

ABSTRACT: The distinct conformational dependence of chemical shifts caused by α-helices and β-sheets renders NMR chemical shift analysis a powerful tool for the structural determination of proteins. However, the time scale of NMR experiments can make a secondary structure assignment of highly flexible peptides or proteins, which may be converting between conformational substates, problematic. For instance the amyloid-β monomer, according to NMR chemical shifts, adopts a predominately random coil structure in aqueous solution (with <3% α-helical content). Molecular dynamics simulations, on the other hand, suggest that α-helical content can be significant (10−25%). In this paper, we explore the possible reasons for this discrepancy and show that the different results from experiments and theory are not necessarily mutually exclusive but may reflect a general problem of secondary structure assignment of conformationally flexible biomolecules.

1. INTRODUCTION

One of the hallmarks of Alzheimer's disease is the deposition of fibrils containing the amyloid-β peptide (Aβ) in the extracellular space of the limbic and association cortices (for recent reviews see refs 1 and 2). Aβ is a 39 to 43 amino acid protein derived from the normal metabolism of the transmembrane amyloid precursor protein (APP).[3−5] The 43 amino acid fragment is characterized by the sequence: DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIAT. The remaining alloforms are derived by deleting residues from the C-terminal end.

The most abundant form is the 40 amino acid alloform denoted Aβ(1−40).[6] However, the most toxic form has been identified as the 42 amino acid peptide Aβ(1−42).[6] There is growing evidence that the toxic agents in Alzheimer's disease are small soluble oligomeric structures of Aβ; in this model, the fibril is a symptom of the disease rather than a causative agent itself.[7−9] In order to understand the pathology of Alzheimer's disease and to be in a position to effectively develop drugs, a detailed atomistic knowledge of the conformational transitions connecting the native form (transmembrane and APP-incorporated) to the fibril is desirable.

Under normal physiological conditions, Aβ aggregates quickly, rendering standard experimental tools of biochemistry ineffective for characterizing the intermediate species. As a consequence, the bulk of our structural knowledge refers to the two end-points of the conformational transition. Although no direct structure of APP-incorporated Aβ is available, a number of experimental model studies,[10−14] which simulate transmembrane conditions, have shown that Aβ exists primarily as two α-helices connected through a kink spanning residues Gly25−Asn27. A longer helix contains residues 10−25, and a shorter one contains residues 27−35 (see Figure 1). The N-terminal end (residues 1−9) adopts variable conformations depending on the exact experimental conditions (i.e., pH, temperature, and stabilizing agents) but is described mainly as "unstructured".

The fibril containing Aβ(1−40) monomers has been characterized by solid-sate NMR[15] and is present as a pair of antiparallel β-sheets spanning residues 10 to 40 with a turn located between Glu22 and Lys28. Again, the N-terminal end (residues 1−9) is unstructured (not shown in Figure 1). In addition, an important turn-stabilizing salt bridge between Asp23 and Lys28 has also been identified.

The Aβ monomer in aqueous solution at a pH of ∼7 can be considered as the first stage of the transition from membrane bound Aβ to the fibril, making a structural analysis of this situation highly desirable. However, the rapid in vitro aggregation of Aβ makes a direct experimental evaluation of the monomeric structure difficult to obtain. Despite this difficulty, a number of CD spectra and solution NMR studies of the full-length species, i.e., Aβ(1−40) and Aβ(1−42), have been carried out.[16−19] In addition, the solution structure of the fragment containing residues 10−35 has been characterized by NMR.[20,21] This fragment has a higher solubility in aqueous solution and has been suggested as a good structural model for the full-length peptides. The consensus of these studies is that the Aβ monomer adopts predominately a random coil structure in solution with little or no well-defined secondary structure motifs. One particular spectrum based on CD suggests, for Aβ(1−42), 79% random coil, 13% β-turn/β-strand, and 3% α-helix.[18] An NMR study of Aβ(1−42),[16] which analyzed the secondary structure using chemical shifts, indicated that the α-helical content is 0% (derived from Cα chemical shifts) and the β-strand content is 20% (derived from Hα chemical shifts), in good agreement with the CD spectrum. This study employed the use of the chemical shift index (CSI) method,[22] which is an important tool for identifying secondary structure elements of large biomolecules. The method is based on the observation made in the early 1980s[23] that Cα protons experience a relative upfield shift when the residue in question is incorporated into an α-helix and a relative downfield shift when incorporated into a β-sheet.

Received: March 4, 2011
Published: April 22, 2011

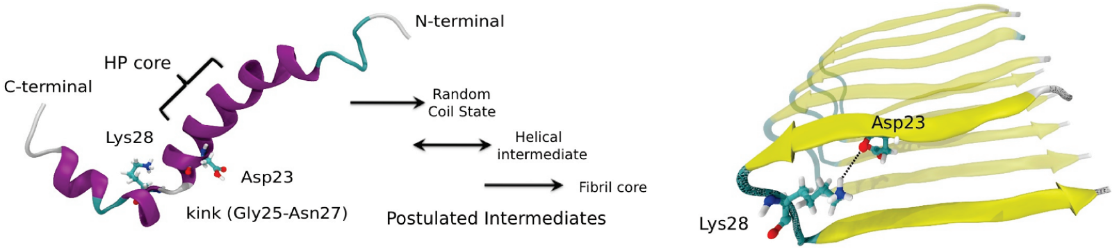


Figure 1. Key features of the conformations associated with transmembrane amyloid-β and when incorporated in a fibril.

Since then, similar chemical shift-to-structure correlations have been reported for other nuclear centers, including Cαs, Cβs, and C's, (for reviews, see refs 24 and 25); however, the correlation may be reversed. The identification process requires the determination of a set of "random coil" (rc) values that lie in the middle of the α-helix and β-sheet extremes. The reference values give the conformational dependence on the chemical shift as

$$\Delta\delta_s^i = \delta_{obs}^i - \delta_{rc}^i \qquad (1)$$

Secondary structure is then assigned through a set of well-defined rules.[25]

Molecular dynamics (MD) simulations do not suffer from the same difficulties as experimental studies in that the solution structure of the monomer can be probed directly. However, caution needs to be exercised with MD simulations of Aβ because it has been shown that kinetic trapping may occur;[26] therefore, long time scales and/or enhanced sampling is required in order to effectively sample the conformational space. From the plethora of MD studies in the literature, only a handful satisfy either or both of these criteria.[26−30] The general consensus of these studies is that Aβ has a predominately random coil and turn structure, in agreement with experimental results, but the α-helical content differs significantly from the experimental findings. Depending on the MD protocol used, the total helical content ($\alpha + 3_{10} + \pi$) is at least 10% (using the AMBER99SB[31] force field) and can be as high as 25% (using the CHARMM19[32] force field). However, a number of force fields have been shown to overstabilize helical content by up to a factor of 2.[31] Taking the possible overstabilization into account, a helical content of no less than 10−12% is predicted by theory, which is *at least* 4 times greater than the experimental findings.

In the present work, we investigate this discrepancy by carrying out enhanced sampling molecular dynamics simulations and evaluating the secondary structure content of Aβ(1−42) via the conformational dependencies of chemical shifts. Our results suggest that because both downfield and upfield shifts may occur, conformationally flexible biomolecules may on average give NMR signals that can be interpreted as random coil or turn elements. However, in certain extreme cases, the average chemical shift of a particular nuclear center may be sampling from both helical and β structures with little or no "random coil" structure present, even though the average NMR signal suggests that a random coil structure dominates. In other less extreme cases, the average chemical shift is sampling from any of a number different motifs during the course of the simulation, but one single averaged chemical shift does not capture this complexity. We then show how the helical content depends on its definition, and for a flexible biomolecule, a single "static" content is not the appropriate measure for helicity.

The idea that a single dominant configuration may lead to inconsistencies in the interpretation of spectroscopic data for flexible biomolecules is not new. In particular, van Gunsteren and co-workers have shown how the correct interpretation of NOE data and *J*-coupling constants requires the incorporation of additional conformations along with the dominant structure.[33,34] Further examples where a conformational distribution is important for the correct interpretation of spectroscopic quantities have been reviewed elsewhere.[35]

2. COMPUTATIONAL DETAILS

Replica exchange molecular dynamics (REMD)[36] has emerged as a method that explores a significantly larger portion of phase space than what a single-temperature simulation of the same (aggregate) length can achieve.[37] However, a major drawback of conventional (temperature) replica exchange molecular dynamics is that the number of replicas needed to span a given temperature range increases as the square root of the number of degrees of freedom in the system. Numerous techniques have been developed in order to deal with this problem, which include optimizing the allocation of replicas,[38] using a perturbation on the Hamiltonian (rather than temperature) when defining each replica (H-REMD),[39−41] and coupling replicas to reservoirs of pregenerated ensembles.[42,43]

Another novel approach is to retain the conventional (temperature) REMD protocol but use a modified Hamiltonian when attempting exchanges. In the formulation by Simmerling and coworkers,[44,45] the exchange attempts are made with a Hamiltonian that reduces the total number of degrees of freedom of the system. Instead of using the entire periodic box for exchange attempts, a predetermined subset of explicit water molecules is retained around the solute of interest; this subsystem is then immersed in a dielectric continuum. We emphasize that only the exchange attempts use this hybrid Hamiltonian; the dynamics are propagated with a full periodic box of explicit waters. With a reduced number of degrees of freedom for the exchange attempts, the "hybrid-REMD" approach uses fewer replicas for a given temperature range. Promising results have been published showing that this scheme can reproduce properties derived from full REMD simulations.[44]

There is no recipe for choosing the number of explicit water molecules when swaps are carried out. It has been recommended that the first solvation shell be used because in some instances increasing the number of waters to include the second solvation shell gave less satisfactory results. This indicates that convergence with respect to the number of explicit water molecules may not be a monotonically decreasing function. In light of this, we chose to saturate the system with as many explicit water molecules as we could computationally afford and thus settled on 5000. This number may be compared with approximately 350 and 600 water molecules in the first and second solvation shells, respectively, of Aβ(1−42). To determine the temperature spacing between replicas, we chose initial temperatures using an online predicting algorithm.[46,47] After propagating the dynamics for approximately 2 ns, energetic data from the exchange attempts using the hybrid Hamiltonian were plotted against the temperature. An exponential curve was then fitted to the data to give an expression for the temperature as a function of the energy. The resultant curves were used to solve iteratively the Monte Carlo swap condition for a given acceptance ratio (p), i.e.,

$$p = \exp[(E_2 - E_1)(\beta_2 - \beta_1)] \quad (2)$$

and thus obtain an approximation to the ideal temperature spacing.

The final numbers of replicas, theoretical swapping probability, and temperature distributions were set to the following: 16 replicas, 8.5% probability, with temperatures set to 282.0, 291.0, 300.0, 310.0, 320.0, 331.0, 342.0, 354.0, 366.0, 380.0, 393.0, 408.0, 424.0, 440.0, 457.0, and 476.0 K. In addition to these parameters, the simulations were carried out using the AMBER 9 software package[48] using the ff03 force field,[49] which has been shown to give good population distributions for secondary structure elements in biomolecular simulations compared to the other Amber force fields but may overpopulate helical structures.[50] The replica exchange simulations were set up by first surrounding the Aβ(1−42) monomer (coordinates were obtained from the PDB structure ID 1Z0Q) in a box of 16 777 TIP3P water molecules. The protonation states of titratable residues were adjusted according to a pH of approximately 7.0 achieved by using empirical structural rules to determine the pK_a of each residue.[51] This resulted in standard protonation states for all residues, giving the protein an overall charge of −3.0. His6, His13, and His14 were found to have pK_a's of 6.4, 7.0, and 6.4, respectively, and are all surface exposed. It is likely that the protonation states of these residues are in rapid equilibrium; therefore, we arbitrarily chose to protonate the ring in the δ position. To neutralize the −3.0 charge, we chose to use a background jellium rather than explicit counterions to avoid spurious coordination effects. This decision was made because the structure of Aβ seems to be acutely sensitive to external factors, and we wanted to draw conclusions in the absence of any other extraneous elements that may bias the statistics. Following energy minimization to remove close contacts, NPT simulations were run for a single system using a 2 fs time step and Langevin dynamics with a collision frequency of 1.0 ps^{-1} to couple to the constant target temperature of 300 K and a Berendsen barostat to control pressure using a coupling constant of 2.0 ps to the target pressure of 1 bar. A 10.0 Å cutoff for nonbonded interactions was used in combination with the particle mesh Ewald procedure for long-range electrostatics, while H−X bond lengths were constrained using the SHAKE algorithm.[52] This single simulation was run until both the pressure and temperature stabilized (∼100 ps) and was then used as the starting point for each replica of the REMD simulations, which were carried out in the NVT ensemble. Temperatures for each replica were defined in the manner described above, and swaps were attempted every 2 ps. After a further equilibration of approximately 10 ns per replica, structural data were accumulated for 103 ns per replica, with structural data sampled every 2 ps (resulting in ∼45 000 configurations per replica). This gave an aggregate simulation time of 1.65 μs. The theoretical swapping probability, based on the temperature spacing, is determined to be 8.5%; the actual swapping probability was found to be 6.3%. Although this is a low probability, the frequency of swap trials is high, and also the length of the simulation is long. This means that the number of *successful* swaps for our simulation was relatively high (9857), indicating that the REMD simulation is effective. In order to illustrate the efficiency of the REMD simulation, we ran a straight NPT molecular dynamics simulation at 300 K of the same system for approximately 500 ns, with data sampled every 12 ps (∼40 000 configurations), and compared this simulation to the REMD 300 K temperature ensemble. To make the comparison, we projected all configurations from the 300 K REMD and constant temperature ensembles onto their first two principal components, calculated by combining the two trajectories together and analyzing their covariance matrix. The total combined configurations number approximately 80 000. The projections are shown in Figure 2.

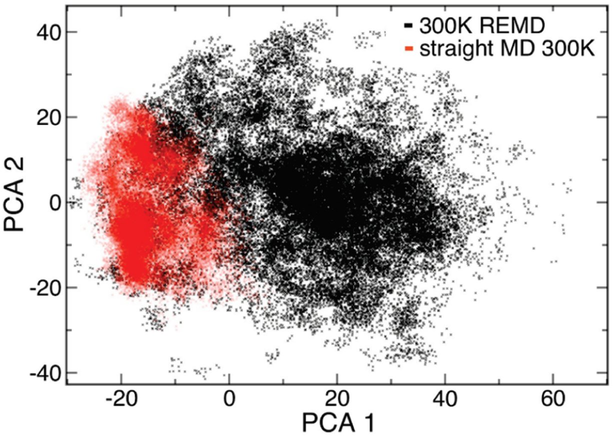


Figure 2. Comparison of replica exchange and constant temperature molecular dynamics at 300 K for Aβ(1−42). Each point represents one configuration generated either at 300 K from the REMD simulation (black) or at 300 K from a straight constant temperature simulation (red). The configurations are projected onto the plane of the first two principal components (PCA-1, PCA-2) calculated by analyzing the covariance matrix of the combined configurations.

The figure clearly shows that the REMD simulation at 300 K covers the PCA-1/PCA-2 plane much more extensively than the constant temperature simulation.[53] This is particularly important since the constant temperature simulation is run for effectively 5 times the length of the REMD simulation (500 ns versus 100 ns per replica), indicating that enhanced sampling is required for simulations of Aβ(1−42) because of kinetic trapping. This has been alluded to in a previous study.[26]

In addition to the simulations carried out for the Amyloid-β peptide, we also ran a small number of pentapeptide simulations of peptides having the general formula GGXGG. These simulations are discussed in the context of analyzing the change in chemical shift (see eq 1 and section 3.2), as evaluated from our theoretical calculations. Initially, we ran these simulations in 15 Å boxes of TIP3P water using the NVT ensemble; after a short equilibration period (∼200 ps), we switched to the NPT ensemble to equilibrate the density of water. Statistics were then gathered using the same control parameters given above for the constant temperature simulation of Aβ(1−42) with the following differences: Acetyl and methyl end-caps were used to prevent charge interactions affecting the equilibrium population of the peptide. In addition, we used the Amber ff99SB force field.[31] These simulations were not run for a fixed time but rather until

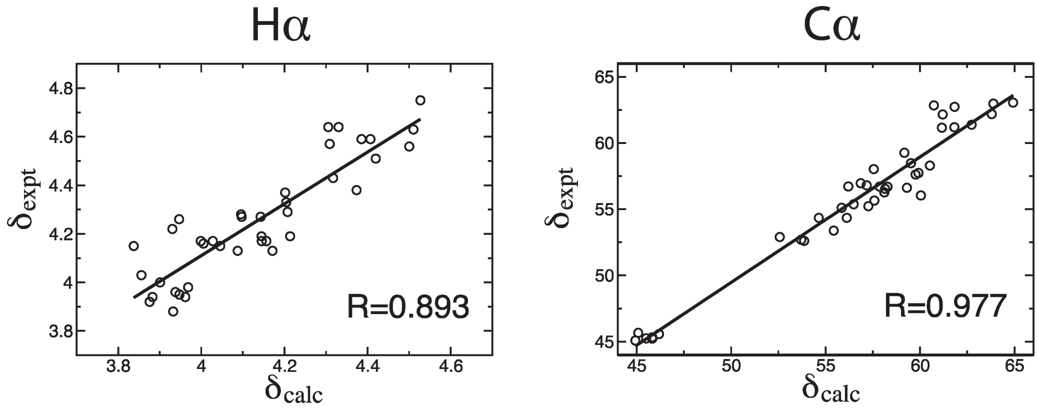


Figure 3. Correlation of average theoretical chemical shifts with experimentally determined chemical shifts. The left panel compares 41 experimentally reported Hα chemical shifts with those determined from the simulation; the right panel compares 41 experimentally reported Cα chemical shifts with those determined from simulation. The lines are linear regressions, and the *R* values are the corresponding Pearson correlation coefficients.

the difference in the percentage populations of dihedral substates between the entire simulation and of the second half of the simulation was less than 5% (using a sampling frequency of 8 ps). This usually resulted in the simulations being approximately 200-ns-long.

3. RESULTS AND DISCUSSION

3.1. Comparison to NMR Chemical Shift Data. Hα and Cα chemical shifts were calculated using an empirical approach as implemented in the SHIFTS program.[54] In order to validate the quality of the theoretically determined chemical shifts, we compared calculated average Hα and Cα NMR shifts from our ensemble of structures with the experimentally determined values for Aβ(1−42).[16] Figure 3 plots calculated Hα and Cα chemical shifts (δ_{calc}, pH ∼ 7, 9 °C) versus experimentally determined chemical shifts (δ_{exp}, pH ∼ 7, 5 °C) along with a least-squares line and the corresponding Pearson coefficient (*R*).

For the Hα chemical shifts, the Pearson coefficient (0.893) is relatively high, suggesting that our simulations reproduce the structural features associated with the chemical shifts. The inherent spread of the data may be due to a number of associated errors including insufficient sampling of configuration space, inaccuracies of the force field, and errors in the theoretical methodology used to calculate the chemical shift. The standard deviation of the theoretical methodology is estimated to be 0.23 ppm,[54] while the mean residual from the current regression analysis is 0.01 ppm. This indicates that error from the theoretical methodology for calculating the chemical shift eclipses any other source of error, i.e. sampling and force-field inaccuracies. Similar results are found for the calculated versus experimental chemical shifts of α-carbons. In this case, the Pearson coefficient is even higher, 0.977. The estimated standard deviation for the theoretical model is 0.97 ppm,[55] whereas the mean residual of the current regression analysis is 1.35 ppm, suggesting again that the largest source of error in our analysis comes from the chemical shift calculation, although in this case it seems that other sources also contribute to the error in the regression analysis.

The high Pearson coefficients indicate that the calculated average chemical shifts from our ensemble of structures are in good agreement with experimental results. This result lends itself to a more detailed analysis of the secondary structure of Aβ(1−42) based on the change in chemical shifts as given in eq 1.

3.2. Analysis of the Brookhaven PDB. Before undertaking a detailed analysis of Aβ(1−42), we wanted to test whether a theoretical protocol can capture the secondary structure shift for globular proteins with well-defined secondary structure. To do this, we analyzed the Brookhaven Protein Data Bank (PDB)[56] and selected all protein structures with a resolution of better than or equal to 1.50 Å and calculated their Cα and Hα chemical shifts. This resulted in analyzing 4456 structures and producing approximately 8 000 000 chemical shifts. This analysis also removes the possibility that errors are resulting from incorrectly parametrized force fields.

In order to automate the process and minimize errors associated with incorrect placement of protons, we first stripped the structures of all their protons and then re-added them with protons according to residue templates in Amber's LeAP module.[48] When adding protons, we made the assumption that all histidine residues were protonated at the ε position and all other titratable residues were protonated assuming an environmental pH of 7. Although this may lead to the incorrect placement for some side-chain protons, especially those buried within a protein, we are assuming that these protons are far enough removed from the α-centers that they do not affect their chemical shifts to a large degree. As it turns out, which we note below, this may not be a justifiable assumption in the case of histidine.

To calculate the change in the chemical shifts ($\Delta\delta$), we used random coil reference values taken from experimental results.[57] However, this presents problems for cysteine and proline residues because they each have two possible random coil references. In the case of cysteine, this is because it is commonly found in two different oxidation states and, for proline, because it may adopt a cis or trans conformation. In order to speed up the automated process, we decided to leave all cysteine residues out of the analysis and assume that all proline residues were in a trans conformation. After calculating the chemical shifts, we accumulated statistics according to what secondary structure element the residue originated from. As there is no unique means to do this, we compared two popular methods, viz, Ramachandran dihedral angles and the DSSP[58] (define secondary structure of proteins) protocol.

For the Ramachandran analysis, we used the following definitions (in degrees) of secondary structure. Right-handed α-helix (alpha): $-100 \leq \varphi \leq -30$; $-80 \leq \psi \leq -5$. Near right-handed α-helix (alpha N): $-175 \leq \varphi \leq -100$; $-55 \leq \psi \leq -5$. Left-handed

Table 1. Average (av), Standard Deviation (std), and Median Absolute Deviation (MAD) for the Change in the Chemical Shift (Δδ) for Cα and Hα Atoms for Various Secondary Structure Elements as Defined Using Ramachandran Angles and the DSSP Protocol for the Brookhaven PDB

	Ramachandran				DSSP		
Cα	av	std	MAD	Cα	av	std	MAD
alpha[a]	2.69	1.18	0.61	α-helix[b]	2.87	1.15	0.58
alpha N[a]	0.28	1.00	0.63	π-helix[b]	1.91	1.55	0.91
alpha L[a]	0.72	0.83	0.40	β-sheet (para)[b]	−1.10	1.36	0.83
PPII[a]	−0.01	1.26	0.87	β-sheet (anti)[b]	−1.11	1.34	0.86
beta[a]	−1.66	1.16	0.78	turn[b]	0.84	1.55	1.02
other[c]	−0.70	1.45	0.98	other[c]	−0.08	1.79	1.15
	Ramachandran				DSSP		
Hα	av	std	MAD	Hα	av	std	MAD
alpha[a]	−0.29	0.29	0.12	α-helix[b]	−0.30	0.27	0.12
alpha N[a]	0.18	0.24	0.13	π-helix[b]	−0.12	0.38	0.26
alpha L[a]	−0.50	0.36	0.10	β-sheet (para)[b]	0.27	0.33	0.19
PPII[a]	0.01	0.38	0.23	β-sheet (anti)[b]	0.27	0.35	0.22
beta[a]	0.26	0.34	0.21	turn[b]	−0.18	0.37	0.21
other[c]	0.00	0.38	0.19	other[c]	−0.08	0.37	0.21

[a] The following definitions for secondary structure were used. Right-handed α-helix (alpha): $-100 \leq \varphi \leq -30$; $-80 \leq \psi \leq -5$. Near right-handed α-helix (alpha N): $-175 \leq \varphi \leq -100$; $-55 \leq \psi \leq -5$. Left-handed α-helix (alpha L): $100 \leq \varphi \leq 30$; $80 \leq \psi \leq 5$. Polyproline II (PPII): $-110 \leq \varphi \leq -50$; $120 \leq \psi \leq 180$. Extended β sheet (beta): $-170 \leq \varphi \leq -110$; $80 \leq \psi \leq 180$ and $-170 \leq \varphi \leq -110$; $-180 \leq \psi \leq -170$. [b] For DSSP definitions, see ref 58. [c] Defined as being any other element that does not fall within the definitions of the prototcol in question, viz, Ramachandran or DSSP.

α-helix (alpha L): $100 \leq \varphi \leq 30$; $80 \leq \psi \leq 5$. Polyproline II (PPII): $-110 \leq \varphi \leq -50$; $120 \leq \psi \leq 180$. Extended β sheet (beta): $-170 \leq \varphi \leq -110$; $80 \leq \psi \leq 180$ and $-170 \leq \varphi \leq -110$; $-180 \leq \psi \leq -170$.

The DSSP algorithm is more sophisticated in that it first identifies hydrogen bonds by using an electrostatic-energy function and then assigns secondary structure either by the topology of hydrogen bonds in a repetitive sequence (for helices and sheets) or by angle restraints (for turns).

After assigning each Δδ to a secondary structure element, either according to Ramachandran angles or the DSSP protocol, we calculated their averages (av), standard deviations (std), and median absolute deviations (MAD). These results are summarized in Table 1. Although there is no one-to-one correspondence between secondary structure elements defined by Ramachandran angles and the DSSP protocol, we can still directly compare the results for α-helices and β-sheets (the β region in the case of Ramachandran angles and the parallel and antiparallel structures in the case of DSSP).

The average Δδ value of α-helical residues for Cα atoms was found to be 2.69 ppm and 2.87 ppm when using Ramachandran angles and the DSSP protocol for assigning secondary structure, respectively. Experimentally, the secondary structure shift is usually reported as 2.50 ppm for α-helical residues. This indicates that the combination of either Ramachandran angles or DSSP for secondary structure assignment with the empirical method for calculating Cα chemical shifts is able to capture the Δδ for α-helical residues reasonably well. For β-sheets, the average Δδ for Cα chemical shifts was found to be −1.66 ppm when using the Ramachandran angle definitions and −1.10 and −1.11 ppm for parallel and antiparallel β-sheets, respectively, when using the DSSP definitions. Experimentally, the Δδ for Cα β-sheets is reported as −2.00 ppm. Again, this indicates that both Ramachandran angles and the DSSP protocol are capturing the Δδ effect. However, using the DSSP definitions moves the average value in the positive direction when compared with experimental results.

A similar analysis of the Hα chemical shifts finds the average Δδ of α-helical residues to be −0.29 ppm and −0.30 ppm when using the Ramachandran and DSSP definitions, respectively. These are in good agreement with the experimentally determined value of −0.30 ppm. For β-sheets, the Ramachandran definition gave an average Δδ of 0.26 ppm, while the DSSP definitions of parallel and antiparallel β-sheets gave averages of 0.27 ppm. Again, these results are in good agreement with the experimental value of 0.30 ppm.

To further quantify the distributions of calculated Δδ values, we have also computed the standard deviation (std) and median absolute deviations (MAD) for each secondary structure element; these results are also presented in Table 1. In addition, we have plotted the Δδ distributions by calculating their histograms. Figure 4 presents the Δδ distributions for all Cα and Hα chemical shifts (upper panels). These distributions are then split according to secondary structure elements using the Ramachandran definitions (middle panels) or the DSSP protocol (lower panels).

The MADs of Table 1 are smaller compared with their respective standard deviations for both Cα and Hα chemical shifts using either Ramachandran or DSSP definitions. This indicates that the calculated distributions of Δδ have long tails because the MADs give them less weight. Figure 4 supports this conclusion. In addition, the upper panels of Figure 4 show that the total Δδ distributions (independent of secondary structure) are unsymmetrical and, in the case of Cα atoms, multimodal.

Comparing the standard deviations of α-helical and β-sheet values with their averages indicates that these distributions overlap more significantly for Hα chemical shifts than Cα chemical shifts; this is also confirmed by the middle and lower panels of Figure 4 (see for example the overlap between the solid black and dashed red lines in the middle panels and the solid black and solid red lines in the lower panels). More importantly, if other secondary structure elements are included in the analysis, it is difficult to decompose both the Cα and Hα distributions into areas that do not include contributions from other secondary structure elements; in particular, it would be difficult to assign a β-sheet element to any residue solely on the basis of Δδ if the protein or peptide is dynamically exploring conformational space. This certainly has consequences when using the CSI method to assign secondary structure to flexible biomolecules. That being said, the Ramachandran decomposition of the Cα distribution shows that the α-helical area of the Δδ distribution is relatively devoid of contributions from other secondary structure elements. This is not the case for the DSSP decomposition if π-helices are present in addition to α-helices.

The distributions in Figure 4 also highlight another important aspect of the random-coil reference values. Frequently, these references are derived from small unstructured peptides or denatured proteins.[57,59−61] The "random coil" or more appropriately the statistical coil is not random but based on maintaining predefined equilibrium populations of each substate through

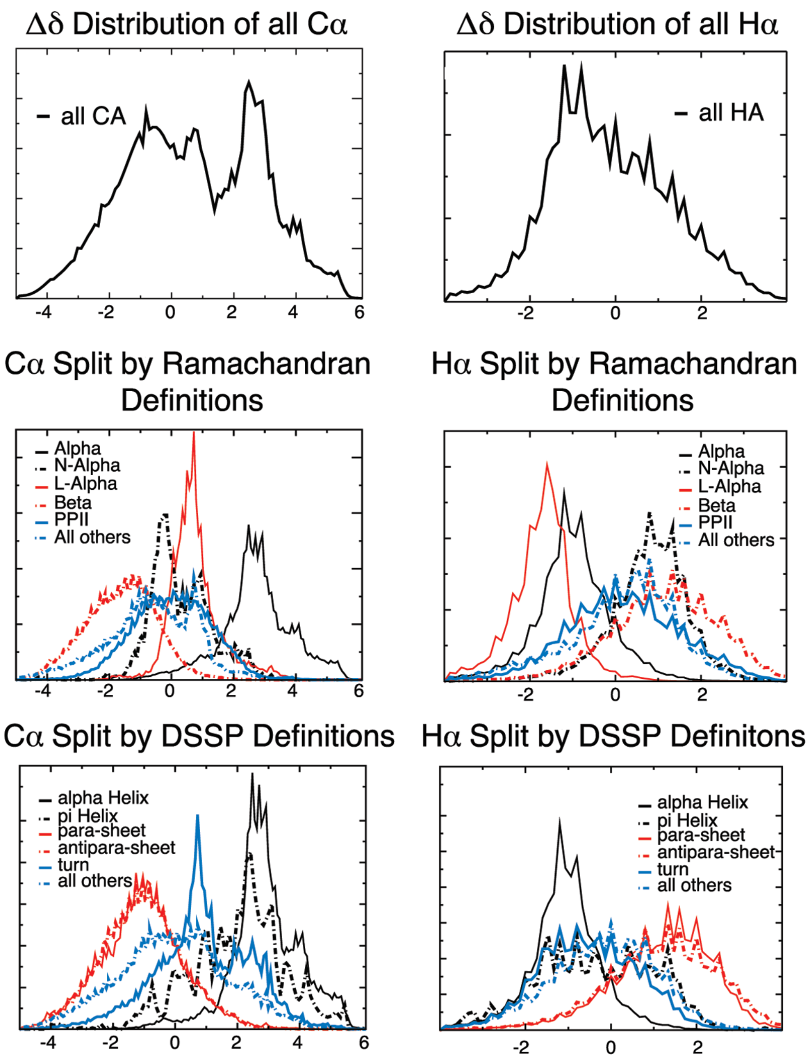


Figure 4. Distributions of $\Delta\delta$ for Cα and Hα chemical shifts derived from structures in the PDB database with a resolution equal to or better than 1.50 Å. Panels show the full distributions (upper) and then these distributions divided according to secondary structure elements according to either Ramachandran angles (middle) or the DSSP protocol (lower).

rapid conformational switching. The equilibrium populations depend primarily on the residue in question[62] but also on its sequence-dependent context.[63] Small peptides are one means of obtaining random-coil values; another way is to look at nonhomologous protein structures from X-ray experiments.[64] In essence, this is what we have done in generating Figure 4; therefore, the average $\Delta\delta$ over all residues for both Cα and Hα should be zero if

(a) the proteins we selected are sufficiently nonhomologous as to represent the inherent random coil structure

(b) the empirical model used to calculate the chemical shifts does not contain systematic errors

(c) the reference random coil values are appropriate

The average of $\Delta\delta$ for all shifts was found to be 0.64 ppm and −0.07 ppm for Cα and Hα shifts, respectively, indicating that the three conditions are being met to a reasonable degree. By the same arguments above, the residue-by-residue values should also average to zero. These results are given in Table 2.

Table 2 clearly demonstrates that while the overall averages of $\Delta\delta$ for Cα and Hα chemical shifts of 0.64 ppm and −0.07 ppm, respectively, are reasonably good, their signs are resulting from systematic errors in the by-residue estimates of $\Delta\delta$. In addition to the systematic errors that are apparent from Table 2, there are a small number of residues for which the total average $\Delta\delta$ deviates significantly from zero. For the Cα chemical shifts, these include Ala (1.10 ppm), His (1.71 ppm), and Arg (1.17 ppm). For the Hα shifts, the residues are Ala (−0.18 ppm), Glu (−0.18 ppm), and His (−0.16 ppm) and to a lesser degree Lys (−0.12 ppm) and Arg (−0.12 ppm).

In order to probe the origin of the systematic errors and a small number of large deviations, we ran molecular dynamics simulations for a selection of pentapeptides of the general formula GGXGG and calculated the average $\Delta\delta$ for the center residue X. These results are also given in Table 2.

For alanine, the large deviations of 1.10 ppm and −0.18 ppm are reduced to 0.11 ppm and −0.10 ppm for Cα and Hα centers, respectively, when using molecular dynamics simulations. This indicates that these deviations are originating from a bias in the statistics generated from the PDB analysis. The known prevalence of alanine in α-helical structures and the respective average downfield and upfield shifts of the Hα and Cα atoms suggests that the PDB structures are over-representing alanine in

Table 2. Per Residue Averages for the Change in the Chemical Shift ($\Delta\delta$) for Cα and Hα Atoms from an Analysis of the PDB Databank and a Selection of GGXGG Peptides from Molecular Dynamics Simulations

	Cα Ave.		Cα Ave.		Hα Ave.	Hα	Ave.
Brookhaven PDB[a]							
Ala	1.10	Met	0.75	Ala	−0.18	Met	−0.07
Asp	0.35	Asn	0.64	Asp	−0.08	Asn	−0.07
Glu	0.71	Gln	0.87	Glu	−0.18	Gln	−0.10
Phe	0.40	Arg	1.17	Phe	−0.05	Arg	−0.12
His	1.71	Ser	0.29	His	−0.16	Ser	−0.06
Pro	0.66	Thr	0.39	Pro	−0.07	Thr	0.10
Ile	0.60	Val	0.65	Ile	0.05	Val	0.05
Lys	0.85	Trp	0.22	Lys	−0.12	Trp	−0.11
Leu	0.79	Tyr	−0.07	Leu	−0.08	Tyr	0.00
Molecular Dynamics[b]							
Ala	0.11	Pro	0.11	Ala	−0.10	Pro[b]	−0.03
Asp	0.11	Lys	−0.32	Asp	−0.12	Lys	−0.06
His	1.30	Thr	−0.02	His	−0.15	Thr	−0.02

[a] Statistics obtained from an analysis of the PDB; see text for details. [b] Statistics obtained from molecular dynamics simulations; see text for details.

α-helices compared with the inherent random coil state. In fact, 57.8% of alanine residues from the PDB analyzed structures are found in an α-conformation using Ramachandran angle definitions, whereas the alanine of GGAGG spends only 17.2% of the time in the α region.

The similar large deviations for histidine of 1.71 ppm and −0.16 ppm for Cα and Hα, respectively, would immediately suggest a similar bias for this residue. However, the molecular dynamics simulations also produce large deviations (1.30 ppm and −0.15 ppm). The small reduction in the $\Delta\delta$ values is because the PDB structures have an α-region population of 33.4%, whereas the molecular dynamics simulations produce an α-region population of 18.2%. The remaining large error is most likely because of the pH effect on Cα and Hα chemical shifts of histidine, which can be up to 2.5 ppm for the Cα atom.[65,66] Therefore, the large errors seen for histidine are probably a combination of assuming that the side chain is permanently protonated at the ε position and the random coil reference values that we have used to generate the $\Delta\delta$ statistics being incorrect for this permanent protonation state. Modest improvements in the average $\Delta\delta$ are also seen for the other residues by using molecular dynamics simulations. This further indicates that even with a large data set from the PDB, biasing is introduced to a small degree. However, because the dynamics simulations do not produce exactly zero deviations, it is suggested that there is also a modest systematic error in the theoretical methodology of on average approximately 0.10 ppm for both Cα and Hα atoms.

In summary, we have analyzed the empirical methodology for calculating the secondary structure shift ($\Delta\delta$) using a large data set of structures from the PDB. In conjunction with these chemical shifts, we examined two popular means of dividing the conformational subspace into secondary structure elements. We found that the secondary structure distributions of $\Delta\delta$ as divided by Ramachandran angles and the DSSP protocol results in substantial overlap of the distributions for both Cα and Hα chemical shifts. However, if the conformational subspace is divided by Ramachandran angles, then the α region of the Cα distribution is not significantly contaminated by other secondary structure elements. In addition, we found that modest systematic errors are introduced into the average $\Delta\delta$'s because of the theoretical methodology. In some cases, however, large errors are introduced because of either incorrect statistics from the data set (Ala) or from misrepresenting the correct dynamic protonation states (His).

3.3. Distributions of Chemical Shifts for Aβ(1−42).

In the previous sections, we have shown that the mean chemical shifts for Aβ(1−42), as calculated from theory, agree well with experimental results. We also demonstrated by analyzing the PDB that the theoretically determined change in the chemical shift ($\Delta\delta$) is also reproduced well. If the protein in question does not vary significantly from its equilibrium position, then the $\Delta\delta$ distributions for each residue would be expected to have minor fluctuations around their means, making secondary structure assignment unambiguous. However, in the case of a flexible biomolecule like Aβ(1−42), we will show below that the distributions of $\Delta\delta$ are broad like in Figure 4, where we combined many peptide conformations from the PDB. We will also investigate the consequences this has for assigning its secondary structure. In the discussion below, we will refer to secondary structure elements defined through Ramachandran dihedral angles because the PDB analysis showed that the $\Delta\delta$ effect is in closer agreement with experimental results using these definitions.

It is important to state from the outset the differences between the inherent random-coil state of a residue and the random-coil region of chemical-shift distributions. The random-coil state of a residue is a statistical combination of all of the dihedral angles the backbone explores when it is unconstrained by contextual influences. As the unconstrained residue flips from β to α regions, the Hα chemical shift changes from being relatively upfield to being relatively downfield; the reverse is true for Cα shifts. The PDB analysis above demonstrated that other structural elements may also move the shift upfield or downfield and in some cases leaves it unchanged. The mean chemical shift of the random-coil state is then the average of these competing shifts, and the change in chemical shift for this state is defined to be zero. The random-coil region of the distribution of changes in chemical shifts is thus a zone centered on zero. When ascribing a random coil structural element to a residue, it is assumed that a zero overall shift is resulting from the residue in question, sampling from the same populations of dihedral angles that originally defined its random coil state.

In order to demonstrate the relationship between the mean chemical shift and the underlying distributions for a flexible peptide, we plotted the $\Delta\delta$ distributions of each residue from our simulations of Aβ(1−42) and compared these distributions against typical random coil, α-helical, and β-sheet $\Delta\delta$ values. Eight examples of these plots are shown in Figures 5 and 6 (plots for all residues can be found in the Supporting Information, Figures S1 and S2). In addition to the distributions and typical values of $\Delta\delta$, each plot indicates the mean value from our simulations and the observed value from experiments. Random coil values were taken from tabulated data, as were helical and sheet values; the domains of the $\Delta\delta$ zones shown in Figures 5 and 6 were derived from the first standard deviation of $\Delta\delta$ values calculated from the analysis of the PDB described above.

For the Hα chemical shifts, we find a number of consistent key features, which are illustrated by the distributions calculated for

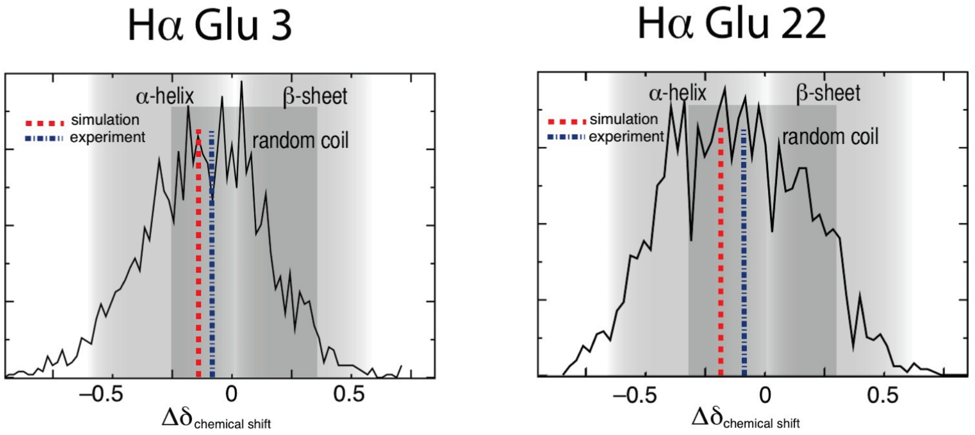


Figure 5. Δ*δ* Distributions of Hα chemical shifts of Glu3 and Glu22 in A*β*(1−42) at 300 K. Shaded areas give the domains encompassing typical random coil, helix, and *β*-sheet values. The sizes of the shaded areas have been determined by using one standard deviation from the mean values found from analyzing the PDB. The mean values from simulations and experiments are indicated with dashed lines. See also Figure S1, Supporting Information.

Glu3 and Glu22 as shown in Figure 5a and b. First, like the analysis of the PDB, the regions spanning the secondary structure domains of Δ*δ* values overlap one another significantly. The significant overlap between different structural elements causes the resultant distributions to be Gaussian-like (often skewed to one side) with the mean value aligned near the distribution's maximum value. This situation is clearly shown in Figure 5. For example, because of the overlap in Δ*δ* values from different structural elements, it is difficult to discern whether or not the similar average Δ*δ* values for Glu3 and Glu22 are resulting from the same combination of structural elements. The mean values for Glu3 and Glu22 from experiments and simulations are located in the random coil region. Naturally, one would infer from this information that both residues are sampling from the random-coil state. However, when the α-content is calculated using Ramachandran angles, we find that Glu3 spends 29.3% of the time in this configuration, while Glu22 spends 61.4% of the time in this configuration. Therefore, Glu3 and Glu22 cannot both be sampling from the inherent random-coil population of glutamic acid.

The large overlap of domains causes 34 of the 38 Hα distributions plotted in this study (see also Figure S1 in the Supporting Information) to have their mean values from simulations and experiments lying in the region combining α-helical and random coil values, even though each residue may be sampling from distinctly different dihedral populations.

The Cα chemical shift distributions show some significant differences compared with the Hα distributions. Unlike the Hα distributions, they are not simply Gaussian-like and centered on the mean value but may be complicated bi- or trimodal distributions, see for example the plots in Figure 6. From this figure, it can be clearly seen that the average value is derived from combinations of three distinct and separated distributions, viz, α-helical, random coil, and *β*-sheet. Because each zone is clearly observable, the Cα shifts may lend themselves to a more robust means of identifying if distributions with similar averages have distinct populations. To investigate this, we have plotted in Figure 6 the distributions for three pairs of residues with the same side chain from the A*β*(1−42) sequence.

In the top panels, we have plotted the distributions of Ala2 and Ala21. The mean value from experimental results and theory lie in the random coil region, and the underlying distributions look fairly similar. That is, they have large peaks in the random coil and α-regions with little signal in the *β*-region. However, like the situation for glutamic acid above, the secondary structure populations calculated from Ramachandran angles are quite different; for example, Ala2 spends 25.7% of the time in the α-region, whereas Ala21 spends 42.8% of the time in the α-region. The simulation of GGAGG we performed in conjunction with the PDB analysis indicates that alanine spends approximately 26.8% of the time in the α-region when in a random coil configuration, which is supported by another study.[62] Therefore, even though both distributions average to a random coil value the underlying distribution of Ala2 seems to be much more closely related to the random coil then does the distribution of Ala21.

For Val18 and Val36, the distributions are again divided into distinct areas. Val18 has large peaks in the α- and random coil regions with a smaller peak in the *β*-region. Val36 on the other hand has large peaks in the *β*- and random coil regions with little or no signal in the α-region. Again, however, their means from simulations and experiments lie in the random coil region. In this case, the distributions are representative of their secondary structure content. For instance, Val18 spends 35.2% of the time in the α-region, and Val36 spends 6.3% of the time in this region. The sizes of the peaks in the *β*-regions are commensurate with their calculated *β*-content of 43.1% and 51.4% for Val18 and Val36, respectively. The large middle peaks are derived from predominately the PII structure, 14.4% in the case of Val18 and 22% in the case of Val36. Although the distributions can be related to the calculated secondary structure and their mean values lie in the random coil region, the random coil distribution of valine has approximately 21.9% helical content and 30.4% *β*-content;[62] therefore, neither Val18 nor Val36 seem to be representing this population exactly.

The final pair of distributions we plot in Figure 6 are for Phe19 and Phe20. Like the situation for Ala2 and Ala21, these distributions share similar shapes. Both have large α-peaks and a large peak on the edge of the *β*-/random coil region. Their mean values lie in the random coil region. Again, the peaks are representative of their calculated secondary structure; Phe19 and Phe20 spend 29.7% of the time in the α-region and 65.5 and 62.3% of the time in the *β*-region, respectively. This indicates that in the case of phenylalanine the chemical shift for *β*-sheets is shifted relatively less downfield compared with other residues.

The differences observed between the six residues above highlight a number of important points. First, signals that result

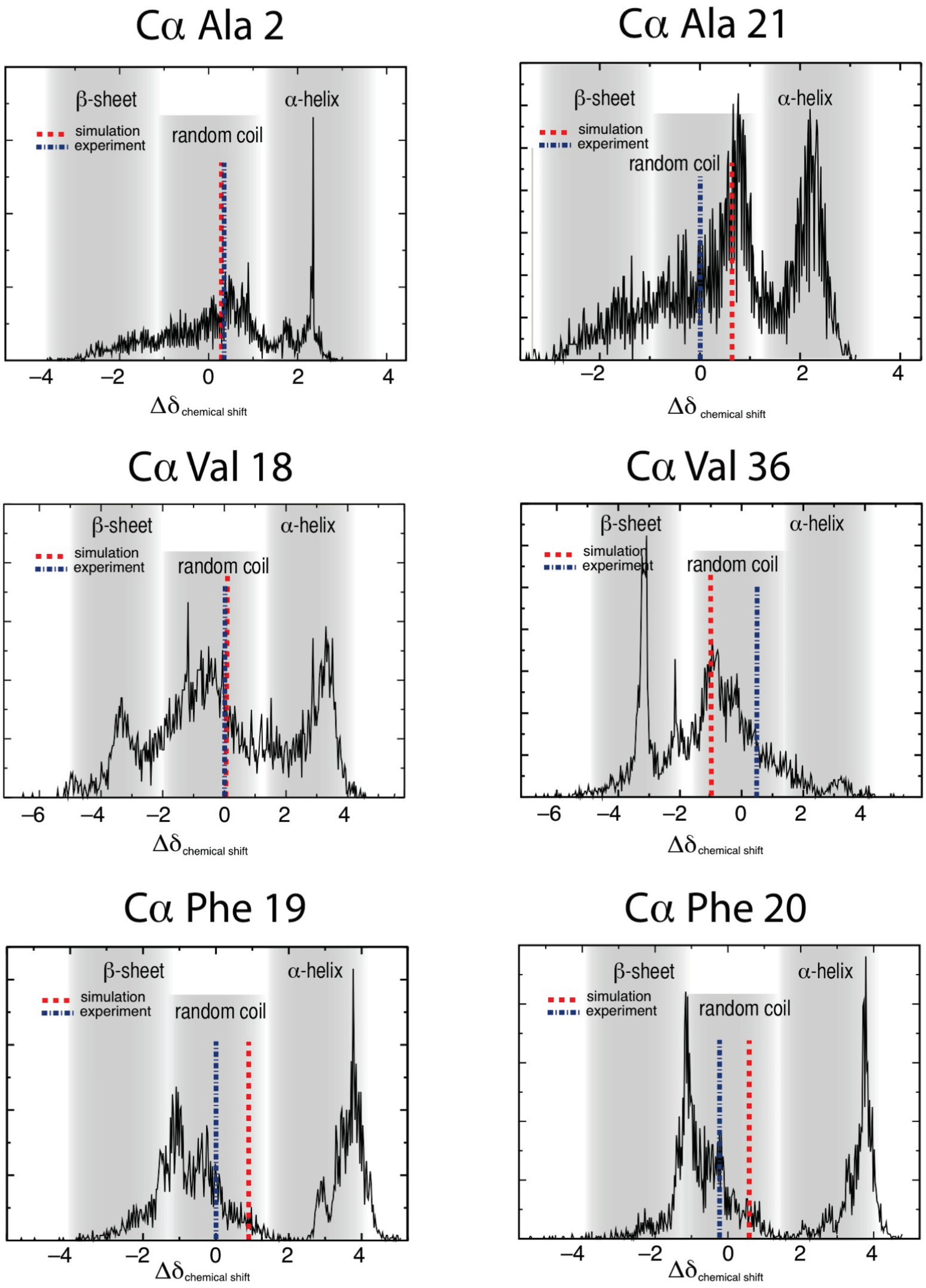


Figure 6. Distributions of Cα chemical shifts of Ala2, Ala21, Val18, Val36, Phe19, and Phe20 from Aβ(1−42) at 300 K. Shaded areas give the domains encompassing typical random coil, helix, and β-sheet values. The sizes of the shaded areas have been determined by using one standard deviation from the mean values found in an analysis of the PDB data bank (see test). The mean values from simulations and experiments are indicated with dashed lines. See also Figure S2, Supporting Information.

in random coil mean values are a result of sampling from combinations of all three regions. Second, the dynamically explored secondary structure elements can change by up to 30% between two residues with the same side chain but still result in random coil averages. Finally, the random coil region of the chemical shift distribution does not necessarily have to contain a signal for the mean value to reside there.

From the distributions given in Figure 6, one might be tempted to estimate the secondary structure content of a particular residue by integrating the signals in each area and comparing them against the total signal. However, as shown in Figure 4, the β-region, random coil region, and to a lesser degree the α-region may contain a signal from more than one secondary structure element when defined through Ramachandran angles. To demonstrate how this may influence the secondary structure populations from a histogram analysis, we plot in Figure 7 the α- and β- Ramachandran populations for all nonglycine residues of Aβ(1−42) against the estimated populations by counting the

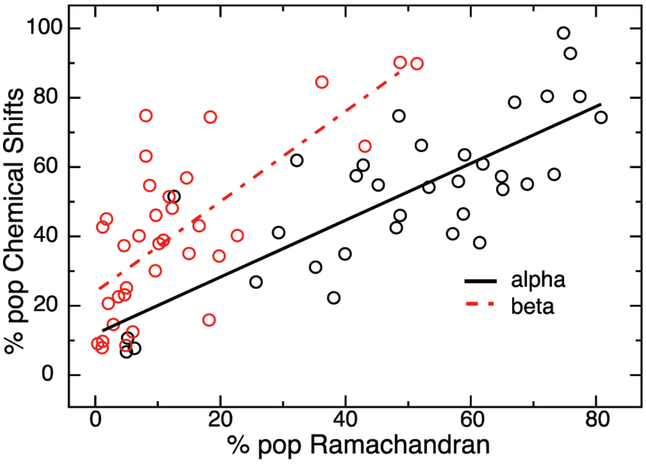


Figure 7. Correlation of % population of α- an β-structure calculated with Ramachandran angles with the % population of α- and β-structure determined by counting the number of chemical shifts found in the α- and β-areas of the chemical shift distributions. Data points are from all nonglycine residues in Aβ(1−42).

total signal within two standard deviations of the mean β- and α-chemical shifts.

The figure shows that both α- and β-Ramchandran populations are correlated to the areas of the chemical shift distributions. However, simply because they correlate does not infer that the model is valid. If it were valid, then the slope of the regression lines would pass through zero and have a slope of one. The line correlating the β-populations has a slope larger than one (1.3) and crosses the *y* intercept at 22.0. The *y* intercept indicates that on average the β-area of the chemical shift distribution contains at least 22% additional structure not associated with β dihedral angles. The slope indicates that, as the percentage of β-sheet structure increases, the contamination from other elements in the β-area of the chemical shift plot also increases. The line correlating the size of the α-area to the α-populations is a better model for predicting the α-content with a slope of 0.82 and a *y* intercept of 11.0. An intercept larger than zero is again an effect associated with contamination from other secondary structure elements, albeit to a lesser degree than the β-area contamination. With a slope smaller than one, this model infers that as the percentage of α-helical structures increases, the number of those that are captured in the α-area of the chemical shift plot decreases.

3.4. Static and Dynamic Helical Content of Aβ(1−42). As stated in the Introduction, molecular dynamics simulations usually report a much higher helical content for Aβ(1−42) than experimental studies. In both types of reporting, the "static" helical content is being calculated, that is, the average over the course of the simulation or over the course of the experiment. In this section, we report the possible underlying reasons for the discrepancy between theory and experimental results. We conclude that it is ambiguous to assign a static content to a peptide that does not constitute one specific conformational fold. This is because the final average helical content depends on how a helix per residue is defined, and as a result, we may calculate a wide range of percent content (from 0% to 45.7% in this case) even when using the same data set.

First, using the change in Cα chemical shift as an indicator of helical content, we can calculate the static content in the same way that the CSI method does. Of the 40 Cα chemical shift distributions (see Figure S2 of the Supporting Information), only four of them have mean values from simulation in the α-helical area (Ser8, Tyr10, Leu17, and Ser26); an additional three lie in the β-sheet area (Val36, Val39, and Val40). In order to define an α-helix, the chemical shift index method requires that both the mean chemical shift is in the α-helical area, and at least four consecutive residues are found in this area. Using these rules to interpret our chemical shift data, Aβ(1−42) has 0% α-helical content, reproducing the results found for Aβ(1−42) using the CSI method in ref 16.

Normally, helical content is reported from simulation by averaging the helical content from theoretical calculations; popular methods are to average the content as determined by DSSP, Ramachandran angles, or STRIDE,[67] which is defined similarly to DSSP. Although for DSSP we found less satisfactory results for the change in chemical shift and, in addition, the DSSP structural decomposition leads to an overlap of α-helices and π-helices, the vast body of literature supports DSSP as a valid method for finding helical structures. Using this methodology, we find an average helical content of 25.9%, in agreement with most molecular dynamics studies. The STRIDE algorithm similarly finds a 29.1% helical content. For both algorithms, the residues contributing most to these numbers are Tyr10−Leu17, which for approximately 50% of the simulation are in a helix, and residues Val18−Ser26, which spend approximately 30% of the simulation as a helix (see Supporting Information Table S1). However, these numbers may overestimate the *static* helical content with respect to experimental results because residues that form part of the boundary between the helical and nonhelical structure are counted even though they may not be part of a persistent helix.

More simply than the DSSP and STRIDE estimates, we could average the Ramachandran angles that fall inside of the α region. For our simulation, this results in a helical content of 45.7%. However, this is a rather naive approach for two reasons: First, a Ramachandran angle pair residing in the α-region does not equate to an α-helix. In essence, this is what the DSSP and STRIDE algorithms are correcting for. Second, the random coil state of an amino acid samples significantly from the α-region, see ref 62. To correct for the random coil populations, we may define the helical "excess" of the species by taking the average of the differences in the helical content per residue from the simulation with the helical population of the random coil state and normalize (see eq 3).

$$\text{helical excess} = \frac{\sum_{i=1}^{N} (\%\text{helix}^{\text{res}(i)}_{\text{simulation}} - \%\text{helix}^{\text{res}(i)}_{\text{randomcoil}})}{\sum_{i=1}^{N} (1 - \%\text{helix}^{\text{res}(i)}_{\text{randomcoil}})} \qquad (3)$$

If we take the average random coil helical populations[68] for each residue from ref 62, the helical excess of Aβ(1−42) is found to be 27.5%, which is interestingly in agreement with the DSSP and STRIDE results. However, this formula can be potentially misleading because it may include negative values in the average, much like the chemical shift assignment does.

The preceding discussion demonstrates that it is difficult to ascribe a single static content to Aβ(1−42) even when the data set is the same, as it depends on how the helical content is defined. A more appropriate way of describing the helical content is to look at its dynamic structure. This can be demonstrated through the joint probability of finding a sequence of *n* residues (starting from

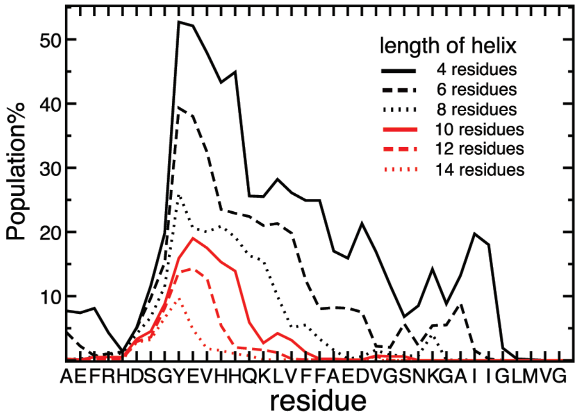


Figure 8. Population of helical structure as a function of helix length and residue number calculated from the joint probability function of Ramachandran angles in the α region, where the joint probability function is $P(i,i+1,..,i+n-1)$, i is the starting residue, and n is the length of the helix (see text for further details).

residue i) that reside in the helical basin of the Ramachandran plot. In Figure 8, we have plotted this function for helix lengths of 4, 6, 8, 10, 12, and 14 residues. As expected the probability of finding longer helices is smaller than finding shorter ones. The plot in Figure 8 suggests that there is a strong probability of finding short helices four residues in length starting from Tyr10. This short sequence forms the start of longer helices that extend toward the C-terminal end, which even for a length of 14 residues has a probability of approximately 10% from our simulation. Short helical sequences of four to six residues also have a 10 or 20% probability at the N- and C-terminal ends, respectively.

This description of helcity makes sense when comparing our knowledge of Aβ in transmembrane and apolar environments. If we imagine gradually making the environment around Aβ-(1−42) more apolar, then the probability of finding the helix starting at Tyr10 increases. In addition, the probability of finding longer helices starting at this residue increases until the permanent helix spanning residues 10−25 is established, as found from experimental studies in apolar environments. The helix that starts at residue 27 in Figures 1 and 8 behaves in a similar fashion.

4. CONCLUSION

In summary, we have performed a comparison of experimental chemical shift data with theoretically determined chemical shifts for Aβ(1−42). First, we investigated the correlation between chemical shifts and protein secondary structure in the case of structures in the PDB. We found that for Hα chemical shifts, the theoretical chemical shift distributions of secondary structure elements overlap one another significantly, making it difficult to discern if the calculated mean was composed of one element or another. For Cα chemical shifts, the distributions for each secondary structure element investigated are more well separated, making them in principle more suitable for secondary structure assignment. However, we showed that residues with the same side chain could produce similar means and distributions but could have differences in secondary structure of up to 30%. Finally, we showed that a single static helical content could vary from 0% to 45.7%, depending on the method used and how the helical content is defined, even when the same data set is used. For flexible molecules, it is thus more appropriate to look at the probability of helical persistence as a function of length.

■ ASSOCIATED CONTENT

S Supporting Information. Plots of the distributions of Hα and Cα chemical shifts for all residues can be found in Figures S1 and S2, respectively. Percentage populations of secondary structure elements calculated using Ramachandran angles are given in Table S1. This information is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: geoffrey.wood@epfl.ch, ursula.roethlisberger@epfl.ch.

■ ACKNOWLEDGMENT

This work was supported by the Swiss National Science Foundation through grant 200020-116294. We also acknowledge generous computing time from the HPC department at the EPF Lausanne, and we thank Professor M. G. Zagorski for providing us with the experimentally determined chemical shifts of Aβ(1-42).

■ REFERENCES

(1) Selkoe, D. J. *Physiol. Rev.* **2001**, *81*, 741–766.

(2) Hardy, J.; Selkoe, D. J. *Science* **2002**, *297*, 353–356.

(3) Shoji, M.; Golde, T. E.; Ghiso, J.; Cheung, T. T.; Estus, S.; Shaffer, L. M.; Cai, X. D.; McKay, D. M.; Tintner, R.; Frangione, B.; Younkin, S. G. *Science* **1992**, *258*, 126–129.

(4) Selkoe, D. J. *Trends Cell Biol.* **1998**, *8*, 447–453.

(5) Kang, J.; Lemaire, H. G.; Unterbeck, A.; Salbaum, J. M.; Masters, C. L.; Grzeschik, K. H.; Multhaup, G.; Beyreuther, K.; Mullerhill, B. *Nature* **1987**, *325*, 733–736.

(6) Gregory, G. C.; Halliday, G. M. *Neurotox. Res.* **2005**, *7*, 29–41.

(7) Lambert, M. P.; Barlow, A. K.; Chromy, B. A.; Edwards, C.; Freed, R.; Liosatos, M.; Morgan, T. E.; Rozovsky, I.; Trommer, B.; Viola, K. L.; Wals, P.; Zhang, C.; Finch, C. E.; Krafft, G. A.; Klein, W. L. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 6448–6453.

(8) Walsh, D. M.; Hartley, D. M.; Kusumoto, Y.; Fezoui, Y.; Condron, M. M.; Lomakin, A.; Benedek, G. B.; Selkoe, D. J.; Teplow, D. B. *J. Biol. Chem.* **1999**, *274*, 25945–25952.

(9) Hartley, D. M.; Walsh, D. M.; Ye, C. P.; Deihl, T. D.; Vasquez, S.; Vassilev, P. M.; Teplow, D. B.; Selkoe, D. J. *J. Neurosci.* **1999**, *19*, 8876–8884.

(10) Barrow, C. J.; Zagorski, M. G. *Science* **1991**, *253*, 179–182.

(11) Shao, H. Y.; Jao, S. C.; Ma, K.; Zagorski, M. G. *J. Mol. Biol.* **1999**, *285*, 755–773.

(12) Ma, K.; Clancy, E. L.; Zhang, Y. B.; Ray, D. G.; Wollenberg, K.; Zagorski, M. G. *J. Am. Chem. Soc.* **1999**, *121*, 8698–8706.

(13) Crescenzi, O.; Tomaselli, S.; Guerrini, R.; Salvadori, S.; D'Ursi, A. M.; Temussi, P. A.; Picone, D. *Eur. J. Biochem.* **2002**, *269*, 5642–5648.

(14) Tomaselli, S.; Esposito, V.; Vangone, P.; van Nuland, N .A. J.; Bonvin, A. M. J. J.; Guerrini, R.; Tancredi, T.; Temussi, P. A.; Picone, D. *Chembiochem* **2006**, *7*, 257–267.

(15) Petkova, A. T.; Ishii, Y.; Balbach, J. J.; Antzutkin, O. N.; Leapman, R. D.; Delaglio, F.; Tycko, R. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 16742–16747.

(16) Hou, L. M.; Shao, H. Y.; Zhang, Y. B.; Li, H.; Menon, N. K.; Neuhaus, E. B.; Brewer, J. M.; Byeon, I. J. L.; Ray, D. G.; Vitek, M. P.; Iwashita, T.; Makula, R. A.; Przybyla, A. B.; Zagorski, M. G. *J. Am. Chem. Soc.* **2004**, *126*, 1992–2005.

(17) Danielsson, J.; Jarvet, J.; Damberg, P.; Graslund, A. *FEBS J.* **2005**, *272*, 3938–3949.

(18) Kirkitadze, M. D.; Condron, M. M.; Teplow, D. B. *J. Mol. Biol.* **2001**, *312*, 1103–1119.

(19) Bitan, G.; Kirkitadze, M. D.; Lomakin, A.; Vollers, S. S.; Benedek, G. B.; Teplow, D. B. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 330–335.

(20) Lee, J. P.; Stimson, E. R.; Ghilardi, J. R.; Mantyh, P. W.; Lu, Y. A.; Felix, A. M.; Llanos, W.; Behbin, A.; Cummings, M.; Vancriekinge, M.; Timms, W.; Maggio, J. E. *Biochemistry* **1995**, *34*, 5191–5200.

(21) Zhang, S.; Iwata, K.; Lachenmann, M. J.; Peng, J. W.; Li, S.; Stimson, E. R.; Lu, Y.; Felix, A. M.; Maggio, J. E.; Lee, J. P. *J. Struct. Biol.* **2000**, *130*, 130–141.

(22) Wishart, D. S.; Sykes, B. D.; Richards, F. M. *Biochemistry* **1992**, *31*, 1647–1651.

(23) Dalgarno, D. C.; Levine, B. A.; Williams, R. J. P. *Biosci. Rep.* **1983**, *3*, 443–452.

(24) Deslauriers, R.; Smith, I. C. P. In *Topics in Carbon-13 NMR Spectroscopy*; Levy, G. C., Eds.; John Wiley & Sons: New York, 1976; Vol 2, pp 1−80.

(25) Szilagyi, L. *Proc. Nucl. Magn. Reson. Spectrosc.* **1995**, *27*, 325–443.

(26) Raffa, D. F.; Rauk, A. *J. Phys. Chem B* **2007**, *111*, 3789–3799.

(27) Sgourakis, N. G.; Yan, Y.; McCallum, S. A.; Wang, C.; Garcia, A. E. *J. Mol. Biol.* **2007**, *368*, 1448–1457.

(28) Yang, M. F.; Teplow, D. B. *J. Mol. Biol.* **2008**, *384*, 450–464.

(29) Takeda, T.; Klimov, D. K. *J. Phys. Chem B* **2009**, *113*, 6692–6702.

(30) Vitalis, A.; Caflisch, A. *J. Mol. Biol.* **2010**, *403*, 148–165.

(31) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65*, 712–725.

(32) Ferrara, P.; Apostolakis, J.; Caflisch, A. *Proteins: Struct. Funct. Bioinf.* **2002**, *46*, 24–33.

(33) Trzesniak, D.; Glattli, A.; Bernhard Jaun, B.; van Gunsteren, W. F. *J. Am. Chem. Soc.* **2005**, *127*, 14320–14329.

(34) Glattli, A.; van Gunsteren, W. F. *Angew. Chem., Int. Ed. Engl.* **2004**, *43*, 6312–6312.

(35) van Gunsteren, W. F.; Dolenc, J.; Mark, A. E. *Curr. Opin. Struct. Biol.* **2008**, *18*, 149–153.

(36) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141–151.

(37) Sanbonmatsu, K. Y.; Garcia, A. E. *Proteins: Struct. Funct. Bioinform.* **2002**, *46*, 225–234.

(38) Rathore, N.; Chopra, M.; de Pablo, J. J. *J. Chem. Phys.* **2005**, *122*, 024111.

(39) Fukunishi, H.; Watanabe, O.; Takada, S. *J. Chem. Phys.* **2002**, *116*, 9058–9067.

(40) Hritz, J.; Oostenbrink, C. *J. Chem. Phys.* **2008**, *128*, 144121.

(41) Sugita, Y.; Kitao, A.; Okamoto, Y. *J. Chem. Phys.* **2000**, *113*, 6042–6051.

(42) Okur, A.; Roe, D. R.; Cui, G. L.; Hornak, V.; Simmerling, C. *J. Chem. Theory Comput.* **2007**, *3*, 557–568.

(43) Roitberg, A. E.; Okur, A.; Simmerling, C. *J. Phys. Chem. B* **2007**, *111*, 2415–2418.

(44) Okur, A.; Wickstrom, L.; Layten, M.; Geney, R.; Song, K.; Hornak, V.; Simmerling, C. *J. Chem. Theory Comput.* **2006**, *2*, 420–433.

(45) Okur, A.; Wickstrom, L.; Simmerling, C. *J. Chem. Theory Comput.* **2008**, *4*, 488–498.

(46) Temperature generator for REMD-simulations. http://folding.bmc.uu.se/remd/index.php (accessed March, 2008).

(47) Patriksson, A.; van der Spoel, D. *Phys. Chem. Chem, Phys.* **2008**, *10*, 2073–2077.

(48) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J., Duke, R. E., Luo, R., Merz, K. M., Pearlman, D. A., Crowley, M.; Tsui, V.; Gohlke, H.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P.; Kollman, P. A. *AMBER 9*; University of California: San Francisco, CA, 2006.

(49) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J. M.; Kollman, P. *J. Comput. Chem.* **2003**, *24*, 1999–2012.

(50) Wang, T.; Wade, R. C. *J. Chem. Theory Comput.* **2006**, *2*, 140–148.

(51) Li, H.; Robertson, A. D.; Jensen, J. H. *Proteins* **2005**, *61*, 704–721.

(52) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 372–341.

(53) Although the sampling frequencies for the two simulations are different (2 ps versus 12 ps), the figure does not change significantly if the number of configurations is adjusted to reflect equal sampling frequencies (12 ps). The total number of configurations (~80 000) is kept in this analysis for purposes of clarity.

(54) Osapay, K.; Case, D. A. *J. Am. Chem. Soc.* **1991**, *113*, 9436–9444.

(55) Xu, X. P.; Case, D. A. *J. Biomol. NMR* **2001**, *21*, 321–333.

(56) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne., P. E. *Nucleic Acids Res.* **2000**, *28*, 235–242.

(57) Wishart, D. S.; Case, D. A. *Methods Enzymol.* **2001**, *338*, 3–34.

(58) Kabsch, W.; Sander, C. *Biopolymers.* **1983**, *22*, 2577–2637.

(59) Braun, D.; Wider, G.; Wuthrich, K. *J. Am. Chem. Soc.* **1994**, *116*, 8466–8469.

(60) Merutka, G.; Dyson, H. J.; Wright, P. E. *J. Biomol. NMR* **1995**, *5*, 14–24.

(61) Wishart, D. S.; Bigam, A.; Holm, R. S.; Sykes, B. D. *Biomol. NMR* **1995**, *5*, 67–81.

(62) Beck, D. A. C.; Alonso, D. O. V.; Inoyama, D.; Daggett, V. *Proc. Natl. Acad. Sci. U.S.A* **2008**, *105*, 12259–12264.

(63) Ting, D.; Wang, G. L.; Shapovalov, M.; Mitra, R.; Jordan, M. I.; Dunbrack, R. L. *PLOS Comput. Biol.* **2010**, *6*, e1000763.

(64) Smith, L. J.; Fiebig, K. M.; Schwalbe, H.; Dobson, C. M. *Curr. Biol.* **1996**, *1*, R96–R106.

(65) Wasylishen, R. E.; Tomlinson, G. *Biochem. J.* **1975**, *147*, 605–607.

(66) Wang, L. Y.; Eghbalnia, H. R.; Markley, J. L. *J. Biomol. NMR* **2006**, *35*, 155–165.

(67) Frishman, D; Argos, P. *Proteins* **1995**, *23*, 566–579.

(68) We do not include the first and last residues in this calculation because the Ramachandran angles are not defined for these residues.

Assessing Protein Loop Flexibility by Hierarchical Monte Carlo Sampling

Jerome Nilmeier,[†,§] Lan Hua,[†,§] Evangelos A. Coutsias,[‡] and Matthew P. Jacobson[*,†]

[†]Department of Pharmaceutical Chemistry, University of California in San Francisco, San Francisco, California 94158-2517, United States

[‡]Department of Mathematics and Statistics, University of New Mexico, Albuquerque, New Mexico 87131, United States

ABSTRACT: Loop flexibility is often crucial to protein biological function in solution. We report a new Monte Carlo method for generating conformational ensembles for protein loops and cyclic peptides. The approach incorporates the triaxial loop closure method, which addresses the inverse kinematic problem for generating backbone move sets that do not break the loop. Side chains are sampled together with the backbone in a hierarchical way, making it possible to make large moves that cross energy barriers. As an initial application, we apply the method to the flexible loop in triosephosphate isomerase that caps the active site and demonstrate that the resulting loop ensembles agree well with key observations from previous structural studies. We also demonstrate, with three other test cases, the ability to distinguish relatively flexible and rigid loops within the same protein.

1. INTRODUCTION

A great deal of effort has been directed toward the development of computational methods for predicting the conformations of protein loops, which is a critical task in comparative protein modeling and in computational protein design.[1−4] The success of these methods has been evaluated primarily by comparing the results of the loop predictions with the loop conformations observed in crystal structures. That is, the focus is predicting the structure of the loop—a specific conformation—rather than the ensemble of conformations populated under biologically relevant conditions. Although these loop prediction methods can be used to identify multiple low-energy conformations, it is challenging to determine populations of the conformations, i.e., to relate energies of individual conformations to free energies of micro- or macrostates in the ensemble, although significant progress in this regard has been made by Meirovitch and co-workers.[5−7]

The flexibility of loops, i.e., the ability to adopt multiple conformations at relevant temperatures, is often critical to biological function, by playing an important role in molecular recognition. For example, the active site loop of the triosephosphate isomerase (TIM barrel) changes its conformation from an open to a closed state after binding of the ligands.[8,9] In kinases, two critical loops near the active site are flexible, with important implications for drug discovery: the glycine-rich loop (also called the P-loop) and the activation loop, including the DFG motif, which can adopt at least two major conformations in some kinases, referred to as "out" and "in". For example, while c-Src generally adopts the DFG-in conformation, the unfavorable DFG-out conformation can be induced by binding small molecules.[10] Loop flexibility can also play an important role in antibody−antigen recognition. The H3 loop in the complementarity-determining region of antibodies, which has the most diversity in sequence and is the most critical loop for antigen affinity and specificity, frequently demonstrates evidence of conformational flexibility.[11−13]

More broadly, there are many cases where loops adopt different conformations in different crystal structures, e.g., holo vs apo, or even different crystal unit cells for the same protein.[14] Although the B factors in crystal structures provide some information about conformational flexibility, each structure is best viewed as a snapshot from the equilibrium ensemble. NMR experiments can provide some direct information about conformational equilibria but generally cannot provide complete information about the ensemble of interconverting structures.

Molecular dynamics (MD) has been widely used to study protein flexibility, including loop dynamics.[15,16] The main liability of MD is that the time scales for interconverting between loop conformations can be long relative to the femtosecond time steps used, such as the millisecond time scale for the TIM capping loop to interconvert between the open and closed states.[17] Although such time scales may soon become accessible by MD simulation, they will remain extremely computationally expensive. Methods like replica exchange MD can be used to accelerate convergence but are likewise computationally expensive.

Here, we describe a Monte Carlo method for generating ensembles of loop conformations and cyclic peptides. It is related to classes of loop prediction methods that use torsion-angle sampling of backbone and side chain degrees of freedom (DoF), which makes it possible to make large conformational moves that cross energy barriers. Specifically, it builds on loop prediction methods that exploit "inverse kinematics" methods for creating move sets that do not "break" the loop.[18−24] The new contribution here is implementing these moves in a Monte Carlo scheme that also samples side chain DoF.[25] We apply the method to a number of proteins with flexible loops, including the well-known case of TIM. We also evaluate our ability to distinguish between (relatively) rigid and flexible loops within the same protein.

Received: November 18, 2010
Published: April 05, 2011

2. THE MOVE SET: TORSIONAL PERTURBATIONS VIA INVERSE KINEMATICS

2.1. Torsions and Sterics. It is widely accepted that the essential dynamics of a protein backbone can be captured by moves involving only the torsions ϕ,ψ with the other internal variables (bond lengths, bond angles, and ω torsions) being kept close to their canonical values, although not necessarily rigid.[19,23,24]

Compared to the high energy associated with ω angle deformation, ϕ and ψ angles are relatively free to rotate, but their range is restricted by steric interactions. Ramachandran regions in the (ϕ,ψ) coordinates for each peptide ensure intrapeptide steric avoidance, and additional restrictions are imposed by more distant clashes. Clashes involving backbone atoms (or atoms bonded to them) are completely determined from the backbone angles. On the other hand, atoms further along side chains (from the γ position out) are not completely determined from the backbone, although their placement may be restricted by it. Significantly, side chains may interact with other side chains so that their placement must be accomplished as a whole. Given a backbone conformation, a separate search is required to determine sterically acceptable or otherwise energetically viable side chain conformations. Reciprocally, backbone moves may be restricted by fixed side chain geometry.

2.2. MC Move and State Variables. To design a Monte Carlo move for reversibly exploring the torsion space, we must herefore consider the state space as the set of all torsions, $\{t_i;\chi_j\}$ where the t_i are backbone torsions and χ_j are side chain torsions, with the indices running respectively over all of the backbone and side chain DoF. A chain of {N,Cα,C} triplets (a standard backbone) is one possibility, but chains through, e.g., cysteine bridges, or other macromolecules, such as nucleic acids, could also be considered. In the following, we will assume the standard case (protein backbone loops) exclusively. For the case of a loop of N residues bridging two fixed ends, the essential backbone DoF would be $M = 2N - 6$. Here, six backbone DoF are involved in placing the end of the loop in a fixed rotation/translation relationship to the beginning. We call these DoF, labeled arbitrarily as t_i (i = 1, ..., 6), the compensators. The remaining M DoF, labeled as t_i, (i = 7, ..., 2N), are the controls. This separation in controls and compensators is arbitrary and may change from one move to the next. We could assume that the end residues 0 and $N + 1$ act as hinges; i.e., the ϕ_0 and ψ_{N+1} torsions are fixed, but ψ_0,ϕ_{N+1} are free, adding two DoF to the backbone. The treatment is essentially the same, replacing M by $M + 2$ and redefining some indices. We will only discuss the first case (no hinge mobility). It will be assumed that there are K side chain DoF in the set $\mathcal{S}$ of side chains interacting with the loop; we may only wish to include in $\mathcal{S}$ those side chains on the loop and hinges. The placement for those depends on the loop conformation. We may also include side chains on residues in some sphere of influence about the loop. Or we may simply include all of the side chains in the protein. We make no distinction at this stage.

Then, to design a reversible MC move that involves only the loop backbone DoF as well as the selected group $\mathcal{S}$ of side chains coupled to the loop, we must establish the Metropolis criterion for acceptance of a move of the form

$$\{t_i, \chi_j\} \rightarrow \{t_i + \delta t_i; \chi_j + \delta\chi_j\}, i \in [7, 2N], j \in [1, K] \qquad (1)$$

The shape space geometry accessible via our formulation characterizes our moves: assume that the L (= 2N) torsions for a loop kinematic chain are divided into the $L - 6$ controls and 6 compensators. The method used here employs the ϕ,ψ pairs of three amino acids (the *pivots*). These can be chosen at arbitrary locations along the loop, breaking it into three subfragments for kinematic purposes. To each value of the $L - 6$ controls there correspond up to 16 distinct conformations satisfying the closure conditions, each characterized by a unique set of values of the compensators. As discussed in our earlier work,[26] the 16 alternative solutions represent different orientations of the three subfragments between successive pivots in a reference frame attached to the three pivot Cα atoms about the three axes joining each pair of pivots. Thus, we refer to the method as Triaxial Loop Closure (TLC). The basic idea in the TLC method (discussed more in detail in the next section) is to construct a loop with arbitrary internal degrees of freedom, taking advantage of the fact that the inverse kinematic problem can be solved by determining appropriate values of six torsions. Thus, any variation in the remaining DoF's—other torsions, including Ω's, bond angles, and even bond lengths—can be considered, if so desired. Here, we treated only $\phi-\psi$ variations, as these are the most "flexible" DoF's, but we could have included all other DoF's in the MC scheme in any combination desired. The conformational variability of the constitutive pieces for loop closure, i.e., the three subfragments, is of course an important factor for solving the closure problem. We see that this variability can be decomposed into two types: the end-to-end variability of the individual fragments and the inherent variability of the loop closure problem, i.e., relative locations and orientations of the ends of the loop as well as the environment in the loop vicinity.

The first is a direct problem: compute the fragment (in practice, we do not check that the fragment is indeed sterically feasible until the assembly is successful). The individual fragment assembly, being subject to no end constraints, is only limited by the Ramachandran and other steric restrictions. However, for purposes of assembling the three subfragments into a self-consistent loop, each individual fragment of length L_i residues with i = 1−3, is encoded by four variables: the overall geometric length of the virtual bond joining first and last atoms, d_i; the angles θ_i and ξ_i made by the two end bonds to the virtual bond; and the torsion of the two end bonds about the virtual bond, δ_i. The variability of the closure problem is governed by these 12 parameters (d_i, θ_i, ξ_i, δ_i; i = 1−3). The equations expressing closure depend on these parameters smoothly; small changes cause usually small changes in the number and disposition of solutions except that, for certain arrangements, solutions could spontaneously appear or disappear (pairs of polynomial roots may join and become complex, or the converse, see the discussion of the inverse kinematic problem below).

We now search the nearby conformation space by perturbing one of the control torsions. This will result in perturbing the overall structure of one of the chains, leading to a perturbed set of solutions. These changes may lead to overall large motions, see e.g. ref 27 for a discussion of the end conditions and their constraining of various inner DoF. However, a reasonable acceptance ratio for the method can be more or less guaranteed by varying the controls and restricting the step size. Below, we discuss a two-stage scheme, splitting the move into a pure backbone and a pure side chain stage.

2.3. Solving the Inverse Kinematic Problem. Many methods for finding solutions that satisfy the closure conditions have been proposed, both exact[18,22,26,28−32] and approximate.[6,21,33−37] Exact methods address the inverse kinematic problem by searching

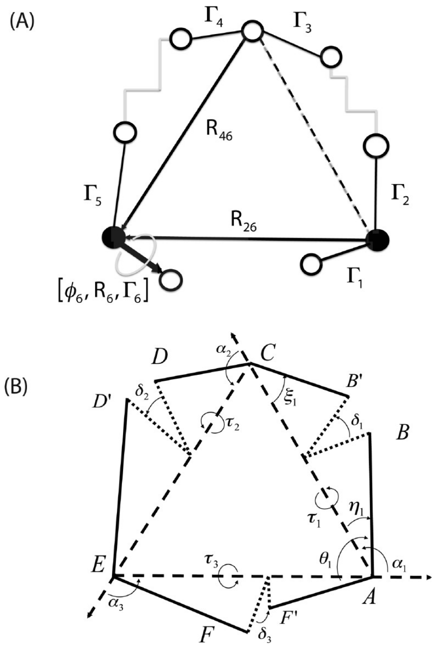


Figure 1. (A) The atoms and parameters defining triaxial loop closure (TLC). (B) The generalized 6R/3A kinematic chain.

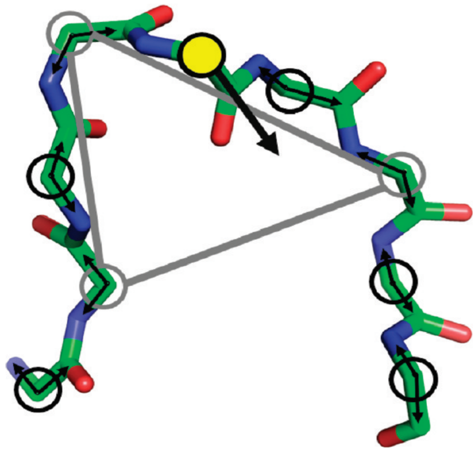


Figure 2. Construction of a tripeptide move. A node consists of a ϕ/ψ pair at each α carbon of the loop (with only backbone shown). The yellow filled circle is the α carbon, whose dihedral angle serves as a driver angle (the wide black arrow). A randomly constructed triaxial closure is shown as the gray triangle in which each gray circle represents the randomly selected pivot.

for the values of a certain torsion, say τ, in terms of which all other torsions can be determined. Go and Scheraga[18] pursued a direct solution in the original angle variables. This involves finding the zeros of a certain transcendental expression, a process that may require substantial computation to adequately resolve the entire domain. Subsequent works employ standard techniques from the robotics literature to convert to a more tractable polynomial form in the variable $u = \tan \tau/2$. All of the real roots of this 16th degree polynomial can be found efficiently and stably by the use of the method of Sturm chains.[38] All other torsions can be recovered readily, and therefore such methods are capable of finding all backbone solutions for any given combination of control torsion values. On the other hand, approximate methods typically use an iterative procedure to find a solution. As a result, they are not guaranteed to find all solutions consistent with a given set of control values, and the same is true for the approach in ref 18, which is also followed in refs 20, 23, and 24, although for this class of methods the issues are mainly related to the computational sensitivity of multiple roots.

In previous applications the *conrot* algorithm has been used.[20] It places the rotatable bonds on six consecutive bonds plus a driver. A generalization by Wu and Deem[22] uses one driver on either end. A weakness of the *conrot* approach is that a change on either side of the short compensator segment may make the closure problem unsolvable.[24] A generalization from robotics removes that restriction.[29] Our own method for solving the tripeptide closure problem, explained in detail in ref 26, has the advantage of mathematical simplicity, speed, and robustness. It also allows for a straightforward generalization for longer chains of arbitrary geometry. Its simplicity comes from taking advantage of the natural pairing up of rotatable bonds in amino acids to reduce the closure problem to three rotations, and we refer to this as the TLC method.[26] Referring to Figure 1b, we note that each Cα,C,N,Cα unit is identified by four variables: the overall geometric length of the virtual bond joining first and last atoms, d_i; the angles θ_i and ξ_i made by the two end bonds to the virtual bond; and the torsion of the two end bonds about the virtual bond, δ_i (actually, the formulation uses the angles α_i of the triangle formed with edges d_i). These definitions remain unchanged even if an arbitrary structure exists between the two end pairs (Figure 1a). We may produce multiple conformations for a long closed chain by partitioning into three subsegments and mapping each to a simple kinematic generalization of the tetrad Cα,C,N,Cα (Figure 1a,b).

In brief, three Cα atoms are selected (the pivots). The chain between any two of these, containing L atoms including the end points, is determined to within a rotation/translation (i.e., in its own body frame) by its own internal coordinates: $L - 3$ torsions, $L - 2$ angles, $L - 1$ lengths. With fixed (to any prescribed value) bond lengths and bond angles, each chain can be completely described by its $L - 3$ internal torsions. Below, we will index the residues of the three pivots as 1, 2, and 3, and we will index their backbone atoms as N_i, $C\alpha_i$, and C_i, $i = 1-3$, accordingly. Below, we use the atom names interchangeably with their Cartesian coordinates; e.g., N_1 can be thought of as equivalent to the vector $\mathbf{R}_1$ etc (see eq 5).

As is explained in ref 26 and somewhat more at length in ref 39 (see also the Supporting Information discussion in ref 40), the three fragments, respectively between pivots 1−2, 2−3, and 3−1, form a triangle with edges d_i, $i = 1-3$. The parameters necessary for setting up and solving the TLC equations can be extracted from knowledge of only the first two and last two atoms of each chain (Figure 2). Once the three four-atom fragments have been assembled into a triangle, the relative rotation of each fragment about the triangle must place the end atoms relative to those on each neighboring fragment so that the angles ($N_iC\alpha_iC_i$, $i = 1-3$) assume prescribed values (Figure 1). In this way, loop closure is accomplished when an appropriate rotation for each piece has been found. It turns out that the problem overlays the solution of a 16th degree polynomial, so that to each real root there corresponds a possible backbone loop geometry (subject, of course, to overall steric viability) to a total of, at most, 16 solutions possible for a given collection of state variables, the control $2N - 6$ torsions.

2.4. Jacobian.

Since fixing the end of the chain (the *Closure Conditions*) implies relationships among the torsions, we seek solution of these relationships such that specifying M torsions along the loop leads to complete determination of all $2N$ torsions and unambiguous Cartesian coordinates for all loop backbone atoms that are sterically self-consistent. In general, for any feasible value of the controls, there may exist multiple sets of compensators that allow the loop to close. They are functions of the controls, and their values solve the loop closure problem.

As a result, the element of volume in torsion space, initially uniform in these variables

$$d\mathcal{V} = d\phi_1 d\psi_1 \ldots d\phi_N d\psi_N d\chi_1 \ldots d\chi_K$$

will need to be modified by

$$dt_1 \ldots dt_6 = \frac{\partial(t_1, \ldots, t_6)}{\partial(R_6, \Gamma_6, t_6)} dR_6\, d\Gamma_6\, dt_6$$

leading to the well-known expression (e.g., see formula 23 in ref 23) for the inverse of the above Jacobian:

$$J_i = \det \frac{\partial(\mathbf{R}_6, \Gamma_6, t_6)}{\partial t}$$

$$= \begin{vmatrix} \frac{\partial \mathbf{R}_6}{\partial t_1} & \frac{\partial \mathbf{R}_6}{\partial t_2} & \frac{\partial \mathbf{R}_6}{\partial t_3} & \frac{\partial \mathbf{R}_6}{\partial t_4} & \frac{\partial \mathbf{R}_6}{\partial t_5} & \frac{\partial \mathbf{R}_6}{\partial t_6} \\ \frac{\partial \Gamma_6}{\partial t_1}\cdot \mathrm{e}_1 & \frac{\partial \Gamma_6}{\partial t_2}\cdot \mathrm{e}_1 & \frac{\partial \Gamma_6}{\partial t_3}\cdot \mathrm{e}_1 & \frac{\partial \Gamma_6}{\partial t_4}\cdot \mathrm{e}_1 & \frac{\partial \Gamma_6}{\partial t_5}\cdot \mathrm{e}_1 & \frac{\partial \Gamma_6}{\partial t_6}\cdot \mathrm{e}_1 \\ \frac{\partial \Gamma_6}{\partial t_1}\cdot \mathrm{e}_2 & \frac{\partial \Gamma_6}{\partial t_2}\cdot \mathrm{e}_2 & \frac{\partial \Gamma_6}{\partial t_3}\cdot \mathrm{e}_2 & \frac{\partial \Gamma_6}{\partial t_4}\cdot \mathrm{e}_2 & \frac{\partial \Gamma_6}{\partial t_5}\cdot \mathrm{e}_2 & \frac{\partial \Gamma_6}{\partial t_6}\cdot \mathrm{e}_2 \\ \frac{\partial t_6}{\partial t_1} & \frac{\partial t_6}{\partial t_2} & \frac{\partial t_6}{\partial t_3} & \frac{\partial t_6}{\partial t_4} & \frac{\partial t_6}{\partial t_5} & \frac{\partial t_6}{\partial t_6} \end{vmatrix}$$

Since

$$\frac{\partial \mathbf{R}_k}{\partial t_j} = \Gamma_j \times \mathbf{R}_{jk}, \frac{\partial \Gamma_6}{\partial t_j} = \Gamma_j \times \Gamma_6, \frac{\partial t_i}{\partial t_j} = \delta_{ij} \tag{2}$$

this Jacobian can assume the simpler, 5 × 5 form

$$\begin{aligned} J_i : &= \mathbf{J}(\mathbf{R}_6, \Gamma_6, t_6; t_1, \ldots, t_6) \\ &= \begin{vmatrix} \Gamma_1 \times \mathbf{R}_{16} & \Gamma_2 \times \mathbf{R}_{26} & \Gamma_3 \times \mathbf{R}_{36} & \Gamma_4 \times \mathbf{R}_{46} & 0 \\ (\Gamma_1 \times \Gamma_6)\cdot \mathrm{e}_1 & (\Gamma_2 \times \Gamma_6)\cdot \mathrm{e}_1 & (\Gamma_3 \times \Gamma_6)\cdot \mathrm{e}_1 & (\Gamma_4 \times \Gamma_6)\cdot \mathrm{e}_1 & (\Gamma_5 \times \Gamma_6)\cdot \mathrm{e}_1 \\ (\Gamma_1 \times \Gamma_6)\cdot \mathrm{e}_2 & (\Gamma_2 \times \Gamma_6)\cdot \mathrm{e}_2 & (\Gamma_3 \times \Gamma_6)\cdot \mathrm{e}_2 & (\Gamma_4 \times \Gamma_6)\cdot \mathrm{e}_2 & (\Gamma_5 \times \Gamma_6)\cdot \mathrm{e}_2 \end{vmatrix} \end{aligned} \tag{3}$$

Here

$$\mathbf{R}_{ij} = \mathbf{R}_j - \mathbf{R}_i, \Gamma_i = \frac{\mathbf{R}_i^{'} - \mathbf{R}_i}{||\mathbf{R}_i^{'} - \mathbf{R}_i||} \tag{4}$$

and e_i, $i = 1-3$, are the usual unit vectors along axes x, y, and z of an arbitrary reference frame (the *Lab frame*). The atoms associated with closure are

$$\mathbf{R}_{2k-1} = N_k, \mathbf{R}_{2k} = C\alpha_k(= \mathbf{R}_{2k-1}^{'}), \mathbf{R}_{2k}^{'} = C_k; \quad k = 1, 2, 3 \tag{5}$$

We note that the term $\Gamma_5 \times \mathbf{R}_{56} = 0$ and was omitted. In the general case, the three pivot residues are indexed by $1 \le n_1 < n_2 < n_3 \le N$, and this reindexing will be implied where appropriate.

It is well-known[22] that the Jacobian in the form first proposed by Dodd et al.[20] is incomplete and lacks frame invariance. In a rigorous derivation of the Jacobian from the configuration integral, Wu and Deem[22] show that the correct, frame invariant form is

$$J^{-1} = \frac{1}{\Gamma_6 \cdot \mathrm{e}_3} J_i \tag{6}$$

However, since the acceptance criterion involves ratios of Jacobians computed at the same frame, the additional factors cancel and the relative probabilities remain unchanged.

Although the latter form 6 is indeed invariant if all vectors are changed by an arbitrary affine transformation, it has the undesirable feature that it involves a projection to an arbitrary frame. Consequently, the factor $\Gamma_6 \cdot \mathrm{e}_3$ may accidentally vanish (in which case J_i will also vanish), necessitating a random reorientation of the frame to break the degeneracy. Thus, it is desirable to eliminate this superfluous dependence and derive a form that depends only on intrinsic (body frame) coordinates, for which invariance is easily seen. This can be accomplished by carrying out an expansion of this determinant in complementary minors; indeed, the top three rows are expressed in terms of intrinsic coordinates, while the last two involve projections to the space frame. We thus expand the determinant as

$$J_i = \sum_{i=1}^{4} (-1)^i \begin{vmatrix} (\Gamma_i \times \Gamma_6)\cdot \mathrm{e}_1 & (\Gamma_5 \times \Gamma_6)\cdot \mathrm{e}_1 \\ (\Gamma_i \times \Gamma_6)\cdot \mathrm{e}_2 & (\Gamma_5 \times \Gamma_6)\cdot \mathrm{e}_2 \end{vmatrix} \left| \Gamma_j \times \mathbf{R}_{j6} \quad \Gamma_k \times \mathbf{R}_{k6} \quad \Gamma_l \times \mathbf{R}_{l6} \right| \tag{7}$$

where the indices (i,j,k,l) are a cyclic permutation of (1,2,3,4).

Applying the well-known identity (e.g., in ref 41, eq 25, p.76)

$$\begin{vmatrix} \mathbf{A}\cdot\mathbf{C} & \mathbf{B}\cdot\mathbf{C} \\ \mathbf{A}\cdot\mathbf{D} & \mathbf{B}\cdot\mathbf{D} \end{vmatrix} = \mathbf{A}\cdot\mathbf{C}\mathbf{B}\cdot\mathbf{D} - \mathbf{B}\cdot\mathbf{C}\mathbf{A}\cdot\mathbf{D} = (\mathbf{A} \times \mathbf{B})\cdot(\mathbf{C} \times \mathbf{D}) \tag{8}$$

to the first of the 2 × 2 minors in eq 7, we have

$$\begin{vmatrix} (\Gamma_1 \times \Gamma_6)\cdot \mathrm{e}_1 & (\Gamma_5 \times \Gamma_6)\cdot \mathrm{e}_1 \\ (\Gamma_1 \times \Gamma_6)\cdot \mathrm{e}_2 & (\Gamma_5 \times \Gamma_6)\cdot \mathrm{e}_2 \end{vmatrix} = (\Gamma_1 \times \Gamma_6) \times (\Gamma_5 \times \Gamma_6)\cdot \mathrm{e}_3 = (\Gamma_1 \cdot \Gamma_5 \times \Gamma_6)(\Gamma_6 \cdot \mathrm{e}_3)$$

The remaining 2 × 2 minors result in analogous expressions. Substituting these into eq 7, we have

$$\frac{J_i}{\Gamma_6 \cdot \mathrm{e}_3} = \sum_{i=1}^{4} (-1)^i (\Gamma_i \cdot \Gamma_5 \times \Gamma_6) \left| \Gamma_j \times \mathbf{R}_{j6} \quad \Gamma_k \times \mathbf{R}_{k6} \quad \Gamma_l \times \mathbf{R}_{l6} \right| \tag{10}$$

(as above, the indices (i,j,k,l) are a cyclic permutation of (1,2,3,4)), which can be recombined to give the expression for the inverse Jacobian

$$J^{-1} = \frac{1}{\Gamma_6 \cdot \mathrm{e}_3}$$

$$J_i = \begin{vmatrix} \Gamma_1 \times \mathbf{R}_{26} & \Gamma_2 \times \mathbf{R}_{26} & \Gamma_3 \times \mathbf{R}_{46} & \Gamma_4 \times \mathbf{R}_{46} \\ (\Gamma_1 \cdot \Gamma_5 \times \Gamma_6) & (\Gamma_2 \cdot \Gamma_5 \times \Gamma_6) & (\Gamma_3 \cdot \Gamma_5 \times \Gamma_6) & (\Gamma_4 \cdot \Gamma_5 \times \Gamma_6) \end{vmatrix} \tag{11}$$

where we took advantage of the fact that $\Gamma_i \times \mathbf{R}_{i6} = \Gamma_i \times \mathbf{R}_{i+1,6}$ with $i = 1$ and 3 due to the fact that the axes Γ_i and Γ_{i+1}, $i = 1$ or 3, are coterminal. Figure 1a shows all quantities that enter in the Jacobian.

This 4 × 4 determinant is the frame invariant form of the inverse Jacobian for the TLC method. It has the advantage that it is expressed entirely in terms of body coordinates, and thus it is free from degeneracies and can be evaluated without projecting to an ad hoc coordinate system. It is numerically equivalent to the Wu and Deem form 6, when the latter is defined. The Jacobian 11 can be easily expressed in terms of the intrinsic parameters $(d_i,\theta_i,\xi_i,\delta_i)$, $i = 1-3$, entering in the TLC algorithm,[42] a feature that it shares with reduced Jacobians derived by other authors.[22,43] However, such expressions lack the simplicity and geometrical appeal of eq 11.

2.5. Backbone Perturbation Procedure. The loop closure algorithm described in the previous section, while perfectly general, is currently implemented as a strategy for perturbing only the backbone coordinates. The side chain coordinates perturbation procedure, as well as the strategy for combining these perturbations in a way such that detailed balance is maintained, will be outlined in the next two sections. An important design feature of this approach is that the backbone and side chain perturbations are generated independently.

An important feature of both the backbone selection probability and the side chain selection probability is that they are reversible, or

$$\alpha(\mathbf{t} \rightarrow \mathbf{t}') = \alpha(\mathbf{t}' \rightarrow \mathbf{t}) \tag{12}$$

where $\mathbf{t}' = \mathbf{t} + \delta\mathbf{t}$ is the trial move starting from the torsion state $\mathbf{t}$ and $\delta\mathbf{t}$ is the perturbation vector to the loop of interest. For the purposes of this work, we require the selection probability to be uniform to enforce eq 12. For this to be true, we need to establish the procedure which ensures that a uniform distribution of torsions over the entire loop can be generated.

The procedure for generating a trial move $\delta\mathbf{t}$ closely follows that of refs 20, 22, 23, and 29. Since the algorithm currently solves for $2N - 6$ torsions, and we wish to have a procedure that is valid for loops of arbitrary length, we must select a subset of $2N - 6$ torsions. There is some flexibility in how this could be done, but the present implementation is as follows (see Figure 2):

(1) From the designated loop torsions, a single torsion angle i is selected uniformly and identified as a driver angle coordinate (the yellow circle in Figure 2), as has been described in previous work.[26]

(2) For torsion t_i, a random variate δt_i is generated, with a maximum value of up to π.

(3) A randomly constructed triaxial closure is generated by randomly selecting three α carbons as pivots from the loop (excluding the α carbon on which the driver angle resides) and assigning the ϕ/ψ angles as the torsions (the gray triangle in Figure 2).

(4) A set of torsions for the stationary solution $\mathbf{t}_k$, $k \in [1, K]$, is generated, resulting in up to $K = 16$ solutions. For this case, only the alternative sets of pivot coordinates are considered, with the driver angle held at t_i. For each solution, a Jacobian term $J(\mathbf{t}_k)$ is computed.

(5) A set of torsions for the perturbed solution $\mathbf{t}_l$, $l \in [1, L]$, is similarly generated, with associated $J(\mathbf{t}_l)$ terms.

(6) A trial solution $\mathbf{t}'$ is selected from the solutions $(\mathbf{t}_k,\mathbf{t}_l)$ with the following probability:

$$\alpha(\mathbf{t} \rightarrow \mathbf{t}') = \frac{J(\mathbf{t}')}{\sum_{k=1}^{K} J(\mathbf{t}_k) + \sum_{l=1}^{L} J(\mathbf{t}_l)} \tag{13}$$

To show that this procedure generates a uniform distribution, the ϕ/ψ angles of an 11-residue polypeptide is sampled with no potential. Half of the time, the loop closure procedure is applied as described above, and the other half of the time, only a driver angle is perturbed uniformly, with the remaining Cartesians updated accordingly (with no closure condition enforced). The second procedure is required so that the full space of dihedral angles is accessible. Every move is accepted, with no potential applied or steric exclusion. This procedure generates a uniform distribution of torsions, as is shown in Figure 3. It shows a distribution of an 11-residue peptide sampled with the loop closure procedure described above. Only backbone DoF are sampled, and no force field is applied. The end points are constrained to fixed positions. This control closely follows previous work.[20,23] Figure 3a shows the distribution of angles with no Jacobian selection term applied, and Figure 3b shows the distribution with the reweighting term applied. The Jacobian term clearly improves the uniformity of the sampling.

2.6. Side Chains. The efficient sampling of side chains[22] is important since side chain conformations often determine the biological function of proteins. In the current work, the side chain χ angles are not taken from the rotamer library due to their nonuniform distribution. Instead, to generate the side chain trial moves, a single side chain is randomly selected, and each χ angle is perturbed by a value which is randomly and uniformly distributed in a defined domain $[-d/2,d/2]$.[25,44] The polar hydrogens for the selected residue are sampled as well over the domain $[-\pi,\pi]$.

To improve the sampling efficiency, no energy is computed for the states with steric clashes, which are defined on the basis of the distances between heavy atoms. Specifically, a steric clash is defined when pairs of heavy atoms are closer than 0.7 times the sum of their Lennard-Jones radii. Rapid identification of steric clashes (using neighbor lists) avoids computationally expensive energy evaluations, for conformations that will result in very high energies and negligible acceptance probabilities.

The most expensive term in energy evaluation is the solvation energy in which the time-consuming step is the computation of Born radii. Since the Born radii and the long-range energy terms generally vary slowly for relatively small, local conformational changes, less frequent evaluation of these terms will contribute more to the sampling performance. For this purpose, the multiple time-step Monte Carlo sampling (MTSMC) procedure[45] is incorporated in the present method, in a scheme based on that in ref 44. The Born radii and the long-range interactions are held fixed at the latent state of the original coordinates during the

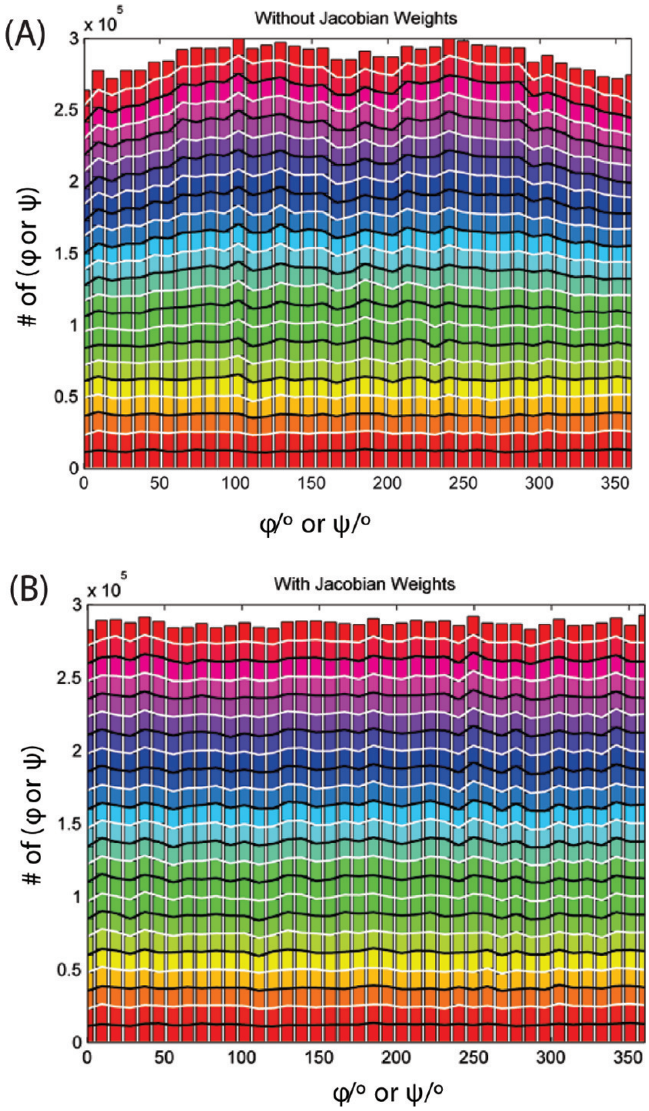


Figure 3. Distribution of ϕ/ψ angles without (A) and with (B) Jacobian weighting of selection for an 11-residue peptide. A total of 4.5×10^5 trial moves were generated. No force field is used in the selection probability, and all trial moves are accepted.

inner loop sampling and only updated in every outer loop calculation. The final configuration from the inner loop is then taken to be a trial move and subjected to the MTSMC acceptance criterion (see eq 20 in ref 25).

2.7. The POSH Monte Carlo Method. Both the TLC method for determining the backbone moves of loop residues and the side chain sampling via perturbation have been incorporated in the POSH (port out, starboard home) Monte Carlo method introduced in a previous work.[25] The application of this method on small peptide systems has shown reasonable agreement with experiments.[25] In the present work, we are interested in its performance in more complicated protein systems with flexible loops.

Briefly, the move sets in this approach consist of two steps: an initial trial (1 → 2) move with large perturbation followed by a series of annealing moves consisting of smaller perturbation within the inner loop of length N_I (2 → 3). The generalized Metropolis acceptance probability for this series of moves is given by

$$acc(1 \rightarrow 3) = \min\left(1, \frac{p_3 T_{41}}{p_1 T_{23}}\right) \qquad (14)$$

where p_1 and p_3 are the probabilities of being in the original and final trial state, respectively. T_{41} and T_{23} are transition probabilities. T_{23} is the normal forward transition probability, as would be given in the usual derivation of detailed balance, but T_{41} is a reverse transition probability that is constructed using an alternative reverse path through configuration space that is constructed by taking the final state (state 3) and subtracting the perturbation (1 → 2) from state 3 to arrive at state 4. Further details are given in ref 25.

The trial moves are generated by a perturbation that uniformly varies over some domain $[-d/2,d/2]$ with a different magnitude for the initial and annealing steps. In this work, for both types of

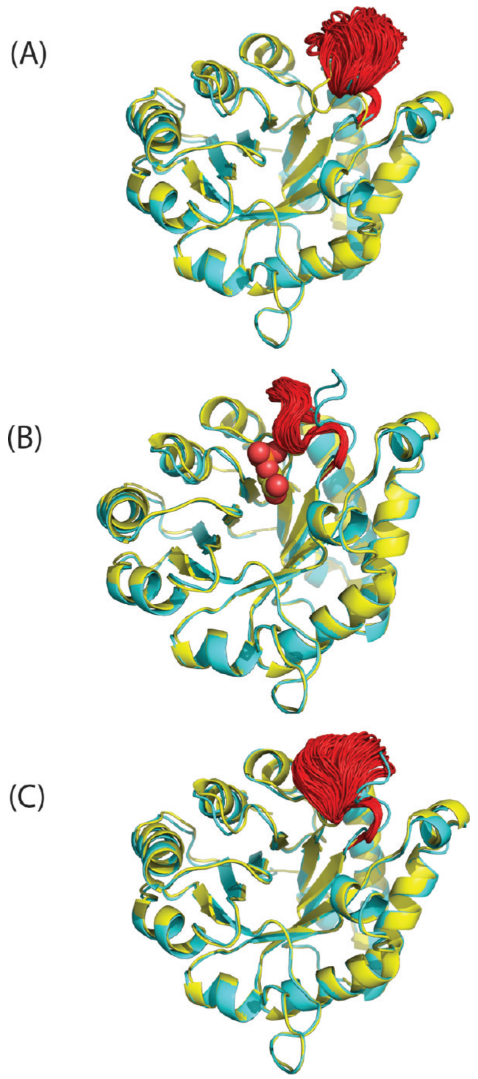


Figure 4. The ensemble structures (red) for the flexible loop (residues 165−178) of yeast TIM were taken from the equilibrium simulation with initial structures of (A) the apo (open) conformation, (B) the bound (closed) conformation, and (C) the closed conformation with the ligand PGA removed. The X-ray structure of apo yeast TIM (PDB 1YPI) is shown in yellow and the bound state (PDB 2YPI) in cyan. The ligand PGA is depicted by spheres.

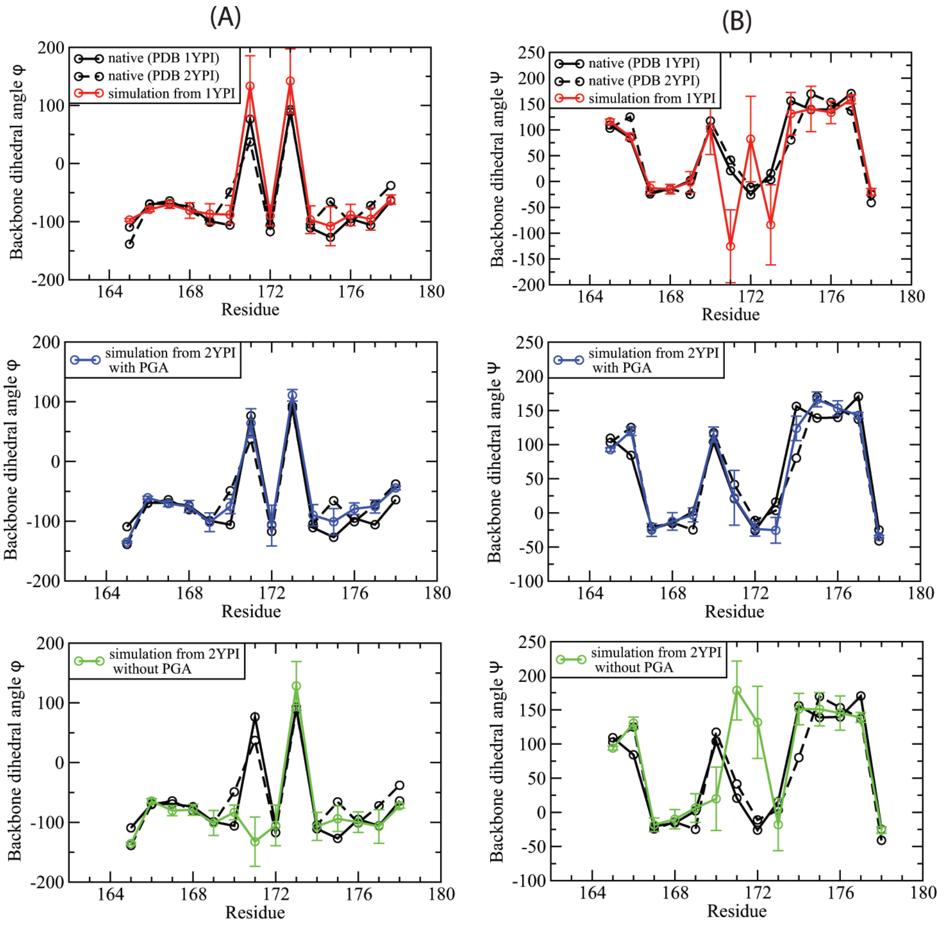


Figure 5. Comparison of the calculated backbone dihedral angles, ϕ (A) and ψ (B), with those measured in the X-ray structures. The black solid line is for apo TIM (PDB 1YPI) and the dashed line for the ligand-bound TIM (PDB 2YPI). The calculated dihedral angles were averaged over the equilibrium ensemble simulated from the initial structure of apo (red), ligand-bound (blue), and closed forms with the ligand PGA removed (green).

trial moves, either the backbone or side chain is allowed to be perturbed with equal probability. For backbone perturbations, the ϕ or ψ dihedral angle can vary over the domain of $[-2\pi,2\pi]$ for initial steps and $[-\pi/4,\pi/4]$ for annealing steps. For side chain χ angles, the domains are $[-\pi,\pi]$ and $[-\pi/9,\pi/9]$, respectively, for the initial and inner step trial moves. The number of inner steps N_I is set to 20, which was reported as the upper bound of inner steps for generating precise distribution. For all protein systems studied in this work, a mixture of 50% POSH and 50% standard MC sampling, followed by the MTSMC procedure, is used due to its better performance as studied in the previous work.[25]

3. SIMULATIONS

We applied the loop Monte Carlo method described above to several proteins with flexible loops. The first is the enzyme triosephosphate isomerase (TIM), which has been used as a model system for studying loop flexibility, primarily by NMR. This enzyme catalyzes the reversible isomerization of dihydroxyacetone phosphate (DHAP) to D-glyceraldehyde 3-phosphate (GAP). The active site loop 6 (residues 167–176) undergoes conformational changes upon ligand binding and is believed to be flexible in the absence of ligand binding, transitioning between "open" and "closed" states. To assess the capability of our method to capture the dynamical properties of this flexible loop, three sets of simulations were performed. The first one started from the apo yeast TIM (PDB ID 1YPI) with an open loop conformation (we call this SIM1). The second started from the 2-phosphoglycolate (PGA)-bound TIM (PDB ID 2YPI) with the closed loop conformation (SIM2), and the third is the same as the second except that the ligand PGA was removed from the initial structure (SIM3).

The titratable residues in the starting structures were predicted according to the experimental conditions. Specifically, in all simulations, His95 was treated as neutral and protonated on the N $\in$ 2. Glu165 is protonated in SIM2 in order to maintain the strong interaction with ligand PGA[9] but was unprotonated in the other simulations. Residues within 8 Å of the active site loop were included for the side chain sampling, and the flexible loop was extended to include residues 165–178 in the simulations for both the backbone and side chain sampling. The force field OPLS-AA[46,47] was used for the protein TIM and ligand PGA except that the partial charges for the phosphate group of PGA were adjusted on the basis of the previous work by Wong et al.[48] The surface generalized Born (SGB)[49,50] model was used for implicit solvent with the treatment of nonpolar terms.[50] To prevent the sampling from being trapped in local minima, all simulations were performed at a temperature of 600 K. Each simulation has a length of $N_o = 2 \times 10^5$ up to 5×10^5 outer steps. Data analyses were performed over

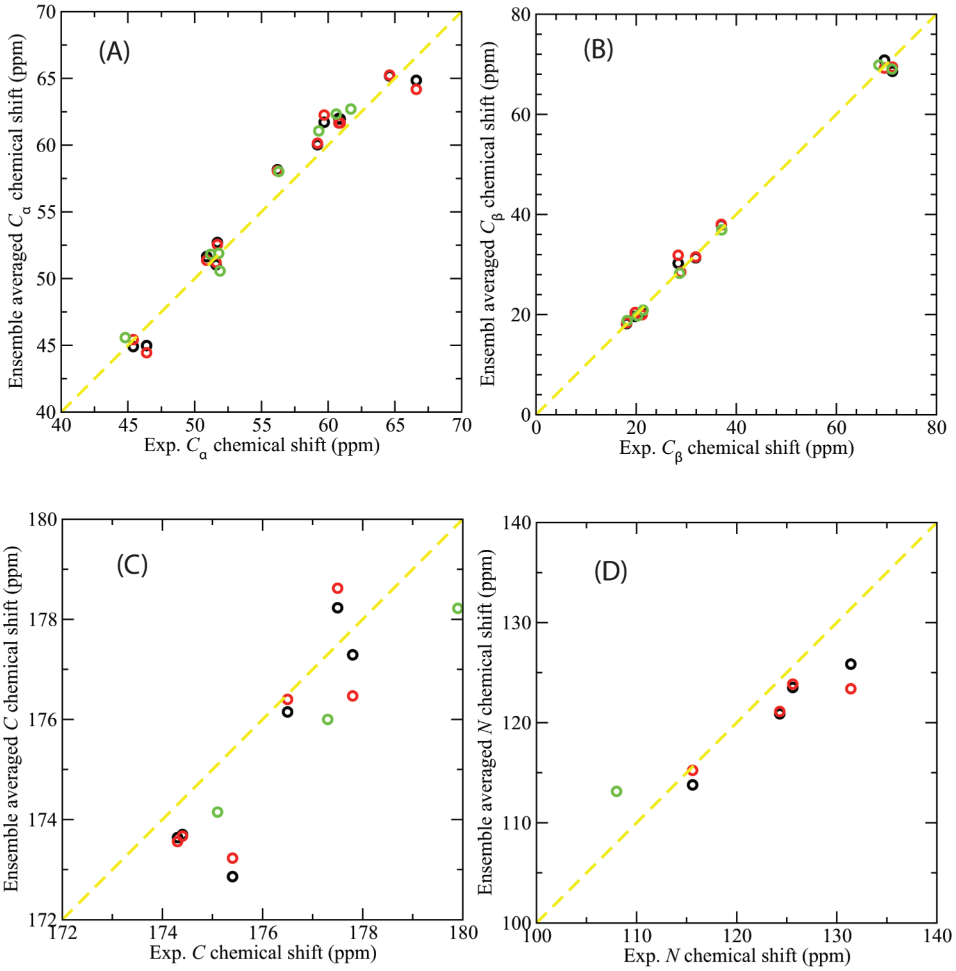


Figure 6. Ensemble-averaged chemical shifts (ppm) versus the NMR experimental measurements for C_α (A), C_β (B), carbonyl C (C), and amide N (D) atoms of the flexible loop 6 of yeast TIM. SHIFTX[56] was used to calculate chemical shifts, which were then averaged over an ensemble of 1000 structures from the equilibrated MC simulations. The starting PDB structures for the simulations are 1YPI (black), 2YPI with the ligand PGA removed (red), and 2YPI with PGA bound (green). The experimental chemical shift data are those for apo yeast TIM in the NMR experiment[57] (for comparison with the apo simulations), and for yeast TIM with ligand G3P[57] (for comparison with the holo simulation). Experimental chemical shifts are not available for some atoms, and these are omitted.

the equilibrium simulations (roughly after 10^5 outer steps) during which the potential energy is relatively stable.

The same protocol was also applied to other protein systems which have been studied by NMR experiments, specifically those with PDB ID 1H2O, 1XWE, and 1Q9P. By choosing NMR structures, we eliminate any concerns about crystal packing influencing the loop conformation or flexibility. These specific proteins were chosen because each has two loops consisting of 5−8 residues, one of which has multiple conformations with large variation among the various NMR models (flexible loop) and the other has a narrow range of loop conformations among the NMR models (rigid loop). Both the flexible and rigid loops were simulated using the same sampling protocol and the same parameter settings in order to compare with the experimental data since both loops within the same protein were measured under the same experimental conditions. The titratable residues in the starting structures were protonated at the experimental pH = 7.0 for 1H2O, 6.0 for 1XWE, and 5.8 for 1Q9P. The flexible loops consist of residues 59−64 for 1H2O, 1609−1616 for 1XWE, and 48−53 for 1Q9P; the residues in the rigid loops are 46−51 for 1H2O, 1536−1540 for 1XWE, and 78−82 for 1Q9P.

4. RESULTS AND DISCUSSION

As an initial illustration of the utility of our loop MC method for sampling the conformation space of protein loops, we applied this method to the well-studied enzyme triosephosphate isomerase (TIM). The active site loop undergoes large-scale motions interconverting between open and closed conformations. This conformational transition occurs on the time scale of milliseconds,[17] making it a challenge for molecular dynamics simulations in previous studies.[51,52]

In the current work, multiple transitions between open and closed loop conformations of yeast TIM have been observed in the simulation of the apo protein, but only at 600 K (vide infra). Figure 4a and c, which start from the open and closed state, respectively, show sampled loop conformations from the equilibrium ensemble, spanning both the open and closed form. In the simulation with the ligand PGA bound, the active site loop stays in the closed conformation, as can be seen in Figure 4b. These results agree qualitatively with NMR experiments, which found that the loop samples open and closed conformations whether or not a ligand was bound, but ligand binding shifted the

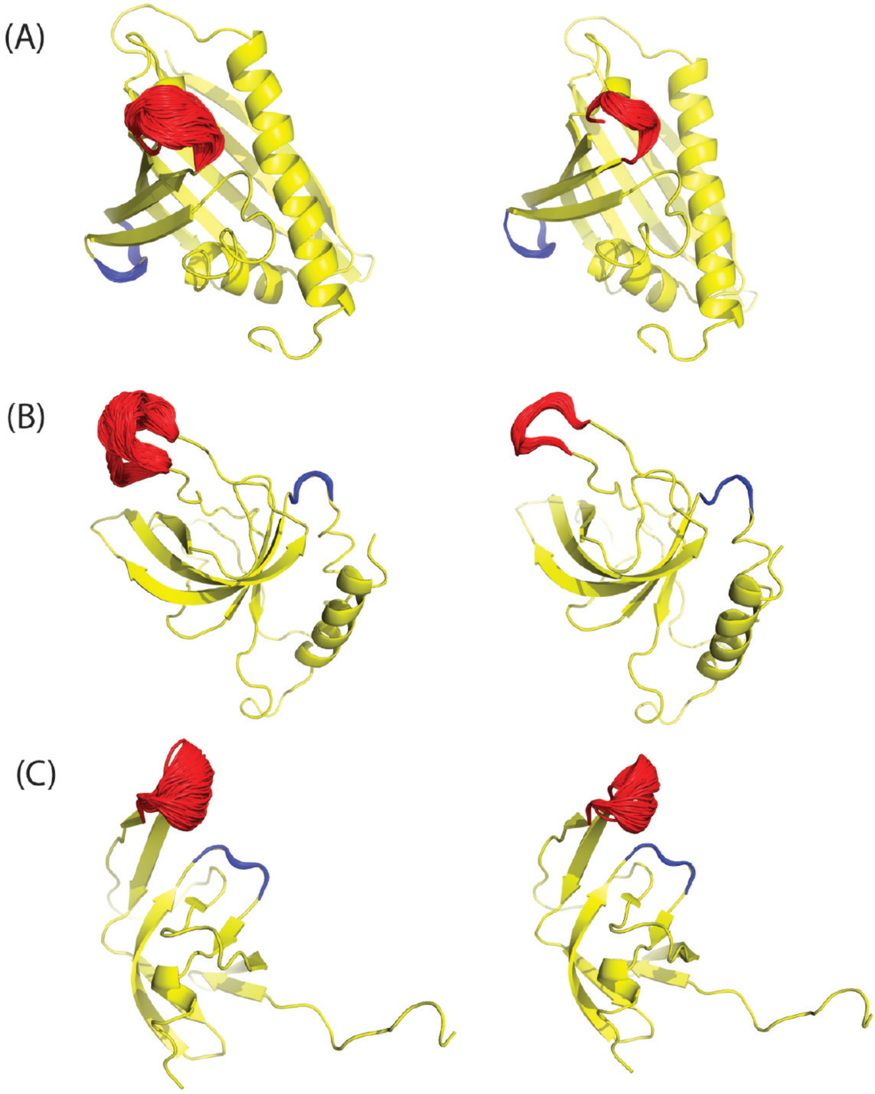


Figure 7. Ensembles of loop structures from equilibrium simulations using MC sampling for proteins with PDB ID (A) 1H2O, (B) 1XWE, and (C) 1Q9P sampled at T = 600 K (left) and T = 300 K (right). The sampled flexible loops ("floppy") which have a large fluctuation in the NMR models are shown in red, and the rigid loops with very small fluctuations are in blue. The structures in yellow are taken from MODEL 1 of the PDB file.

equilibrium strongly toward the closed conformation.[17,53] Upon PGA binding, the carboxylate of the ligand protonates residue Glu165, making it hydrogen bonded with PGA instead of with Ser96 in the apo structure, such that the closed loop conformation is preferred in the presence of a ligand.

It has been known that the active site loop of TIM moves largely as a rigid unit.[51,54] Figure 5 shows that the backbone dihedral angles of the flexible loop in the X-ray structure of apo TIM are very similar to those in the structure of ligand-bound TIM. The ensembles generated by the loop MC method largely agree with the experimental data in this regard. We calculated the backbone ϕ and ψ angles and averaged them over the equilibrium ensemble for each of the three simulations. For the holo simulations, the ensemble averaged ϕ and ψ angles agree well with those measured in the X-ray structures, as shown in Figure 5a and b (blue lines). Similar agreement was also found for the apo simulations started from both the open and closed conformations, except that residues 170−173 have relatively large deviations and fluctuations, which is consistent with the findings in previous simulation studies[17,52] (red and green lines in Figure 5a and b).

NMR spectroscopy can provide information on both the structure and dynamics of proteins in physiologically relevant environments.[55] The chemical shift is NMR's most ubiquitous parameter, the variation of nuclear magnetic resonance frequencies of the same kind of nucleus being due to variations in the electron distribution. To directly compare with the experimental data, ensemble averaged chemical shifts were calculated for each equilibrium ensemble using SHIFTX[56] to calculate chemical shifts for the residues of the flexible loop in each conformation and then averaging over all of the conformations in the ensemble. For the apo simulations, starting from either the open or closed structures, the ensemble-averaged chemical shifts were compared with NMR measurements of apo yeast TIM.[57] For the simulation of the ligand-bound, closed structure, NMR data measured for G3P-bound yeast TIM[57] were used. [The chemical shifts for the closed loop of the enzyme bounded with G3P and GPA are very similar (Yimin Xu, personal communication).]

Table 1. Root-Mean-Squared Fluctuation (RMSF; Å) of Heavy Atoms of Both Floppy and Rigid Loops in the Equilibrium Ensemble Simulated by POSH MC Method with Initial Structure of the First Model of NMR Structures[a]

	1H2O			1Q9P			1XWE		
		POSH			POSH			POSH	
heavy atom RMSF	NMR	600 K	300 K	NMR	600 K	300 K	NMR	600 K	300 K
flexible loop	2.75	1.64	0.75	3.51	1.27	1.06	4.83	2.50	1.10
rigid loop	0.49	0.40	0.15	1.25	0.38	0.18	1.03	0.50	0.31

[a] For comparison, the RMSFs over all NMR models for each protein are also computed at both 600 and 300 K.

A strong linear correlation was found between the ensemble-averaged and experimentally measured chemical shifts for C_α (Figure 6a) and C_β (Figure 6b) atoms with a correlation coefficient r of 0.98 or higher in all cases. For carbonyl C and amide N atoms of the flexible loop, although there are fewer experimental chemical shifts available, the calculated ensemble averages have small variations from experimental values (Figure 6c and d). The agreement with the NMR chemical shifts provides additional evidence that the ensembles generated by the loop MC sampling are reasonable.

We note that the experimental chemical shifts were measured at 300 K, while our simulations were performed at 600 K. This is because at 300 K it is difficult to observe the conformational transitions between the open and the closed state. We suspect, but cannot prove, that this occurs in part due to (1) the well-known tendency of generalized Born implicit solvent models to overstabilize salt bridges, (2) the effect of constraining the Ω angles, as well as the bond angles and lengths, in addition to the loop closure condition, and (3) sampling only the loop and not the remainder of the protein. Using a higher temperature overcomes all of these effects, and reasonable ensembles are generated which agree with the NMR chemical shifts. Because the Monte Carlo sampling scheme does not perturb degrees of freedom outside the loop, such that the overall structure is preserved, a higher temperature sampling protocol can still provide physical insights. The efficiency gained by sampling a lower dimensional space, while still obtaining a reasonable estimate of ensemble properties, motivates the use of this set of approximations.

As a second initial application, we also applied our sampling method to other protein structures, solved by NMR, which have loops with differing flexibilities in order to evaluate our ability to distinguish the flexible and rigid loops within the same protein. The conformational ensembles from equilibrium simulations for both the flexible and rigid loops are shown in Figure 7 for three proteins with PDB ID 1H2O (a), 1XWE (b), and 1Q9P (c) sampled at 600 K (left) and 300 K (right). These results clearly show that the loop residues which are flexible in the experimentally derived structures consistently are more floppy in the sampled ensemble at either temperature than the loop residues, which are relatively rigid in the same NMR structures. To further quantify these results, root-mean-square fluctuations (RMSF) of the heavy atoms in both loops were calculated for the sampled and NMR models, as shown in Table 1. We recognize that the set of NMR models for each protein cannot be viewed as a true ensemble, but the qualitative agreement is nonetheless encouraging. Thus, for studying protein loop flexibility, our method is a viable alternative to molecular dynamics simulations, which have also been used successfully to obtain ensembles in quantitative agreement with NMR data. In the cases examined here, the differences in rigidity appear to be related simply to the level of solvent exposure; i.e., floppy loops are more solvent exposed and have less interaction with their neighbors. For simulations of all studied protein systems, three NMR targets and TIM, the average acceptance ratio is about 14%.

Our current approach only varies $\phi-\psi$ angles, as they are most flexible, but actually it is possible to include all other DoF in the MC scheme in any desired combination. We are working on a further version of the algorithm that will incorporate sampling which allows Ω angles, as well as bond lengths and angles, to fluctuate more freely, which may allow for lower temperature sampling of systems of this type. Although in the present study we have considered solvation effects implicitly only, including water molecules explicitly in the simulation is possible in principle. However, water molecules in the immediate vicinity of a loop would lead to steric clashes whenever a large backbone move was attempted, which would reduce the efficiency of the present approach.

AUTHOR INFORMATION

Corresponding Author

*Tel.: (415) 514-9811. E-mail: Matt.Jacobson@ucsf.edu.

Author Contributions

[§]Jerome Nilmeier and Lan Hua contributed equally to this work.

ACKNOWLEDGMENT

This work was supported in part by grants from NIH-NIGMS, GM081710 (to M.P.J. and E.A.C.), GM086602 (to M.P.J.), and R01-GM090205 (E.A.C.). M.P.J. is a consultant to Schrodinger LLC.

REFERENCES

(1) Jones, D. *Curr. Opin. Struct. Biol.* **1997**, *7*, 377.
(2) Fiser, A.; Do, R.; Sali, A. *Protein Sci.* **2000**, *9*, 1753.
(3) Al-Lazikani, B.; Jung, J.; Xiang, Z.; Honig, B. *Curr. Opin. Struct. Biol.* **2001**, *5*, 51.
(4) Jacobson, M. P.; Pincus, D.; Rapp, C.; Day, T.; Honig, B.; Shaw, D.; Friesner, R. *Proteins* **2004**, *55*, 351.
(5) Meirovich, H. *Chem. Phys. Lett.* **1977**, *45*, 389.
(6) Baysal, C.; Meirovich, H. *J. Phys. Chem. A* **1997**, *101*, 2185.
(7) Mihailescu, M.; Meirovitch, H. *J. Phys. Chem. B* **2009**, *113*, 7950.
(8) Lolis, E.; Abler, T.; Davenport, R.; Rose, D.; Hartman, F.; Petsko, G. *Biochemistry* **1990**, *29*, 6609.
(9) Lolis, E.; Petsko, G. *Biochemistry* **1990**, *29*, 6619.
(10) Dar, A.; Lopez, M.; Shokat, K. *Chem. Biol.* **2008**, *15*, 1015.
(11) Padlan, E. *Adv. Protein Chem.* **1996**, *49*, 57.
(12) Xu, J.; Davis, M. *Immunity* **2000**, *13*, 37.
(13) Wong, S.; Jacobson, M. P. *Proteins* **2010**in press.
(14) Rapp, C.; Pollack, R. *Proteins* **2005**, *60*, 103.
(15) Wong, S.; Jacobson, M. P. *Proteins* **2008**, *71*, 153.
(16) Yi, M.; Tjong, H.; Zhou, H. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 8280.
(17) Massi, F.; Wang, C.; Palmer, A. G. *Biochemistry* **2006**, *45*, 10787.
(18) Go, N.; Scheraga, H. *Macromolecules* **1970**, *3*, 178.
(19) Bruccoleri, R. E.; Karplus, M. *Macromolecules* **1985**, *18*, 2767.
(20) Dodd, L. R.; Boone, T. D.; Theodorou, D. N. *Mol. Phys.* **1993**, *78*, 961.

(21) Deem, M.; Bader, J. *Mol. Phys.* **1996**, *87*, 1245.
(22) Wu, M. G.; Deem, M. W. *Mol. Phys.* **1999**, *97*, 559.
(23) Dinner, A. *J. Comput. Chem.* **2000**, *21*, 1132.
(24) Ulmschneider, J. P.; Jorgensen, W. L. *J. Chem. Phys.* **2003**, *118*, 4261.
(25) Nilmeier, J.; Jacobson, M. P. *J. Chem. Theory Comput.* **2009**, *5*, 1968.
(26) Coutsias, E. A.; Seok, C. L.; Jacobson, M. P.; Dill, K. A. *J. Comput. Chem.* **2004**, *25*, 510.
(27) Hayward, S.; Kitao, A. *Biophys. J.* **2010**, *98*, 1976.
(28) Wedemeyer, W. J.; Scheraga, H. A. *J. Comput. Chem.* **1999**, *20*, 819.
(29) Wu, M. G.; Deem, M. W. *J. Chem. Phys.* **1999**, *111*, 6625.
(30) Cortes, J.; Simeon, T.; Remaud-Simeon, M.; Tran, V. *J. Comput. Chem.* **2004**, *25*, 956.
(31) Noonan, K.; O'Brien, D.; Snoeyink, J. *Int. J. Robotics Res.* **2005**, *24*, 971.
(32) Milgram, R.; Liu, G.; Latombe, J. *J. Comput. Chem.* **2008**, *29*, 50.
(33) Favrin, G.; Irbäck, A.; Sjunnesson, F. *J. Chem. Phys.* **2001**, *114*, 8154.
(34) Wang, L.-C. T.; Chen, C. C. *IEEE Trans Robot. Autom.* **1991**, *7*, 489.
(35) Cahill, S.; Cahill, M.; Cahill, K. *J. Comput. Chem.* **2003**, *24*, 1364.
(36) Canutescu, A.; Dunbrack, R. *Protein Sci.* **2003**, *12*, 963.
(37) Lee, A.; Streinu, I.; Brock, O. *Phys Biol* **2005**, *2*, 108.
(38) Stoer, J.; Bulirsch, R. *Numerical Analysis*, 2nd ed.; Springer: Berlin, 1991.
(39) Coutsias, E. A.; Seok, C.; Wester, M. J.; Dill, K. A. *Int J. Quant. Comp.* **2006**, *106*, 176.
(40) Mandell, D. J.; Coutsias, E. A.; Kortemme, T. *Nature Methods* **2009**, *6*, 551.
(41) Gibbs, J. W.; Wilson, E. B. *Vector Analysis*, 1st ed.; Yale University Press: New Haven, CT, 1901.
(42) Pollock, S. N.; Coutsias, E. A. *Numerical Analysis of Inverse Kinematic Algorithms*, preprint, 2011.
(43) Hoffman, D.; Knapp, E.-W. *Eur. Biophys. J.* **1996**, *24*, 387.
(44) Nilmeier, J.; Jacobson, M. P. *J. Chem. Theory Comput.* **2008**, *4*, 835.
(45) Hetenyi, B.; Bernacki, K.; Berne, B. *J. Chem. Phys.* **2002**, *117*, 8203.
(46) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Phys. Chem. B.* **2001**, *105*, 6474.
(47) Jorgensen, W.; Maxwell, D.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225.
(48) Wong, S.; Bernacki, K.; Jacobson, M. P. *J. Phys. Chem. B* **2005**, *109*, 5249.
(49) Ghosh, A.; Rapp, C.; Friesner, R. *J. Phys. Chem. B* **1998**, *102*, 10983.
(50) Gallicchio, E.; Zhang, L.; Levy, R. *J. Comput. Chem.* **2002**, *23*, 517.
(51) Joseph, D.; Petsko, G.; Karplus, M. *Science* **1990**, *249*, 1425.
(52) Derreumaux, P.; Schlick, T. *Biophys. J.* **1998**, *74*, 72.
(53) Williams, J. C.; McDermott, A. E. *Biochemistry* **1995**, *34*, 8309.
(54) Davenport, R.; Bash, P.; Seaton, B.; Karplus, M.; Petsko, G.; Ringe, D. *Biochemistry* **1991**, *30*, 5821.
(55) Teng, Q. Protein Structure Determination from NMR Data. In *Structural Biology: Practical NMR Applications*, 1st ed.; Lee, W., Ed.; Springer: Berlin, 2005.
(56) Neal, S.; Nip, A.; Zhang, H.; Wishart, D. *J. Biomol. NMR.* **2003**, 3, 215.
(57) Xu, Y.; Lorieau, J.; McDermott, A. E. *J. Mol. Biol.* **2010**, *397*, 233.