JCTC Journal of Chemical Theory and Computation

Practically Efficient QM/MM Alchemical Free Energy Simulations: The Orthogonal Space Random Walk Strategy

Donghong Min,[†] Lianqing Zheng,[†] William Harris,[‡] Mengen Chen,[†] Chao Lv,[‡] and Wei Yang*,[†,‡]

Institute of Molecular Biophysics, Florida State University, Tallahassee, Florida 32306, Department of Chemistry and Biochemistry, Florida State University, Tallahassee, Florida 32306

Received January 19, 2010

Abstract: The difference between free energy changes occurring at two chemical states can be rigorously estimated via alchemical free energy (AFE) simulations. Traditionally, most AFE simulations are carried out under the classical energy potential treatment; then, accuracy and applicability of AFE simulations are limited. In the present work, we integrate a recent second-order generalized ensemble strategy, the orthogonal space random walk (OSRW) method, into the combined quantum mechanical/molecular mechanical (QM/MM) potential based AFE simulation scheme. Thereby, within a commonly affordable simulation length, accurate QM/MM alchemical free energy simulations can be achieved. As revealed by the model study on the equilibrium of a tautomerization process of hydrated 3-hydroxypyrazole and by the model calculations of the redox potentials of two flavin derivatives, lumichrome (LC) and riboflavin (RF) in aqueous solution, the present OSRW-based scheme could be a viable path toward the realization of practically efficient QM/MM AFE simulations.

I. Introduction

In the recent decades, the alchemical free energy (AFE) simulation approach[1−12] has evolved to be a key technique of predicting free energy changes associated with various chemical and biophysical events, such as solvation, protein−ligand binding, protein−protein binding, protonation/deprotonation, electron transfer, etc. AFE simulations allow rigorous estimation of the difference between free energy changes occurring at two chemical states via calculations of free energy values along alchemical directions rather than physical directions. For instance, the solvation free energy difference between two chemical species A and B [$\Delta\Delta G = \Delta G_{\text{gas}\rightarrow\text{solution}}(\text{B}) - \Delta G_{\text{gas}\rightarrow\text{solution}}(\text{A})$] can be obtained as the difference of the alchemical free energy changes in two target environments [$\Delta\Delta G = \Delta G_{\text{solution}}(\text{A}\rightarrow\text{B}) - \Delta G_{\text{gas}}(\text{A}\rightarrow\text{B})$].[3] Thereby, visits of the phase regions that could be challenging to sample, such as the ones associated with gas-solution interface transitions, are naturally avoided. In an AFE simulation, an expanded ensemble is usually constructed so as to allow end chemical states (for instance, the states A and B) to be connected in the same phase space; for the above purpose, in a prebuilt hybrid potential energy function,

$$U_o = U_s(\lambda) + U_e \tag{1}$$

the constraints of $U_s(0) = U_s^A$ and $U_s(1) = U_s^B$ should be set to realize the energy terms (U_s^A and U_s^B) unique in two end chemical states (respectively represented by $\lambda = 0$ and $\lambda = 1$). In eq 1, the common environmental energy terms shared by two end states are labeled as U_e. One of the simplest forms of eq 1 is the linear function

$$U_o = (1 - \lambda)U_s^A + \lambda U_s^B + U_e \tag{2}$$

On the basis of a constructed hybrid potential energy function, molecular simulations (molecular dynamics simulations or Monte Carlo simulations) are required to collect

* Corresponding author e-mail: yyang2@fsu.edu.
[†] Institute of Molecular Biophysics.
[‡] Department of Chemistry and Biochemistry.

10.1021/ct100033s © 2010 American Chemical Society
Published on Web 07/08/2010

samples for the estimation of the free energy difference between two end states A ($\lambda = 0$) and B ($\lambda = 1$); it should be noted that specific procedures of AFE simulations vary with the employed free energy simulation algorithms.[13−35]

In AFE simulations, target systems are commonly described molecular mechanically (MM). As generally known, classical force fields are limited in treating intricate molecular interactions in terms of accuracy and applicability; for instance, redox potential calculations, tautomerization free energy predictions, and metal-associated ligand binding affinity estimations are classical examples. Therefore, developing AFE simulation methods[23,35−59] in the combined quantum mechanical and molecular mechanical (QM/MM) paradigm, where the interaction center is treated quantum mechanically and the rest are treated classically,[60−64] is a natural next step with the evolution of the field. Despite almost two decades' development, due to the apparent efficiency issue, there are only a few nontrivial AFE simulations that have been performed under the QM/MM treatment. Very recently, more attention has been paid to the practical issue on how to improve AFE simulation techniques so as to achieve converged QM/MM AFE simulations using commonly available computing resources.[23,50−54,56−59] Synergistic with our general interest in improving free energy simulation efficiency, we have been actively exploring the above issue in the framework of the generalized ensemble (GE) sampling.[64−67] Particularly motivated by a recent development, the orthogonal space random walk (OSRW) algorithm,[34,68] which can lead to significant efficiency improvements in both free energy simulation and conformational sampling, we generalized the usage of the OSRW strategy in QM/MM AFE simulations, and the presently reported studies reveal the fact that this OSRW strategy could be a viable path to the realization of practically efficient QM/MM AFE simulations.

In the present paper, the OSRW-based QM/MM AFE simulation method is introduced. Its efficiency is demonstrated by a model study on the equilibrium of a tautomerization process of hydrated 3-hydroxypyrazole and a model calculation of the redox potentials of two flavin derivatives: lumichrome (LC) and riboflavin (RF) in aqueous solution.

II. Theoretical Designs

II.A. Brief Introduction to QM/MM AFE Simulation Methods. For the purpose of calculating the free energy difference $\Delta G_{\mathrm{A,QM/MM\rightarrow B,QM/MM}}$ between two chemical states (for instance, A and B), both of which are described by the QM/MM potential, there have been two schemes of the related method developments: the "direct" scheme and the "indirect" scheme.

In the "direct" scheme, each QM/MM AFE calculation is performed via alchemical changes directly from the starting state $\mathrm{A_{QM/MM}}$ to the ending state $\mathrm{B_{QM/MM}}$. Here, alchemical transitions can be realized directly through mechanical switching, where alchemical mixing is implemented on the basis of two independent QM/MM electronic structural calculations as reflected in the following equation:

$$U_{\mathrm{o}} = (1 - \lambda)U_{\mathrm{s}}^{\mathrm{A,QM/MM}} + \lambda U_{\mathrm{s}}^{\mathrm{B,QM/MM}} + U_{\mathrm{e}} \qquad (3)$$

where $U_{\mathrm{s}}^{\mathrm{A}}$ and $U_{\mathrm{s}}^{\mathrm{B}}$ in eq 2 are described by the corresponding QM/MM energy terms $U_{\mathrm{s}}^{\mathrm{A,QM/MM}}$ and $U_{\mathrm{s}}^{\mathrm{B,QM/MM}}$; the other common environmental energy terms shared by two end states are labeled as U_{e}. As a comparison, if alchemical mixing is implemented at the level of the QM/MM electronic structural description,[42,55] where intermediate states are defined in electronic structural Hamiltonians, simulation efficiency could be doubled for the fact that, with electronic structural switching, only one QM/MM force calculation is needed at each time step. The recent fractional electron strategy is a promising example of such a design.[55] The "direct" scheme allows rigorous calculation of AFE values when two end states share a similar molecular structure, for instance, in the redox potential calculations.[36,37,44,46,49−51,55,57−59] When two end states possess different numbers of atoms (even the same number of atoms but with distinct structures), the chaperoned approach has been proposed to realize the "direct" scheme.[44,47−49,52,59] Here, two sets of structures are required so that for each of the end states, there is one structure to take care of the physical nature of the corresponding chemical state, while there is a second one to maintain the structural integrity of the vanishing atoms. It should be noted that when two end states are chemically distinct, it is extremely challenging to rigorously apply the "direct" scheme due to the numerical singularity problem that occurs in the energy derivative calculations at the atom-annihilation end state; because of the complication of the QM/MM treatment and the numerical singularity problem, in this situation, end point contributions can only be empirically estimated via an extrapolation strategy. Clearly, the "direct" scheme is more appropriate for the redox potential or the excitation free energy[69] type of calculations, where two end states have very similar nucleus configurations but different electronic configurations; in this case, eq 3 can be directly employed. Notably, in pK_{a} calculations, if the deprotonated states do not interact strongly with the surrounding environment, an effective approximation is to keep the van der Waals interactions associated with the acidic proton in QM/MM free energy simulations and then annihilate the acidic proton in a separate set of simulations.[70]

In comparison with the "direct" scheme, the "indirect" scheme,[38−41,43,52,54] which was pioneered by Gao et al., can robustly deal with the cases where target end states are drastically different. In the "indirect" scheme, the target free energy difference $\Delta G_{\mathrm{A,QM/MM\rightarrow B,QM/MM}}$ is typically calculated through three steps which are respectively responsible for estimating $\Delta G_{\mathrm{A,MM\rightarrow A,QM/MM}}$, $\Delta G_{\mathrm{A,MM\rightarrow B,MM}}$, and $\Delta G_{\mathrm{B,MM\rightarrow B,QM/MM}}$; thereafter, $\Delta G_{\mathrm{A,QM/MM\rightarrow B,QM/MM}}$ can be calculated via the following equation:

$$\Delta G_{\mathrm{A,QM/MM\rightarrow B,QM/MM}} = \Delta G_{\mathrm{B,MM\rightarrow B,QM/MM}} - \Delta G_{\mathrm{A,MM\rightarrow A,QM/MM}} + \Delta G_{\mathrm{A,MM\rightarrow B,MM}} \qquad (4)$$

The advantage of the "indirect" scheme lies in the fact that QM/MM calculations are only involved in the transitions between two resolutions of energy function descriptions on the same molecule, such as the one based on the hybrid potential function

$$U_o = (1 - \lambda)U_s^{A,MM} + \lambda U_s^{A,QM/MM} + U_e \quad (5)$$

where the other common energy terms shared by two end states are labeled as U_e, for the calculation of $\Delta G_{A,MM\rightarrow A,QM/MM}$, and then possible atom annihilation or large chemical configuration change is only required in the $\Delta G_{A,MM\rightarrow B,MM}$ calculation, where the end point singularity problem can be readily taken care of by the soft-core potential treatment.[71−73] Clearly, the "indirect" scheme is more appropriate for the free energy calculations where two end states have different chemical configurations.

II.B. The Sampling Issue in QM/MM AFE Simulations. Due to the fact that QM/MM force calculations are time-demanding, how to achieve decent free energy convergence within a short simulation length has been a bottleneck problem in the QM/MM AFE simulation method development. The sampling issue in QM/MM AFE simulations is by nature the same as that in classical AFE simulations. It concerns two interrelated aspects:[35,74] (1) how to efficiently collect samples to fill the phase space gap between two target states and (2) how to accurately collect samples to achieve certain ensemble averages as required by the employed free energy theory. The former aspect is usually called the "overlap sampling" issue, and the latter aspect is usually called the "conformational sampling" issue. Apparently, the difficulty level of an AFE simulation depends on (1) how well the phase regions of two target chemical states overlap and (2) how rough the energy landscape is in the region connecting two target states. Specifically, in the "direct" scheme of QM/MM AFE simulations (for instance, in the redox potential calculations), the sampling challenge lies in the fact that in the environmental portion (for instance, solvent or protein), nontrivial structural reorganization might occur between chemical state transitions. In the "indirect" scheme of QM/MM AFE simulations, any discrepancy between the employed MM and QM/MM potentials could lead to nontrivial environmental reorganization; it is particularly true when "generic" or any fast-made MM parameter is utilized, although theoretically, employed MM intermediate states in eq 5 should not influence the target free energy difference $\Delta G_{A,QM/MM\rightarrow B,QM/MM}$, provided that adequate sampling is ensured.

Recently, the sampling issue in QM/MM AFE simulations has attracted several research efforts.[23,50−54,56−59] Most of these developments were carried out in the framework of the generalized ensemble sampling.[23,52−54] Up to now, two specific strategies have been used to realize the first-order generalized ensemble based QM/MM AFE simulations; one is the replica exchange based strategy,[52,54] and the other is the simulated scaling based strategy.[23,53] In the simulated scaling strategy, a biasing potential (or a biasing weight function) is employed as in the following equation:

$$U_m = U_o + f_m(\lambda) = U_s(\lambda) + U_e + f_m(\lambda) \quad (6)$$

where $f_m(\lambda)$ is targeted as $-G_o(\lambda)$; here, $G_o(\lambda)$ is the λ-dependent free energy profile in the canonical ensemble with U_o as the potential energy function. The simulated scaling strategy requires λ to be dynamically coupled with the system motions; for this purpose, we can employ either the hybrid Monte Carlo based methods[21,22] or the λ-dynamics approach.[19,20] As shown in our recent works, the simulated scaling strategy can alleviate the above sampling problems in QM/MM AFE simulations in both the "direct" scheme,[23] where the hybrid potential is constructed according to eq 3, and the "indirect" scheme,[52] where the hybrid potential as described by eq 4 is used.

As discussed above, the sampling issue in QM/MM AFE simulations largely occurs in the phase region orthogonal to the order parameter λ, namely, the environmental portion; to ensure efficient free energy convergence, environmental relaxation needs to be synergistically sampled with the λ move. From this point of view, the first-order generalized ensemble treatment is limited because of its sole focus on the removal of explicit free energy barriers along λ. Recently, the concept of a second-order generalized ensemble was introduced to additionally remove the hidden free energy barriers that are responsible for slow environmental relaxation. In this novel scheme, the developed orthogonal space random walk (OSRW) technique has shown intriguing capability in both free energy simulation and general conformational sampling.[35,68] In the present work, we are adopting the OSRW strategy to deal with the sampling issue in QM/MM AFE simulations. The OSRW method and its generalization for QM/MM AFE simulation are detailed in II.C.

II.C. The Orthogonal Space Random Walk Based QM/MM AFE Simulation Method. As discussed in our previous work, on the top of the first-order generalized ensemble treatment [namely, with a biasing energy term $f_m(\lambda)$], a second-order generalized ensemble simulation method can be designed with one more biasing energy term $F_m[\lambda,h(\lambda)]$ so as to further flatten free energy surfaces along λ and $h(\lambda)$. Ideally, for every target λ state, $h(\lambda)$ represents the corresponding order parameter that describes necessary structural relaxations in the space orthogonal to the λ direction; thereby, required environmental sampling can be accelerated. In our earlier works,[35] $\partial U_o/\partial\lambda$ was identified as such a function form; this order parameter function is particularly robust for the systems in which orthogonal space structural transitions strongly couple with the move of the order parameter λ. The second-order generalized ensemble method, where the "hidden barrier" order-parameter function $h(\lambda)$ is represented as $(\partial U_o)/(\partial\lambda)$, was named as the orthogonal space random walk (OSRW) algorithm.

Accordingly, in the OSRW-based QM/MM AFE simulations, the target energy function should be

$$U_m = U_o + f_m(\lambda) + F_m\left[\lambda, \frac{\partial U_o}{\partial\lambda}\right] = U_s(\lambda) + U_e + f_m(\lambda) + F_m\left[\lambda, \frac{\partial U_o}{\partial\lambda}\right] \quad (7)$$

where U_o is the employed QM/MM alchemical hybrid potential based on either eq 3 (in the direct scheme) or eq 5 (in the indirect scheme), $f_m(\lambda)$ is recursively updated toward $-G_o(\lambda)$ [$G_o(\lambda)$ is the λ-dependent free energy profile in the canonical ensemble with U_o as the potential energy function], and $F_m(\lambda,\partial U_o/\partial\lambda)$ is recursively updated toward $-G'_o(\lambda,\partial U_o/\partial\lambda)$

[$G'_o(\lambda,\partial U_o/\partial\lambda)$ is the $(\lambda,\partial U_o/\partial\lambda)$ space free energy profile corresponding to the canonical ensemble with $U_o - G_o(\lambda)$ as the potential energy function].

To obtain the biasing energy terms, $f_m(\lambda)$ [toward $-G_o(\lambda)$] and $F_m(\lambda,\partial U_o/\partial\lambda)$ [toward $-G'_o(\lambda,\partial U_o/\partial\lambda)$ (eq 7), the same simulation design as in the original OSRW algorithm is employed here. First, the dynamic coupling of the scaling parameter λ with the rest of the system is realized via the λ-dynamics treatment,[19,20] e.g., the scaling parameter λ is treated as a one-dimension particle moving between 0 and 1, as commonly applied in the extended Hamiltonian framework.[75] Second, the recursion procedure, by which $f_m(\lambda)$ and $F_m(\lambda,\partial U_o/\partial\lambda)$ are adaptively achieved, has two components: a "recursion kernel" to estimate $F_m(\lambda,\partial U_o/\partial\lambda)$ directly from simulations and a "recursion slave" to estimate $f_m(\lambda)$ based on $F_m(\lambda,\partial U_o/\partial\lambda)$.

Specifically on the "recursion kernel", we employ a similar recursion strategy as the one in the metadynamics method.[76,77] $F_m(\lambda,\partial U_o/\partial\lambda)$ is obtained as the sum of a relatively small Gaussian-shaped repulsive potential:

$$h\exp\left(-\frac{|\lambda-\lambda(t_i)|^2}{2w_1^2}\right)\exp\left(-\frac{\left|\frac{\partial U_o}{\partial\lambda}-\frac{\partial U_o}{\partial\lambda}(t_i)\right|^2}{2w_2^2}\right) \quad (8)$$

which is centered at $[\lambda(t_i),(\partial U_o/\partial\lambda)(t_i)]$, and thereby discourages the system from often visited configurations. Here, t_i represents the time step when the ith update is scheduled. With this procedure repeated, the overall biasing potential

$$F_m\left(\lambda,\frac{\partial U_o}{\partial\lambda}\right)=\sum_{t_i}h\exp\left(-\frac{|\lambda-\lambda(t_i)|^2}{2w_1^2}\right)\times \exp\left(-\frac{\left|\frac{\partial U_o}{\partial\lambda}-\frac{\partial U_o}{\partial\lambda}(t_i)\right|^2}{2w_2^2}\right) \quad (9)$$

builds up and eventually flattens the underlying curvature of the target free energy surface. Then, like in traditional metadynamics simulations,[76] the free energy profile along the orthogonal space $(\lambda,\partial U_o/\partial\lambda)$ can be estimated as $-F_m(\lambda,\partial U_o/\partial\lambda)$. Thus, for any state λ', the free energy profile along its generalized force space can be estimated as $-F_m(\lambda',(\partial U_o/\partial\lambda)|_{\lambda'})$; correspondingly, the energy derivative probability should be proportional to $\exp[\beta F_m(\lambda',(\partial U_o/\partial\lambda)|_{\lambda'})]$. Then, the free energy derivative, as the energy derivative ensemble average at the state λ', should be estimated as

$$\left.\frac{\partial G_o}{\partial\lambda}\right|_{\lambda'}=\left\langle\frac{\partial U_o}{\partial\lambda}\right\rangle_{\lambda'}=\frac{\int_{\frac{\partial U_o}{\partial\lambda}}\frac{\partial U_o}{\partial\lambda}\exp\left\{\beta\left[F_m\left(\lambda,\frac{\partial U_o}{\partial\lambda}\right)\right]\right\}\delta(\lambda-\lambda')}{\int_{\frac{\partial U_o}{\partial\lambda}}\exp\left\{\beta\left[F_m\left(\lambda,\frac{\partial U_o}{\partial\lambda}\right)\right]\right\}\delta(\lambda-\lambda')} \quad (10)$$

Following the thermodynamic integration formula,[13] the free energy change between the initial state with λ_i, which is the lower bound of the scaling parameter range, and any target state with the scaling parameter λ can be estimated as a function of λ:

$$G_o(\lambda)=\int_{\lambda_i}^{\lambda}\left.\frac{\partial G_o}{\partial\lambda}\right|_{\lambda'}d\lambda' \quad (11)$$

Therefore, eqs 9 and 10 can act as the "recursion slave" to obtain $-G_o(\lambda)$, which is the recursion target of $f_m(\lambda)$.

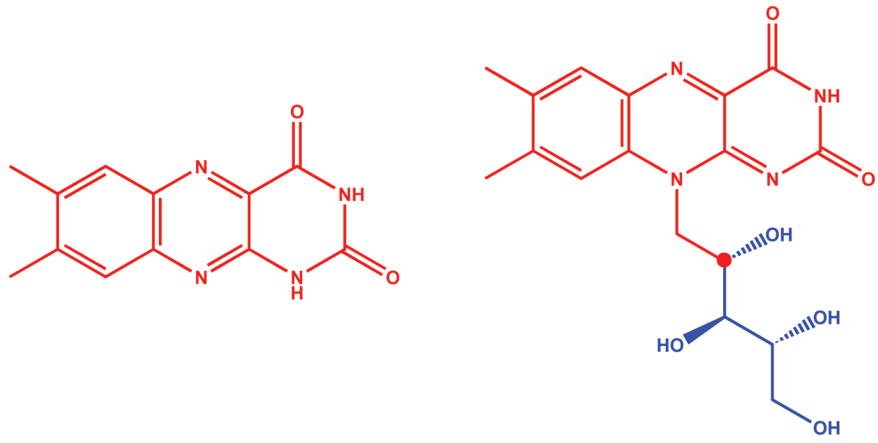
In summary, our recursion procedure for the purpose of obtaining a decent biasing potential $f_m(\lambda) + F_m(\lambda,\partial U_o/\partial\lambda)$ has two components. First, $f_m(\lambda)$ is updated via the "recursion slave" based on eqs 10 and 11. It should be noted that in eqs 10 and 11, the estimation of $G_o(\lambda)$ for the $f_m(\lambda)$ update is not directly dependent on the instantaneous $f_m(\lambda)$ but only dependent on individual one-dimensional (along $\partial U_o/\partial\lambda$) free energy profiles $-F_m(\lambda',(\partial U_o/\partial\lambda)|_{\lambda'})$ at each state λ'. Therefore, the above procedure can ensure the convergence of $f_m(\lambda)$ to the target function of $-G_o(\lambda)$. When $f_m(\lambda)$ converges to $-G_o(\lambda)$, this recursion procedure effectively becomes a two-dimensional metadynamics simulation along the directions of λ and $\partial U_o/\partial\lambda$ with $U_o - G_o(\lambda)$ as the potential energy function. Thereafter, through the "recursion kernel" described by eq 10, $F_m(\lambda,\partial U_o/\partial\lambda)$ will converge to the target $-G'_o(\lambda,\partial U_o/\partial\lambda)$. To understand how the above procedure works, model study 1 in ref 35 can be a good illustrative example.

It should be noted that, in all of the above equations, $\partial U_o/\partial\lambda$ is equal to $(U_s^{B,QM/MM} - U_s^{A,QM/MM})$ for the direct scheme based QM/MM AFE simulations (eq 3), and $\partial U_o/\partial\lambda$ is equal to, for instance, $U_s^{A,QM/MM} - U_s^{A,MM}$ to calculate $\Delta G_{A,QM/MM\rightarrow A,MM}$, in the indirect scheme (eq 5).

III. Computational Details

The present OSRW-based QM/MM AFE simulation method was implemented in our customized CHARMM program.[78,79] Presently, it is implemented with the SCCDFTB (self-consistent charge density functional tight binding)/CHARMM treatment.[80] We are actively working on the implementation of the present method for higher level electronic structural descriptions. Notably, the present direct scheme implementation is realized via the mechanical mixing method rather than the electronic structural switching method;[55] the latter scheme requires only one electronic structural force calculation at each time step, and therefore it is another target for our future development. To illustrate the presently proposed algorithm in both the direct scheme and the indirect scheme, two model studies were performed.

In OSRW simulations, the introduction of the energy term $F_m(\lambda,\partial U_o/\partial\lambda)$ may bring some computational overhead compared with canonical MD simulations. However, our implementation regarding this energy term benefit from two unique treatments: (1) like the implementation in a presently employed metadynamics package,[81] unit Gaussian functions are stored in grids; (2) the force vectors related to $\partial U_o/\partial\lambda$ are stored in the middle of the force calculations on U_o. Therefore, the force calculation in each step has less than 0.3% overhead, even for small-sized systems like ones in model study 1. The other source of the overhead involves the recursion slave estimations based on eqs 9 and 10; it is also less than 0.1% of 100 step force calculations, which



(a) Lumichrome (LC) (b) Riboflavin (RF)

Figure 1. Molecular structures of two flavin derivatives: lumichrome (a) and riboflavin (b). The portions colored in red are described quantum mechanically; the portion colored in blue is described molecular mechanically. The atom labeled by the red dot represents the interface atom described by the generalized hybrid orbital (GHO) treatment.

are commonly used by us as the interval of the recursion slave update.

To illustrate the presently proposed algorithm in both the direct scheme and the indirect scheme, two model studies were performed.

III.A. Model Study 1: The Calculations of the Redox Potentials of Two Hydrated Flavin Derivatives. To illustrate the direct scheme, model study 1 was performed to calculate the redox potentials of two flavin derivatives: lumichrome (LC) and riboflavin (RF) (Figure 1), in aqueous solution. The model calculation on RF was designed to illustrate a situation when an across-bond boundary between the quantum mechanical (QM) portion and the molecular mechanical (MM) portion is defined (as shown in Figure 1b, where the carbon atom labeled by the red dot is the interface atom described by the generalized hybrid orbital (GHO) treatment[82−84]). It is noted that such a boundary definition is not obligatory for predicting the RF redox potential; it was set up only to demonstrate our implementation, which naturally permits the GHO boundary treatment. Moreover, the long-range electrostatics were treated by the recent integrated SCCDFTB-based lattice Ewald method.[49,85]

Seven sets of independent OSRW-based QM/MM AFE simulations, which were started with different random seeds, were performed on two model systems to assess the efficiency of the present method. The statistical error values were calculated among seven simultaneously estimated free energy values, which were obtained at the same simulation length. In each of these alchemical switchings, which are based on eq 3, the starting state ($\lambda = 0$) was set as the neutral state (LC or RF), and the ending state ($\lambda = 0$) was set as the reduced radical state ($LC^{\bullet-}$ or $RF^{\bullet-}$). Therefore, the obtained free energy changes correspond to the reduction potentials of two model systems. It is noted that the utilization of these two model systems was motivated by a recent work, which was to develop an efficient minimum energy path based redox potential estimation algorithm.[56]

In order to compare the present method with the classical approaches, a set (11) of canonical ensemble MD simulations was performed on the RF model system; these MD simulations were based on the energy function described by eq 3 with the scaling parameter λ set as 0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, and 1.0. Starting with $\lambda = 0.0$, sequentially, each simulation was run with a simulation length of 280 ps. The reduction potential of the RF model system was estimated on the basis of an advanced thermodynamic integration (TI) procedure; specifically based on the samples collected in the 11 canonical ensemble simulations, the free energy derivatives of 101 equally spaced λ states were estimated via the maximal likelihood based λ-WHAM method, the details of which can be found in ref 86. It should be noted that the λ-WHAM approach can ensure better free energy convergence than the classical TI method.

III.B. Model Study 2: The Equilibrium of a Tautomerization Process of Hydrated 3-Hydroxypyrazole. To illustrate the indirect scheme, model study 2 was performed to calculate the free energy change of a tautomerization process of hydrated 3-hydroxypyrazole (Figure 2), which was a target in a recent prediction contest. Following the related discussions in section II.A (eq 5), two indirect-scheme-based QM/MM AFE simulations are required to respectively calculate $\Delta G_{B,MM\rightarrow B,QM/MM}$ and $\Delta G_{A,MM\rightarrow A,QM/MM}$, and one classical AFE simulation is required to calculate $\Delta G_{A,MM\rightarrow B,MM}$. As shown in Figure 2, $\Delta G_{A,MM\rightarrow B,MM}$ can be estimated by the following:

$$\Delta G_{A,MM\rightarrow B,MM} = \Delta G_{A,MM\rightarrow A',MM} + \Delta G_{A',MM\rightarrow B',MM} + \Delta G_{B',MM\rightarrow B,MM} \quad (12)$$

where the A′ and B′ states are defined as the modified A and B states with one extra dummy atom in each; the addition of these dummy atoms allows the calculation of $\Delta G_{A',MM\rightarrow B',MM}$ to be more efficiently handled. Specifically, in the A′ state, a dummy atom, which is alchemically switched to a hydroxyl hydrogen during the $\Delta G_{A',MM\rightarrow B',MM}$ AFE simulations, is linked to the corresponding oxygen via a bond and an angle energy term; in the B′ state, a dummy atom, which is alchemically switched to a hydrogen in a five-member ring during the $\Delta G_{A',MM\rightarrow B',MM}$ AFE

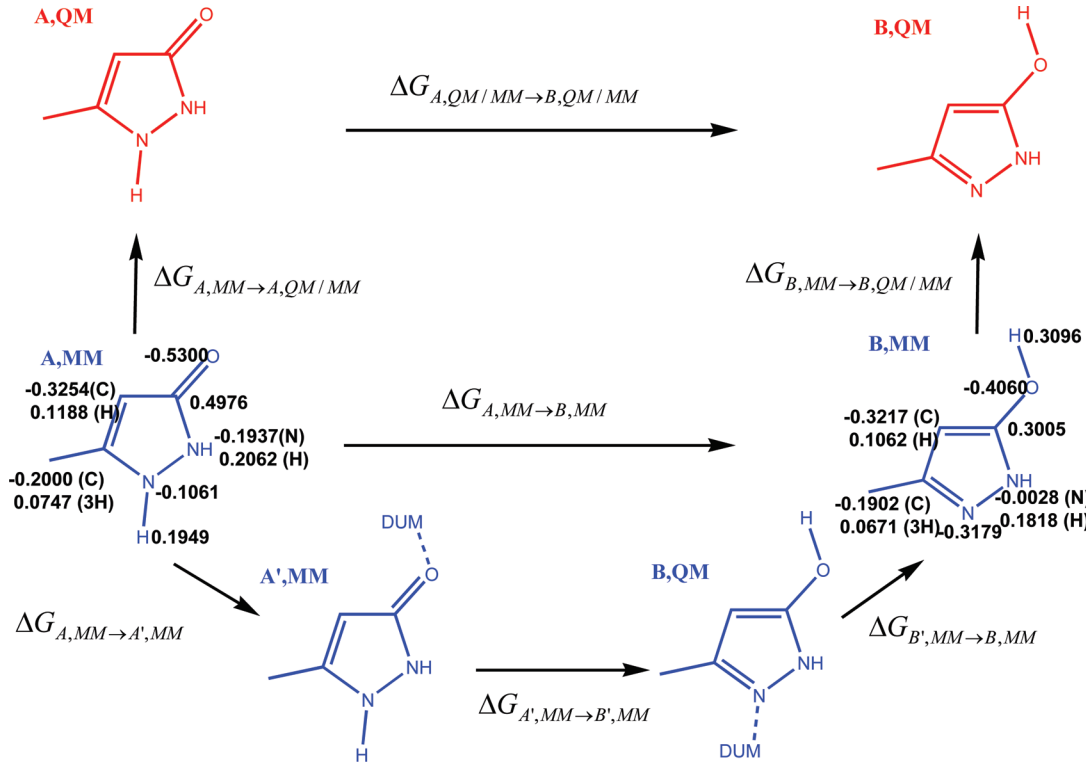


Figure 2. Thermodynamic cycle employed to estimate the free energy change of a tautomerization process of hydrated 3-hydroxypyrazole based on the QM/MM potential. The portions colored in red are described quantum mechanically; the portions colored in blue are described molecular mechanically. The charges in the charge set 1 of the MM parameters are labeled around their corresponding atoms.

Table 1. Restraint Setups for the Calculation of $\Delta G_{A',MM\rightarrow B',MM}$ in Model Study 2[a]

	bond restraints (Å, kcal/mol/Å^2)	angle restraints (rad, kcal/mol/rad^2)	torsional restraints (rad, kcal/mol/rad^2)
dummy atom in A′	(0.96, 1090.0)	(0.52, 242.0)	NA
dummy atom in B′	(1.03, 900.0)	(2.12, 60.0)	(3.14, 5)

[a] All the restraint potentials are based on the harmonic equation, $U_{Restraint} = 1/2K_\xi(\xi - \xi_o)^2$, where ξ_o represents a reference value and K_ξ represents the corresponding force constant. The dummy atom in B′ is torsionally restrained on the same plane as the heterocyclic ring.

simulations, is linked to the corresponding nitrogen atom via a bond, an angle, and a dihedral energy restraint energy term. The dummy atom setup is based on the general principle of the "virtual bond algorithm" (VBA):[87] (1) Each dummy atom has only one bond, angle, and dihedral term so that in the regime of the harmonic oscillator and the rigid rotor approximations, its free energy contribution is separable from the rest of the system and only analytically depends on these restraint terms. (2) The restraint energy terms are set as the same to the corresponding ones at the other end state of $\Delta G_{A',MM\rightarrow B',MM}$ AFE simulations to maximize the phase space overlap between two end chemical states. In the present model study, the restraint energy term parameters of the A′ and B′ state dummy atoms are listed in Table 1. Moreover, the dummy atom in the $\Delta G_{A',MM\rightarrow B',MM}$ AFE simulations is not torsionally restrained; this design is to take advantage of the generalized ensemble tunneling mechanism so as to facilitate the conformational sampling of its alchemical partner.[23,24,75] $\Delta G_{A',MM\rightarrow B',MM}$ was directly calculated via the MM-based OSRW AFE simulation approach.[35] On the basis of the VBA analysis,[87] $\Delta G_{A,MM\rightarrow A',MM} + \Delta G_{B',MM\rightarrow B',MM}$ is estimated to be −0.02 kcal/mol. Due to the introduction of the dummy atoms, the separation-shifted type[71] of the soft-core potentials was employed. The same as in our previous AFE simulations,[35,74] the soft-core potential treatments are simultaneously applied to the van der Waals portion and the real-space electrostatic portion. It should be noted that this soft-core electrostatic potential switching has never been published; however, this treatment has been in the CHARMM program for many versions and has been extensively utilized.[79] A similar simultaneous soft-core treatment can also be found

in the recent literature.[88] The soft-core shifting parameters are set as 5 for both the van der Waals switching and the electrostatic switching.

Two sets of QM/MM AFE simulations on this model system were performed to assess the sampling robustness of the indirect scheme based method. All of the MM state bond, angle, dihedral, and van der Waals parameters were set up according to the CHARMM parameter set. In the first simulation set, the MM charge of each atom was set the same as the charge derived from the SCCDFTB calculation based on the corresponding gas phase minimum energy structure. These charges are shown in Figure 2. In the second

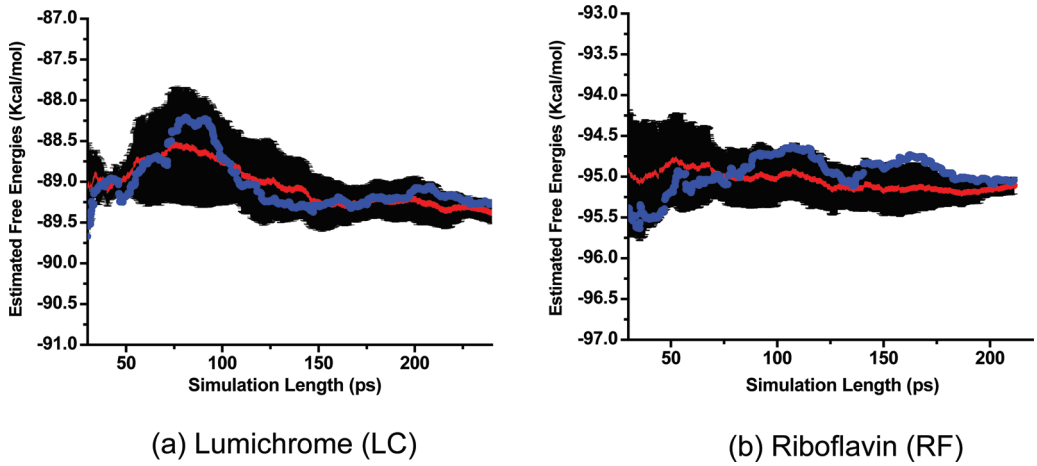


Figure 3. Time-dependent estimated reduction potentials on lumichrome (a) and riboflavin (b). The averages of seven simulation results are shown by the red lines. On each system, the results of a randomly picked free energy simulation are shown by the blue dots, and the statistical errors among seven simulation results are shown around their average results.

simulation set, the MM charge of each atom was set as a half of the charge of the corresponding atom in the first simulation set. Between two simulation sets, the first simulation set represents a case that the MM parameters are well-prepared, and the other simulation set represents the case that the MM parameters are casually or badly prepared.

In all of the OSRW setups, the height of the Gaussian function h (eq 9) was set as 0.01 kcal/mol; the widths of the Gaussian function, ω_1 and ω_2, were set as 0.02 and 4 kcal/mol, respectively. In the recursion kernel, $F_m(\lambda,\partial U_o/\partial\lambda)$ was updated every 10 time steps; in the recursion slave, $f_m(\lambda)$ was updated every 100 time steps. Regarding the choice of the unit Gaussian function, like that discussed in the metadynamics method,[89,90] a greater Gaussian height (or smaller widths) may lead to a larger error of the estimated $-G'_o(\lambda,\partial U_o/\partial\lambda)$; however, a lower Gaussian height (or smaller widths) may cost extra recursion time.[89,90] The set of the Gaussian parameters employed in this work has been extensively tested in various earlier OSRW-based alchemical free energy simulations and robustly applied in our earlier work. Certainly they can be optimized to further improve the efficiency for the present model systems, for instance, via the Wang−Laudau updating scheme;[90] we will leave such discussions for future work.

III.C. General Simulation Setups. In terms of the molecular dynamics simulation setup, all the target systems are respectively solvated in the cubic water (with the TIP3P water model[91]) box with an initial size of 31 × 31 × 31 Å^3, and the linear Ewald method[85] was applied to take care of the long-range columbic interactions while the short-range interactions were switched starting at 8 Å and were totally off at 12 Å. The Nose−Hoover method[92] was employed to maintain a constant temperature at 300 K, and the Langevin piston algorithm[93,94] was used to maintain the constant pressure at 1 atm. The time step was set as 1 fs.

IV. Results and Discussions

IV.A. Model Study 1: The Calculations of the Redox Potentials of Two Hydrated Flavin Derivatives. All of the OSRW AFE simulations on LC and RF show robust and efficient convergence behaviors. As shown in Figure 3, the average of seven independent free energy simulation results in LC reached a superior convergence with a statistical error of 0.13 kcal/mol in about 240 ps; then, the reduction potential of LC was estimated as −89.37 kcal/mol (−3.88 eV). The average of seven independent free energy simulations on RF also reached a superior convergence with a statistical error of 0.10 kcal/mol in about 210 ps; then the reduction potential of RF was estimated as −95.11 kcal/mol (−4.12 eV). As introduced in section III.A, each statistical error value was calculated among seven simultaneously estimated free energy values that were obtained at the same simulation length. Therefore, such statistical error should reflect the magnitude of the divergence of a single free energy prediction from the average at each particular simulation length; i.e., it reflects the robustness of a single OSRW AFE simulation prediction. In terms of the absolute values, the estimated reduction potential of LC, −3.88 eV, is in good agreement with the experimental range −4.3 to −3.7 eV,[95] and the estimated reduction potential of RF, −4.12 eV, is also in a good agreement with the experimental range −4.5 to −3.9 eV.[95] The uncertainty of the experimental values is due to the limitation of the reference reduction potential measurement, which was carried out on the basis of the standard hydrogen electrode (SHE).[96−100] The difference between two estimated reduction potentials of RF and RL is −0.24 eV (−5.53 kcal/mol), which also agrees well with the corresponding measured difference −0.21 eV (−4.84 kcal/mol).[95] Besides the superior convergence, such an excellent agreement should also be attributed to many other factors, for instance, the QM/MM treatment, the inclusion of the EWALD treatment, the possible error cancelation, and even the employment of the GHO boundary condition etc. Although, in an earlier section, the possible limitation of the employed SCCDFTB description was acknowledged and the emphasis of the present work on the efficiency of AFE simulations was clarified, the quantitative prediction power on the LC and RF redox potentials based on the present package of the technical treatments, as demonstrated here, is still practically encouraging.

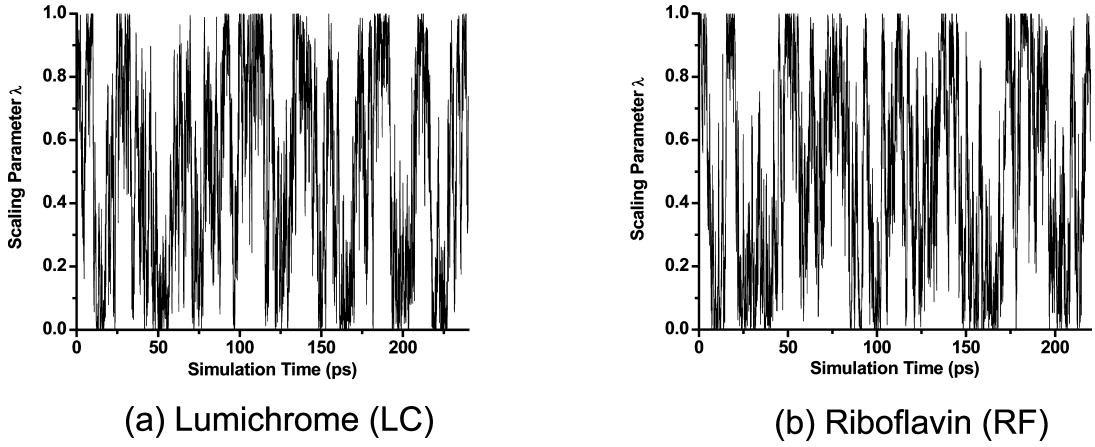


Figure 4. Time-dependent scaling parameter changes of the representative runs on lumichrome (a) and riboflavin (b).

To understand the convergence behavior and assess the efficiency of a single OSRW alchemical FES, let us take a close look at one randomly chosen case from each of the two simulation sets (on LC and RF). As shown in Figure 4a, which reveals the time-dependent scaling parameter changes in a representative run on LC, a complete round trip of the scaling parameter between two end states was realized in about 30 ps. As shown in Figure 3a, at a simulation length of 30 ps, the free energy value (shown by the blue dots) estimated by this single OSRW simulation is −89.64 kcal/mol, which is different from the finally converged value −89.37 kcal/mol only by 0.27 kcal/mol; at the same time, the average free energy value of seven OSRW simulations is −89.07 kcal/mol and the statistical error is 0.52 kcal/mol. As the matter of fact, in the following 20 ps, the convergence robustness of a single OSRW simulation was further improved; this can be reflected by the smaller magnitudes of the time-dependent free energy value changes (the blue dots in Figure 3a) and the reduced statistical errors (Figure 3a). It should be noted that there are two major types of environmental relaxation that strongly couple with the electron transfer process in LC: the water molecule reorganization and the bending and stretching motions of the LC ring. In about 50 ps of simulation, the necessary environmental water reorganization was sufficiently activated by the energy surface flattening along the orthogonal space order parameter $\partial U_o/\partial\lambda$ so that a fast random walk of the scaling parameter was achieved (Figure 4a). As shown by the first plot of Figure 5a, although at 50 ps the height of $-F_m(\lambda,\partial U_o/\partial\lambda)$ was only about 0.6 kcal/mol, the free energy derivative $\partial G_o/\partial\lambda$ curve (the black line) had then been well obtained. Obviously, the free energy result obtained at around 50 ps has been quite acceptable for quantitative prediction purposes. When this OSRW simulation was continued, further free energy flattening started accelerating the LC high-frequency bending and stretching motions and in particular their coupled water reorganization, which could also influence the equilibrium of the target electron transfer process. Then, the contribution of the second type of environmental relaxation, which commonly requires nontrivial canonical sampling to capture, is included as well. The additional recursion, by which the biasing potential term $-F_m(\lambda,\partial U_o/\partial\lambda)$ was further updated to realize possible acceleration of such environmental reorganization, took about another 25 ps of simulation length; as shown in Figure 4a, in this relaxation activation recursion period, another complete scaling parameter round trip ended at 75 ps, when the statistical error among seven independent predictions increased to a peak value, 0.71 kcal/mol. Thereafter, the further recursion allowed the refinement of the following free energy estimations. At a simulation length of 100 ps, when the height of $-F_m(\lambda,\partial U_o/\partial\lambda)$ was increased to 1.2 kcal/mol (the second plot of Figure 5a), the statistical error was reduced to 0.52 kcal/mol (Figure 3a). At a simulation length of 150 ps, when the height of $-F_m(\lambda,\partial U_o/\partial\lambda)$ was increased to 1.6 kcal/mol (the third plot of Figure 5a), the statistical error was reduced to 0.34 kcal/mol; then, the averaged estimated value was −89.25 kcal/mol, which is already very close to the finally estimated value −89.37 kcal/mol. As shown by the fourth plot of Figure 5b, at around 200 ps, a high resolution kernel biasing potential $-F_m(\lambda,\partial U_o/\partial\lambda)$ was well obtained; on the basis of eqs 11 and 12, such a high resolution kernel biasing potential naturally leads to a nicely converged free energy value.

In terms of the RF case, Figure 4b reveals the time-dependent scaling parameter changes of a representative run, where a complete round trip of the scaling parameter between two end states was realized in about 40 ps. As shown in Figure 3b, at a simulation length of 40 ps, the free energy value (shown by the blue dots) estimated by this single OSRW simulation is −95.52 kcal/mol, which is different from the finally converged value −95.11 kcal/mol by only 0.41 kcal/mol; at the same time, the average free energy value from seven OSRW simulations is −94.98 kcal/mol, and the statistical error is 0.62 kcal/mol. When this OSRW simulation was continued, the following recursion simulation allowed the free energy estimations to be further refined. For instance, at a simulation length of 50 ps, when the height of $-F_m(\lambda,\partial U_o/\partial\lambda)$ was around 1.0 kcal/mol (the first plot of Figure 5b), the statistical error was reduced to 0.54 kcal/mol (Figure 3a). At a simulation length of 100 ps, when the height of $-F_m(\lambda,\partial U_o/\partial\lambda)$ was increased to 1.6 kcal/mol (the second plot of Figure 5b), the statistical error was reduced to 0.30 kcal/mol. At a simulation length of 200 ps, the

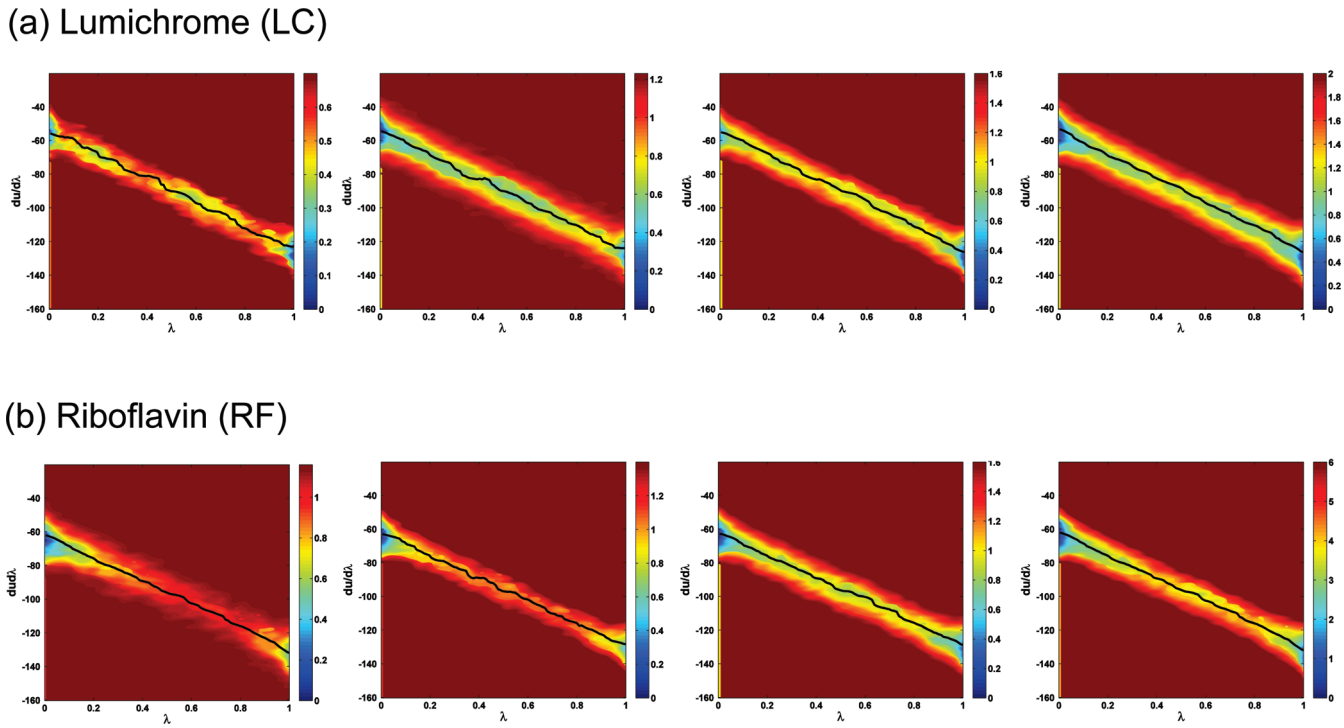


Figure 5. Contour plots on the estimated free energy potential $G'_o(\lambda,\partial U_o/\partial\lambda)$ (colored contours) and the plots on the λ-dependent free energy derivatives (black lines). The depth of each estimated free energy potential $G'_o(\lambda,\partial U_o/\partial\lambda)$ contour plot reveals the activation level in the $\partial U_o/\partial\lambda$ direction. These results were obtained from the representative simulations on lumichrome (a) and riboflavin (b). On each system, the results from the 50 ps simulation are shown as the first, the results from the 100 ps simulation are shown as the second, the results from the 150 ps simulation are shown as the third, and the results from the 200 ps simulation are shown as the fourth.

statistical error was reduced to 0.11 kcal/mol, and then, the averaged estimated value was −95.11 kcal/mol, which is identical to the finally estimated value.

IV.B. The Comparison Study on the RF System Based on the λ-WHAM TI Method. As described in the Computational Details section, the comparison study on the RF system is based on 11 canonical ensemble MD simulations, which were sequentially performed with a length of 280 ps on each λ state; all of the free energy analysis is based on an improved TI (λ-WHAM) procedure.[86]

The lines in Figure 6 show the changes of the estimated free energies with an increase of the simulation length on each single λ state. It is noted that the computing cost should be calculated on the basis of the overall sampling length; for instance, when the simulation length on each λ state is 10 ps, the overall simulation length is 110 ps (from all 11 state simulations). The red line reveals the fact that it took each of the simulations about 180 ps (with overall 1.9 ns simulation length) to reach a "pseudo-convergence", and then the free energy value was estimated as −95.28 kcal/mol. Considering possible contamination of the nonequilibrium data, the samples generated during the first 50 ps of each simulation (with overall 550 ps simulation length) were removed; then, the estimated free energy value shifts up to −95.24 kcal/mol, which converges around 180 ps (also with overall 1.9 ns simulation length; blue line). If the samples generated during the first 150 ps of each simulation (with overall 1.65 ns simulation length) are removed, then the estimated free energy value further shifts to −95.13 kcal/mol, which converges around 220 ps (with overall 2.42 ns simulation length) with a statistical error of 0.52 kcal/mol (green line). If the samples generated during the first 200 ps of each simulation (with overall 2.2 ns simulation length) are removed, then the estimated free energy value further

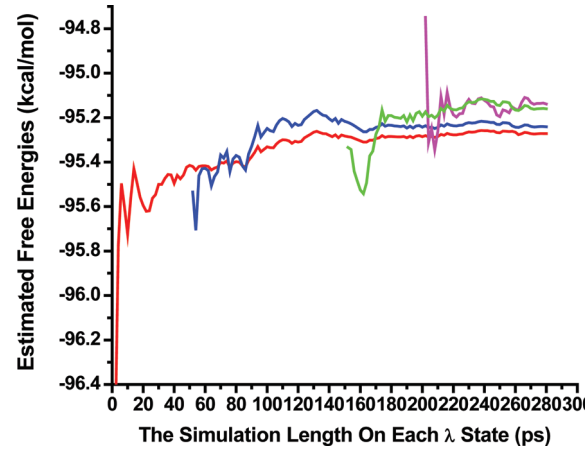


Figure 6. Free energy simulation results from the comparison study on the riboflavin model system. The lines show the changes of the estimated free energies with the increase of the simulation length on each single λ state. The red line shows the time-dependent results based on the samples without the removing of any equilibration data. The blue line shows the time-dependent results when the first 50 ps samples in each λ state simulation are removed. The green line shows the time-dependent results when the first 150 ps samples in each λ state simulation are removed. The pink line shows the time-dependent results when the first 200 ps samples in each λ state simulation are removed.

Table 2. Free Energy Components for the Calculation of $\Delta G_{A,QM/MM\rightarrow B,QM/MM}$ in Model Study 2[a]

	$\Delta G_{A',MM\rightarrow B',MM}$ (kcal/mol)	$\Delta G_{A,MM\rightarrow B,MM}$ (kcal/mol)	$\Delta G_{A,MM\rightarrow A,QM/MM}$ (kcal/mol)	$\Delta G_{B,MM\rightarrow B,QM/MM}$ (kcal/mol)	$\Delta G_{A,QM/MM\rightarrow B,QM/MM}$ (kcal/mol)
charge set 1	6.72	6.70	−10 761.54	−10 771.91	−3.67
charge set 2	1.90	1.88	−10 794.92	−10 800.55	−3.75

[a] $\Delta G_{A,MM\rightarrow B,MM}$ was estimated on the basis of eq 12; on the basis of the VBA analysis, $\Delta G_{A,MM\rightarrow A',MM} + \Delta G_{B',MM\rightarrow B',MM}$ is estimated to be −0.02 kcal/mol. $\Delta G_{A,QM/MM\rightarrow B,QM/MM}$ was estimated on the basis of eq 5. The experimental value of $\Delta G_{A,QM/MM\rightarrow B,QM/MM}$ is −3.40 kcal/mol, which was measured on the basis of the pK_a value difference between two target molecules.

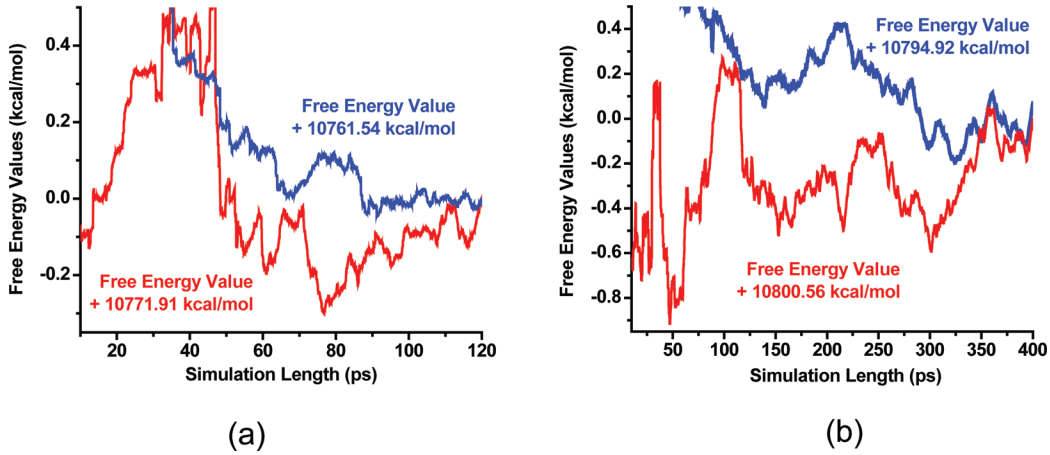


Figure 7. Time-dependent predicted free energy values in model study 2. The free energy results from the simulations where the charge set 1 of the MM parameters was employed are shown in a, and the free energy results from the simulations where the charge set 2 of the MM parameters was employed are shown in b. The free energy results on molecule A are shown in red, and the free energy results on molecule B are shown in blue. To show the results clearly, the values from each simulation are shown in the shifted fashion; for instance, the red line in Figure 6a shows the change of the free energy value plus 10771.91 kcal/mol.

shifts to −95.12 kcal/mol, which also converges around 220 ps (with overall 2.42 ns simulation length) with a statistical error of 0.61 kcal/mol (pink line). It should be noted that, certainly, each of the 11 simulations may have different equilibration lengths. However, the present analysis is performed only for a qualitative efficiency comparison purpose, and a quantitative efficiency comparison requires more elaborate analysis.

Clearly, the comparison study demonstrates the fact that the OSRW approach and the classical methods produce the same free energy results, and the OSRW approach has superior sampling efficiency over the employed λ-WHAM-based TI procedure. Notably, in this study on the RF system, the intrinsic environmental (water molecules) relaxation is relatively fast; it can be anticipated that in more complex systems, where environmental responses are slower, more efficiency gains based on the OSRW method could be obtained.

IV.C. Model Study 2: The Equilibrium of a Tautomerization Process of Hydrated 3-Hydroxypyrazole. As shown in Table 2, two sets of the indirect scheme based AFE simulations led to quantitatively the same predictions (−3.67 and −3.75 kcal/mol) on the equilibrium free energy change of the tautomerization process of the hydrated 3-hydroxypyrazole that is illustrated in Figure 2. These estimations are in great agreement with the experimental measured result, −3.40 kcal/mol; this experimental value was obtained on the basis of the pK_a difference (2.50) of two target molecules.[101] As emphasized in section IV.A, although the purpose of this work is to demonstrate the efficiency of the present QM/MM AFE scheme, the quantitative prediction power of the present package of implementations, which include the OSRW sampling strategy, the EWALD treatment, and even the coupling of the constant pressure treatment etc., shows the necessity of an elaborate development of QM/MM-based AFE simulation facilities.

In the indirect scheme of the simulations based on charge set 1 of the MM parameters, both of the QM/MM AFE simulations reached decent convergence within 50 ps (Figure 7a), when the predicted values were within 0.3 kcal/mol from the converged free energy values (obtained after 500 ps); this is consistent with the fact that at around 50 ps, the first scaling parameter round-trips had been realized (as shown in Figure 7a). The free energy refinements, which were obtained by more frequent scaling parameter space random walks after the initial scaling parameter round-trips (Figure 8a), resulted in smaller deviation of the estimated results from the converged free energy values; for instance, at round 100 ps, the predicted values are only within 0.1 kcal/mol from the converged free energy results. The $\Delta G_{A',MM\rightarrow B',MM}$ prediction was achieved via a 300 ps classical OSRW simulation. It is noted that from the sampling viewpoint, the QM/MM AFE simulations on charge set 1 of the MM parameters are less challenging for the fact that charge set 1 was obtained on the basis of the SCCDFTB gas phase calculations; therefore the QM/MM and the MM end states in each QM/MM AFE simulation have reasonable phase-space overlaps. In comparison, the QM/MM AFE simulations on

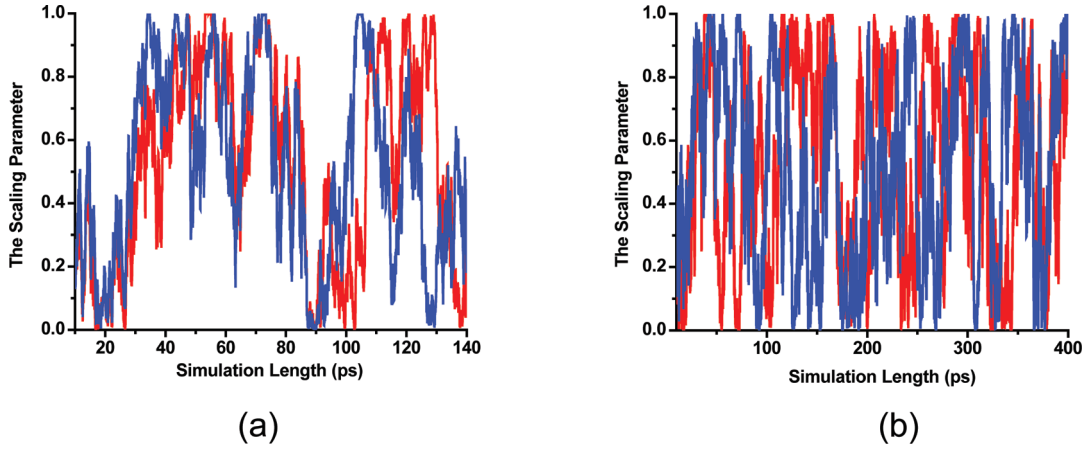


Figure 8. Time-dependent scaling parameter changes in model study 2. The results from the simulations where the charge set 1 of the MM parameters was employed are shown in a, and the results from the simulations where the charge set 2 of the MM parameters was employed are shown in b. The results on molecule A are shown in red, and the results on molecule B are shown in blue.

charge set 2 of the MM parameters show slower convergence behaviors; this observation is also consistent with the fact that charge set 2 was set to represent a poorly prepared charge set with each atomic charge equal to the half of its corresponding value in charge set 1. Their initial scaling parameter round trips were obtained until about 100 ps (Figure 8b), after the predicted values could be 0.5 kcal/mol from the converged free energy values (Figure 7b). More frequent scaling parameter random walks after the first 100 ps of simulation length (Figure 8b) naturally led to better free energy convergence; for instance, at 350 ps, the predicted values are only within 0.1 kcal/mol from the converged free energy results (Figure 7b). On charge set 2, the $\Delta G_{A',MM\rightarrow B',MM}$ prediction was achieved via a 300 ps classical OSRW simulation as well.

V. Concluding Remarks

The difference of free energy changes occurring at two chemical states can be rigorously estimated via alchemical free energy (AFE) simulations. Traditionally, most AFE simulations are carried out under the classical energy potential treatment; then, accuracy and applicability of AFE simulations could be limited; for instance, tautomerization free energy estimations and metal associated ligand binding affinity estimations are outstanding examples.

Due to the fact that QM/MM force calculations are time-demanding, how to achieve decent free energy convergence within a short simulation length has been a bottleneck problem in the QM/MM AFE simulation method development. In the present work, we integrate a recent second-order generalized ensemble strategy, the orthogonal space random walk (OSRW) method, into the combined quantum mechanical/molecular mechanical (QM/MM) calculation based AFE simulation scheme. Thereby, within a commonly affordable simulation length of time, accurate QM/MM alchemical free energy simulations can be achieved. As revealed by the model study on the equilibrium of a tautomerization process of hydrated 3-hydroxypyrazole and by the model calculations of the redox potentials of two flavin derivatives: lumichrome (LC) and riboflavin (RF) in aqueous solution, the present OSRW based scheme could be a viable path toward the realization of practically efficient QM/MM AFE simulations.

Acknowledgment. This paper is dedicated to the 80th birthday of Professor Martin Karplus, whose encouragement on the development of practically efficient QM/MM AFE simulation methods has been pivotal to the success of the present work. We would also like to thank Drs. Jingzhi Pu, Kwangho Nam, Guishan Zheng, Qiang Cui, and Darrin York for many helpful discussions. We would like to acknowledge the National Science Foundation (MCB 0919983) for the funding support. We also thank the Florida State University High Performance Computing Center and the Institute of Molecular Biophysics computing facility for their computing support.

References

(1) Warshel, A. Simulating the energetics and dynamics of enzymatic reactions. *Pont. Acad. Sci. Scr. Var.* **1983**, *55*, 59–81.

(2) Tembe, B. L.; McCammon, J. A. Ligand receptor interactions. *Comput. Chem.* **1984**, *8*, 281–283.

(3) Jorgensen, W. L.; Ravimohan, C. Monte-Carlo simulation of differences in free-energies of hydration. *J. Am. Chem. Soc.* **1985**, *83*, 3050–3054.

(4) Bash, P. A.; Singh, U. C.; Langridge, R.; Kollman, P. A. Free-energy calculations by computer-simulation. *Science* **1987**, *236*, 564–568.

(5) Gao, J.; Kuczera, K.; Tidor, B.; Karplus, M. Hidden thermodynamics of mutant proteins - A molecular-dynamics analysis. *Science* **1989**, *244*, 1069–1072.

(6) Jorgensen, W. L. Free-energy calculations - A breakthrough for modeling organic-chemistry in solution. *Acc. Chem. Res.* **1989**, *22*, 184–189.

(7) Beveridge, D. L.; Dicapua, F. M. Free-energy via molecular simulation - Applications to chemical and biomolecular systems. *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 431–492.

(8) Straatsman, T. P.; McCammon, J. A. Computational alchemy. *Annu. Rev. Phys. Chem.* **1992**, *43*, 407–435.

(9) Kollman, P. Free-energy calculations - Applications to chemical and biochemical phenomena. *Chem. Rev.* **1993**, *93*, 2395–2417.

(10) Simonson, T.; Archontis, G.; Karplus, M. Free energy simulations come of age: Protein-ligand recognition. *Acc. Chem. Res.* **2002**, *35*, 430–437.

(11) Gilson, M. K.; Zhou, H. X. Calculation of protein-ligand binding affinities. *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 21–42.

(12) Jorgensen, W. L.; Thomas, L. L. Perspective on free-energy perturbation calculations for chemical equilibria. *J. Chem. Theory Comput.* **2008**, *4*, 869–876.

(13) Kirkwood, J. G. Statistical mechanics of fluid mixtures. *J. Chem. Phys.* **1935**, *3*, 300–313.

(14) Zwanzig, R. W. High-temperature equation of state by a perturbation method. 1. Nonpolar gases. *J. Chem. Phys.* **1954**, *22*, 1420–1426.

(15) Bennett, C. H. Efficient estimation of free-energy differences from Monte-Carlo data. *J. Comput. Phys.* **1976**, *22*, 245–268.

(16) Souaille, M.; Roux, B. Extension to the weighted histogram analysis method: combining umbrella sampling g with free energy calculations. *Comput. Phys. Commun.* **2001**, *135*, 40–57.

(17) Shirts, M. R.; Bair, E.; Pande, V. S. Equilibrium free energies from nonequilibrium measurements using maximum-likelihood methods. *Phys. Rev. Lett.* **2003**, *91*, 140601.

(18) Lu, N. D.; Kofke, D. A.; Woolf, T. B. Improving the efficiency and reliability of free energy calculations using overlap sampling methods. *J. Comput. Chem.* **2004**, *25*, 28–39.

(19) Kong, X. J.; Brooks, C. L. Lambda-dynamics: A new approach to free energy calculations. *J. Chem. Phys.* **1996**, *105*, 2414–2423.

(20) Knight, J. L.; Brooks, C. L. Lambda-dynamics free energy simulation methods. *J. Comput. Chem.* **2009**, *30*, 1692–1700.

(21) Tidor, B. Simulated annealing on free-energy surfaces by a combined molecular-dynamics and Monte-Carlo approach. *J. Phys. Chem.* **1993**, *97*, 1069–1073.

(22) Pitera, J.; Kollman, P. Designing an optimum guest for a host using multimolecule free energy calculations: Predicting the best ligand for Rebek's "tennis ball". *J. Am. Chem. Soc.* **1998**, *120*, 7557–7567.

(23) Li, H. Z.; Fajer, M.; Yang, W. Simulated scaling method for localized enhanced sampling and simultaneous "alchemical" free energy simulations: A general method for molecular mechanical, quantum mechanical, and quantum mechanical/molecular mechanical simulations. *J. Chem. Phys.* **2007**, *126*, 024106.

(24) Min, D. H.; Yang, W. Energy difference space random walk to achieve fast free energy calculations. *J. Chem. Phys.* **2008**, *128*, 191102.

(25) Darve, E.; Pohorille, A. Calculating free energies using average force. *J. Chem. Phys.* **2001**, *115*, 9169–9183.

(26) Bitetti-Putzer, R.; Yang, W.; Karplus, M. Generalized ensembles serve to improve the convergence of free energy simulations. *Chem. Phys. Lett.* **2003**, *377*, 633–641.

(27) Fasnacht, M.; Swendsen, R. H.; Rosenberg, M. Adaptive integration method for Monte Carlo simulations. *Phys. Rev. E* **2004**, *69*, 056704.

(28) Ytreberg, F. M.; Swendsen, R. H.; Zuckerman, D. M. Comparison of free energy methods for molecular systems. *J. Chem. Phys.* **2006**, *126*, 184114.

(29) Pomes, R.; Eisenmesser, E.; Post, C. B.; Roux, B. Calculating excess chemical potentials using dynamic simulations in the fourth dimension. *J. Chem. Phys.* **1999**, *111*, 3387–3395.

(30) Sugita, Y.; Kitao, A.; Okamoto, Y. Multidimensional replica-exchange method for free-energy calculations. *J. Chem. Phys.* **2000**, *113*, 6042–6051.

(31) Woods, C. J.; Essex, J. W.; King, M. A. The development of replica-exchange-based free-energy methods. *J. Phys. Chem. B* **2003**, *107*, 13703–13710.

(32) Lu, N. D.; Wu, D.; Woolf, T. B.; Kofke, D. A. Using overlap and funnel sampling to obtain accurate free energies from nonequilibrium work measurements. *Phys. Rev. E* **2004**, *69*, 057702.

(33) Christ, C. D.; van Gunsteren, W. F. Enveloping distribution sampling: A method to calculate free energy differences from a single simulation. *J. Chem. Phys.* **2007**, *126*, 184110.

(34) Abrams, J. B.; Rosso, L.; Tuckerman, M. E. Efficient and precise solvation free energies via alchemical adiabatic molecular dynamics. *J. Chem. Phys.* **2006**, *125*, 074115.

(35) Zheng, L. Q.; Chen, M. G.; Yang, W. Random walk in orthogonal space to achieve efficient free-energy simulation of complex systems. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 20227–20232.

(36) Warshel, A. Dynamics of reactions in polar-solvents - Semiclassical studies of electron-transfer and proton-transfer reactions. *J. Phys. Chem.* **1982**, *86*, 2218–2224.

(37) King, G.; Warshel, A. Investigation of the free-energy functions for electron-transfer reactions. *J. Chem. Phys.* **1990**, *93*, 8682–8692.

(38) Gao, J. L.; Xia, X. F. A priori evaluation of aqueous polarization effects through Monte Carlo QM-MM simulations. *Science* **1992**, *258*, 631–635.

(39) Luzhkov, V.; Warshel, A. Microscopic models for quantum-mechanical calculations of chemical processes in solutions - LD/AMPAC and SCAAS/AMPAC calculations of solvation energies. *J. Comput. Chem.* **1992**, *13*, 199–213.

(40) Gao, J. L.; Luque, F. J.; Orozco, M. Induced dipole-moment and atomic charges based on average electrostatic potentials in aqueous-solution. *J. Chem. Phys.* **1993**, *98*, 2975–2982.

(41) Wesolowski, T.; Warshel, A. Ab-initio free-energy perturbation calculations of solvation free-energy using the frozen density-functional approach. *J. Phys. Chem.* **1994**, *98*, 5183–5187.

(42) Stanton, R. V.; Little, L. R.; Merz, K. M. Quantum free-energy perturbation study within a PM3-MM coupled potential. *J. Phys. Chem.* **1995**, *99*, 483–486.

(43) Gao, J. L.; Freindorf, M. Hybrid ab initio QM/MM simulation of N-methylacetamide in aqueous solution. *J. Phys. Chem. A* **1997**, *101*, 3182–3188.

(44) Li, G. H.; Zhang, X. D.; Cui, Q. Free energy perturbation calculations with combined QM/MM potentials: Complications, simplifications, and applications to redox potential calculations. *J. Phys. Chem. B* **2003**, *107*, 8643–8653.

(45) Li, G. H.; Cui, Q. pKa calculations with QM/MM free energy perturbations. *J. Phys. Chem. B* **2003**, *107*, 14521–14528.

(46) Olsson, M. H. M.; Hong, G. Y.; Warshel, A. Frozen density functional free energy simulations of redox proteins: Computational studies of the reduction potential of plastocyanin and rusticyanin. *J. Am. Chem. Soc.* **2003**, *125*, 5025–5039.

(47) Yang, W.; Bitetti-Putzer, R.; Karplus, M. Chaperoned alchemical free energy simulations: A general method for QM, MM, and QM/MM potentials. *J. Chem. Phys.* **2004**, *120*, 9450–9453.

(48) Hu, H.; Yang, W. T. Dual-topology/dual-coordinate free-energy simulation using QM/MM force field. *J. Chem. Phys.* **2005**, *123*, 041102.

(49) Riccardi, D.; Schaefer, P.; Yang, Y.; Yu, H. B.; Ghosh, N.; Prat-Resina, X.; Konig, P.; Li, G. H.; Xu, D. G.; Guo, H.; Elstner, M.; Cui, Q. Development of effective quantum mechanical/molecular mechanical (QM/MM) methods for complex biological processes. *J. Phys. Chem. B* **2006**, *110*, 6458–6469.

(50) Blumberger, J.; Tavernelli, I.; Klein, M. L.; Sprik, M. Diabatic free energy curves and coordination fluctuations for the aqueous Ag^{+}/Ag^{2+} redox couple: A biased Born-Oppenheimer molecular dynamics investigation. *J. Chem. Phys.* **2006**, *124*, 064507.

(51) Blumberger, J.; Sprik, M. Quantum versus classical electron transfer energy as reaction coordinate for the aqueous Ru^{2+}/Ru^{3+} redox. *Theor. Chem. Acc.* **2006**, *115*, 113–126.

(52) Li, H. Z.; Yang, W. Sampling enhancement for the quantum mechanical potential based molecular dynamics simulations: A general algorithm and its extension for free energy calculation on rugged energy surface. *J. Chem. Phys.* **2007**, *126*, 114104.

(53) Zheng, L.; Li, H.; Yang, W. In *From Computational Biophysics to Systems Biology (CBSB08)*; Hansmann, U. H. E., Meinke, J., Mohanty, S., Nadler, W., Zimmerman, O., Eds.; NIC: Jülich, Germany, 2008; NIC Series Vol. 36, pp 57−64.

(54) Woods, C. J.; Manby, F. R.; Mulholland, A. J. An efficient method for the calculation of quantum mechanics/molecular mechanics free energies. *J. Chem. Phys.* **2008**, *128*, 014109.

(55) Zeng, X. C.; Hu, H.; Hu, X. Q.; Cohen, A. J.; Yang, W. T. Ab initio quantum mechanical/molecular mechanical simulation of electron transfer process: Fractional electron approach. *J. Chem. Phys.* **2008**, *128*, 124510.

(56) Rosta, E.; Haranczyk, M.; Chu, Z. T.; Warshel, A. Accelerating QM/MM free energy calculations: Representing the surrounding by an updated mean charge distribution. *J. Phys. Chem. B* **2008**, *112*, 5680–5692.

(57) Kamerlin, S. C. L.; Haranczyk, M.; Warshel, A. Progress in Ab initio QM/MM free-energy simulations of electrostatic energies in proteins: Accelerated QM/MM studies of pKa, redox reactions and solvation free energies. *J. Phys. Chem. B* **2009**, *113*, 1253–1272.

(58) Zeng, X. C.; Hu, H.; Hu, X. Q.; Yang, W. T. Calculating solution redox free energies with ab inito quantum mechanical/molecular mechanical minimum free energy path method. *J. Chem. Phys.* **2009**, *130*, 164111.

(59) Cheng, J.; Sulpizi, M.; Sprik, M. Redox potentials and pKa for benzoquinone from density functional theory based molecular dynamics. *J. Chem. Phys.* **2009**, *131*, 154504.

(60) Warshel, A.; Levitt, M. Theoretical studies of enzymic reactions - Dielectric, electrostatic and steric stabilization of carbounium-ion in reaction of lysozyme. *J. Mol. Biol.* **1976**, *103*, 227–249.

(61) Field, M. J.; Bash, P. A.; Karplus, M. A combined quantum-mechanical and molecular mechanical potential for molecular-dynamics simulations. *J. Comput. Chem.* **1990**, *11*, 700–733.

(62) Gao, J. L. Hybrid quantum and molecular mechanical simulations: An alternative avenue to solvent effects in organic chemistry. *Acc. Chem. Res.* **1996**, *29*, 298–305.

(63) Bakowies, D.; Thiel, W. Hybrid models for combined quantum mechanical and molecular mechanical approaches. *J. Phys. Chem.* **1996**, *100*, 10580–10594.

(64) Monard, G.; Merz, K. M. Combined quantum mechanical/molecular mechanical methodologies applied to biomolecular systems. *Acc. Chem. Res.* **1999**, *32*, 904–911.

(65) Berne, B. J.; Straub, J. E. Novel methods of sampling phase space in the simulation of biological systems. *Curr. Opin. Struct. Biol.* **1997**, *7*, 181–189.

(66) Mitsutake, A.; Sugita, Y.; Okamoto, Y. Generalized-ensemble algorithms for molecular simulations of biopolymers. *Biopolymers* **2001**, *60*, 96–123.

(67) Okamoto, Y. Generalized-ensemble algorithms: Enhanced sampling techniques for Monte Carlo and molecular dynamics simulations. *J. Mol. Graphics Modell.* **2004**, *22*, 425–439.

(68) Zheng, L. Q.; Chen, M. G.; Yang, W. Simultaneous escaping of explicit and hidden free energy barriers: Application of the orthogonal space random walk strategy in generalized ensemble based conformational sampling. *J. Chem. Phys.* **2009**, *130*, 234105.

(69) Gao, J. L.; Li, N. Q.; Freindorf, M. Hybrid QM/MM simulations yield the ground and excited state pK(a) difference: Phenol in aqueous solution. *J. Am. Chem. Soc.* **1996**, *118*, 4912–4913.

(70) Riccardi, D.; Schaefer, P.; Yang, Y.; Yu, H. B.; Ghosh, N.; Prat-Resina, X.; Konig, P.; Li, G. H.; Xu, D. G.; Guo, H.; Elstner, M.; Cui, Q. Development of effective quantum mechanical/molecular mechanical (QM/MM) methods for complex biological processes. *J. Phys. Chem. B* **2006**, *110*, 6458–6469.

(71) 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.

(72) 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.

(73) Steinbrecher, T.; Mobley, D. L.; Case, D. A. Nonlinear scaling schemes for Lennard-Jones interactions in free energy calculations. *J. Chem. Phys.* **2007**, *127*, 214108.

(74) Min, D. H.; Li, H. Z.; Li, G. H.; Bitetti-Putzer, R.; Yang, W. Synergistic approach to improve "alchemical" free energy calculation in rugged energy surface. *J. Chem. Phys.* **2007**, *126*, 144109.

(75) Martyna, G. J.; Tuckerman, M. E.; Tobias, D. J.; Klein, M. L. Explicit reversible integrators for extended systems dynamics. *Mol. Phys.* **1996**, *87*, 1117–1157.

(76) Laio, A.; Parrinello, M. Escaping free-energy minima. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12562–12566.

(77) Ensing, B.; De Vivo, M.; Liu, Z. W.; Moore, P.; Klein, M. L. Metadynamics as a tool for exploring free energy landscapes of chemical reactions. *Acc. Chem. Res.* **2006**, *39*, 73–81.

(78) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM - A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **1983**, *4*, 187–217.

(79) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Calfisch, 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*, 1545–1614.

(80) Cui, Q.; Elstner, M.; Kaxiras, E.; Frauenheim, T.; Karplus, M. A QM/MM implementation of the self-consistent charge density functional tight binding (SCC-DFTB) method. *J. Phys. Chem. B* **2001**, *105*, 569–585.

(81) Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; Raiteri, P.; Donadio, D.; Marinello, F.; Pietrucci, F.; Broglia, R. A.; Parrinello, M. PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Comput. Phys. Commun.* **2009**, *10*, 1961–1972.

(82) Gao, J. L.; Amara, P.; Alhambra, C.; Field, M. J. A generalized hybrid orbital (GHO) method for the treatment of boundary atoms in combined QM/MM calculations. *J. Phys. Chem. A* **1998**, *102*, 4714–4721.

(83) Amara, P.; Field, M. J.; Alhambra, C.; Gao, J. L. The generalized hybrid orbital method for combined quantum mechanical/molecular mechanical calculations: Formulation and tests of the analytical derivatives. *Theor. Chem. Acc.* **2000**, *104*, 336–343.

(84) Pu, J. Z.; Gao, J. L.; Truhlar, D. G. Combinding self-consistent-charge density-functional tight-binding (SCC-DFTB) with molecular mechanics by the generalized hybrid orbital (GHO) method. *J. Phys. Chem. A* **2004**, *108*, 5454–5463.

(85) Nam, K.; Gao, J. L.; York, D. M. An efficient linear-scaling Ewald method for long-range electrostatic interactions in combined QM/MM calculations. *J. Chem. Theory Comput.* **2005**, *1*, 2–13.

(86) Li, H.; Yang, W. Forging the missing link in free energy estimations: λ-WHAM in thermodynamic integration, overlap histogramming, and free energy perturbation. *Chem. Phys. Lett.* **2007**, *440*, 155–159.

(87) Boresch, S.; Tettinger, F.; Leitgeb, M.; Karplus, M. Absolute binding free energies: A quantitative approach for their calculation. *J. Phys. Chem. B* **2003**, *107*, 9535–9551.

(88) Anwar, J.; Heyes, D. M. Robust and accurate method for free-energy calculation of charged molecular systems. *J. Chem. Phys.* **2005**, *122*, 224117.

(89) Min, D. H.; Liu, Y. S.; Carbone, I.; Yang, W. On the convergence improvement in the metadynamics simulations: A Wang-Landau recursion approach. *J. Chem. Phys.* **2007**, *126*, 194104.

(90) Laio, A.; Rodriguez-Fortea, A.; Gervasio, F. L.; Ceccarelli, M.; Parrinello, M. Assessing the accuracy of metadynamics. *J. Phys. Chem. B* **2005**, *109*, 6714.

(91) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(92) Nose, S. A unified formulation of the constant temperature molecular-dynamics methods. *J. Chem. Phys.* **1984**, *81*, 511–519.

(93) Nose, S.; Klein, M. L. Constant pressure molecular-dynamics for molecular-systems. *Mol. Phys.* **1983**, *50*, 1055–1076.

(94) Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant-pressure molecular-dynamics algorithms. *J. Chem. Phys.* **1994**, *101*, 4177–4189.

(95) Wardman, P. Reduction potentials of one-electron couples involving free-radicals in aqueous-solution. *J. Phys. Chem. Ref. Data* **1989**, *18*, 1637–1755.

(96) Donald, W. A.; Leib, R. D.; O'Brien, J. T.; Bush, M. F.; Williams, E. R. Absolute standard hydrogen electrode potential measured by reduction of aqueous nanodrops in the gas phase. *J. Am. Chem. Soc.* **2008**, *130*, 3371–3381.

(97) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. Aqueous solvation free energies of ions and ion-water clusters based on an accurate value for the absolute aqueous solvation free energy of the proton. *J. Phys. Chem. B* **2006**, *110*, 16066–16081.

(98) Reiss, H. The Fermi level and the redox potential. *J. Phys. Chem.* **1985**, *89*, 3783–3791.

(99) Gomer, R.; Tryson, G. Experimental-determination of absolute half-cell EMFS and single ion free-energies of solvation. *J. Chem. Phys.* **1977**, *66*, 4413–4424.

(100) Marcus, Y. Thermodynamics of solvation of ions. 5. Gibbs free-energy of hydration at 298.1 K. *J. Chem. Soc. Faraday Trans.* **1991**, *87*, 2995–2999.

(101) Parchment, O. G.; Green, D. V. S.; Taylor, P. J.; Hillier, I. H. The prediction of tautomer equilibria in hydrated 3-Hydroxypyrazole: A challenge to theory. *J. Am. Chem. Soc.* **1993**, *115*, 2352–2356.

CT100033S

JCTC Journal of Chemical Theory and Computation

Solute Partitioning into Lipid Bilayers: An Implicit Model for Nonuniform and Ordered Environment

Giulia Parisio and Alberta Ferrarini*

Dipartimento di Scienze Chimiche, Università degli Studi di Padova Via Marzolo, 1 - 35131 Padova, Italy

Received April 26, 2010

Abstract: We have developed a theoretical and computational methodology to evaluate the coupled orientational−positional distribution of solutes in lipid bilayers. Four different contributions to the solute free energy are considered, which can be traced back to (i) electrostatic and (ii) dispersion interactions between the solute and environment, (iii) work for the formation of a solute-shaped cavity, and (iv) anisotropic interactions with the ordered acyl chains. An atomistic representation of the solute is adopted, which includes the conformational degrees of freedom, whereas an implicit model is used for the water/bilayer environment. The highly nonuniform and anisotropic nature of this is introduced through the profiles of density, dielectric permittivity, lateral pressure, and acyl chain order parameters, which can be derived from experiments or simulations. Effects of chemical composition and physical state of the bilayer can be accounted for by a proper form of these profiles. The methodology which we propose is suitable for the integrated calculation of spectroscopic observables for probes in membranes, for the estimate of partition and permeability coefficients of solutes, and for the implicit modeling of the membrane environment in molecular dynamics and Monte Carlo simulations. Here, the method is presented, and the underlying assumptions are discussed. Cholesterol in the liquid crystalline DPPC bilayer is taken as a case study, to illustrate the capabilities of the proposed approach. Free energy maps, distribution profiles, and orientational properties are shown; they compare well with those obtained from all-atom molecular dynamics simulations, as well as with available experimental data, suggesting that the model used is able to capture the subtle effects of the interplay between intermolecular interaction and nanoscale architecture of the lipid bilayer. The detailed picture provided by our calculations appears suitable to investigate the determinants of the behavior of solutes in lipid membranes, highlighting even nonstraightforward issues, which may have biophysical implications.

Introduction

Partitioning of solutes into lipid bilayers has important implications: beside playing a role in cellular processes, it underlies the permeation of drugs across biomembranes. This has motivated along the years a steady effort to develop models able to catch the features of the membrane environment.[1] Traditionally, lipid bilayers have been approximated as slabs of apolar solvents in water, and their high nonuniformity and orientational order have been neglected.[2] Actually, these are the most peculiar features of lipid bilayers, which are characterized by depth-dependent anisotropic stresses and by enormous gradients of properties over a nanometer length scale. The profiles of density, lateral forces, and order are strictly related to the lipid composition and the physical state of the bilayer: they depend on the acyl chain saturation and the headgroup structure and exhibit significant differences between the liquid crystal and the gel phase. The importance of these aspects was recognized early by Dill and co-workers, who singled out the peculiar nature

* Corresponding author phone: +39 049 8275682; fax: +39 049 8275239; e-mail: alberta.ferrarini@unipd.it.

10.1021/ct100210u © 2010 American Chemical Society
Published on Web 07/06/2010

of partitioning when "interfacial phases" are involved.[3,4] However, the structural features of lipid bilayers were generally ignored in subsequent studies, most of which focused on the effects of the dielectric discontinuity and tried to extend to its description the theoretical and computational methods developed for liquid solvents.[5−7] The introduction of bilayer properties into solvation theories is a difficult task: due to the inherent chemical and structural complexity, modeling of bilayers remains a challenge for statistical theories of liquids.[8,9] Only in a few exceptions were the structural features of the bilayer included in models for solute partitioning between water and lipid membranes. Xiang and Anderson evaluated the free energy change upon formation of a molecular cavity in the bilayer as the reversible work against the lateral pressure, including in this way the configurational entropy and the change in conformational energy of lipid chains.[10] They performed calculations for ellipsoidal solutes, with a lateral pressure profile obtained from coarse-grained molecular dynamics simulations, and could predict the dependence of solute orientational order parameters on the molecular volume and on the ratio of the long to short molecular axis. Mitragotri and co-workers developed a model for the cavity work for ellipsoidal solutes in the ordered chain region.[11] It was based on a 2D form of scaled particle theory and related partition coefficients to solute size and bilayer properties, like lipid density and lipid chain order parameters. Kessel and co-workers assumed a decomposition of the solute-bilayer interaction free energy in a set of contributions including, in addition to electrostatic and nonpolar terms, customary in solvation theories, two further terms accounting for elastic deformations of the bilayer and conformational restrictions of lipid chains induced by the solute.[12] The former contributions were described by continuum solvent models with atomistic representation of the solute. For the latter, the elastic continuum theory and a liquid crystal model, respectively, were used, and for simplicity, the solute was represented as a cylinder.

We present here a theoretical and computational method for modeling the free energy landscape of solutes in lipid bilayers, wherein the order and nonuniformity characteristics of this environment are consistently taken into account. The free energy of solutes is expressed as the superposition of the work required to form a molecular shaped cavity and other contributions associated with the switching on of solute−solvent interactions. The cavity work is evaluated on the basis of the lateral pressure profile.[13] A nonpolar and an electrostatic contribution are evaluated as a function of the density and permittivity gradients in the bilayer,[14] respectively; a further term, explicitly accounting for the anisotropy of intermolecular interactions due to lipid chain order, is derived from liquid crystal theory.[15] The environment enters with its collective properties: dielectric permittivity, lateral pressure, and density profile, in addition to acyl chain order parameters. Values appropriate for the physical state and composition of the bilayer can be obtained from molecular dynamics simulations or from experiments. An atomistic representation of the solute is used, comprising its geometry, charge distribution, polarizability, and flexibility. Free energy maps are obtained as a function of molecular position and orientation, whose coupling is intrinsic to the bilayer structure.

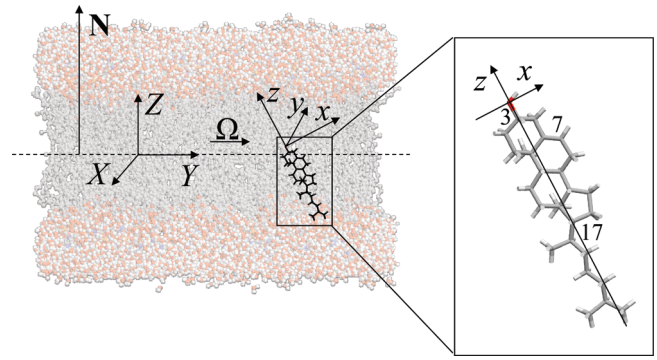


Figure 1. Cholesterol and DPPC bilayer, with the bilayer and the molecular reference frames, (*X*,*Y*,*Z*) and (*x*,*y*,*z*), respectively. The laboratory frame has its origin in the bilayer midplane and the *Z* axis parallel to the bilayer normal (**N**); the latter is directed from the midplane toward water. The molecular frame has its origin at the nuclear position of the oxygen atom and the *x*, *y*, and *z* axes anchored to the fused ring core, with the *z* axis parallel to the C_3-C_{17} segment and the *x* axis coplanar with C_7 (atom numbering as in Figure S5, Supporting Information). The Euler angles $\Omega = (\alpha, \beta, \gamma)$ specify the orientation of the molecular frame with respect to the laboratory frame.

In the next section, the theoretical model is presented. First, the different contributions to the solute free energy are described; then suitable distribution functions and average values are introduced. In the third section, computational details are reported. In the fourth section, cholesterol in liquid crystalline DPPC (1,2-dipalmitoyl-*sn*-glycero-3-phosphatidylcholine; see Abbreviations) is taken as an example to illustrate the features of the proposed methodology. The results are discussed in relation to experimental data and other theoretical−computational investigations. In the final section, capabilities and limits of the proposed approach are summarized, and future developments are outlined.

Theory

The lipid organization in the bilayer defines an alignment axis, the so-called director (**n**), which in the liquid crystal phase is parallel to the bilayer normal (**N**). So, the degrees of freedom of a rigid solute within the lipid bilayer are its position along the normal and its orientation with respect to it; the former is defined, in the bilayer frame, by the *Z* coordinate of the origin of the molecular frame, the latter by the Euler angles, $\Omega = (\alpha, \beta, \gamma)$, which specify the orientation of the molecular frame in the bilayer frame. (Actually, due to the axial symmetry of the environment, the *X*, *Y* coordinates of the origin of the molecular frame, as well as the Euler angle α, do not need to be specified.) Reference frames and transformations are sketched in Figure 1. Our implicit model assumes that the solute positional and orientational distribution is determined by a mean field potential $U(\Omega,Z)$: this can be viewed as the energy of the "solute microstate" (Ω,Z) in the solvent, under the implicit assumption that all degrees of freedom of the latter are averaged out, which gives $U(\Omega,Z)$ a free energy character.

As customary in solvation models,[16] we decompose the mean field potential experienced by a solute into a set of contributions:

$$U = U_{cav} + U_{el} + U_{disp} + U_{ord} \quad (1)$$

The first term, U_{cav}, is related to the work required to form a molecular cavity in the medium, whereas the next two terms, U_{el} and U_{disp}, account for the switching on of electrostatic and dispersion interactions between the solute and environment, respectively. In addition to these terms, which usually appear in solvation models, we have included a further contribution, U_{ord}, which can be traced back to the anisotropy of intermolecular interactions, due to acyl chain order, in the bilayer environment. In the following, the individual contributions to the mean field potential will be considered, and the model adopted for each of them will be presented.

Cavity Work. The cavity contribution to the mean field potential is evaluated as the work against the external pressure to create a cavity containing the solute. As a consequence of the system organization, an axially symmetric pressure tensor can be defined in liquid crystalline bilayers; one of its components, the normal pressure P_N, is parallel to the normal and everywhere equal to the bulk water pressure, whereas the other component, the lateral pressure P_L, is a function of depth, due to the position dependence of forces in the interior of the bilayer. The difference $\pi(Z) = P_L(Z) - P_N$, generally denoted as the lateral pressure profile, is identified with the opposite of a surface tension and describes the local tendency of the bilayer to contract, becoming thicker ($P_L - P_N < 0$), or to expand, becoming thinner ($P_L - P_N > 0$).[13,17–19] Though not easily measurable, the lateral pressure profile across a lipid bilayer can be calculated. General forms can be derived by statistical-thermodynamics models,[9] whereas more specific profiles can be obtained from atomistic simulations.[20] The lateral pressure should be intended as a force per unit area perpendicular to the bilayer plane and varies dramatically with the position across the bilayer, ranging from very high positive to very low negative values within a few nanometers. Three regions are generally identified: (1) the hydrophilic headgroup region, where $P_L > 0$ due to electrostatic and steric interactions and hydration repulsion; (2) the membrane−water interface, where $P_L < 0$ as a consequence of the tendency to decrease the lipid−water interfacial area; and (3) the bilayer interior, where again $P_L > 0$, due to steric repulsions between hydrophobic chains. The stability of the bilayer is determined by the balance of these contributions. The lateral pressure profile is a specific property of the bilayer, which depends on its chemical composition and physical state.[20,21] The integral of the lateral pressure profile across a monolayer is equal to the opposite of its surface tension.[22] In a tensionless bilayer, the surface tension vanishes: repulsion in hydrophobic and hydrophilic regions, mainly due to short-range lipid−lipid interactions and headgroup hydration, is balanced by cohesion, ascribed to the hydrophobic effect, at the hydrophobic/hydrophilic interface. The first moment of the lateral pressure profile is related to the bending elastic modulus and the second moment to the saddle splay modulus.[23–25]

The work required to create a molecular shaped cavity containing a solute at the position Z with orientation Ω is calculated as the integral:[13]

$$U_{cav}(\Omega, Z) = \int_{\text{cavity length}} P_L(Z + z')\, \sigma(\Omega, Z + z')\, dz' \quad (2)$$

where $\sigma(\Omega,Z + z')$ is the surface area of the section of the cavity at the coordinate $(Z+z')$ in the bilayer frame and the integral is over the length of the cavity, along the z' axis of a molecular frame parallel to the bilayer frame, with its origin in Z. $P_L(Z+z')$ is the lateral pressure at the same coordinate $(Z+z')$; in water, it takes a constant value, equal to bulk pressure. Thus, changes in the sign of the lateral pressure inside the bilayer correspond to regions where solute insertion is favorable ($P_L < 0$, i.e., attractive interactions in the bilayer) and unfavorable ($P_L > 0$, i.e., repulsive interactions in the bilayer).

Electrostatic Interactions. The electrostatic contribution to the mean field potential is the work required to charge the solute in its environment. Within the continuum approximation, the molecule is treated as an assembly of charges in a low dielectric molecular cavity embedded in a dielectric medium, and the electrostatic free energy is then evaluated by solving the Poisson equation.[26,27] Approximate expressions for the electrostatic free energy are provided by the so-called generalized Born models.[28,29] They can account for the chemical structure of the solute at a low computational cost, which may be important when repeated calculations must be performed, as is our case, since several solute positions and orientations must be sampled. A difficulty in the case of bilayers derives from the nanoscale nonuniformity, which makes the definition of the dielectric permittivity not obvious.[30] Forms suitable for the membrane environment have been proposed.[31–36] In the approach developed by Tanizaki and Feig, the phospholipid bilayer is modeled as a multilayer composed of a small number of dielectric slabs.[14] When the molecular charge distribution is described as a set of atomic charges located at the nuclear positions, the electrostatic free energy of the solute, in the Ω orientation at the Z position along the bilayer normal, is expressed as

$$U_{el}(\Omega, Z) = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{1}{\varepsilon_{in}} - \frac{1}{\dfrac{\varepsilon^i_{out}(Z_i) + \varepsilon^j_{out}(Z_j)}{2}}\right) \frac{q_i q_j}{\sqrt{r_{ij}^2 + R_i^B R_j^B \exp(-r_{ij}^2/F R_i^B R_j^B)}} \quad (3)$$

where summations are extended to all solute atoms (N), q_i is the atomic charge of the ith atom at the position Z_i, $\varepsilon^i_{out}(Z_i)$ is the local solvent (outer) dielectric constant defined on the basis of the ith atom van der Waals radius,[14] ε_{in} is the solute (internal) dielectric constant, r_{ij} is the distance between the ith and the jth charges, F is an empirical parameter, and R_i^B is the effective Born radius of the ith atom. This is calculated as

$$R_i^{\mathrm{B}} = \frac{1}{C_0 A_{4i} + C_1\left(\frac{3\varepsilon_{\mathrm{out}}(Z_i)}{3\varepsilon_{\mathrm{out}}(Z_i) + 2\varepsilon_{\mathrm{in}}}\right)A_{7i}} + D + \frac{E}{\varepsilon_{\mathrm{out}}(Z_i) + 1} \tag{4}$$

where C_0, C_1, D, and E are empirical parameters and A_{4i} and A_{7i} are atom specific quantities defined as

$$A_{ni} = \left(\frac{1}{(n-3)(R_i^{\mathrm{vdW}})^{n-3}} - \frac{1}{4\pi}\int_{\substack{\text{molecular}\\ \text{volume}\\ r'>R_i^{\mathrm{vdW}}}} \frac{1}{r'^n}\,\mathrm{d}\mathbf{r}'\right)^{1/(n-3)} \tag{5}$$

with R_i^{vdW} being the van der Waals radius of the ith atom.

Dispersion Interactions. The dispersion contribution to the mean field potential is evaluated on the basis of the London expression, which for a pair of polarizable spherical particles, i and j, with volume polarizabilities α_i and α_j, reads $U_{ij}^{\mathrm{disp}} = -^3/_4 I(\alpha_i\alpha_j)/(r_{ij}{}^6)$. Here, r_{ij} is the distance between the centers of the spheres and I is the ionization potential, which is assumed to be equal for the two particles. We represent solute and solvent molecules as assemblies of polarizable spheres and introduce the solvent nonuniformity through the position-dependent density of its components; thus, the dispersion contribution to the solvation free energy of an arbitrary solute in the Ω orientation at the Z position along the bilayer normal is calculated as

$$U_{\mathrm{disp}}(\Omega, Z) = -\frac{3}{4} I \sum_{i=1}^{N}\alpha_i \sum_{j=1}^{M}\alpha_j \int_{\substack{\text{solvent}\\ r'>R_i^{\mathrm{d}}+R_j^{\mathrm{vdW}}}} \frac{\rho_j(Z_i + z')}{r'^6}\,\mathrm{d}\mathbf{r}' \tag{6}$$

where N is the number of atoms in the solute and M is the number of polarizable units in the solvent. A coarse-grained representation of this has been used, whose details are reported in Figure S2 (Supporting Information). The integral extends over the whole solvent volume outside a sphere centered on the ith solute atom, with a radius equal to the sum of an effective radius (R_i^{d}) and the van der Waals radius of the jth solvent component (R_j^{vdW}); $\mathbf{r}' = (x',y',z')$ is a vector position in a local frame, parallel to the bilayer frame, with its origin in the center of the ith atom, located at Z_i in the bilayer frame. In the integral, ρ_j is the number density of the jth polarizable unit of solvent, which is a function of the position along the bilayer normal, defined by the sum $Z_i + z'$. The atomic effective radius R_i^{d} is defined as[37]

$$R_i^{\mathrm{d}} = 1/\sqrt[3]{3}A_{6i} \tag{7}$$

with A_{6i} calculated according to eq 5.

Anisotropic Interactions with Acyl Chains. The contributions to the mean field potential considered so far retain the nonuniformity of the lipid bilayer but neglect the presence of orientational order in the lipid chain region. In the liquid crystal phase, chains are preferentially aligned along an axis (the director **n**), which coincides with the normal **N**, and all properties are axially symmetric with respect to this axis. The degree of alignment is quantified by the chain order parameters, S_{CD}^i, which take their name from the fact that they are obtained from NMR quadrupole splittings for deuteriated acyl chains.[38] They are defined as

$$S_{\mathrm{CD}}^i = \frac{3\langle\cos^2\theta_i\rangle - 1}{2} \tag{8}$$

where θ_i is the angle between the C_iD bond and the bilayer normal (see Figure S1, Supporting Information); the angular bracket denotes a statistical average. If acyl chains were perfectly aligned to the bilayer normal in the all-*trans* conformation, S_{CD}^i would be equal to -0.5 for any i; for randomly oriented chains (as in isotropic liquids), S_{CD}^i would vanish for any i. The S_{CD}^i order parameters are generally negative, since the C_iD bonds preferentially lie perpendicular to the bilayer normal. They exhibit typical profiles, which depend on the phospholipid structure and the physical state of the system; for saturated lipids, $|S_{\mathrm{CD}}^i|$ decreases from relatively high values close to the headgroup to small values on moving toward the chain end.[38]

As a consequence of the chain order, anisotropic interactions are not totally washed out by the orientational average in the bilayer environment. It is important to realize that the interactions which we are considering here have axial symmetry, resulting from the partial alignment of chain segments, though they do not have the polar character originated by the presence of an interface. A related property is birefringence of lipid bilayers; the refractive index parallel to the bilayer normal (n_{e}) is higher than the perpendicular one (n_{o}),[39,40] which means that the average polarizability is higher perpendicular than parallel to the bilayer. To account for the anisotropy of the interactions of solutes with acyl chains, we have extended to bilayers a model which was originally proposed for thermotropic nematic liquid crystals.[15] This is a phenomenological molecular field model which has been proven to be able to capture the relation between molecular structure and orientational order in nematics and has been successfully used to predict the dependence on molecular structure of diverse properties of nematic liquid crystals.[41−43] The model, which can be considered an extension of the Maier−Saupe theory[44] allowing for the account of the molecular shape, bases on the assumption that each element of the surface of a molecule in the nematic phase (of symmetry $D_{\infty h}$) preferentially aligns to the director. The orienting potential, dU_{ord}, acting on each infinitesimal element dS of the molecular surface, is a function of the angle θ between its normal **s** and the director **n** and can be expanded on the basis of Legendre polynomials.[45] Given the symmetry of the system, the first nonvanishing term of this expansion is proportional to the second Legendre polynomial, thus

$$dU_{\mathrm{ord}} = k_{\mathrm{B}}T\xi P_2(\cos\theta)dS \tag{9}$$

where k_{B} is the Boltzmann constant, T is the temperature, and ξ a positive parameter, which quantifies the orienting strength of the environment; it depends on the degree of orientational order and vanishes in the isotropic phase. The overall orienting potential experienced by a molecule in the nematic phase is then obtained by integrating eq 9 over the molecular surface, S:

$$U_{ord}(\Omega) = k_B T \xi \int_S P_2(\cos\theta)\, dS \quad (10)$$

It can be easily seen that, in the model case of a cylindrical particle with lateral area S_{cyl}, the Maier−Saupe form[44] is recovered, $U_{ord}(\beta) = -c_2 P_2(\cos\beta)$, where β is the angle between the cylinder axis and nematic director and $c_2 = k_B T \xi S_{cyl}/2$:[46] the particle preferentially orients with its axis along the nematic director, with no distinction between parallel and antiparallel orientation. According to the molecular field theory, the orienting strength can be related to the order parameters of the nematic phase;[47] if mesogens are approximated as cylindrical particles, the relationship $\xi \propto \langle P_2 \rangle$ holds, where $\langle P_2 \rangle$ is the second rank order parameter, which specifies the degree of order of cylinders.[15] It may be worth remarking that the mean field potential, eq 10, is experienced by all molecules in the nematic system, both mesogens and solutes, if present. The specific form of this potential will depend on the molecular structure, through the molecular surface. On the contrary, the orienting strength ζ is a phase property; at a low concentration of solutes, it is nearly the same as for the pure nematic solvent and is proportional to solvent order parameters.

Extension of the model to lipid bilayers requires the introduction of a position-dependent orienting strength, $\xi = \xi(Z)$; so the orienting contribution to the mean field potential experienced by a solute in a lipid bilayer is expressed as

$$U_{ord}(Z, \Omega) = k_B T \int_S \xi(Z + z') P_2(\cos\theta)\, dS \quad (11)$$

where z' is the position of the surface element dS along an axis parallel to the normal **N**, in a molecular frame parallel to the bilayer frame, with its origin in Z. The profile of the orienting strength, $\xi(Z)$ is a specific function of the bilayer, which depends on the lipid composition and temperature. Extending to the bilayer the same considerations valid for nematics, we can assume a linear relationship between orienting strength ξ and S^i_{CD} order parameters: the orienting strength at the coordinate $\langle Z_i \rangle$, equal to the average position, along the bilayer normal, of the C_i atom, is expressed as $\xi(\langle Z_i \rangle) = -(A/k_B T) S^i_{CD}$, where A is a positive constant. Given the S^i_{CD} order parameters, which are available from experiments or simulations, the values of A and $\langle Z_i \rangle$ can be determined as explained in the Supporting Information.

Density Probability for Solute Position and Orientation. For any molecular position along the bilayer normal and any molecular orientation in the bilayer frame of reference, the mean field potential $U(\Omega,Z)$ is calculated as the sum of the contributions expressed by eqs 2, 3, 6, and 11. Under the conditions of the canonical ensemble, the coupled positional-orientational distribution function can then be calculated as

$$P(\Omega, Z) = \frac{\exp(-U(\Omega, Z)/k_B T)}{Q} \quad (12)$$

with the partition function

$$Q = \int_0^L dZ \int \exp(-U(\Omega, Z)/k_B T)\, d\Omega \quad (13)$$

where T is the temperature, k_B is the Boltzmann constant, and $\int ... d\Omega = \int_0^{2\pi} d\gamma \int_0^{\pi} ... \sin\beta\, d\beta$. The normalization condition $\int_0^L dZ \int P(\Omega,Z)\, d\Omega = 1$ holds. The integral over the Z variable is taken over the bilayer half-thickness, under the assumption that the midplane is a symmetry plane. The upper integration limit, L, depends on the system under investigation. For the calculation of all those properties which only depend on the solute behavior within the bilayer, but are unaffected by the relative amount of solute in the water and bilayer, we can take as a reasonable choice for L a value large enough to ensure that for $Z = L$ the solute behavior can be assumed to be the same as in bulk water.

We can then define the reduced probability distribution function:

$$P_\Omega(Z) = \frac{Q_\Omega(Z)}{Q} \quad (14)$$

where the orientational partition function is defined as

$$Q_\Omega(Z) = \int \exp(-U(\Omega, Z)/k_B T)\, d\Omega \quad (15)$$

so that the normalization condition $\int_0^L P_\Omega(Z)\, dZ = 1$ holds. The reduced distribution function $P_\Omega(Z)$ can be expressed in the Boltzmann form:

$$P_\Omega(Z) = \frac{\exp(-u(Z)/k_B T)}{\int_0^L \exp(-u(Z)/k_B T)\, dZ} \quad (16)$$

with

$$u(Z) = \text{const} - k_B T \ln Q_\Omega(Z) \quad (17)$$

which can be interpreted as the transfer free energy of the solute from water to the position Z in the bilayer, if the constant is taken equal to the value of $k_B T \ln Q_\Omega$ in bulk water.

Once the coupled positional-orientational distribution is known, the average value $\langle f \rangle$ of any arbitrary function $f(\Omega,Z)$ is defined as

$$\langle f \rangle = \int_0^L dZ \int f(\Omega, Z)\, P(\Omega, Z)\, d\Omega \quad (18)$$

Partially averaged functions may also be meaningful; for instance, the orientationally averaged value of the function $f(\Omega,Z)$ can be calculated as

$$\langle f \rangle_\Omega(Z) = \frac{1}{P_\Omega(Z)} \int f(\Omega, Z)\, P(\Omega, Z)\, d\Omega \quad (19)$$

Distribution Functions for Flexible Solutes. The model can be easily extended to flexible solutes. Let us denote as χ the torsional degrees of freedom of the solute; then, the mean field potential, eq 1, will become a function of the molecular conformation, $U(\Omega,Z,\chi)$. If $V(\chi)$ is the torsional potential of the molecule in a vacuum, the coupled orientational-positional-torsional density probability of the solute is defined as

$$P(\Omega, Z, \chi) = \frac{\exp(-V(\chi)/k_BT)\exp(-U(\Omega, Z, \chi)/k_BT)}{Q} \quad (20)$$

with the partition function

$$Q = \int_0^L dZ \int d\Omega \int \exp(-V(\chi)/k_BT)\exp(-U(\Omega, Z, \chi)/k_BT)\, d\chi \quad (21)$$

and the normalization condition $\int_0^L dZ \int d\Omega \int P(\Omega,Z,\chi)\, d\chi = 1$.

In analogy with what is shown above, we can define a reduced position distribution function $P_{\Omega,\chi}(Z)$:

$$P_{\Omega,\chi}(Z) = \frac{Q_{\Omega,\chi}(Z)}{Q} \quad (22)$$

with

$$Q_{\Omega,\chi}(Z) = \int d\Omega \int \exp(-V(\chi)/k_BT)\exp(-U(\Omega, Z, \chi)/k_BT)\, d\chi \quad (23)$$

and the normalization condition $\int_0^L P_{\Omega,\chi}(Z)\, dZ = 1$. In analogy with eq 17, we can define the free energy:

$$u(Z) = \text{const} - k_BT \ln Q_{\Omega,\chi}(Z) \quad (24)$$

which includes entropic effects deriving from rotational and torsional restrictions in the bilayer environment.

The effect of the environment on the conformational distribution can be analyzed in terms of the reduced conformational distribution functions, obtained by averaging over the orientational and positional degrees of freedom:

$$P_{\Omega,Z}(\chi) = \frac{Q_{\Omega,Z}(\chi)}{Q} \quad (25)$$

with

$$Q_{\Omega,Z}(Z, \chi) = \exp(-V(\chi)/k_BT) \int_0^L dZ \int e^{-U(\Omega,Z,\chi)/k_BT} d\Omega \quad (26)$$

and the normalization condition $\int P_{\Omega,Z}(\chi) d\chi = 1$. Equation 25 can also be expressed as:

$$P_{\Omega,Z}(\chi) = \frac{\exp(-V'(\chi)/k_BT)}{\int \exp(-V'(\chi)/k_BT)\, d\chi} \quad (27)$$

with

$$V'(\chi) = \text{const} - k_BT \ln Q_{\Omega,Z}(\chi) \quad (28)$$

If the constant is taken such that $V'(\chi) = V(\chi)$ for $U(\Omega,Z,\chi) = 0$, eq 28 defines an effective torsional potential in the water/bilayer environment. Given the nonuniform character of this system, a more meaningful property should retain the position dependence. A depth-dependent effective torsional potential $V'(Z,\chi)$ can be defined, from the reduced distribution function:

$$P_\Omega(Z, \chi) = \frac{Q_\Omega(Z, \chi)}{Q} \quad (29)$$

with

$$Q_\Omega(Z, \chi) = \exp(-V(\chi)/k_BT) \int \exp(-U(\Omega, Z, \chi)/k_BT)\, d\Omega \quad (30)$$

and the normalization condition $\int_0^L dZ \int P_\Omega(Z,\chi)\, d\chi = 1$. The effective torsional potential is then defined as

$$V'(Z, \chi) = c(Z) - k_BT \ln Q_\Omega(Z, \chi) \quad (31)$$

with $c(Z)$ chosen in such a way that $V'(Z,\chi) = V(\chi)$ for $U(\Omega,Z,\chi) = 0$.

The average value $\langle f \rangle$ of any arbitrary function $f(\Omega,Z,\chi)$ can be calculated as

$$\langle f \rangle = \int_0^L dZ \int d\Omega \int f(\Omega, Z, \chi)\, P(\Omega, Z, \chi)\, d\chi \quad (32)$$

The partially averaged function can also be defined, in analogy with eq 19. In particular, the positional-orientational average for a given conformation can be calculated as

$$\langle f \rangle_{\Omega,Z}(\chi) = \frac{1}{P_{\Omega,Z}(\chi)} \int_0^L dZ \int f(\Omega, Z)\, P(\Omega, Z, \chi)\, d\Omega \quad (33)$$

This is the property that would be calculated for the molecule frozen in the conformation specified by the χ dihedrals. The full average, eq 32 can be obtained from partial averages, eq 33, as $\langle f \rangle = \int \langle f \rangle_{\Omega,Z}(\chi)\, P_{\Omega,Z}(\chi)\, d\chi$.

If the torsional potential is characterized by localized minima, separated by barriers higher than few k_BT units, the molecular flexibility can be described in terms of a finite number of conformers, corresponding to minima of the potential energy.[48] Then, the single conformer orientational-positional distribution function

$$P_J(\Omega, Z) = \frac{\exp(-V_J/k_BT)\exp(-U_J(\Omega, Z)/k_BT)}{Q} \quad (34)$$

can be introduced, where the index J is used to indicate functions calculated for the Jth conformer ($\chi = \chi_J$), and integration over the χ dihedrals can be replaced by summation over conformers; for instance,

$$Q = \sum_J \int_0^L dZ \int \exp(-V_J/k_BT)\exp(-U_J(\Omega, Z)/k_BT)\, d\Omega$$

Numerical Methods

Computation of each of the contributions to the mean field potential, eq 1, requires a definition of the solute molecular surface. The solvent-excluded surface,[49] calculated as a triangulated surface according to the Sanner algorithm,[50] is used. The expressions used to calculate eqs 2, 5, 6, and 11 are reported in the following. In these expressions N_f represents the number of triangular faces, $\mathbf{s}_J$, with components $s_{J,X}$, $s_{J,Y}$, and $s_{J,Z}$ in the bilayer frame, is the outward pointing unit vector normal to the Jth face, and S_J is the surface area of this face.

The volume integral in eq 2 is calculated as the sum over N_s slices of thickness $\Delta z'$ perpendicular to the bilayer normal:

$$\int_{\substack{\text{molecular}\\ \text{length}}} \pi(Z+z')\,\sigma(\Omega, Z+z')\,dz' = \sum_{i=1}^{N_s} \pi(Z+z'_i)\,\sigma(\Omega, Z+z'_i)\,\Delta z' \quad (35)$$

The section area, $\sigma(\Omega,Z + z'_i)$, is calculated from the perimeter of the λ-sided polygon, defined by the intersection of the triangulated surface with the ith plane, according to Surveyor's formula for non self-intersecting polygons:[51]

$$\sigma(Z+z'_i) = \left|\frac{1}{2}\sum_{\alpha=1}^{\lambda}(x'_\alpha y'_{\alpha+1} - x'_{\alpha+1}y'_\alpha)\right| \quad (36)$$

where (x'_α,y'_α), with $\alpha = 1-\lambda$, are the coordinates of the polygon vertices, ordered according to consecutive sides, with the boundary condition $(x'_{\lambda+1},y'_{\lambda+1}) = (x'_1,y'_1)$.

The volume integrals A_{ni}, defined in eq 5, are conveniently rewritten in the form of surface integrals, using Gauss's theorem:[52]

$$A_{ni} = \left[\frac{1}{n-3}\left(\frac{1}{(R_i^{\mathrm{vdW}})^{n-3}} - \frac{1}{4\pi}\int_{\substack{\text{molecular}\\ \text{volume}\\ r>R_i^{\mathrm{vdW}}}}\left(-\nabla\cdot\frac{\mathbf{r}}{r^n}\right)d\mathbf{r}\right)\right]^{1/(n-3)} = \left[\frac{1}{4(n-3)\pi}\int_S \frac{\mathbf{r}}{r^n}\cdot\mathbf{s}\,dS\right]^{1/(n-3)} \quad (37)$$

The integral is then calculated as a sum over the triangular faces:

$$A_{ni} = \left(\frac{1}{4(n-3)\pi}\sum_{J=1}^{N_f}\frac{\mathbf{r}_{iJ}}{r_{iJ}^n}\cdot\mathbf{s}_J S_J\right)^{1/(n-3)} \quad (38)$$

where $\mathbf{r}_{iJ}$ is the vector from the center of the ith atom to the center of mass of the Jth face.

Due to the axial symmetry of the bilayer, the volume integral in eq 6 is conveniently calculated in cylindrical coordinates (z',q,ϕ):

$$\begin{aligned}\int_{\text{solvent } r>R_{ij}} \frac{\rho_j(Z_i+z')}{r'^6}\,d\mathbf{r}' &= \int_{R_{ij}}^{z'_{\max}}\rho_j(Z_i+z')\,dz'\int_0^\infty \frac{q}{(q^2+z'^2)^3}\,dq\int_0^{2\pi}d\phi\\ &+ \int_{-R_{ij}}^{R_{ij}}\rho_j(Z_i+z')\,dz'\int_{\sqrt{R_{ij}^2-z'^2}}^\infty \frac{q}{(q^2+z'^2)^3}\,dq\int_0^{2\pi}d\phi\\ &+ \int_{-z'_{\max}}^{-R_{ij}}\rho_j(Z_i+z')\,dz'\int_0^\infty \frac{q}{(q^2+z'^2)^3}\,dq\int_0^{2\pi}d\phi\\ &= \frac{\pi}{2}\left[\int_{R_{ij}}^{z'_{\max}}\frac{\rho_j(Z_i+z')}{z'^4}\,dz' + \frac{1}{R_{ij}^4}\int_{-R_{ij}}^{R_{ij}}\rho_j(Z_i+z')\,dz' + \int_{-z'_{\max}}^{-R_{ij}}\frac{\rho_j(Z_i+z')}{z'^4}\,dz'\right]\end{aligned} \quad (39)$$

where $R_{ij} = R_i^d + R_j^{\mathrm{vdW}}$. Under the assumption that the size of the sample is much larger than R_{ij}, the integrals over the q variable are calculated by extending to infinity the upper integration limit. In our calculations, $z'_{\max} = 4R_{ij}$ was taken. The three integrals in the final expression are calculated with the composite midpoint formula.[53]

The surface integral in eq 11 is simply calculated as the sum over all triangular faces:

$$\int_S \xi(Z+z')\,P_2(\cos\theta)\,dS = \sum_{J=1}^{N_f}\xi(Z+z'_J)\frac{3s_{J,Z}^2-1}{2}S_J \quad (40)$$

where the position of the Jth face is identified with that of its center of mass.

Free Energy Profiles and Average Quantities. Gauss−Legendre and Gauss−Chebyshev quadrature algorithms were used for numerical integration over the orientational variables β and γ, respectively.[53] Numerical integration over the position variable Z was simply performed according to the composite midpoint formula; calculations for cholesterol in DPPC were performed giving the upper integration limit the value L = 45 Å. Convergence of numerical integrals was checked for free energy profiles and other average properties. A typical calculation, using 12 β, 24 γ, and 91 Z points, takes on the order of hundreds of seconds of CPU time on a 2 GHz desktop PC.

Cholesterol in DPPC

Bilayer Parametrization. Figure 2 shows the profiles of bilayer properties needed for the calculations. Lateral pressure,[54] mass density,[55] and S_{CD} order parameter[56] profiles were taken from molecular dynamics simulations of a DPPC bilayer in the liquid crystal phase at the temperature T = 323 K. The definition of polarizable units and the corresponding mass density profiles, appearing in eq 6, are reported in Figure S2 (Supporting Information); the value $I = 2\times10^{-18}$ J[57] was assumed in this equation. The outer dielectric constant profiles, used in eqs 3 and 4, were calculated on the basis of a three-dielectric model, as proposed in ref 14. The empirical parameters appearing in eq 4 were determined by matching electrostatic free energies, calculated according to eq 3, to those calculated by the finite difference solution of the Poisson equation, for a test set of molecules (see SI-3, Supporting Information). Parts of the same molecules were used to check calculated water/*n*-octanol transfer free energies against experimental values (see SI-4, Supporting Information). Given the S_{CD} order parameters, the profile of the orienting strength, $\xi(Z)$, entering eq 11, was obtained by nonlinear fitting of the values calculated at the average acyl chain carbon positions, $\xi(\langle Z_i\rangle)$ (see SI-1, Supporting Information).

Solute Structural Data. Nine conformers of cholesterol differing in the alkyl chain conformation were taken for our calculations (see Figure S6, Supporting Information). Conformer energy and geometry were obtained by DFT optimization in a vacuum at the B3LYP/6-31g** level.[58] A common molecular frame was chosen for all conformers, with the origin at the nuclear position of the oxygen atom

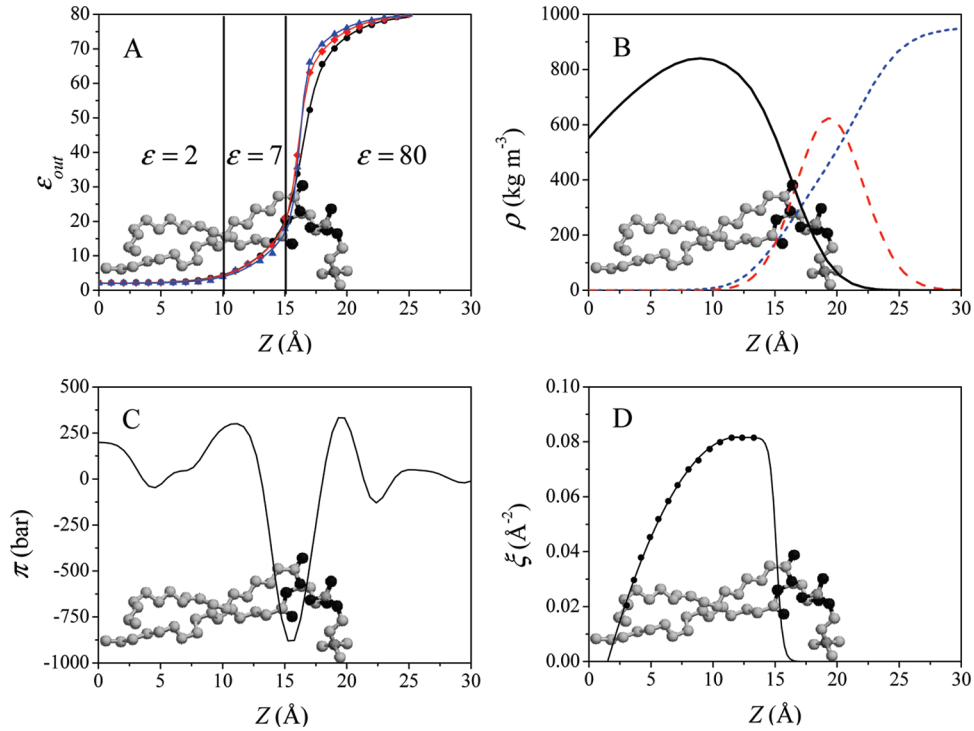


Figure 2. DPPC properties used in the calculations, shown as a function of the distance from bilayer midplane. (A) Local dielectric constant profile calculated with a three-dielectric model, as described in ref 14, for ions of van der Waals radius equal to 1 Å (triangles), 1.85 Å (circles), and 1.5 Å (squares). (B) Mass density profile for lipid tails (solid line), lipid headgroups (dashed line), and water (dotted line).[55] (C) Lateral pressure profile.[54] (D) Orienting strength profile from S_{CD} order parameters for the *sn*1 acyl chain.[56] All of the data refer to the temperature $T = 323$ K. A phospholipid molecules is superimposed on each plot, to indicate different bilayer regions.

Table 1. Parameters Used in Calculations for Cholesterol

van der Waals radii	H, 1.00 Å; C, 1.85 Å; O, 1.50 Å[59]
Connolly surface	probe sphere radius =3 Å; vertex density =13 Å^{-2}
atomic charges	OPLS force field[60] (see Table S1, Supporting Information)
atomic polarizabilities ($4\pi\varepsilon_0$ Å^3)	from bond polarizabilities[57] [a] H−C: 0.65; C−C: 0.57; C=C: 1.65; C−O: 0.64; O−H: 0.73
internal dielectric constant	$\varepsilon_{in} = 2$[14]
Born empirical parameters	$F = 8$; $C_0 = 0.3028$; $C_1 = 1.009$; $D = 0.00$ Å^{-1}; $E = -0.16$ Å^{-1} [b]

[a] Half of a bond polarizability is assigned to each of the two connected atoms; hence the polarizability of an atom is obtained as half the sum of the polarizabilities of all the bonds in which it is involved. [b] Obtained from optimization with respect to the finite difference solution of the Poisson equation, for a test set of benzene derivatives (see SI-3, Supporting Information), according to the procedure reported in ref 61.

and the (x,y,z) axes fixed to the rigid core, as shown in Figure 1. Solute parameters used for calculations are summarized in Table 1.

Results

Free Energy and Distributions. Figure 3 displays the transfer free energy profile and the corresponding position distribution function, calculated by averaging over all nine conformers of cholesterol. The transfer free energy profile exhibits a marked position dependence, not only near the water/bilayer interface but also in the hydrocarbon region. The free energy minimum, which is about 13 kcal/mol lower than the free energy in bulk water, is reached when cholesterol is inserted in the bilayer, with the oxygen at a distance of about 18 Å from the midplane. The position distribution function $P_{\Omega\chi}$, shown in Figure 3, indicates that cholesterol, though confined with the oxygen atom near the hydrophobic/hydrophilic interface, may undergo wide-amplitude fluctuations parallel to the bilayer normal. Both the average position and the vertical distribution of cholesterol are in agreement with experimental data. By recent neutron scattering experiments, the center of mass of deuteriated sites of [2,2,3,4,4,6-$^2H_6$]-labeled cholesterol, in the liquid crystalline phase of different lipid bilayers, was found to be confined in a region a few Ångstroms wide, centered at approximately 16 Å from the midplane.[62] Vertical motions of cholesterol wider than 5 Å were also revealed by quasielastic neutron scattering experiments in 40% cholesterol:DPPC bilayers in the liquid ordered phase.[63]

The tail flexibility has negligible effects on the free energy profile of cholesterol (see Figure S7, Supporting Information), which is not surprising, given the small size of the alkyl chain in comparison to the broad ring region and the reduced conformational space explored by the tail, due to steric constraints. For this reason in the following, until differently stated, we will simply refer to the results obtained

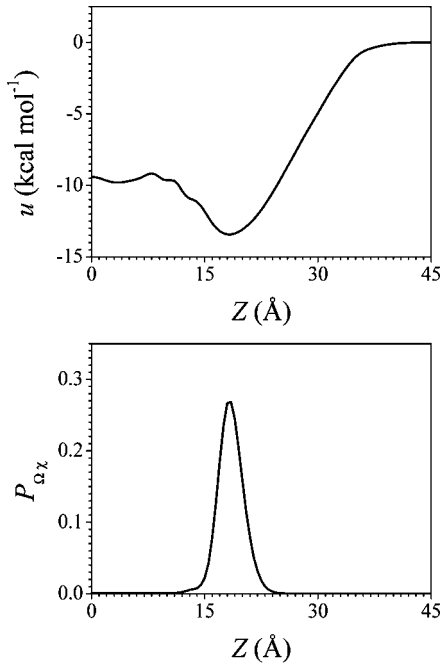


Figure 3. Transfer free energy (top) and position distribution (bottom) profiles of cholesterol as a function of distance of the oxygen atom from the bilayer midplane. The results shown here were obtained as averages over all conformers at the temperature $T = 323$ K.

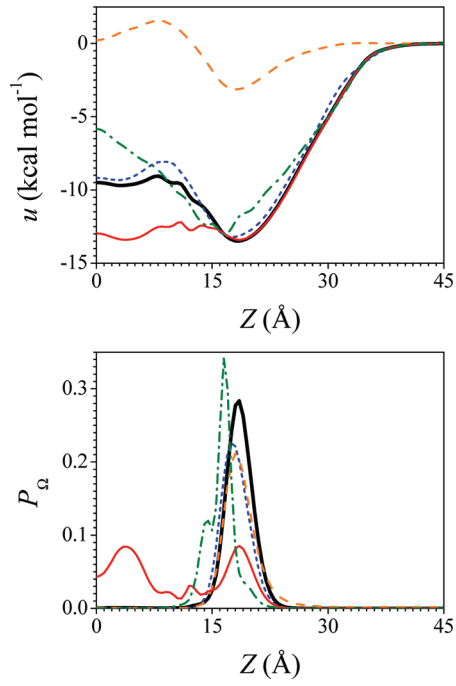


Figure 4. Transfer free energy (top) and position distribution (bottom) profiles calculated for cholesterol as a function of the distance of the oxygen atom from the bilayer midplane, using an incomplete form of the mean field potential $U(\Omega,Z)$. Profiles were obtained by excluding in turn the single cavity (thin dotted line), the electrostatic (thin solid line), the dispersion (thin dashed line), and the chain order (thin dash-dotted line) contributions in eq 1. The profiles from the complete form of the potential (thick solid line) are also shown for comparison. The results shown here were obtained for the B conformer, at the temperature $T = 323$ K.

for the B conformer of cholesterol (see Figure S6, Supporting Information).

With the purpose of disentangling the role of the different contributions to the mean field potential, we show in Figure 4 the free energy profiles and the corresponding position distribution functions, obtained by excluding a single contribution at a time in eq 1. The figure clearly shows that dispersion interactions bear the main responsibility for the free energy decrease within the bilayer, in agreement with the simple picture of cholesterol brought into the bilayer by favorable interactions with lipid chains. Electrostatic interactions, which in the case of cholesterol mostly involve the OH group, have the effect of destabilizing upside-down configurations, with this group located deep in the bilayer interior. The effects of the order and cavity contribution to the mean field potential will be clearer below, when the order parameters of cholesterol are discussed; however, something can be anticipated in relation to the transfer free energy and the distribution profiles shown in Figure 4. The order contribution promotes alignment of the long axis of cholesterol to the director, i.e., the average orientation of acyl chains, which in the liquid crystalline phase of DPPC coincides with the bilayer normal **N**. The small shift of the oxygen atom toward the lipid headgroups, which we can observe in Figure 4 upon inclusion of the order contribution in the mean field potential, can be explained by the fact that it allows a larger portion of the broad and thick core of cholesterol to lie in the most ordered lipid region (it may be useful to look at the molecular surface of cholesterol, which is shown in Figure S8, Supporting Information). Given the strong orienting effect of the ordering term, the cavity contribution, which would bring about the insertion of cholesterol flat at the hydrophobic/hydrophilic interface, where there is a deep well in the lateral pressure, seems to play a minor role.

The free energy profile shown in Figure 3 compares well with that reported in Figure 1b of ref 64, which was obtained from long all-atom molecular dynamics simulations (3-μs-long trajectories) of cholesterol in DPPC. Perhaps more interestingly, the relatively small differences between the free energy profiles obtained from all-atom simulations of cholesterol and its analog missing the hydroxyl group in DPPC, displayed in Figure S5 of the same reference,[64] closely resemble the differences arising from removing electrostatic interactions in our model, as shown in Figure 4. There is however a scale factor: the transfer free energy calculated with our model is about 30% lower, in absolute value, than that obtained from simulations, which probably reflects a different parametrization of interactions. In fact, discrepancies of this magnitude are not uncommon even between simulations using different force fields.

A major feature of our model is the inclusion of orientation−position coupling inside the lipid bilayer. Figure 5 shows the map of the mean field potential, calculated in correspondence of the maximum of the position distribution probability, as a function of the Euler angles (β,γ), defining the molecular orientation. The figure shows a strong dependence of the potential on the Euler angle β between the long

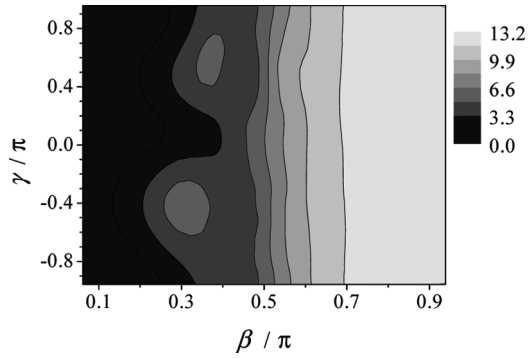


Figure 5. Contour plot of the mean field potential *U*, eq 1, as a function of the Euler angles (β,γ) defining the solute orientation in the bilayer frame (Figure 1), calculated in correspondence of the maximum of the position distribution. The results shown here were obtained for the B conformer of cholesterol at the temperature $T = 323$ K. Energies are expressed in kcal mol^{-1}, and a value equal to zero is given to the minimum of the mean field potential in the plot.

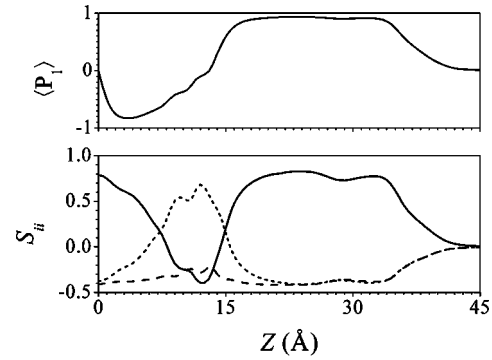


Figure 6. Local orientational order parameters calculated for cholesterol as a function of the distance of the oxygen atom from the bilayer midplane. The plot at the bottom shows the diagonal elements of the Saupe tensor, calculated in the molecular frame (Figure 1): S_{xx} (dotted line), S_{yy} (dashed line), and S_{zz} (solid line). The results shown here were obtained for the B conformer, at the temperature $T = 323$ K.

axis of cholesterol and the bilayer normal (see Figure 1), with the energy increasing as cholesterol moves from parallel to antiparallel. The dependence of the mean field potential on the γ angle indicates that cholesterol has some preference to keep the molecular plane (i.e., the *xz* plane) parallel to the bilayer normal.

A complete insight into the orientational behavior of the solute is provided by orientational order parameters. The order parameter $\langle P_1 \rangle = \langle \cos \beta \rangle$ describes the polarity of the orientational distribution; it can take values ranging from -1 to $+1$, with the extremes corresponding to the molecular *z* axis antiparallel and parallel to the bilayer normal, respectively, going through 0 for an apolar distribution. The second rank order parameters $S_{ii} = (3\langle \cos^2 \theta_{iZ} \rangle - 1)/2$, with θ_{iZ} denoting the angle between the *i*th molecular axis ($i = x,y,z$) and the bilayer normal, quantify the degree of alignment of molecular axes to the bilayer normal. Unlike $\langle P_1 \rangle$, the S_{ii} order parameters do not distinguish between parallel and antiparallel orientations; they range between -0.5 (*i* axis perpendicular to the director) and $+1$ (*i* axis parallel to the director). Figure 6 shows the profile of the order parameters $\langle P_1 \rangle$ and S_{ii} calculated for the B conformer of cholesterol. Here, we can distinguish at least three zones: (i) a first region where cholesterol is in upright orientations, well aligned to the bilayer normal **N** with its head in water or in the hydrophilic region, pointing toward water; (ii) a second region where the oxygen atom is close to the midplane and the long axis is again aligned to **N**; (iii) an intermediate region, where cholesterol lies inside the bilayer, with its long axis preferentially perpendicular to **N** and the molecular plane perpendicular to the bilayer plane. The result is a clear indication of the peculiar effects of the nanoscale organization of lipid bilayers. It must be remarked that the probability that cholesterol resides in the ii or in the iii regions is negligible, as shown in Figure 3. This is in agreement with the absence of experimental evidence for cholesterol lying flat inside DPPC bilayers. However, we can speculate that such orientations may be favored under different conditions, e.g., when the nature of phospholipids leads to a different balance of forces within the bilayer. Indeed, the existence of the horizontal orientation of cholesterol in polyunsaturated phospholipids was recently demonstrated by neutron scattering investigations.[62,65] Some preference for orientations parallel to the bilayer plane was also found in atomistic simulations of cholesterol, sitting in scarcely populated regions inside DPPC or other lipid bilayers.[64]

With the help of Figure 7, which shows the order parameters calculated by excluding one contribution at a time, we can now distinguish the different contributions to the orientational order of cholesterol. From inspection of the plots in this figure, we can draw the following picture. The anisotropy of interactions of cholesterol with acyl chains bears the main, although not the only, responsibility for the orientational order of cholesterol across the bilayer. In its absence (Figure 7D), cholesterol would be significantly ordered in two zones, by virtue of dispersion and the cavity contributions. In the former region, part of cholesterol, comprising the oxygen atom, lies in water, and alignment of the long axis to the normal allows the molecule to insert as much as possible of its body into the bilayer. In the latter region, the drop in lateral pressure at approximately 15 Å from the midplane and the optimization of the dispersion interaction when a large portion of the solute resides in the region around 10 Å promote horizontal orientation of the long molecular axis. Inclusion of the anisotropic interactions with acyl chains increases the degree of alignment of the long axis of cholesterol, when this is parallel to the bilayer normal, and differently stabilizes the orientations of the molecular plane, when the long axis lies perpendicular to the normal: orientations with this plane parallel are favored over those with the plane perpendicular to lipid chains. Electrostatic interactions, which have the main responsibility for keeping the polar group of cholesterol in the hydrophilic region, are not very important for the orientational order of this molecule.

Accurate information on orientational order is obtained from quadrupole splittings in NMR spectra of specifically deuteriated samples. We have calculated the order parameters

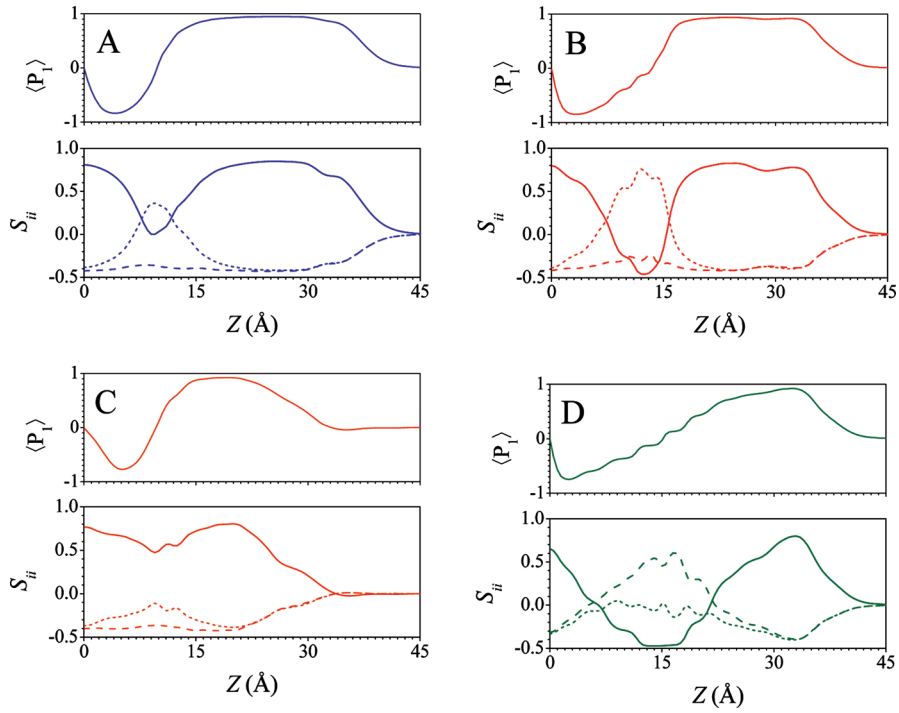


Figure 7. Local orientational order parameters calculated for cholesterol as a function of the distance of the oxygen atom from the bilayer midplane, using an incomplete form of the potential $U(\Omega,Z)$ in eq 1 of the main text. The results shown here were obtained by excluding in turn the cavity (A), the electrostatic (B), the dispersion (C), and the chain order (D) contributions in eq 1. For each case, the plot at the top displays the orientational average of the first Legendre polynomial, $\langle P_1 \rangle$, whereas that at the bottom shows the diagonal elements of the Saupe tensor, calculated in the molecular frame (Figure 1): S_{xx} (dotted line), S_{yy} (dashed line), and S_{zz} (solid line). The results shown here were obtained for the B conformer at the temperature $T = 323$ K.

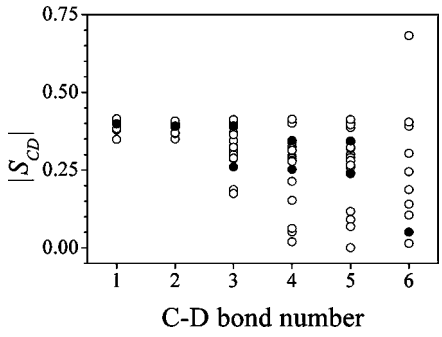


Figure 8. $|S_{CD}|$ order parameters, calculated for the alkyl chain of cholesterol at the temperature $T = 323$ K. Full symbols: averages over conformers. Open symbols: individual conformers. Bonds are numbered as follows: 1 for C_{17}−D, 2 for C_{20}−D, 3 for C_{22}−D, 4 for C_{23}−D, 5 for C_{24}−D, and 6 for C_{25}−D (atom numbering as in Figure S5, Supporting Information).

for the C_3H bond and the tail CH bonds (see Figure S5, Supporting Information), defined as in eq 8, which can be directly compared with experimental data. The value $S_{CD} = -0.33$ was obtained for the C_3H bond after averaging over all nine conformers; the negative sign indicates the tendency of this bond to lie perpendicular to the bilayer normal. Figure 8 shows the position dependence of the $|S_{CD}|$ order parameters calculated for the alkyl chain CH bonds of cholesterol. In addition to the order parameters obtained as averages over all nine conformers, those calculated for the single conformers are shown; the former and the latter would correspond to experimental data for fast and slow conformational motions (on the NMR time scale), respectively. We can see that, for bonds between C_{22} and C_{24}, significantly different order parameters are predicted for the nine conformers, due to their different chain geometry. We can also see that, even after the conformational average, different order parameters are obtained for the pairs of deuterons at C_{22}, C_{23}, and C_{24}; this means that deuterons linked to the same carbon atom have a different average orientation in the bilayer. Quadrupole spectra of cholesterol specifically deuterated at C_{24} were reported for a wide range of temperatures in the liquid crystal phase of 3:7 (molar ratio) cholesterol/DMPC mixtures.[67] From the quadrupole splittings, 33.2 and 42 kHz, measured at about 10 °C above the melting temperature, $|S_{CD}|$ values equal to 0.26 and 0.33, respectively, can be estimated, assuming a deuterium quadrupole coupling constant equal to 170 kHz.[68] The average order parameters calculated for the two C_{24}−D bonds, S_{CD}=−0.24 and $S_{CD} = -0.34$, are in line with experimental results. As for the C_3H bond, the quadrupole splitting $\Delta\nu \sim 84$ kHz was reported in ref 66 for 5 mol % cholesterol in DPPC at the temperature $T = 47$ °C, from which the value $|S_{CD}| \sim 0.33$ is derived, assuming again a deuterium quadrupole coupling constant equal to 170 kHz. The agreement between predicted and experimental order parameters for different sites is a good check for the ability of our model to predict the orientational behavior of cholesterol in the bilayer environment.

Finally, we report in Table 2 a list of properties calculated for cholesterol by averaging over all degrees of freedom (translational, orientational and conformational), according to eq 32. Such averages make sense in the case of cholesterol,

Table 2. Average Properties Calculated for Cholesterol in DPPC Bilayer at the Temperature $T = 323$ K[a]

$\langle Z \rangle$	$\langle P_1 \rangle$	S_{XX}	S_{YY}	S_{ZZ}
18.5 Å	0.86	−0.31	−0.45	0.76

[a] $\langle Z \rangle$ is the average distance of the oxygen atom from the midplane. $\langle P_1 \rangle$ is the average value of the cosine of the angle between the bilayer normal and the z molecular axis (Figure 1), and S_{ii} ($i = x,y,z$) are the principal values of the Saupe matrix.

which is predicted to be mostly confined in a region with uniform orientational behavior within the bilayer. The properties reported in Table 2 are close to those calculated for the single A or B conformers (see Figure S6, Supporting Information); these are very similar to each other and are predicted to be significantly more probable than the other conformers. In fact, not only A and B are, according to DFT calculations, the most stable in a vacuum, but they are further stabilized by the bilayer environment, as shown by the values of the difference $\Delta V_J = V'_J - V_J$, calculated according to eq 28 and reported in Figure S6. The very high $\langle P_1 \rangle$ value, close to 1, indicates the strong preference of cholesterol for orientations parallel to the bilayer normal with the hydroxyl group pointing toward water, over antiparallel orientations. The preference for parallel over perpendicular orientation of the long molecular axis is quantified by the $S_{z'z'}$ value. The principal axis system of the Saupe matrix (x',y',z'), shown in Figure S9 (Supporting Information), does not exactly coincide with the molecular frame (x,y,z); in particular, the major alignment axis (z') is tilted by about 14° with respect to the molecular z axis (shown in Figure 1). Of the two other principal alignment axes, y', nearly perpendicular to the molecular plane, is predicted to have a higher tendency to lie perpendicular to the bilayer normal than x'.

Conclusions

In a recent review, the need of new implicit models of membranes, combining an atomistic solute representation with a better account of the bilayer structure, was stressed.[2] The present work is an effort in this direction; the main new features, which are generally ignored by the existing models, are summarized as follows.

The anisotropy and nonuniformity of the bilayer environment are included, in a way that allows for a realistic account of the chemical composition and physical state of the phospholipid system; they are introduced through the profiles of lateral pressure, density, order parameters and dielectric permittivity. Appropriate values of these for a given membrane can be derived from experiments or from atomistic simulations. The elastic response of the bilayer is, at least in part, implicitly contained in the cavity contribution to the mean field potential.

The full solute distribution function, accounting for the coupling between molecular position, orientation, and conformation, is calculated. Detailed information is provided, not only on the preferred configurations but also on fluctuations about them. This is important to get a realistic description of the behavior in the fluid bilayer environment where solutes, though confined, may undergo wide displacements and reorientations, which are likely to be essential for their biophysical role.

In its present form, the model ignores the membrane perturbation induced by the solute; therefore it is not suitable at high solute concentrations. However, the theoretical framework presented here allows the inclusion of the bilayer response, and future effort in this direction is planned.

To illustrate the capability of our model, we have presented the case study of cholesterol in DPPC. We have calculated a variety of properties, showing that the method proposed here is suitable for a general description of solutes in the bilayer environment. The use of the model allows us to disentangle the different contributions to solute properties in the bilayer and to shed light on the often unexpected effects of the anisotropy and nonuniformity of this environment. Small conformational effects are found for cholesterol, which is not surprising, in view of the small size of the flexible tail, if compared with the broad rigid core. Comparison of our results with experiments and simulations is very encouraging. The free energy profile calculated for cholesterol across the bilayer exhibits the same features as the profile obtained from all-atom molecular dynamics simulations.[63] Strong position/orientation couplings emerge: cholesterol is predicted to be mostly buried in the bilayer, with its head in the hydrophilic region and its long axis nearly perpendicular to the surface. When embedded deeper in the bilayer, cholesterol is predicted to change its average orientation and to preferentially keep its long axis parallel to the surface. According to our calculations, the latter configurations should have a very low probability for cholesterol in liquid crystalline DPPC; however, they may become more significant in other lipid systems. Interestingly, cholesterol lying flat, close to the bilayer center, was recently revealed by neutron scattering studies in unsaturated lipids.[64,65]

The method proposed here is computationally cheap. A detailed insight into the properties of a given solute/bilayer pair can be reached at very low cost, so we think that it could be usefully exploited for different purposes. It can provide the free energy landscapes needed to model solute permeability through membranes. Moreover, it may be suitable to mimic the bilayer environment in Monte Carlo and molecular dynamics simulations and to be integrated with quantum mechanical calculations of spectroscopic observables in lipid membranes.[69]

Abbreviations

DPPC, 1,2-dipalmitoyl-*sn*-glycero-3-phosphatidylcholine; DMPC, 1,2-dimyristoyl-*sn*-glycero-3-phosphatidylcholine; DFT, density functional theory.

Acknowledgment. This work was supported by the University of Padova (ex 60%). We are grateful to Dr. Matteo Stocchero for stimulating discussions and insightful comments during the course of this work. We gratefully acknowledge an anonymous reviewer for valuable criticism and suggestions.

Supporting Information Available: Derivation of the relationship between orienting strength ξ and S_{CD} order

parameters; definition of polarizable units in DPPC and corresponding mass density profiles; list of molecules used to parametrize the electrostatic contribution to the mean field potential; calculated and experimental water/*n*-octanol transfer free energies; atom labels and OPLS charges for cholesterol; list of cholesterol conformers considered in calculations with their torsional energy; representation of the molecular surface of cholesterol and of the principal axis system of its Saupe ordering tensor; plot of transfer free energy calculated for nine conformers of cholesterol in DPPC. This information is available free of charge via the Internet at http://pubs.acs.org/.

References

(1) White, S. H.; Wimpley, W. C. Membrane protein folding and stability: physical principles. *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 319–365.

(2) Grossfield, A. Implicit modeling of membranes. *Curr. Topics Membr.* **2008**, *60*, 131–157.

(3) De Young, L. R.; Dill, K. A. Solute partitioning into lipid bilayer membranes. *Biochemistry* **1988**, *27*, 5281–5289.

(4) Marqusee, J. A.; Dill, K. A. Solute partitioning into chain molecule interphases: monolayers, bilayer membranes, and micelles. *J. Chem. Phys.* **1986**, *85*, 434–444.

(5) Tobias, D. J. Electrostatic calculations: recent methodological advances and applications to membranes. *Curr. Opin. Struct. Biol.* **2001**, *11*, 253–261.

(6) Feig, M.; Brooks, C. L. Recent advances in the development and application of implicit solvent models in biomolecule simulations. *Curr. Opin. Struct. Biol.* **2004**, *14*, 217–224.

(7) Chen, J.; Brooks, C. L.; Khandogin, J. Recent advances in implicit solvent-based methods for biomolecular simulations. *Curr. Opin. Struct. Biol.* **2008**, *18*, 140–148.

(8) Fattal, D. R.; Ben-Shaul, A. A molecular model for lipid-protein interactions in membranes: the role of hydrophobic mismatch. *Biophys. J.* **1993**, *65*, 1795–1809.

(9) Fattal, D. R.; Ben-Shaul, A. Mean-field calculations of chain packing and conformational statistics in lipid bilayers: comparison with experiments and molecular dynamics studies. *Biophys. J.* **1994**, *67*, 983–995.

(10) Xiang, T.; Anderson, B. Molecular distributions in interphases: statistical mechanical theory combined with Molecular Dynamics simulation of a model lipid bilayer. *Biophys. J.* **1994**, *66*, 561–572.

(11) Mitragotri, S.; Johnson, M. E.; Blankschtein, D.; Langer, R. An analysis of the size selectivity of solute partitioning, diffusion, and permeation across lipid bilayers. *Biophys. J.* **1999**, *77*, 1268–1283.

(12) Kessel, A.; Ben-Tal, N.; May, S. Interactions of cholesterol with lipid bilayers: the preferred configuration and fluctuations. *Biophys. J.* **2001**, *81*, 643–658.

(13) Cantor, R. S. Lateral pressures in cell membranes: a mechanism for modulation of protein function. *J. Phys. Chem.* **1997**, *101*, 1723–1725.

(14) Tanizaki, S.; Feig, M. A generalized Born formalism for heterogeneous dielectric environments: application to the implicit modeling of biological membranes. *J. Chem. Phys.* **2005**, *122*, 124706:1–13.

(15) Ferrarini, A.; Moro, G. J.; Nordio, P. L.; Luckhurst, G. R. A shape model for molecular ordering in nematics. *Mol. Phys.* **1992**, *77*, 1–15.

(16) Mennucci, B.; Cammi, R. *Continuum Solvation Models in Chemical Physics*; Wiley: Chichester, U. K., 2007.

(17) Cantor, R. S. Lipid composition and the lateral pressure profile in bilayers. *Biophys. J.* **1999**, *76*, 2625–2639.

(18) Cantor, R. S. The influence of membrane lateral pressures on simple geometric models of protein conformational equilibria. *Chem. Phys. Lipids* **1999**, *101*, 45–56.

(19) Marsh, D. Lateral pressure in membranes. *Biochim. Biophys. Acta* **1996**, *1286*, 183–223.

(20) Patra, M. Lateral pressure profiles in cholesterol-DPPC bilayers. *Eur. Biophys. J.* **2005**, *35*, 79–88.

(21) Baoukina, S.; Marrink, S. J.; Tieleman, D. P. Lateral pressure profiles in lipid monolayers. *Faraday Disc.* **2010**, *144*, 393–409.

(22) Rowlinson, J. S.; Widom, B. *Molecular Theory of Capillarity*; Clarendon: Oxford, U. K., 1982.

(23) Helfrich, W. In *Physics of Defects*; Balian, R., Kleman, M., Poirier, J. P., Eds.; North-Holland Publishing: Amsterdam, 1981; pp 716−755.

(24) Szleifer, I.; Kramer, D.; Ben-Shaul, A.; Gelbart, W. M.; Safran, S. A. Molecular theory of curvature elasticity in surfactant films. *J. Chem. Phys.* **1990**, *92*, 6800–6817.

(25) Kozlov, M. M. In *Soft Condensed Matter Physics in Molecular and Cell Biology*; Poon, W. C. H., Andelman, D., Eds.; Taylor & Francis: Boca Raton, FL, 2006; pp 79−96.

(26) Honig, B.; Sharp, K.; Yang, A.-S. Macroscopic models of aqueous solutions: biological and chemical applications. *J. Phys. Chem.* **1993**, *97*, 1101–1109.

(27) Baker, N. A. Poisson-Boltzmann methods for biomolecular electrostatics. *Methods Enzymol.* **2004**, *383*, 94–118.

(28) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* **1990**, *112*, 6127–6129.

(29) Bashford, D.; Case, D. A. Generalized Born models of macromolecular solvation effects. *Annu. Rev. Phys. Chem.* **2000**, *51*, 129–152.

(30) Stern, H. A.; Feller, S. E. Calculation of the dielectric permittivity profile for a nonuniform system: application to a lipid bilayer simulation. *J. Chem. Phys.* **2003**, *118*, 3401–3412.

(31) Spassov, V. Z.; Yan, L.; Szalma, S. Introducing an implicit membrane in Generalized Born/Solvent Accessibility continuum solvent models. *J. Phys. Chem. B* **2002**, *106*, 8726–8738.

(32) Lazaridis, T. Effective energy function for proteins in lipid membranes. *Proteins* **2003**, *52*, 176–192.

(33) Im, W.; Feig, M.; Brooks, C. L. An implicit membrane generalized born theory for the study of structure, stability, and interactions of membrane proteins. *Biophys. J.* **2003**, *85*, 2900–2918.

(34) Sigalov, G.; Scheffel, P.; Onufriev, A. Incorporating variable dielectric environments into the generalized Born model. *J. Chem. Phys.* **2005**, *122*, 094511:1−15.

(35) Ulmschneider, M. B.; Ulmschneider, J. P.; Sansom, M. S. P.; Di Nola, A. A generalized Born implicit membrane representation compared to experimental insertion free energies. *Biophys. J.* **2007**, *92*, 2338–2349.

(36) Sengupta, D.; Smith, J. C.; Ullmann, G. Partitioning of side-chain analogues in a five-slab membrane model. *Biochim. Biophys. Acta* **2008**, *1778*, 2234–2243.

(37) Gallicchio, E.; Levy, R. M. AGBNP: an analytic implicit solvent model suitable for Molecular Dynamics simulations and high-resolution modeling. *J. Comput. Chem.* **2003**, *25*, 479–499.

(38) Seelig, J. Deuterium magnetic resonance: theory and application to lipid membranes. *Q. Rev. Biophys.* **1977**, *10*, 353–418.

(39) Mishima, K. In *Advances in Planar Lipid Bilayers and Liposomes*; Liu Leitmannova, A., Tien, H. T., Eds.; Academic Press: London, 2006; Vol. 3, pp 55−84.

(40) Ramsden, J. J. Molecular orientation in lipid bilayers. *Philos. Mag. B* **1999**, *79*, 381–386.

(41) Ferrarini, A.; Moro, G. J.; Nordio, P. L. Simple molecular model for induced cholesteric phases. *Phys. Rev. E* **1996**, *53*, 681–688.

(42) Ferrarini, A. Shape model for the molecular interpretation of the flexoelectric effect. *Phys. Rev. E* **2001**, *64*, 021710:1−11.

(43) Cestari, M.; Bosco, A.; Ferrarini, A. Molecular field theory with atomistic modeling for the curvature elasticity of nematic liquid crystals. *J. Chem. Phys.* **2009**, *131*, 054104:1−16.

(44) Maier, W.; Saupe, A. A simple molecular-statistics theory of the nematic liquid-crystalline phase. Part I. *Z. Naturforsch. A* **1959**, *14A*, 882–900. Ibidem. A simple molecular-statistics theory of the nematic liquid-crystalline phase. Part II. *Z. Naturforsch. A* **1960**, *15A*, 287−292.

(45) Abramowitz, M.; Stegun, I. A. Handbook of Mathematical Functions; Dover: New York, 1972.

(46) Ferrarini, A.; Luckhurst, G. R.; Nordio, P. L.; Roskilly, S. J. Prediction of the transitional properties of liquid crystal dimers. A molecular field calculation based on the surface tensor parametrization. *J. Chem. Phys.* **1994**, *100*, 1460–1469.

(47) Luckhurst, G. R. In *The Molecular Physics of Liquid Crystals*; Luckhurst, G. R., Gray, G. W., Eds.; Academic: London, 1979; pp 85−119.

(48) Flory, P. J. *Statistical Mechanics of Chain Molecules*; Interscience: New York, 1969.

(49) Richards, F. M. Areas, volumes, packing and protein structure. *Annu. Rev. Biophys. Bioeng.* **1977**, *6*, 151–176.

(50) Sanner, M. F.; Olson, A.; Spehner, J.-C. Reduced surface: an efficient way to compute molecular surfaces. *Biopolymers* **1996**, *38*, 305–320.

(51) Stewart, J. *Calculus*, 4th ed.; Brooks/Cole Publishing Company: Pacific Grove, CA, 1999.

(52) Arfken, G. B.; Weber, H. J. *Mathematical Methods for Physicist*, 4th ed.; Elsevier: Boston, 2005.

(53) Quarteroni, A.; Sacco, R.; Saleri, F. *Numerical Mathematics*; Springer: New York, 2000.

(54) Ollila, O. H. S.; Róg, T.; Karttunen, M.; Vattulainen, I. Role of sterol type on lateral pressure profiles of lipid membranes affecting membrane protein functionality: comparison between cholesterol, desmosterol, 7-dehydrocholesterol and ketosterol. *J. Struct. Biol.* **2007**, *159*, 311–323.

(55) Kupiainen, M.; Falck, E.; Ollila, S.; Niemelä, P.; Gurtovenko, A. A. Free volume properties of sphingomyelin, DMPC, DPPC, and PLPC bilayers. *J. Comp. Theory Nano* **2005**, *2*, 401–413.

(56) Falk, E.; Patra, M.; Karttunen, M.; Hyvönen, M. T.; Vattulainen, I. Lessons of slicing membranes: interplay of packing, free area, and lateral diffusion in phospholipid/cholesterol bilayers. *Biophys. J.* **2004**, *87*, 1076–1091.

(57) *CRC Handbook of Chemistry and Physics*, 85th ed.; Lide, D. R., Ed.; CRC: Boca Raton, FL, 2004.

(58) Frisch, M. J. *Gaussian 03*, revision C.02; Gaussian, Inc.: Wallingford CT, 2004.

(59) Bondi, A. Van der Waals volumes and radii. *J. Phys. Chem.* **1964**, *68*, 441–451.

(60) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.

(61) Feig, M.; Im, W.; Brooks, C. L. Implicit solvation based on generalized Born theory in different dielectric environments. *J. Chem. Phys.* **2004**, *120*, 903–911.

(62) Harroun, T. A.; Katsaras, J.; Wassall, S. R. Cholesterol hydroxyl group is found to reside in the center of a polyunsaturated lipid membrane. *Biochemistry* **2006**, *45*, 1227–1233.

(63) Gliss, C.; Randel, O.; Casalta, H.; Sackmann, E.; Zorn, R.; Bayerl, T. Anisotropic motion of cholesterol in oriented DPPC bilayers studied by quasielastic neutron scattering: the liquid-ordered phase. *Biophys. J.* **1999**, *77*, 331–340.

(64) Bennett, W. F. D.; MacCallum, J. L.; Hinner, M. J.; Marrink, S. J.; Tieleman, D. P. Molecular view of cholesterol flip-flop and chemical potential in different membrane environments. *J. Am. Chem. Soc.* **2009**, *131*, 12714–12720.

(65) Kučerka, N.; Marquardt, D.; Harroun, T. A.; Nieh, M.-P.; Wassall, S. R.; Katsaras, J. The functional significance of lipid diversity: orientation of cholesterol in bilayers is determined by lipid species. *J. Am. Chem. Soc.* **2009**, *131*, 16358–16359.

(66) Guo, W.; Kurze, V.; Huber, T.; Afdhal, N. H.; Beyer, K.; Hamilton, J. A. A solid-state NMR study of phospholipid-cholesterol interactions: sphingomyelin-cholesterol binary system. *Biophys. J.* **2002**, *83*, 1465–1478.

(67) Dufourc, E. J.; Parish, E. J.; Chitrakorn, S.; Smith, I. C. P. Structural and dynamical details of the cholesterol-lipid interaction as revealed by deuterium NMR. *Biochemistry* **1984**, *23*, 6062–6071.

(68) In the experiments reported in ref 67, the magnetic field was perpendicular to the bilayer normal.

(69) Pavanello, M.; Mennucci, B.; Ferrarini, A. Quantum-mechanical studies of NMR properties of solutes in liquid crystals: a new strategy to determine orientational order parameters. *J. Chem. Phys.* **2005**, *122*, 064906:1−9.

CT100210U

JCTC Journal of Chemical Theory and Computation

Benchmarking of DFT Functionals for the Hydrolysis of Phosphodiester Bonds

António J. M. Ribeiro, Maria J. Ramos, and Pedro A. Fernandes*

Departamento de Química, Faculdade de Ciências do Porto, Rua do Campo Alegre, 687, 4169-007 Porto, Portugal

Received December 3, 2009

Abstract: Phosphodiester bonds are an important chemical component of biological systems, and their hydrolysis and formation reactions are involved in major steps throughout metabolic pathways of all organisms. In this work, we applied dimethylphosphate as a model for this kind of bonds and calculated the potential energy surface for its hydrolysis at the approximated CCSD(T)/CBS//B3LYP/6-311++G(2d,2p) level. By varying the nucleophile (water or hydroxide) and the medium (vacuum or aqueous implicit solvent) we obtained and described four reaction paths. These structures were then used in a DFT functional benchmarking in which we tested a total of 52 functionals. Furthermore, the performances of HF, MP2, MP3, MP4, and CCSD were also evaluated. This benchmarking showed that MPWB1K, MPW1B95, and PBE1PBE are the more accurate functionals to calculate the energies of dimethylphosphate hydrolysis as far as activation and reaction energies are concerned. If considering only the activation energies, MPWB1K, MPW1B95, and B1B95 give the lowest errors when comparing to CCSD(T). A basis set benchmarking on the same system shows that 6-311+G(2d,2p) is the best basis set concerning the relationship between computational time and accuracy. We believe that our results will be of great help to further studies on related phosphodiester systems. This includes not only pure chemical problems but also biochemical studies in which DNA, RNA, and phospholipids are required to be depicted at a quantum level.

1. Introduction

Phosphodiester bonds are central in biological systems not only because they are an essential part of nucleic acids and phospholipids but also due to their interaction with many enzymes. Their function exceeds simple bonding, since the negatively charged phosphate group adds unique properties to any molecule in which it is present. In nucleic acids, these bonds sustain the polymer covalent backbone by connecting (deoxy)riboses of adjacent nucleotides. Additionally, phosphodiester bonds make these molecules highly charged, a crucial factor in their folding and interaction with proteins, metal ions, and polyamines. Furthermore, the negative charge protects these bonds from nucleophilic attack and thus lowers the phosphodiester rate of hydrolysis: a desirable property when it comes to genetic information.[1,2] In phospholipids, the phosphodiester bond is the bridge between the hydrophilic and hydrophobic portions of the molecule. It is also an important hydrophilic factor and therefore half-responsible for phospholipids chemical behavior.[1] It is then clear the need for understanding the phosphodiester bond and its key properties, namely its geometry and hydrolysis/formation energies.

In this study, we were particularly interested in analyzing how DFT functionals describe the potential energy surface for phosphodiester hydrolysis, using approximated CCSD(T)/CBS energy values as reference (see Theoretical Methods for details). Our objective was to answer the following question: *If we want to study the kinetics and thermodynamics of the breakage/formation of a phosphodiester bond, which functional should we use*. This question is relevant to all sorts of mechanistic studies with biological systems, in which the use of accurate post-Hartree−Fock methods is not yet feasible. That includes a great number of enzyme families such as cyclic nucleotide phosphodiesterases, DNA/RNA

* Corresponding author e-mail: pafernan@fc.up.pt.

10.1021/ct900649e © 2010 American Chemical Society
Published on Web 07/08/2010

Table 1. DFT Functionals and Basis Sets Tested in This Work[a]

functionals			basis sets
LDA	**H-GGA**	**HM-GGA**	**Pople type**
SVWN[42,43]	B1LYP[44,45]	B1B95[45,46]	6-31G(d)
SVWN5[42,43]	B3LYP[44,47]	BB1K[44,46,48]	6-31+G(d)
GGA	B3P86[47,49]	M06[40,41]	6-311G(d)
BLYP[44,46]	B3PW91[47,50]	M06 - 2X[41,40]	6-31+G(d,p)
BP86[46,49]	B97−1[51]	M06 - HF[41,52]	6-311G(d,p)
BPBE[46,53]	B97−2[54]	MPW1B95[45,48,50,55]	6-311+G(d)
BPW91[46,50]	B98[56,57]	MPW1KCIS[50,55,58,59]	6-311++G(d)
G96LYP[44,60]	BhandH[44,46]	MPWB1K[45,48,50,55]	6-311+G(d,p)
HCTH93[51]	BhandHLYP[44,46]	MPWKCIS1K[50,55,58,59]	6-311++G(d,p)
HCTH147[51]	MPW1K[50,55,61]	PBE1KCIS[53,58,62]	6-311G(2d,2p)
HCTH407[51]	MPW1N[50,55,63]	TPSS1KCIS[58,62,64]	6-311+G(2d,2p)
MPWLYP[44,50,55]	MPW1PW91[50,55]	TPSSh[64]	6311++G(2d,2p)
MPWLYP1W[44,50,55,65]	MPW1S[50,55,66]	**M-GGA**	6311++G(3df,3pd)
MPWPW91[50,55]	MPW3LYP[44,50,55]	BB95[45,46]	**correlation-consistent**
OLYP[44,67]	O3LYP[44,67,68]	M06 - L[41,69]	aug-cc-pvdz
PBE1W[53,65]	PBE1PBE[53]	MPWB95[45,50,55]	aug-cc-pvtz
PBELYP1W[44,53,65]		TPSSLYP1W[44,64,65]	aug-cc-pvqz
PBEPBE[53]		TPSSTPSS[64]	
XLYP[44,46,70,71]		VSXC[72]	

[a] Functionals are organized by type: LDA - local density approximation; GGA - generalized gradient approximation; M-GGA - meta generalized gradient approximation; H-GGA - hybrid generalized gradient approximation; HM-GGA - hybrid meta generalized gradient approximation.

polymerases, nucleases, integrases, and phospholipases C/D, among others. We chose dimethylphosphate, the smallest molecule with this kind of bond, to perform our benchmarking calculations. We have tried to keep the system to a minimum to make it as general as possible and to serve as a model for all enzyme classes. Such a small system captures the essential of the chemical reaction, without the influence of a particular enzymatic scaffold. Both water and hydroxide were used as nucleophiles, and vacuum as well as aqueous solvent were employed, leading to four different systems. The effect of different basis sets was also studied, in a basis set benchmarking.

Numerous theoretical works concerning dimethylphosphate hydrolysis and some similar systems have been published in the past.[3−23] These studies describe the stationary states for the hydrolysis reaction path in a very detailed way, and at various levels of theory, namely at the density functional theory (always using B3LYP), second order Moller−Plesset (MP2) and Hartree−Fock (HF), combined with a whole range of split-valence basis sets. As B3LYP is widely used and produces generally good results[24] it is usually preassumed as a good choice for all types of calculations, and hence no effort has been made to justify its use as an alternative to others functionals. We note, however, that neither B3LYP has been properly validated for this so important chemical reaction nor have any others functionals been systematically tested. We therefore present this benchmarking to complement the lack of such information on phosphodiester systems.

2. Theoretical Methods

In order to obtain the structures for the benchmarking studies we started by modeling the molecules of the reactants and products involved in the two main reactions: hydroxide, water, dimethylphosphate, methoxide, methanol, and methyl phosphate. Structures of transition states and intermediates were then created by scanning the proper coordinates of the reactants systems. All these structures were fully optimized in vacuum and solvent at the B3LYP/6-311++G(2d,2p) level. Aqueous solvent systems were treated with the IEFPCM continuum model.[25−28] Frequency calculations were carried out for all structures to confirm their nature (minima/transition state) along the potential energy surface. Zero point energies and entropic corrections were also computed.

Single point energy values at the approximated CCSD(T)/CBS//B3LYP/6-311++G(2d,2p) level were employed as reference to evaluate the accuracy of the functionals. The approximated CCSD(T)/CBS energy was obtained in two main steps. First, the MP2/CBS energy was extrapolated by using the Helgaker scheme: the Hartree−Fock energy was extrapolated by fitting the energy of aug-cc-pVDZ, aug-cc-pVTZ, and aug-cc-pVQZ basis sets to an exponential function and the energy of MP2 by fitting the values of aug-cc-pVTZ and aug-cc-pVQZ basis sets to a polynomial function.[29,30] Second, we summed the correlation energy difference between MP2 and CCSD(T), calculated with the 6-31+G(d) basis set.[31,32] Using the same methodology, we also calculated single point energies at the approximated CCSD/CBS level. We used these values to better understand the importance of triple excitations in CCSD(T), although it is not described in the literature if this kind of approximation can be applied precisely to CCSD.

Basis set benchmarking was used to assess the quality of the basis set indicated in Table 1. Single point energies with all those basis sets were calculated for all our systems using the same functional, B3LYP. MUE (mean unsigned error) and MSE (mean signed error) for these series were based on the comparison with the most complete basis set we have used, aug-cc-pvqz.

We tested the functionals shown in Table 1 by performing single point energy calculations for each one of the geometries mentioned above with the proper functionals and the 6-311++G(2d,2p) basis set. The same scheme

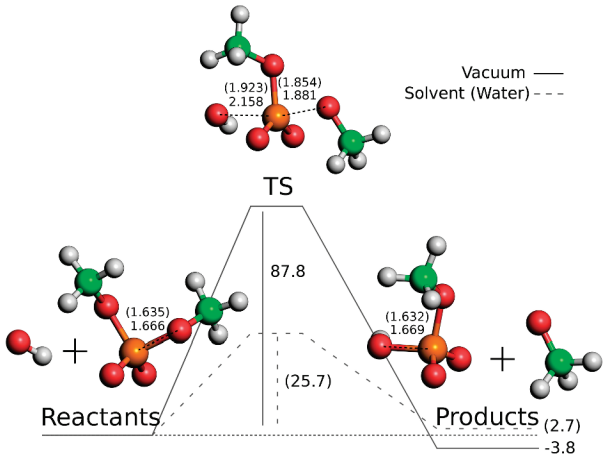


Figure 1. Reaction path for dimethylphosphate hydrolysis with hydroxide attack. Distances for formed and broken bonds are indicated in Å. Energies are in kcal/mol. Numbers inside brackets refer to solvent calculations. Structures shown are for vacuum systems. Solvent structures do not differ significantly from these.

was employed for Hartree−Fock and post-Hartree−Fock methods: MP2, MP3, MP4(SDQ), and CCSD. We have chosen the 6-311++G(2d,2p) basis set because it was the most complete basis set that could be used within most modern hardware capabilities. Calculations with more complete basis sets would become intolerably time-consuming. In addition, according to the basis set analysis we have made, this basis set gives rather reliable results for our system, with a basis set truncation error below 1 kcal/mol for both activation and reaction energies. Single point calculations with M06 functionals were carried out with JAGUAR.[33] As IEFPCM is not included in JAGUAR; all our M06 functionals calculations were performed in vacuum. All other calculations, including scans and optimizations, were done using GAUSSIAN03.[34]

3. Results and Discussion

3.1. Obtained Geometries. Structures obtained at the B3LYP/6-311++G(2d,2p) level are represented in Figures 1 and 2. Tables 2, 3, and 4 show relevant interatomic distances, angles, dihedrals, and Mulliken atomic charges associated with the four studied chemical pathways: dimethylphosphate hydrolysis with hydroxide or water as nucleophiles, both in the gas phase and implicit solvent. Phosphodiester hydrolysis via hydroxide attack is a one step mechanism. Barrier heights for this reaction, calculated at the approximated CCSD(T)/CBS//B3LYP/6-311++G(2d,2p) level, are 87.8 kcal/mol in vacuum and 25.7 kcal/mol in aqueous solvent. The products of the reaction are methylphosphate and methoxide.

The reaction with a water molecule as nucleophile, represented in Figure 2, is more complex. A total of three transition states and two intermediates are necessary. In the first transition state, water attacks the dimethylphosphate molecule and simultaneously loses a proton to it. In the second transition state, this hydrogen atom rotates ca. 180° to form an H-bond with the leaving group. In the last step, a methoxide ion dissociates from the pentacoordinated phosphorus center and simultaneously deprotonates one of the phosphate oxygen atoms, leaving as a neutral methanol molecule. Note that, unlike with the hydroxide mechanism, products and reagents here have favorable interaction energy. Transition states energies, calculated with the same level of theory as above are, respectively, 22.5, 29.0, and 25.4 kcal/mol in vacuum and 36.3, 36.5, and 37.62 kcal/mol in solvent.

When comparing the kinetics of the two nucleophilic reactions in solution, we need to be aware of the difference in the concentration of the two nucleophiles. The entropic cost of confining a molecule to a "solvent cage" in the prereactive state can be given by eq 1 below.[35] This term only affects the hydroxide whose concentration is very low (10^{-7} M) but not the water nucleophile, which is always present at the reactive center.

$$\Delta G^{\circ}_{cage} = -kb \cdot T \cdot \ln\left(\frac{Vf}{Vi}\right) \quad (1)$$

When taking this correction into account, and considering a final restriction volume of six water molecules, the activation energy associated with the hydroxide reaction raises 10.9 kcal/mol. Therefore, the activation energies for hydroxide attack and water attack involved, in solution, are very similar in the two cases. It does not make sense to consider this effect in the system with the water nucleophile since the dimethylphosphate is surrounded by water at all times and no energy cost is needed for a particular water molecule to be in position for attack.

Barrier heights reported in other studies are similar to ours.[8,19,21,22] Not surprisingly, geometries are also equivalent to those of previous studies with smaller basis sets.[7−9,12,17,19,22] It is known that geometric parameters converge very quickly with increasing basis sets and/or methods. Our protocol takes advantage of this, since for the remainder calculations we used these structures without further optimization. Performing optimizations for all functionals and basis sets would be a great computational effort and would bring negligible gains in terms of the calculated hydrolysis energies. A remark needs to be done, however, concerning the hydroxide attack in vacuum: for very small basis sets, STO-3G, 3-21G(d), and 3-21+G(d), this reaction path presents an intermediate and an additional transition state.[4,8,36] However, the depth of the intermediate well diminishes and finally disappears with basis sets of increasing size. Such an intermediate was not a stationary state at the B3LYP/6-311++G(2d,2p) level.

Tables 5, 6, and 7 show the energy values for the four reactions paths calculated with various methods as well as the Gibbs and zero point energy corrections. We will discuss relevant experimental data concerning the dimethyl phosphate hydrolysis, but we will not attempt at evaluating the quality of the various methods by direct comparison with it, as will become clear in the section below.

The B3LYP energy values are not in good agreement with the reference values, and so it is not an adequate functional to study the energetics of phosphodiester systems. Other functionals results will be discussed in the proper section. Pure Hartree−Fock is very unreliable, as the errors in barrier energies are, in most cases, higher than 10 kcal/mol. Regarding post-Hartree−Fock methods, it is very interesting

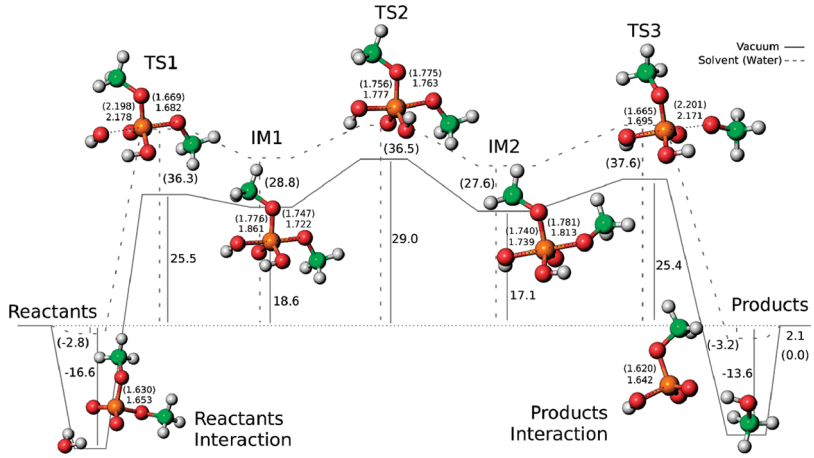


Figure 2. Reaction path for dimethylphosphate hydrolysis with a water molecule attack. Distances for formed and broken bonds are indicated in Å. Energies are in kcal/mol. Numbers inside brackets refer to solvent calculations. Structures shown are for vacuum systems. Solvent structures do not differ significantly from these.

Table 2. Relevant Distances, Angles, Dihedrals, and Atomic Charges for Dimethylphosphate Hydrolysis by Hydroxide Attack in Vacuum and Implicit Solvent[a]

		hydroxide attack					
		vacuum			implicit solvent		
		reactants	TS	products	reactants	TS	products
distances	On−P	---	2.158	1.669	----	1.923	1.923
	P−O1	1.666	1.881	----	1.635	1.854	----
angles	On−P−O1	----	167.6	----	----	170.1	----
dihedrals	O2−P−O1−C1	−73.4	−178.0	----	−66.5	−176.7	----
	C2−O2−P−O4	36.7	−32.5	−38.9	45.2	−20.3	−47.4
Mulliken charges	q1	−1.00	−0.70	−0.45	−1.00	−0.52	−0.41
	q2	−0.51	−0.88	−0.55	−0.55	−0.90	−0.59
	q3	−0.49	−0.41	−1.00	−0.45	−0.48	−1.00

[a] See Figure 3 for atom and charges denomination. Distances are indicated in Å; angles and dihedrals are in degrees. Geometrical parameters were optimized at B3LYP/6-311++G(2d,2p). Mulliken charges were calculated at the CCSD/6-311++G(2d,2p)//B3LYP/6-311++G(2d,2p) level.

Table 3. Important Distances, Angles, Dihedrals, and Charges in Dimethylphosphate Hydrolysis by Hydroxide Attack in Vacuum[a]

		water attack - vacuum								
		R.	R. I.	TS1	IM1	TS2	IM2	TS3	P. I.	P.
distances	On−P	----	3.359	2.198	1.776	1.765	1.740	1.665	1.620	1.632
	P−O1	1.635	1.630	1.669	1.747	1.775	1.781	2.201	3.931	----
	On−Hn2	0.971	0.971	1.709	2.040	2.761	3.209	3.208	3.987	----
	O4−Hn2	----	2.063	0.992	0.973	0.977	0.974	0.999	1.751	----
	O1−Hn2	----	4.046	3.248	3.252	2.751	1.975	1.648	0.981	0.973
angles	O1−P−On	----	128.5	165.7	168.1	166.4	166.6	163.9	118.1	----
dihedrals	O2−P−O1−C1	−66.5	−66.0	−179.7	176.4	179.1	68.2	134.3	−155.2	----
	C2−O2−P−O4	177.6	177.6	130.0	106.8	100.6	117.3	117.8	−72.3	−47.3
	On−P−O4−Hn2	----	0.0	−6.1	−10.7	84.3	174.6	173.2	−124.5	----
Mulliken charges	q1	0.00	−0.01	−0.73	−0.42	−0.32	−0.34	−0.29	−0.36	−0.41
	q2	−0.55	−0.59	−0.05	−0.25	−0.29	−0.27	0.02	0.64	−0.59
	q3	−0.45	−0.41	−0.22	−0.33	−0.39	−0.39	−0.72	−0.01	0.00

[a] See Figure 3 for atom and charges denomination. Distances are indicated in Å; angles and dihedrals are in degrees. Geometrical parameters were optimized at B3LYP/6-311++G(2d,2p). Mulliken charges were calculated at the CCSD/6-311++G(2d,2p)//B3LYP/6-311++G(2d,2p) level. R. - reactants, R. I. - reactants interaction, TS - transition state, IM -intermediate, P. I. - products interaction, P. - products.

to observe the difference between approximated CCSD/CBS and approximated CCSD(T)/CBS results, as an indication that the inclusion of triple excitations is crucial for a good description of our system. It is also important to note that the CCSD/6-311++G(2d,2p) energy is very different from the approximated CCSD/CBS. This means that CCSD energies converge slowly with the improvement of the basis sets, and even a triple-ζ basis set is still far from the complete basis set limit. The same could be said for all post-Hartree−Fock methods.

The thermal and entropic contributions for the Gibbs energy present large values (around 11 kcal/mol) when considering complexes as opposed to structures where the molecular species are set apart. This is due to a major loss

Table 4. Relevant Distances, Angles, Dihedrals, and Atomic Charges for Dimethylphosphate Hydrolysis by Hydroxide Attack in Implicit Solvent[a]

		water attack - implicit solvent								
		R.	R. I.	TS1	IM1	TS2	IM2	TS3	P. I.	P.
distances	On−P	----	3.360	2.178	1.861	1.777	1.739	1.695	1.642	1.669
	P−O1	1.666	1.653	1.682	1.722	1.763	1.813	2.171	3.772	----
	On−Hn2	0.961	0.971	1.668	1.921	2.720	3.213	3.247	4.111	----
	O4−Hn2	----	2.099	0.995	0.966	0.960	0.965	0.999	1.740	----
	O1−Hn2	----	4.049	3.238	3.221	2.751	1.883	1.632	0.985	0.960
angles	O1−P−On	----	129.8	165.5	167.8	166.4	166.9	167.4	121.1	----
dihedrals	O2−P−O1−C1	−73.4	−73.4	176.3	173.6	178.2	74.1	128.5	134.4	----
	C2−O2−P−O4	173.6	172.6	123.1	121.6	113.2	117.7	157.2	−64.2	−38.9
	On−P−O4−Hn2	----	0.3	−5.1	−7.4	82.1	175.2	176.7	−138.4	----
Mulliken charges	q1	0	−0.01	−0.67	−0.50	−0.36	−0.35	−0.30	−0.38	−0.45
	q2	−0.51	−0.64	−0.09	−0.24	−0.31	−0.24	−0.07	−0.61	−0.55
	q3	−0.49	−0.45	−0.24	−0.26	−0.33	−0.41	0.63	−0.02	0.00

[a] See Figure 3 for atom and charges denomination. Distances are indicated in Å; angles and dihedrals are in degrees. Geometrical parameters were optimized at B3LYP/6-311++G(2d,2p). Mulliken charges were calculated at the CCSD/6-311++G(2d,2p)//B3LYP/6-311++G(2d,2p) level. R. - reactants, R. I. - reactants interaction, TS - transition state, IM - intermediate, P. I. - products interaction, P. - products.

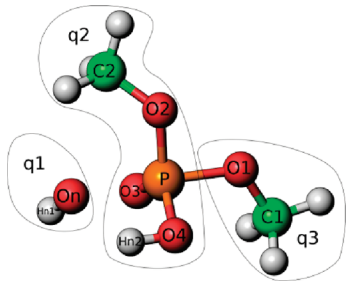


Figure 3. Division of the system in three groups, q1, q2, and q3, to analyze the changes in groups atomic charges along the reaction pathway. Hydroxide attacking systems do not have 'Hn2'. For water attacking systems, in reactants and products 'Hn2' belongs to q1 and q3, respectively. In the others structures belong to q2.

Table 5. Activation and Reaction Energies for Dimethylphosphate Hydrolysis by Hydroxide Attack in Vacuum and Implicit Solvent As Calculated by Various Methods[a]

	vacuum		implicit solvent	
	TS	products	TS	products
DFT - B3LYP	88.1	−8.6	30.1	0.4
Hartree−Fock	101.1	−8.2	41.3	0.2
MP2	86.7	−3.8	27.4	3.5
MP2/CBS	88.7	−2.9	25.0	4.2
MP3	89.1	−6.0	28.8	1.9
MP4(SDQ)	88.8	−5.0	29.2	2.6
CCSD	89.0	−5.3	29.2	2.4
CCSD/CBS	90.6	−4.0	28.1	2.3
CCSD(T)/CBS	87.8	−3.8	25.7	2.7
correction to the Gibbs free energy	9.4	−2.6	−2.0	0.8
zero point energy	0.4	−1.0	1.0	−0.5
solvent cage energy	----	----	10.9	----
total free energy	97.2	−6.4	34.6	3.5

[a] The zero point energy and the correction to the Gibbs free energy are also included. All values are in kcal/mol. In the implicit solvent case, the correction to the Gibbs free energy does not contain the translational and rotational entropic energies. Total free energy refers to the sum of CSSD(T)/CBS energy with the correction to the Gibbs free energy and the solvent cage energy, if applied. The methods not noted with CBS refer to calculations with the 6-311++G(2d,2p) basis set. TS - transition state.

of entropy during the formation of the complex. A detailed thermodynamic analysis of all molecules involved shows that the major contributions come from the translational and rotational entropy associated with the small nucleophiles, which are nearly completely lost in that step. We should not include this correction on the energy of solvated systems because the complex water−dimethylphosphate should always exist in solution and the nucleophile that attacks dimethylphosphate will substitute that water molecule. In this process, no significant energy loss or gain occurs: it remains the same for a water nucleophile and changes only 0.04 kcal/mol for a hydroxide nucleophile.

3.2. Comparison with Experimental Results. Dimethylphosphate hydrolysis in solution occurs mainly by C−O cleavage. The reaction, as we modeled it, accounts for less than 0.5% of total hydrolysis.[37] Thus, the activation energies reported for these reactions of 31.7 kcal/mol for hydroxide attack[38] and 34.9 kcal/mol for water attack[37] could only be seen as a bottom line reference. According to transition state theory, eq 2, we could roughly estimate that a reaction two hundred times slower than the above would have energies 3.1 kcal/mol higher, setting activation energies for P−O dimethylphosphate hydrolysis on 34.9 and 38.0 kcal/mol, respectively.

$$k_v = \frac{k_b \cdot T}{h} e^{-\Delta G^{\neq}/RT} \qquad (2)$$

A recent paper brought new light to this discussion.[39] By using a similar molecule to dimethylphosphate, dineopentyl phosphate, the reaction via P−O cleavage was isolated, as the attack on carbon atoms is avoided by steric hindrance. Experimentally obtained activation energies in this case are 37.7 kcal/mol for hydroxide and 38.1 kcal/mol for water attack.

Even though these values are equivalent to ours of 34.6 and 38.4 kcal/mol, respectively, a direct comparison should not be attempted because the experimental results do not treat exactly the reaction studied in our work. Moreover, we used many approximations, such as the treatment of the translational and rotational entropy and the continuum description of the solvent.

3.3. Functionals Benchmarking. The objective of this study is not to make a full evaluation of DFT functional but instead to evaluate the performance of DFT functionals with

Table 6. Activation and Reaction Energies for Dimethylphosphate Hydrolysis by Water Attack in Vacuum As Calculated by Various Methods[a]

	vacuum							
	R. I.	TS12	IM2	TS23	IM3	TS34	P. I.	P.
DFT - B3LYP	−15.3	25.1	23.5	33.8	22.4	27.3	−14.0	−0.2
Hartree−Fock	−14.4	36.4	32.8	43.1	31.2	38.5	−13.2	0.0
MP2	−16.9	23.0	20.6	31.5	19.4	26.4	−13.4	2.9
MP2/CBS	−16.6	22.2	18.8	29.4	17.5	25.5	−13.4	2.2
MP3	−16.9	25.4	21.7	32.3	20.3	28.3	−14.1	2.1
MP4(SDQ)	−16.5	25.3	22.4	33.0	21.1	28.4	−13.4	2.3
CCSD	−16.6	25.6	22.4	33.0	21.1	28.6	−13.6	2.2
CCSD/CBS	−16.2	24.5	20.2	30.6	20.6	27.4	−13.5	1.6
CCSD(T)/CBS	−16.6	22.5	18.6	29.0	17.1	25.4	−13.6	2.1
correction to the Gibbs free energy	10.2	12.1	13.2	12.0	13.1	10.5	9.0	−0.4
zero point energy	2.3	2.1	3.0	1.9	3.0	1.3	2.3	1.2
total free energy	−6.4	34.6	31.8	41.0	30.2	35.9	−4.6	1.7

[a] The zero point energy and the correction to the Gibbs free energy are also included. All values are in kcal/mol. Total free energy refers to the sum of CSSD(T)/CBS energy with the correction to the Gibbs free energy. The methods not noted with CBS refer to calculations with the 6-311++G(2d,2p) basis set. R. I. - reactants interaction, TS - transition state, IM - intermediate, P. I. - products interaction, P. - products.

Table 7. Activation and Reaction Energies for Dimethylphosphate Hydrolysis by Water Attack in Implicit Solvent As Calculated by Various Methods[a]

	implicit solvent							
	R. I.	TS12	IM2	TS23	IM3	TS34	P. I.	P.
DFT - B3LYP	−0.9	38.8	34.2	41.6	33.3	39.6	−3.8	−2.4
Hartree−Fock	1.0	51.2	44.3	52.3	43.5	52.3	−2.3	−2.8
MP2	−2.6	36.9	31.1	38.9	30.0	38.3	−2.9	0.2
MP2/CBS	−2.4	36.3	29.1	36.8	28.0	37.8	−3.0	0.1
MP3	−2.5	39.6	32.3	40.3	31.3	40.7	−3.5	−0.6
MP4(SDQ)	−2.2	39.2	32.8	40.7	31.8	40.4	−3.1	−0.4
CCSD	−2.2	39.5	32.9	40.8	31.9	40.7	−3.1	−0.5
CCSD/CBS	−2.1	38.5	30.5	38.2	29.5	39.9	−3.1	−0.5
CCSD(T)/CBS	−2.8	36.3	28.8	36.5	27.6	37.6	−3.2	0.0
correction to the Gibbs free energy	−2.2	1.7	−0.7	1.9	−0.7	−2.3	−3.9	1.2
zero point energy	2.2	1.6	2.5	1.3	2.5	1.2	2.0	0.4
total free energy	−5.0	38.0	28.1	38.4	26.9	35.3	−7.1	1.2

[a] The zero point energy and the correction to the Gibbs free energy are also included. All values are in kcal/mol. The correction to the Gibbs free energy does not contain the translational and rotational entropic energies. Total free energy refers to the sum of CSSD(T)/CBS energy with the correction to the Gibbs free energy. The methods not noted with CBS refer to calculations with the 6-311++G(2d,2p) basis set. R. I. - reactants interaction, TS - transition state, IM - intermediate, P. I. - products interaction, P. - products.

regards to phosphodiester bonds hydrolysis reactions. Hence, our results can only be examined in that context, and further conclusions should not be deduced. Due to the large heterogeneity of DFT functionals it is likely that our results can be significantly different from others concerning unrelated reactions.

The ranking criteria for the functional evaluation are the mean unsigned errors (MUE) and mean signed errors (MSE) between the electronic energy of each functional and their approximated CCSD(T)/CBS counterpart. The calculations were carried out on the same geometries. The unsigned error is used to assess the absolute deviation from the reference value, while the signed error measures a systematic deviation. When analyzing these two values we should be aware that a) a positive MSE means an average overestimation of energy values and a negative MSE an average underestimation; b) MUE must be larger or equal to MSE; c) if MSE is equal to MUE the functional always overestimated the energy in considered structures; and d) if MSE is symmetric to MUE the functional always underestimated the energy.

Tables 8, 9, and 10 show the MUE and MSE values for all the functionals tested. Table 8 contemplates all structures, Table 9 refers only to results in vacuum, and Table 10 refers to those in the solvent. Some general trends can be inferred from the analysis of these tables:

• Top ranked functionals are mostly H-GGA (hybrid-generalized gradient approximation) and HM-GGA (hybrid-meta-GGA). As a group, LDA (local density approximation) functionals give the poorest results. It should be noted, however, that the use of a H-GGA or a HM-GGA functional does not guarantee by itself improved estimated energies, since the middle and bottom of the discussed tables are also populated by those.

• When comparing MUEs for energy barriers and reaction energies, an unusual scenario arises. For the top results, reaction energies MUEs are larger than energy barriers MUEs. The probable source for this particularity may lie in the products structures, as a systematic deviation between DFT and CCSD(T) energies could be observed in the reactants energies but not in the activation energies. For the hydroxide attack system in vacuum, for example, the reaction energy calculated with approximated CCSD(T)/CBS is −3.8 kcal/mol, whereas for most DFT functionals is circa −9 kcal/

Table 8. MUEs and MSEs for the Best Studied Functionals for Vacuum and Solvent Systems[a]

		transition state		intermediate		minimum		all	
functional	type	MUE	MSE	MUE	MSE	MUE	MSE	MUE	MSE
MPWB1K	HM-GGA	0.34	−0.10	0.16	−0.04	1.46	1.46	0.75	0.54
MPW1B95	HM-GGA	0.83	−0.79	0.10	−0.02	1.56	1.56	0.97	0.30
PBE1PBE	H-GGA	1.02	−0.67	0.70	0.7	2.01	2.01	1.35	0.68
BB1K	HM-GGA	1.05	1.02	1.25	1.25	1.85	1.85	1.41	1.39
TPSSh	HM-GGA	1.52	−1.44	0.26	0.26	1.93	1.93	1.43	0.25
B3P86	H-GGA	1.06	−0.22	1.37	1.37	2.02	2.02	1.50	0.99
B1B95	HM-GGA	0.88	0.53	1.54	1.54	2.12	−0.83	1.51	0.19
TPSS1KCIS	HM-GGA	1.14	−0.26	1.67	1.67	2.00	2.00	1.59	1.03
MPW1PW91	H-GGA	1.13	0.67	2.15	2.15	2.14	2.14	1.74	1.55
TPSSTPSS	M-GGA	2.40	−2.40	0.18	−0.15	1.97	1.97	1.79	−0.20
MPW1N	H-GGA	1.60	1.60	2.21	2.21	2.06	2.06	1.91	1.91
MPW1S	H-GGA	1.62	−0.79	1.87	1.87	2.26	2.26	1.93	0.96
MPW1K	H-GGA	1.88	1.88	2.21	2.21	2.02	2.02	2.00	2.00
BP86	GGA	1.92	−1.19	1.83	1.83	2.28	2.28	2.05	0.80
TPSSLYP1W	M-GGA	1.61	0.57	3.59	3.59	1.71	1.71	2.05	1.63
B97-1	H-GGA	1.57	1.14	3.23	3.23	1.98	1.98	2.06	1.89
MPWPW91	GGA	2.02	−1.34	1.72	1.72	2.31	2.31	2.07	0.73
MPWB95	M-GGA	3.12	−3.12	0.39	−0.39	1.93	1.93	2.10	−0.56
BB95	M-GGA	1.74	−1.03	1.76	1.76	2.66	2.66	2.11	1.00
PBEPBE	GGA	3.07	−3.07	0.19	−0.12	2.11	2.11	2.11	−0.41
PBE1W	GGA	1.88	−0.71	2.64	2.64	2.19	2.19	2.16	1.12
MPW3LYP	H-GGA	1.80	1.45	3.73	3.73	1.85	1.85	2.20	2.07
MPWLYP	GGA	1.96	0.30	4.05	4.05	2.04	2.04	2.41	1.74
B3PW91	H-GGA	1.92	1.80	3.66	3.66	2.33	2.33	2.43	2.38
PBE1KCIS	HM-GGA	2.11	1.86	3.97	3.97	2.02	2.02	2.44	2.34
B98	H-GGA	2.15	1.97	4.05	4.05	2.06	2.06	2.50	2.42
BPBE	GGA	1.79	0.58	3.66	3.66	2.90	2.90	2.61	2.13
BPW91	GGA	1.83	0.78	3.91	3.91	2.85	2.85	2.65	2.23
MPWLYP1W	GGA	2.13	1.14	4.98	4.98	2.11	2.11	2.69	2.30
MPW1KCIS	HM-GGA	2.56	2.37	5.12	5.12	2.27	2.27	2.96	2.88
B3LYP	H-GGA	2.94	2.94	5.28	5.28	2.02	2.02	3.04	3.04
PBELYP1W	GGA	2.80	2.02	6.08	6.08	2.09	2.09	3.17	2.86
BLYP	GGA	2.88	2.36	6.17	6.17	2.36	2.36	3.33	3.12
B97-2	H-GGA	3.51	3.51	5.58	5.58	2.43	2.43	3.49	3.49
B1LYP	H-GGA	3.78	3.78	5.95	5.95	1.98	1.98	3.49	3.49
MPWKCIS1K	HM-GGA	4.17	4.17	5.26	5.26	2.07	2.07	3.55	3.55
BhandHLYP	H-GGA	4.57	4.57	5.33	5.33	1.74	1.74	3.59	3.59
VSXC	M-GGA	5.15	−5.15	4.14	−4.14	3.40	3.40	4.25	−1.53
HCTH147	GGA	4.13	4.03	7.60	7.6	2.92	2.92	4.34	4.30
HCTH407	GGA	5.12	5.07	8.82	8.82	3.05	3.05	5.03	5.01

[a] All values are in kcal/mol. Functionals are ordered by the "all structures" MUE. M06 functionals are not included in this table, as they were not used for solvent systems.

mol. A bad description of some bond by the DFT functionals is likely to be responsible for this discrepancy.

• If we consider some recent benchmarking studies,[40,41] our ranking of functionals is not entirely expected. Notably, MPWB1K, the first rated functional, is more accurate than any of the M06 family functionals. As these functionals belong to the same lineage, with M06 functionals being the more recent, our results are out of the ordinary. We also fail to observe the assumption that a greater fraction of Hartree−Fock exchange leads to better results for activation energies. Furthermore, PBE1PBE also performs very well, whereas in those other studies that is not the case. This simply shows how DFT methods in general are biased toward certain reactions, due to their intrinsic parametrization and (many times) error cancelation. In these studies, carbon species are prevailing, and it is acceptable that our phosphodiester system does not agree exactly with those. This disagreement by itself is sufficient to prove the significance of this study.

• Excluding the above conclusions, we find it very difficult to make any other assumptions regarding the functionals benchmarking. Due to the great heterogeneity of functionals and their reliance on empirical parametrization, we believe that a systematic analysis would not reach appreciable conclusions. Because of this, we adopted a more descriptive approach, and we will not try to interpret why certain functionals give better or worse results.

According to the MUE for all the systems, MPWB1K is the more accurate functional for studying phosphodiester bond hydrolysis, followed by MPW1B95 and PBE1PBE. The first two of these functionals have chemical accuracy, meaning a MUE smaller than 1 kcal/mol. MPWB1K and MPW1B95 also give the lowest MUE when considering barrier heights, followed by B1B95. In this case, these three functionals have chemical accuracy. For reaction energies, MPWB1K, MPW1B95, and TPSSLYP1W give the best results but with errors larger than 1 kcal/mol. We should also point out that B3LYP did not perform particularly well and, therefore, is not suited to study the energetic of this type of reactions. Performance of the DFT functionals in solution and implicit solvent is very similar as shown in Tables 9 and 10. There is no major reorganization in their

Table 9. MUEs and MSEs for the Best Studied Functionals for Vacuum Systems[a]

		transition state		intermediate		minimum		all	
functional	type	MUE	MSE	MUE	MSE	MUE	MSE	MUE	MSE
MPWB1K	HM-GGA	0.40	−0.40	0.20	−0.20	1.69	−1.61	0.88	−0.84
MPW1B95	HM-GGA	1.10	−1.10	0.10	−0.09	1.83	−1.50	1.19	−1.06
BB1K	HM-GGA	0.78	0.72	1.09	1.09	2.09	−1.33	1.36	−0.02
PBE1PBE	H-GGA	1.37	−1.09	0.56	0.56	2.29	−2.10	1.58	−1.16
B1B95	HM-GGA	0.90	0.23	1.48	1.48	2.44	−1.15	1.63	−0.07
B3P86	H-GGA	1.23	−0.64	1.20	1.20	2.27	−1.89	1.64	−0.77
TPSSh	HM-GGA	1.88	−1.81	0.23	0.23	2.19	−1.54	1.67	−1.30
MPW1N	H-GGA	1.15	1.15	1.93	1.93	2.30	−1.99	1.77	0.05
TPSS1KCIS	HM-GGA	1.39	−0.68	1.60	1.60	2.27	−1.57	1.78	−0.58
MPW1PW91	H-GGA	1.11	0.21	1.97	1.97	2.41	−1.87	1.80	−0.27
MPW1K	H-GGA	1.27	1.27	1.92	1.92	2.29	−2.01	1.80	0.09
M06	HM-GGA	1.82	−1.82	1.53	−1.53	1.96	−1.96	1.82	−1.82
TPSSTPSS	M-GGA	2.79	−2.79	0.19	−0.11	2.28	−1.49	2.06	−1.73
B97-1	H-GGA	1.52	0.65	3.07	3.07	2.26	−1.94	2.12	0.10
MPW3LYP	H-GGA	1.55	0.87	3.49	3.49	2.10	−1.89	2.16	0.29
TPSSLYP1W	M-GGA	1.72	0.01	3.50	3.50	1.97	−1.27	2.18	0.20
MPW1S	H-GGA	2.06	−1.27	1.81	1.81	2.58	−1.79	2.22	−0.86
B3PW91	H-GGA	1.59	1.34	3.48	3.48	2.61	−1.58	2.38	0.60
BP86	GGA	2.41	−1.63	1.82	1.82	2.62	−1.58	2.38	−0.92
PBE1KCIS	HM-GGA	1.84	1.35	3.78	3.78	2.29	−1.85	2.41	0.56
MPWPW91	GGA	2.55	−1.83	1.70	1.70	2.65	−1.79	2.42	−1.11
B98	H-GGA	1.80	1.44	3.84	3.84	2.34	−1.88	2.42	0.59
MPWB95	M-GGA	3.50	−3.50	0.27	−0.27	2.45	−1.39	2.43	−2.01
PBEPBE	GGA	3.53	−3.53	0.23	−0.09	2.47	−2.08	2.45	−2.27
PBE1W	GGA	2.37	−1.24	2.62	2.62	2.54	−1.85	2.49	−0.71
BB95	M-GGA	2.18	−1.38	1.87	1.87	3.19	−0.92	2.52	−0.55
MPWLYP	GGA	2.20	−0.37	3.95	3.95	2.37	−1.77	2.62	−0.07
B3LYP	H-GGA	2.37	2.37	5.03	5.03	2.25	−1.56	2.86	1.33
MPWLYP1W	GGA	2.31	0.46	4.87	4.87	2.44	−1.66	2.87	0.49
BPBE	GGA	2.00	0.12	3.64	3.64	3.40	−1.34	2.89	0.24
BPW91	GGA	2.00	0.30	3.88	3.88	3.34	−1.32	2.91	0.37
MPW1KCIS	HM-GGA	2.36	1.98	4.95	4.95	2.53	−1.60	2.94	1.14
M06-HF	HM-GGA	4.64	−0.53	1.51	1.51	2.08	−1.53	2.99	−0.52
B1LYP	H-GGA	3.20	3.20	5.67	5.67	2.20	−1.49	3.29	1.82
PBELYP1W	GGA	2.89	1.34	5.97	5.97	2.42	−1.65	3.32	1.07
BhandHLYP	H-GGA	3.99	3.99	4.89	4.89	1.94	−1.80	3.35	1.85
MPWKCIS1K	HM-GGA	3.64	3.64	4.91	4.91	2.31	−1.77	3.36	1.73
B97-2	H-GGA	3.00	3.00	5.36	5.36	2.75	−1.63	3.37	1.62
BLYP	GGA	2.77	1.73	6.06	6.06	2.76	−1.30	3.42	1.38
VSXC	M-GGA	5.13	−5.13	3.76	−3.76	2.74	1.37	3.90	−2.26

[a] All values are in kcal/mol. Functionals are ordered by the "all structures" MUE.

positions on these two tables, but, as a general trend, vacuum structures have a slightly larger MUE.

3.4. Basis Set Benchmarking. Basis set benchmarking analysis is a straightforward one: larger basis sets produce better results but are also more computationally demanding. This relationship is not linear though. Usually, when approaching the complete basis set limit, measured properties start converging to a point where computational effort does not justify the gain in terms of accuracy. Note that relative energies converge much faster than absolute energies. Dunning correlation consistent basis sets are especially useful to evaluate this effect, since they produce results in which the energy improves consistently as we go from double-ζ to triple-ζ and beyond, in particular when applied to post-Hartree−Fock methods.

Table 11 shows the results for the basis set benchmarking regarding phosphodiester hydrolysis. At the bottom of the table are the correlation consistent basis sets results. Aug-cc-pvqt was the most complete basis set we could use, conditioned by computational resources and our model size. It is noticeable, however, that in the transition from triple-ζ (aug-cc-pvtz) to quadruple-ζ (aug-cc-pvqz), MUE for energy values change around 1 kcal/mol. This is an acceptable convergence criterion and basically means that calculations with a quintuple-zeta basis set (aug-cc-pz5z) would not be significantly different.

Regarding all others basis sets, we could easily isolate the factors that most influence the results. Diffuse functions on heavy atoms seem to be the most prominent factors. Their inclusion diminishes the total MUE by more than 6 kcal/mol, even in an already large basis set such as 6-311G(2d,2p). This comes with no surprise though, since it is known that diffuse functions are very important in the electronic description of negatively charged molecules with more loosely bound electrons, as in these phosphodiester systems. Accordingly, hydroxide attacking systems originate larger MUEs as opposed to water attacking systems, since they have one extra negative charge. In the transition state of the hydroxide reactions the error due to the noninclusion of diffuse functions is superior to 20 kcal/mol.

The inclusion of polarization functions is very important, both in hydrogen and heavy atoms. The difference between 6 and 311++G(d,p) and 6-311++G(2d,2p) MUEs, for example, is superior to 1 kcal/mol, and it increases to about

Table 10. MUEs and MSEs for the Best Studied Functionals for Implicit Solvent Systems[a]

		transition state		intermediate		minimum		all	
functional	type	MUE	MSE	MUE	MSE	MUE	MSE	MUE	MSE
MPWB1K	HM-GGA	0.27	0.20	0.12	0.12	1.23	−0.86	0.63	−0.24
MPW1B95	HM-GGA	0.56	−0.49	0.09	0.06	1.29	−0.81	0.75	−0.51
PBE1PBE	H-GGA	0.67	−0.26	0.84	0.84	1.74	−1.36	1.13	−0.48
TPSSh	HM-GGA	1.16	−1.07	0.30	0.3	1.67	−0.94	1.19	−0.74
B3P86	H-GGA	0.88	0.19	1.54	1.54	1.77	−1.2	1.37	−0.10
B1B95	HM-GGA	0.85	0.83	1.61	1.61	1.80	−0.5	1.38	0.45
TPSS1KCIS	HM-GGA	0.90	0.17	1.75	1.75	1.73	−0.94	1.40	0.04
BB1K	HM-GGA	1.31	1.31	1.40	1.40	1.60	−0.62	1.45	0.56
TPSSTPSS	M-GGA	2.02	−2.02	0.18	−0.18	1.66	−0.92	1.51	−1.21
MPW1S	H-GGA	1.18	−0.31	1.92	1.92	1.95	−1.1	1.63	−0.18
MPW1PW91	H-GGA	1.15	1.12	2.33	2.33	1.88	−1.13	1.67	0.46
BB95	M-GGA	1.30	−0.69	1.65	1.65	2.14	−0.34	1.70	−0.08
BP86	GGA	1.43	−0.75	1.84	1.84	1.94	−0.97	1.71	−0.32
MPWPW91	GGA	1.49	−0.84	1.74	1.74	1.97	−1.11	1.73	−0.43
MPWB95	M-GGA	2.75	−2.75	0.51	−0.51	1.40	−0.76	1.76	−1.51
PBEPBE	GGA	2.61	−2.61	0.15	−0.15	1.76	−1.4	1.78	−1.64
PBE1W	GGA	1.40	−0.18	2.67	2.67	1.85	−1.17	1.83	−0.01
TPSSLYP1W	M-GGA	1.51	1.14	3.68	3.68	1.44	−0.71	1.92	0.91
B97-1	H-GGA	1.62	1.62	3.39	3.39	1.69	−1.2	2.01	0.85
MPW1N	H-GGA	2.05	2.05	2.50	2.50	1.81	−1.18	2.05	0.85
MPW1K	H-GGA	2.50	2.50	2.51	2.51	1.76	−1.14	2.20	1.04
MPWLYP	GGA	1.73	0.97	4.15	4.15	1.71	−1.09	2.20	0.78
MPW3LYP	H-GGA	2.04	2.04	3.96	3.96	1.59	−1.18	2.25	1.14
BPBE	GGA	1.58	1.04	3.68	3.68	2.41	−0.7	2.33	0.87
BPW91	GGA	1.67	1.25	3.94	3.94	2.36	−0.69	2.40	1.01
PBE1KCIS	HM-GGA	2.37	2.37	4.16	4.16	1.75	−1.12	2.48	1.33
B3PW91	H-GGA	2.26	2.26	3.83	3.83	2.05	−0.87	2.49	1.32
MPWLYP1W	GGA	1.96	1.82	5.10	5.10	1.78	−0.99	2.52	1.35
B98	H-GGA	2.50	2.50	4.26	4.26	1.79	−1.12	2.57	1.40
MPW1KCIS	HM-GGA	2.76	2.76	5.30	5.30	2.01	−0.93	2.97	1.79
PBELYP1W	GGA	2.70	2.70	6.19	6.19	1.77	−0.99	3.03	1.93
B3LYP	H-GGA	3.52	3.52	5.52	5.52	1.78	−0.88	3.22	2.16
BLYP	GGA	3.00	3.00	6.28	6.28	1.96	−0.67	3.24	2.19
B97-2	H-GGA	4.02	4.02	5.79	5.79	2.11	−0.86	3.61	2.42
B1LYP	H-GGA	4.36	4.36	6.23	6.23	1.76	−0.79	3.69	2.67
MPWKCIS1K	HM-GGA	4.71	4.71	5.61	5.61	1.84	−0.96	3.74	2.62
BhandHLYP	H-GGA	5.15	5.15	5.77	5.77	1.55	−0.99	3.83	2.82
HCTH147	GGA	4.66	4.66	7.74	7.74	2.38	−0.83	4.37	3.08
VSXC	M-GGA	5.18	−5.18	4.52	−4.52	4.06	2.47	4.60	−1.99
HCTH407	GGA	5.82	5.82	8.99	8.99	2.53	−0.92	5.14	3.76

[a] All values are in kcal/mol. Functionals are ordered by the "all structures" MUE. M06 functionals are not included in this table, as they were not used for solvent systems.

Table 11. MUEs for the Analyzed Basis Sets[a]

			hydroxide		water	
	basis set	all structures	barrier energies	reaction energies	barrier energies	reaction energies
140	6-31G(d)	8.69	25.22	16.86	6.82	6.08
172	6-31+G(d)	1.98	4.16	0.42	2.39	2.02
176	6-311G(d)	8.34	27.10	16.27	5.94	5.55
196	6-31+G(d,p)	1.98	3.85	0.28	2.61	1.97
200	6-311G(d,p)	8.26	27.06	16.11	6.09	5.29
208	6-311+G(d)	2.23	4.92	1.43	2.43	2.15
216	6-311++G(d)	2.27	5.14	1.63	2.44	2.15
232	6-311+G(d,p)	1.44	3.82	0.29	1.75	1.26
240	6-311++G(d,p)	1.50	4.07	0.51	1.78	1.27
264	6-311G(2d,2p)	6.88	23.61	16.69	4.27	3.91
296	6-311+G(2d,2p)	0.39	1.09	0.41	0.36	0.33
304	6311++G(2d,2p)	0.42	1.37	0.68	0.38	0.26
464	6311++G(3df,3pd)	0.27	0.89	0.83	0.13	0.14
260	aug-cc-pvdz	4.53	7.78	0.02	6.93	4.18
556	aug-cc-pvtz	0.70	1.25	0.01	1.08	0.62
1012	aug-cc-pvqz	0.00	0.00	0.00	0.00	0.00

[a] Values are in kcal/mol. MSEs were not included because they have always the same absolute value as MUEs but with negative sign.

2 kcal/mol when we consider transition state structures. Further improvement of the basis set with superior angular momentum polarizable functions does not justify the computational effort. With the 6-311++G(3df,3pd) basis set the

number of basis functions raises more than 30% in relation to 6-311++G(2d,2p) and the MUE barely improves.

The splitting of the sp shell has lesser significance than we would have expected, as there is little or no improvement in the energy values when shifting from double-ζ to triple-ζ basis sets.

4. Conclusions

During this work we were able to describe the reaction path of the phosphodiester bond hydrolysis of dimethylphosphate at the approximated CCSD(T)/CBS//B3LYP/6-311++G(2d,2p) level. We modeled the reactions with hydroxide and water as nucleophiles, both in the gas phase and in implicit solvent. The geometries and energies obtained are in agreement with previous studies by other groups.

We then used these structures to perform a benchmarking study on 52 density functionals. HF, MP2, MP3, MP4, and CCSD single point energies were also computed. Approximated CCSD(T)/CBS energy values were used as reference. We concluded that, overall, MPWB1K, MPW1B95, and PBE1PBE give the lowest MUEs of all functionals: 0.75, 0.97, and 1.35 kcal/mol, respectively. Taking into account just the activation energies, MPWB1K, MPW1B95, and B1B95 are the best positioned functionals, with MUEs of 0.34, 0.83, and 0.88 kcal/mol, respectively.

A basis set benchmarking, carried out on the same systems, showed us that diffuse functions on heavy atoms are a very important variable when choosing a basis set for the systems. Regarding polarization functions, having two additional functions of higher angular momentum for each atom is the best option. As a result of our calculations, we are able to point out the 6-311+G(2d,2p) basis set as having the best equilibrium between accuracy and computational cost.

We believe that our benchmarking results will be of great help to further studies on related phosphodiester systems. This includes not only pure chemical problems but also biochemical studies in which DNA, RNA, and phospholipids are required to be depicted at a quantum level.

Acknowledgment. The authors acknowledge the FCT (Fundação para a Ciência e Tecnlogia) for financial support through project PTDC/QUI/68302/2006.

References

(1) Nelson, D. L.; Cox, M. M. *Lehninger Principles of Biochemistry*, 4th ed.; Palgrave Macmillan: 2004; pp 276−277.

(2) Westheimer, F. H. Why Nature Chose Phosphates. *Science* **1987**, *235*, 1173–1178.

(3) Lim, C.; Karplus, M. Nonexistence of Dianionic Pentacovalent Intermediates in an Abinitio Study of the Base-Catalyzed-Hydrolysis of Ethylene Phosphate. *J. Am. Chem. Soc.* **1990**, *112*, 5872–5873.

(4) Dejaegere, A.; Lim, C.; Karplus, M. Dianionic Pentacoordinate Species in the Base-Catalyzed-Hydrolysis of Ethylene and Dimethyl-Phosphate. *J. Am. Chem. Soc.* **1991**, *113*, 4353–4355.

(5) Uchimaru, T.; Tanabe, K.; Nishikawa, S.; Taira, K. Abinitio Studies of a Marginally Stable Intermediate in the Base-Catalyzed Methanolysis of Dimethyl-Phosphate and Nonexistence of the Stereoelectronically Unfavorable Transition-State. *J. Am. Chem. Soc.* **1991**, *113*, 4351–4353.

(6) Lim, C.; Tole, P. Concerted Hydroxyl Ion Attack and Pseudorotation in the Base-Catalyzed-Hydrolysis of Methyl Ethylene Phosphate. *J. Phys. Chem.* **1992**, *96*, 5217–5219.

(7) Liang, C. X.; Ewig, C. S.; Stouch, T. R.; Hagler, A. T. Abinitio Studies of Lipid Model Species 0.1. Dimethyl-Phosphate and Methyl Propyl Phosphate Anions. *J. Am. Chem. Soc.* **1993**, *115*, 1537–1545.

(8) Dejaegere, A.; Liang, X. L.; Karplus, M. Phosphate Ester Hydrolysis - Calculation of Gas-Phase Reaction Paths and Solvation Effects. *J. Chem. Soc., Faraday Trans.* **1994**, *90*, 1763–1770.

(9) Landin, J.; Pascher, I.; Cremer, D. Ab-Initio and Semiempirical Conformation Potentials for Phospholipid Head Groups. *J. Phys. Chem.* **1995**, *99*, 4471–4485.

(10) Hu, C. H.; Brinck, T. Theoretical studies of the hydrolysis of the methyl phosphate anion. *J. Phys. Chem. A* **1999**, *103*, 5379–5386.

(11) Mercero, J. M.; Barrett, P.; Lam, C. W.; Fowler, J. E.; Ugalde, J. M.; Pedersen, L. G. Quantum mechanical calculations on phosphate hydrolysis reactions. *J. Comput. Chem.* **2000**, *21*, 43–51.

(12) Murashov, V. V.; Leszczynski, J. A comparison of the B3LYP and MP2 methods in the calculation of phosphate complexes. *J. Mol. Struct. - Theochem* **2000**, *529*, 1–14.

(13) Lopez, X.; Dejaegere, A.; Karplus, M. Solvent effects on the reaction coordinate of the hydrolysis of phosphates and sulfates: Application of Hammond and anti-Hammond postulates to understand hydrolysis in solution. *J. Am. Chem. Soc.* **2001**, *123*, 11755–11763.

(14) Bianciotto, M.; Barthelat, J. C.; Vigroux, A. Reactivity of phosphate monoester monoanions in aqueous solution. 1. Quantum mechanical calculations support the existence of "anionic zwitterion" MeO+(H)PO3 (2-) as a key intermediate in the dissociative hydrolysis of the methyl phosphate anion. *J. Am. Chem. Soc.* **2002**, *124*, 7573–7587.

(15) Bianciotto, M.; Barthelat, J. C.; Vigroux, A. Reactivity of phosphate monoester monoanions in aqueous solution. 2. A theoretical study of the elusive zwitterion intermediates RO+(H)PO32-. *J. Phys. Chem. A* **2002**, *106*, 6521–6526.

(16) Lopez, X.; York, D. M.; Dejaegere, A.; Karplus, M. Theoretical studies on the hydrolysis of phosphate diesters in the gas phase, solution, and RNase A. *Int. J. Quantum Chem.* **2002**, *86*, 10–26.

(17) Kolandaivel, P.; Kanakaraju, R. Structure, stability and interaction studies on nucleotide analogue systems. *Int. J. Mol. Sci.* **2003**, *4*, 486–502.

(18) Wang, Y. N.; Topol, I. A.; Collins, J. R.; Burt, S. K. Theoretical studies on the hydrolysis of mono-phosphate and tri-phosphate in gas phase and aqueous solution. *J. Am. Chem. Soc.* **2003**, *125*, 13265–13273.

(19) Chen, X.; Zhan, C. G. Theoretical determination of activation free energies for alkaline hydrolysis of cyclic and acyclic phosphodiesters in aqueous solution. *J. Phys. Chem. A* **2004**, *108*, 6407–6413.

(20) Range, K.; McGrath, M. J.; Lopez, X.; York, D. M. The structure and stability of biological metaphosphate, phosphate, and phosphorane compounds in the gas phase and in solution. *J. Am. Chem. Soc.* **2004**, *126*, 1654–1665.

(21) Iche-Tarrat, N.; Barthelat, J. C.; Rinaldi, D.; Vigroux, A. Theoretical studies of the hydroxide-catalyzed P-O cleavage reactions of neutral phosphate triesters and diesters in aqueous solution: Examination of the changes induced by H/Me substitution. *J. Phys. Chem. B* **2005**, *109*, 22570–22580.

(22) Imhof, P.; Fischer, S.; Kramer, R.; Smith, J. C. Density functional theory analysis of dimethylphosphate hydrolysis: effect of solvation and nucleophile variation. *J. Mol. Struct. - Theochem* **2005**, *713*, 1–5.

(23) Lopez, X.; Dejaegere, A.; Leclerc, F.; York, D. M.; Karplus, M. Nucleophilic attack on phosphate diesters: A density functional study of in-line reactivity in dianionic, monoanionic, and neutral systems. *J. Phys. Chem. B* **2006**, *110*, 11525–11539.

(24) Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. General performance of density functionals. *J. Phys. Chem. A* **2007**, *111*, 10439–10452.

(25) Cossi, M.; Scalmani, G.; Rega, N.; Barone, V. New developments in the polarizable continuum model for quantum mechanical and classical calculations on molecules in solution. *J. Chem. Phys.* **2002**, *117*, 43–54.

(26) Cossi, M.; Barone, V.; Mennucci, B.; Tomasi, J. Ab initio study of ionic solutions by a polarizable continuum dielectric model. *Chem. Phys. Lett.* **1998**, *286*, 253–260.

(27) Mennucci, B.; Tomasi, J. Continuum solvation models: A new approach to the problem of solute's charge distribution and cavity boundaries. *J. Chem. Phys.* **1997**, *106*, 5151–5158.

(28) Cances, E.; Mennucci, B.; Tomasi, J. A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics. *J. Chem. Phys.* **1997**, *107*, 3032–3041.

(29) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. Basis-set convergence of correlated calculations on water. *J. Chem. Phys.* **1997**, *106*, 9639–9646.

(30) Halkier, A.; Helgaker, T.; Jorgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. Basis-set convergence in correlated calculations on Ne, N-2, and H2O. *Chem. Phys. Lett.* **1998**, *286*, 243–252.

(31) Jurecka, P.; Hobza, P. On the convergence of the (Delta E-CCSD(T)-Delta E-MP2) term for complexes with multiple H-bonds. *Chem. Phys. Lett.* **2002**, *365*, 89–94.

(32) Hobza, P.; Sponer, J. Toward true DNA base-stacking energies: MP2, CCSD(T), and complete basis set calculations. *J. Am. Chem. Soc.* **2002**, *124*, 11802–11808.

(33) Schrodinger. *Jaguar, version 7.5*; New York, 2008.

(34) 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.

(35) Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*; Wiley-Interscience: 1997; pp *136*−140.

(36) Gorenstein, D. G.; Luxon, B. A.; Findlay, J. B. Stereoelectronic Effects in the Reactions of Phosphate Diesters - Abinitio Molecular-Orbital Calculations of Reaction Surfaces. 2. *J. Am. Chem. Soc.* **1979**, *101*, 5869–5875.

(37) Wolfenden, R.; Ridgway, C.; Young, G. Spontaneous hydrolysis of ionized phosphate monoesters and diesters and the proficiencies of phosphatases and phosphodiesterases as catalysts. *J. Am. Chem. Soc.* **1998**, *120*, 833–834.

(38) Kumamoto, J.; Cox, J. R.; Westheimer, F. H. Barium Ethylene Phosphate. *J. Am. Chem. Soc.* **1956**, *78*, 4858–4860.

(39) Schroeder, G. K.; Lad, C.; Wyman, P.; Williams, N. H.; Wolfenden, R. The time required for water attack at the phosphorus atom of simple phosphodiesters and of DNA. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 4052–4055.

(40) 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*, 215–241.

(41) Zhao, Y.; Truhlar, D. G. Density functionals with broad applicability in chemistry. *Acc. Chem. Res.* **2008**, *41*, 157–167.

(42) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate Spin-Dependent Electron Liquid Correlation Energies for Local Spin-Density Calculations - a Critical Analysis. *Can. J. Phys.* **1980**, *58*, 1200–1211.

(43) Slater, J. C. *Quantum Theory of Molecular and Solids. Vol. 4: The Self-Consistent Field for Molecular and Solids*; McGraw-Hill: New York, 1974.

(44) 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: Condens. Matter Mater. Phys.* **1988**, *37*, 785–789.

(45) Becke, A. D. Density-functional thermochemistry 0.4. A new dynamical correlation functional and implications for exact-exchange mixing. *J. Chem. Phys.* **1996**, *104*, 1040–1046.

(46) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic-Behavior. *Phys. Rev. A: At., Mol., Opt. Phys.* **1988**, *38*, 3098–3100.

(47) Becke, A. D. Density-Functional Thermochemistry 0.3. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(48) Zhao, Y.; Truhlar, D. G. Hybrid meta density functional theory methods for thermochemistry, thermochemical kinetics, and noncovalent interactions: The MPW1B95 and MPWB1K models and comparative assessments for hydrogen bonding and van der Waals interactions. *J. Phys. Chem. A* **2004**, *108*, 6908–6918.

(49) Perdew, J. P. Density-Functional Approximation for the Correlation-Energy of the Inhomogeneous Electron-Gas. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1986**, *33*, 8822–8824.

(50) Perdew, J. P. Unified Theory of Exchange and Correlation Beyond the Local Density Aproximation. In *Electronic*

Structure of Solids; Zieche, P., Eschrig, H., Eds.; Berlin, Germany, 1991; pp 11−20.

(51) Hamprecht, F. A.; Cohen, A. J.; Tozer, D. J.; Handy, N. C. Development and assessment of new exchange-correlation functionals. *J. Chem. Phys.* **1998**, *109*, 6264–6271.

(52) Zhao, Y.; Truhlar, D. G. Density functional for spectroscopy: No long-range self-interaction error, good performance for Rydberg and charge-transfer states, and better performance on average than B3LYP for ground states. *J. Phys. Chem. A* **2006**, *110*, 13126–13130.

(53) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

(54) Wilson, P. J.; Bradley, T. J.; Tozer, D. J. Hybrid exchange-correlation functional determined from thermochemical data and ab initio potentials. *J. Chem. Phys.* **2001**, *115*, 9233–9242.

(55) Adamo, C.; Barone, V. Exchange functionals with improved long-range behavior and adiabatic connection methods without adjustable parameters: The mPW and mPW1PW models. *J. Chem. Phys.* **1998**, *108*, 664–675.

(56) Schmider, H. L.; Becke, A. D. Optimized density functionals from the extended G2 test set. *J. Chem. Phys.* **1998**, *108*, 9624–9631.

(57) Becke, A. D. Density-functional thermochemistry 0.5. Systematic optimization of exchange-correlation functionals. *J. Chem. Phys.* **1997**, *107*, 8554–8560.

(58) Krieger, J. B.; Chen, J. Q.; Iafrate, G. J.; Savin, A. Construction of an accurate self-interaction-corrected correlation energy functional based on an electron gas with a gap. *Electron Correl. Mater. Prop., [Proc. Int. Workshop], 1st* **1999**, 463–477.

(59) Zhao, Y.; Gonzalez-Garcia, N.; Truhlar, D. G. Benchmark database of barrier heights for heavy atom transfer, nucleophilic substitution, association, and unimolecular reactions and its use to test theoretical methods. *J. Phys. Chem. A* **2005**, *109*, 2012–2018.

(60) Gill, P. M. W. A new gradient-corrected exchange functional. *Mol. Phys.* **1996**, *89*, 433–445.

(61) Lynch, B. J.; Fast, P. L.; Harris, M.; Truhlar, D. G. Adiabatic connection for kinetics. *J. Phys. Chem. A* **2000**, *104*, 4811–4815.

(62) Zhao, Y.; Lynch, B. J.; Truhlar, D. G. Multi-coefficient extrapolated density functional theory for thermochemistry and thermochemical kinetics. *Phys. Chem. Chem. Phys.* **2005**, *7*, 43–52.

(63) Kormos, B. L.; Cramer, C. J. Adiabatic connection method for X-+RX nucleophilic substitution reactions (X = F, Cl). *J. Phys. Org. Chem.* **2002**, *15*, 712–720.

(64) Tao, J. M.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. Climbing the density functional ladder: Nonempirical meta-generalized gradient approximation designed for molecules and solids. *Phys. Rev. Lett.* **2003**, *91*, 146401.

(65) Dahlke, E. E.; Truhlar, D. G. Improved density functionals for water. *J. Phys. Chem. B* **2005**, *109*, 15677–15683.

(66) Lynch, B. J.; Zhao, Y.; Truhlar, D. G. Effectiveness of diffuse basis functions for calculating relative energies by density functional theory. *J. Phys. Chem. A* **2003**, *107*, 1384–1388.

(67) Handy, N. C.; Cohen, A. J. Left-right correlation energy. *Mol. Phys.* **2001**, *99*, 403–412.

(68) Hoe, W. M.; Cohen, A. J.; Handy, N. C. Assessment of a new local exchange functional OPTX. *Chem. Phys. Lett.* **2001**, *341*, 319–328.

(69) Zhao, Y.; Truhlar, D. G. A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions. *J. Chem. Phys.* **2006**, *125*, 194101.

(70) Xu, X.; Goddard, W. A. The X3LYP extended density functional for accurate descriptions of nonbond interactions, spin states, and thermochemical properties. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 2673–2677.

(71) Perdew, J. P.; Wang, Y. Accurate and Simple Analytic Representation of the Electron-Gas Correlation-Energy. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1992**, *45*, 13244–13249.

(72) Van Voorhis, T.; Scuseria, G. E. A never form for the exchange-correlation energy functional. *J. Chem. Phys.* **1998**, *109*, 400–410.

CT900649E

JCTC Journal of Chemical Theory and Computation

Quantum Chemistry Calculations Provide Support to the Mechanism of the Light-Induced Structural Changes in the Flavin-Binding Photoreceptor Proteins

M. G. Khrenova,*[,†] A. V. Nemukhin,[†,‡] B. L. Grigorenko,[†] A. I. Krylov,[§] and T. M. Domratcheva[||]

Chemistry Department, M.V. Lomonosov Moscow State University, Leninskie Gory 1/3, Moscow, 119991, Russian Federation, N.M. Emanuel Institute of Biochemical Physics, Russian Academy of Sciences, Kosygina 4, Moscow, 119334, Russian Federation, Department of Chemistry, University of Southern California, Los Angeles, California 90089, and Department of Biomolecular Mechanisms, Max Planck Institute for Medical Research, 69120 Heidelberg, Germany

Received April 5, 2010

Abstract: The proposed mechanisms of photoinduced reactions in the blue light using flavin chromophore photoreceptor proteins are primarily based on the results of X-ray crystallography and spectroscopy studies. Of particular value are the observed band shifts in optical and vibrational spectra upon formation of the signaling (light-induced) state. However, the same set of experimental data has given rise to contradictory interpretations suggesting different structures of the dark and signaling states. To verify the specific mechanism of light-induced changes involving the rotation/tautomerization transformations with the conserved Gln residue near the flavin chromophore, we performed accurate quantum chemical calculations of the equilibrium structures, vibrational and absorption bands of the model systems mimicking the BLUF domain of flavoprotein AppA. Geometry optimization and calculations of vibrational frequencies were carried out with the QM(B3LYP/cc-pVDZ)/MM(AMBER) approach starting from the representative molecular dynamics (MD) snapshots. The MD simulations were initiated from the available crystal structures of the AppA protein. Calculations of the vertical excitation energies were performed with the scaled opposite spin configuration interaction with single substitutions SOS-CIS(D) method that enables efficient treatment of excited states in large molecular systems. The computed molecular structures as well as the spectral shifts (the red shift by 12÷16 nm in absorption and the downshift by 25 cm^{-1} for the C4═O flavin vibrational mode) are in excellent agreement with the experimental results, lending a strong support to the mechanism proposed by Domratcheva et al. (*Biophys. J.* **2008,** *94*, 3872).

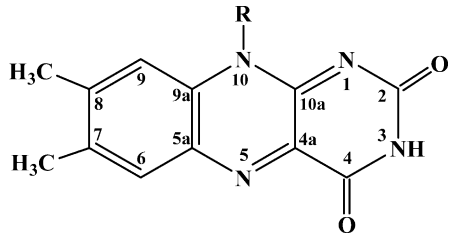
Introduction

Accurate characterization of electronically excited states and vibrational spectra in large molecular systems is a key step in establishing mechanisms of light-induced biophysical processes. Studies of the blue light using flavin adenine dinucleotide (FAD) sensor proteins called BLUF[1] provide prominent examples. The BLUF domains belong to the family of photoreceptor proteins whose function is linked to the subtle changes in the chromophore environment.[2] The primary photochemistry of their activation results in changes in the hydrogen-bond network near the chromophore and presumably in conformational changes involving the nearby

* Corresponding author. Telephone: +74959391096. Fax: +74959390283. E-mail: wasabiko13@gmail.com.
† M.V. Lomonosov Moscow State University.
‡ Russian Academy of Sciences.
§ University of Southern California.
|| Max Planck Institute for Medical Research.

10.1021/ct100179p © 2010 American Chemical Society
Published on Web 07/01/2010

Scheme 1. The Isoalloxazine Ring of FAD and Numbering of Its Atoms



amino acid residues. Upon light excitation the signaling intermediate state of the BLUF photocycle is formed that is characterized by a slight red shift (10÷15 nm) of the intense absorption bands in UV−vis spectra. The light-induced red shifts decay within minutes, and the dark state is recovered. The detailed mechanism of these events at the atomic resolution is still actively debated. All interpretations agree that there is a certain reconstruction of the hydrogen-bond network in the vicinity of the isoalloxazine ring of FAD (Scheme 1), however, a considerable disagreement on the light-induced specific changes in the peptide groups, especially near the N5, C4═O, N3H, and C2═O moieties of the chromophore, persists in the literature.[3−15]

The main focus of this work is on the BLUF domain containing protein AppA whose function is to control photosynthesis gene expression in the purple bacterium *Rhodobacter sphaeroides*.[3] Among other BLUF proteins, AppA is distinguished by the longest recovery time to the dark state. The absorption spectrum of the protein in the dark state consists of the broad peak at 443 nm attributed to the S_0-S_1 electronic transition of flavin buried in the chromophore-containing pocket, while in the light-induced transient form, this band is shifted to 456 nm, i.e., by 13 nm (0.08 eV).[4] The definite change in the vibrational spectrum upon formation of the signaling state is a red shift of about 20−23 cm^{-1} in the flavin C4═O stretching vibration.[4,16]

Such small shifts in the optical and vibrational bands are attributed to rearrangements in the immediate vicinity of the chromophore, but this is the only consensus of opinions among various research groups. The three-dimensional (3D) structures of the protein derived from the X-ray studies[6,17] have not resolved the disagreement but actually initiated the dispute. Figure 1 illustrates the main point of the debate. The molecular groups in the chromophore-containing pocket are from the side chains of Tyr21, Gln63, His44, and Asn45, as shown in Figure 1. The contradictory opinions concern the side chain of Gln63 in the dark state since X-ray crystallography cannot resolve the location of hydrogen atoms and unambiguously distinguish oxygen and nitrogen atoms. Consequently, different forms of Gln63 (in particular, amidic and imidic) and different orientations of its side chain toward Tyr21 have been proposed for the dark and light states of AppA.[6,17] One more amino acid side chain in the vicinity of the isoalloxazine ring of FAD (either Met106 or Trp104) is shown schematically in Figure 1. According to the results of Jung et al.,[17] the side chain of Met106 is located near FAD in the dark state, while Trp104 is located outside this region (the so-called, Trp_{out}/Met_{in} conformation of BLUF). The corresponding crystal structure is described as 2IYG in

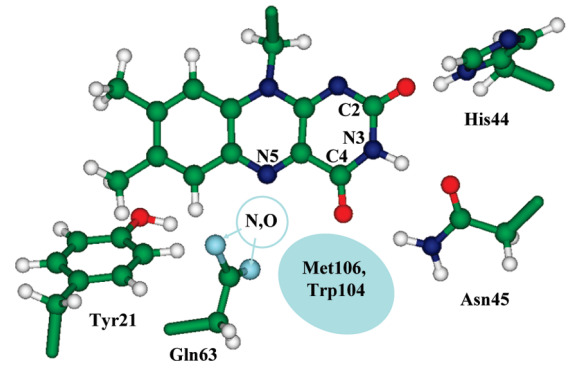


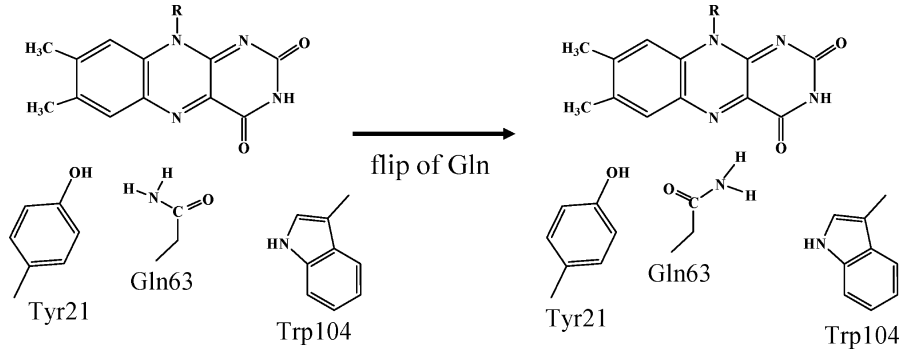
Figure 1. A fragment of the chromophore-containing pocket of AppA BLUF. Here, and below, green, red, and blue colors are reserved for carbon, oxygen, and nitrogen, respectively. The functional head of the side chain of Gln63 is shown schematically as explained in the text. The debated location of either Met106 or Trp104 residues is also schematically indicated.

the PDB archive.[18] In this structure, the side chain of Met106 is hydrogen bonded to Gln63. On the contrary, Anderson et al.[6] assigned the structure (PDB code 1YRX) with the side chain of Trp104 close to the chromophore and with Met106 outside the region (Trp_{in}/Met_{out}) to the dark state. Here, Trp104 forms a hydrogen bond to Gln63. Transition from one structure to another apparently requires considerable conformational changes of the protein.

Our recent paper[12] argues in favor of the Trp_{out}/Met_{in} structure (2IYG) for the dark state and proposes that the (Trp_{in}/Met_{out}) conformation of AppA corresponding to the 1YRX crystal structure should be assigned to the light-induced state. The calculations and careful analysis of the crystal structures indicated that the tautomeric imidic form of Gln63 is present in the light-induced state. The conclusions were supported by quantum chemical calculations of the structures and electronic and vibrational spectra for a series of clusters (i.e., the chromophore and the surrounding residues) as well as quantum mechanical/molecular mechanical (QM/MM) calculations with a relatively low-level description of the QM subsystem. Since then, new theoretical calculations were published,[7,8,19] however, some of their conclusions only partly coincided with our proposal. To provide more solid support to the mechanism formulated in ref 12, we performed here new higher-level calculations of the equilibrium geometries, vibrational spectra, and electronic transitions for the model systems mimicking the BLUF domain of the flavoprotein AppA.

From the computational point of view, accounting for such a small difference in excitation energy between the dark and light states (0.08 eV) poses a great challenge. The effect is due to a subtle reconstruction of the hydrogen-bond network near the chromophore, and, apparently, the model molecular clusters for excited states calculations should be large enough. They should include not only the chromophore but also the nearby amino acid side chains, some of which (Tyr, His, and Trp) donate their π electrons to the model system. Using time-dependent density functional theory (TD-DFT), which is a tempting choice for excited-state calculations in large systems, is questionable owing to the well-known difficulties

Scheme 2. The Mechanism of Light Activation in AppA by Anderson et al[6]



Dark state Light-induced state

with an unphysical description of the charge transfer (CT) states, which may contaminate the valence excited states.[20,21] The work of Sadeghian et al.[8] vividly demonstrated the problematic behavior of TD-DFT precisely for the systems considered in the present work. The application of the complete active space self-consistent field (CASSCF)-based approaches (multiconfigurational second-order perturbation theory, CASPT2, and multiconfigurational quasi-degenerate perturbation theory, MCQDPT2), which are shown to be accurate and efficient for calculating valence transitions in small- and medium-size organic chromophores,[22−32] may also be problematic because of the difficulties with the selection of active space that yields balanced description of the states of interest in such large model systems.

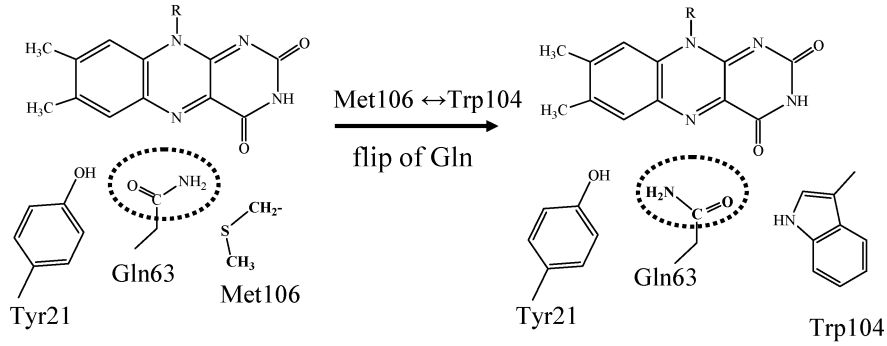
In this context, the scaled opposite-spin configuration interaction with single substitutions SOS-CIS(D) method that has been recently developed[33,34] is an attractive alternative to a more computationally expensive techniques. The approach is closely related to perturbatively corrected configuration interaction singles (CIS), CIS(D),[35] which can be viewed as an approximation of more rigorous equation-of-motion coupled cluster singles and doubles (EOM-CCSD).[36,37] The CIS wave functions are qualitatively correct for the excited states dominated by single-electron excitations including Rydberg, valence, and charge transfer (CT) states as well as interacting states of a mixed character. Due to single-reference formulation and size-intensive properties of the model, the relative energy differences between the excited states and the changes in excitation energies in homological series or isomers are reproduced much better than the absolute values of excitation energies (for which errors often exceed 1 eV). Moreover, the perturbative inclusion of double excitations in CIS(D) yields considerable improvement in accuracy, i.e., the CIS(D) excitation energies for well-behaved cases are very close to that of the EOM-CCSD ones. Overall, CIS(D) is similar to second-order approximate coupled cluster (CC2). The scaling of the method is also N^5, although the computational cost is less than that of CC2 owing to the noniterative character of the correction. Recently, CIS(D) was implemented using the SOS variant of second-order perturbation theory[33,34] in which only the opposite-spin correlation energy is computed, and the same-spin correlation is estimated from the opposite-spin one using an empirical scaling factor. For the ground-state calculations, SOS-MP2 demonstrated more robust behavior and higher accuracy than MP2,[38] suggesting a moderate improvement in accuracy for the excited-state variant. Moreover, the SOS implementation along with RI enabled the reduction of computational scaling to N^4. Thus, SOS-CIS(D) calculations can be performed for relatively large systems. In our recent studies of the green fluorescent protein, we observed an excellent performance of the SOS-CIS(D) method for excitation energies.[39]

Review of the Previous Quantum Chemical Calculations

As mentioned above, different mechanisms of the photocycle were proposed based on the results of experimental and computational studies. Several research groups support the assignment of the Trp_{in} conformation originated from the 1YRX crystal structure[6] to the dark state. An initial suggestion of Anderson et al.[6] was that photoactivation might induce the flip of the Gln63 side chain leading to a new hydrogen bond of Gln with the flavin's O4 (Scheme 2). The recent study from this group[15] concluded that light exposure did not cause noticeable change of Trp104 from Trp_{in} to Trp_{out} conformations.

The structural model[6] illustrated in Scheme 2 was used for interpretation of the fast spectroscopic studies[14,40−42] as well as of NMR studies.[9] Along with the results of mutagenesis and spectroscopy studies, Unno et al.[43] presented B3LYP/6-31G** calculations of the series of molecular clusters. The 7,8-dimethyl-10-glycerylisoalloxazine molecule was used as a model of the chromophore. In addition, water, acetamide, and 4-methylimidazole were included in the calculations to reproduce hydrogen bonds with the active-site residues of Gln63, Asn45, and His44. The fragments were arranged based on the crystal structure of AppA BLUF suggested by Anderson et al.[6] and subsequently optimized. The authors calculated the frequencies for the C4═O stretch in flavin and the electronic excitation energies. No convincing support for the dark protein structure denoted here as Trp_{in} have emerged from these calculations. However, the important conclusion from this work[43] is that there is no hydrogen bond between Gln63 and flavin's C4═O in the dark state, and, therefore, the Gln side chain rotates to make a new hydrogen bond with C4═O in the signaling state. We note

Scheme 3. The Mechanism of Light Activation in AppA by Jung et al[17]



Dark state Light-induced state

that this conclusion may apply to the (Trp_{out}/Met_{in})[17] conformation as well.

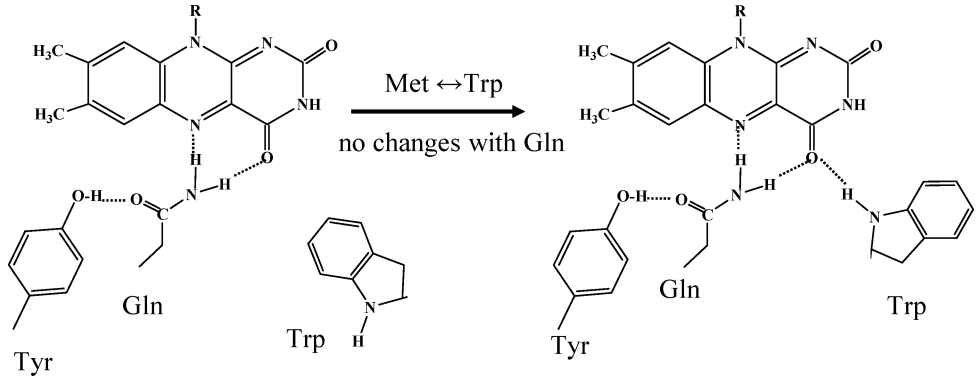
Götze and Saalfrank[19] performed quantum chemical calculations for the cluster composed of the model flavin molecule surrounded by the nearest 10 amino acid side chains (Tyr21, Ser23, Ser41, His44, Asn45, Gln63, Met76, His85, Trp104, and Met106). That was the largest model system (165 atoms) treated at the QM level among all previously considered models.[7,8,12,43] The equilibrium structures were optimized by B3LYP/6-31G* following classical molecular dynamics (MD) simulations initiated from the NMR structure PDBID: 2BUN.[10] Although the latter structure is poorly resolved and cannot be unambiguously assigned to either Trp_{in} or Trp_{out} conformation, the mechanism ultimately advocated by Götze and Saalfrank apparently assumes that the dark state corresponds to the Trp_{in} conformation. The TD-DFT (B3LYP/6-31G*) method was used to calculate excitation energies. The authors applied the following strategy: they calculated the excitation energies for the 20 available NMR structures of AppA BLUF[10] (without MD) and averaged the results. This corresponds to an ensemble consisting of these structures with equal weights. The calculated average UV−vis spectrum has peak maxima at 341 and 450 nm, whereas the experimental spectrum has maxima at 374 and 443 nm for the $S_0 \rightarrow S_2$ and $S_0 \rightarrow S_1$ transitions, respectively. The authors reported a systematic dependence of the computed spectra on the positions of conserved residue Ser41 which, according to their suggestion, is actively involved in the photoactivation.

Now we turn to the supporters of the model in which the dark state is associated with the (Trp_{out}/Met_{in}) conformation. Initially, this idea was formulated by Jung et al.[17] who proposed the photoactivation mechanism illustrated in Scheme 3. According to this model, light induces the exchange between Trp104 and Met106 to form the Trp_{in} conformation in the signaling state. Thus, the interaction of Met106 with Gln63 changes supporting a rotation of the side chain of the latter.

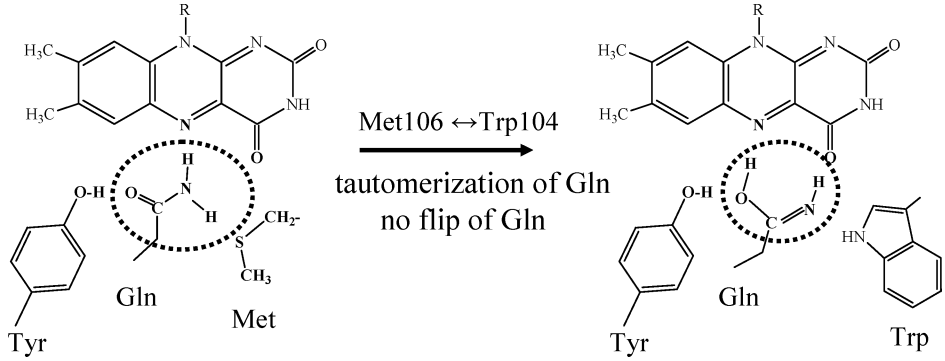
Calculations supporting this model were presented in ref 12 and included QM(RHF/6-31G)/MM(AMBER) optimizations of the equilibrium structures of the model system starting from crystal structures 2IYG and 1YRX. The optimized structure obtained in the calculation initiated using the 2IYG coordinates of heavy atoms agreed well with the parent crystal structure. The minimum energy structure obtained starting from 1YRX was also consistent with the respective crystal structure, however, it suggested the tautomerized form of the Gln63 side chain. This imidic form of Gln63 allows for an additional hydrogen bond of Gln63 with O4 of flavin, thus accounting for the small red shift in absorption maximum of the $S_0 \rightarrow S_1$ transition as well as for the downshift of the C4═O stretch of flavin. The red shift of about 10 nm in absorption was estimated using a relatively low-level approach, QM(CASSCF/6-31)/MM(AMBER). A series of quantum chemical calculations for molecular clusters representing the chromophore-binding pocket yielded qualitatively similar conclusions as the QM/MM simulations. In the cluster calculations, the chromophore was represented by the lumiflavin molecule, and the Tyr21, Gln63, Met106, Trp104 side chains were modeled by phenol, acetamide, dimethylthioether, and pyrrole molecules, respectively. The equilibrium geometry parameters were optimized with B3LYP/6-31G(d). Energies and wave functions of the five lowest singlet states were calculated by CISD in the active space of the 32 highest occupied and 32 lowest virtual Hartree−Fock orbitals. The vertical electron transition energies were obtained using CASSCF/6-31G(d) state averaged over the five lowest roots with the active space chosen based on the CISD solutions. The CASSCF active space consisted of three bonding π, one nonbonding np, and one antibonding π^* molecular orbitals (MOs) localized on the lumiflavin and phenol molecules. The CASSCF energies were corrected by MCQDPT2. The resulting excitation energies, 2.65 (467) and 2.59 eV (477 nm), were in excellent agreement with the experimental results, e.g., the computed red shift was 10 nm, as compared to the experimental value of 13 nm. About 40 cm^{-1} downshift for the C4═O stretch of flavin was obtained.

In another theoretical study, Obanayama et al.[7] also critically analyzed the models shown in Schemes 2[6] and 3.[17] The authors carried out quantum chemical calculations for the cluster composed of the lumiflavin molecule and the truncated side chains of Tyr21, His44, Asn45, Gln63, and Trp104. Geometry was optimized with B3LYP/6-31G** following the MD simulations of the BLUF domains from proteins AppA, BlrB, and Tll0078. The excitation energies were computed using TD-DFT (B3LYP/6-31G**). The calculations supported the model with the Trp_{out} conformation assigned to the dark state and with the Trp_{in} conformation

Scheme 4. The Mechanism of Light Activation in AppA by Obanayama et al[7]



Scheme 5. The Mechanism of Light Activation in BLUF Proteins by Sadeghian et al[8]



to the signaling state. The authors suggested the light-induced transformations illustrated in Scheme 4 assuming that movement of Trp104 from its 'out' to 'in' conformation is sufficient to strengthen the hydrogen bonds between flavin and Gln and Trp to explain the spectroscopic results. However, they failed to reproduce the major features of the signaling state: the computed absorption maximum was 9 nm blue shifted (instead of being red shifted), and the C4═O frequency was blue-shifted by 11 cm^{-1} (instead of being red shifted).

The theoretical work of Sadeghian et al.[8] suggested the mechanism of light activation of BLUF that has common features of the proposal formulated by Domratcheva et al.[12] The simulations were initiated from the crystal structure of another BLUF containing protein BlrB (PDB code 2BYC).[5] Correspondingly, a different numbering system was used for the residues. After the MD relaxation of the model structures, the QM/MM ground-state geometry optimization was performed starting from the representative MD snapshots. The B3LYP and local MP2 methods were employed in the QM subsystem, and the CHARMM force field was used for the MM part. The excited-state potential energy surfaces were described by TD-DFT with various functionals and selected single-point coupled-cluster calculations using the CC2 method. The authors clearly illustrated difficulties in computing excitation energies with the standard functionals (BP, B3LYP, and BHLYP) due to the self-interaction error, which leads to the gross underestimation of the excitation energies of the charge-transfer (CT) states, i.e., Tyr → flavin CT states. Only the relatively expensive CC2 method allowed the authors to obtain the correct ordering of the excited states in the dark state.

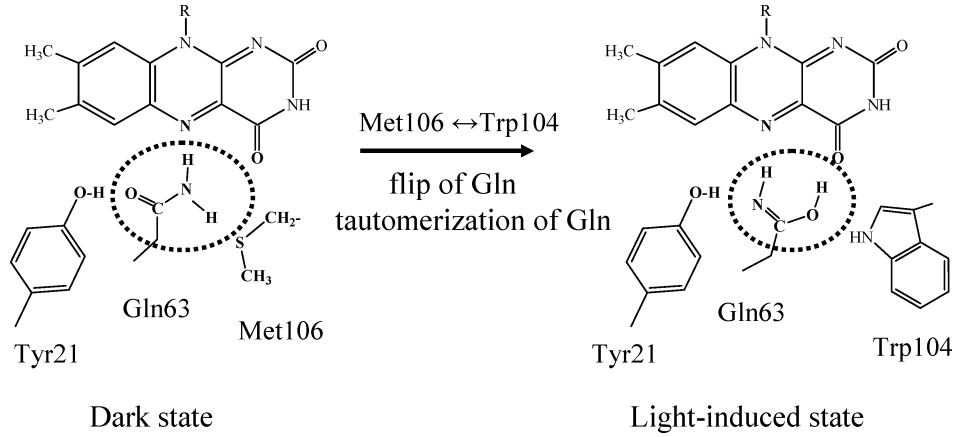
Sadeghian et al.[8] concluded that the (Trp_{out}/Met_{in}) protein structure should be associated with the dark state. For the light-induced state they suggested a variant of the (Trp_{in}/Met_{out}) conformation, in which Gln is present in the tautomerized imidic from, but no flipping of the Gln side chain was assumed (Scheme 5). The best values for the shift in the absorption maximum between the dark and signaling states has been obtained at the CC2 level: 10−15 nm to the red, if one ignores the wrong order of the local excitation and charge-transfer states in the assumed light-induced form (Table 3 of ref 8). The agreement with the experimental shift was deemed fortuitous owing to observed strong dependence of the excitation energies on point charges in the MM subsystem. A small red shift (8 cm^{-1}) in the computed C4═O vibrational frequency was obtained.

Scheme 6 illustrates the model[12] which receives more rigorous support from high-level quantum calculations described below.

Calculations and Results

Classical MD trajectory calculations initiated from the corresponding X-ray structures were executed to provide starting points for optimization of the geometry parameters in the QM/MM calculations. By performing such optimization, coordinates of both QM and MM subsystems were included in calculations. Thus obtained equilibrium geometry parameters were used to compute vibrational frequencies. Quantum chemical calculations of the excited electronic states were performed only for the molecular clusters comprising the QM subsystems at the geometries obtained by the QM/MM optimization.

Scheme 6. The Mechanism of Light Activation in AppA by Domratcheva et al[12]



All MD calculations were performed using the NAMD 2.6 program.[44,45] The initial coordinates of the protein and the flavin mononucleotide (FMN) species were generated from crystal structures PDBIDs: 2IYG and 1YRX to model the dark and light states, respectively. The general AMBER force field (GAFF) parameters[46,47] were used for FMN. The CHARMM22 force field[48] for the protein atoms and the TIP3P model parameters for all water molecules were employed. Parameterization of the imidic form of Gln63 was verified on the basis of ab initio calculations. The protein was solvated using TIP3P water in rectangular 61 × 61 × 66 Å^3 and 62 × 65 × 74 Å^3 boxes in the calculations based on structures 2IYG and 1YRX, respectively. Finally, the systems were neutralized by adding four sodium ions. Periodic boundary conditions were assumed. All long-range electrostatic interactions were computed by using the particle mesh Ewald method.[44] The constant temperature MD simulations were performed for the NVT ensemble at 300 K by using the Langevin thermostat. The simulations were carried out with a 1 fs integration step following the 1500 step energy minimization. No restrictions were imposed on the coordinates of all atoms in the trajectory calculations. The sets of relatively short trajectories (about 2 ns) preceded longer runs of 10 ns. The VMD program[49] was used for the visualization of the results.

The goal of MD simulations was to provide starting coordinates of atoms for optimization of the geometry parameters in the QM/MM calculations. This strategy to employ the MD frames at the beginning and then to proceed to expensive QM/MM optimization is often applied in modeling properties of biomolecular systems, e.g., ref 50. If a MD trajectory does not cover large conformational changes of the protein (as in this application), then selection of a particular frame is not a critical step; the resulting coordinates of QM/MM optimization, at least for the relatively small QM subsystems, are essentially the same for different starting MD frames. As explained below, for our model of the dark state of AppA, we actually obtained four slightly different sets of equilibrium geometry coordinates in QM/MM minimization depending on the selection of MD frames.

The QM/MM geometry optimizations and vibrational analysis were performed within the electronic embedded

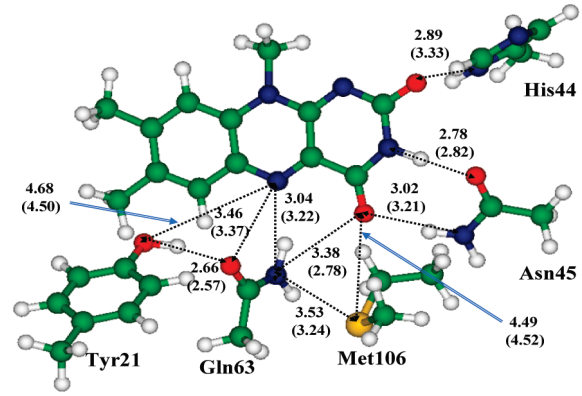


Figure 2. The molecular cluster corresponding to the (Trp_{out}/Met_{in}) conformation: a model for the dark state. The distances between heavy atoms are given in Å. Top: the computed values and bottom (in parentheses): values from the crystal structure 2IYG.[17]

cluster approximation[51] using the NWChem program package.[52] Energies and forces in the QM subsystem were calculated with B3LYP/cc-pVDZ. The MM part was described with the AMBER force field parameters. The cutoff radius of the MM zone allowed to interact with the quantum region both electrostatically and through the van der Waals interactions was 9 Å. For the presumably dark state originated from the 2IYG crystal structure, the quantum subsystem included the chromophore represented by a lumiflavin molecule and the side chains of Tyr21, His44, Asn45, Gln63, and Met106. For the presumably signaling state originated from the 1YRX crystal structure, the QM part consisted of the chromophore and the side chains of Tyr21, His44, Asn45, Gln63, and Trp104. The remaining protein and solvent water molecules were assigned to the MM part.

Vertical excitation energies were calculated by SOS-CIS(D)/cc-pVDZ using the Q-Chem program.[53] These computations were performed for molecular clusters (Figures 2 and 3) comprising the QM subsystems of the preceding QM/MM calculations at the optimized equilibrium geometries.

The Cartesian coordinates of the systems shown in Figures 2 and 3 are given in the Supporting Information to this paper. In these figures, we compare several key intermolecular distances (more precisely, several distances between the

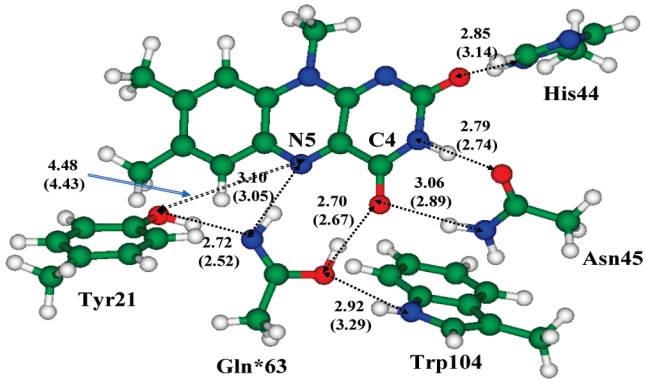


Figure 3. The molecular cluster corresponding to the (Trp_{in}/Met_{out}) conformation: a model for the light state. The distances between heavy atoms are given in Å. Top: the computed values and bottom (in parentheses): values from the crystal structure 1YRX.[6]

heavy atoms from different molecular groups) computed by B3LYP(cc-pVDZ)/MM(AMBER) with those resolved in crystal structures 2IYG[17] and 1YRX.[6] We note that the agreement between the calculated and measured distances is fairly good. Of a special importance is a good agreement of the theoretical model (Figure 3) that assumes a tautomerized form of Gln63 with the experimental data of Anderson et al.,[6] who considered a normal amidic form of this side chain. We repeat the arguments presented in ref 12 that the original assumption of Anderson et al. on the direct contact of the −C═O group from Gln with the −C4═O group from flavin (see the left side of Scheme 2) contradicts the measured distance of 2.7 A. Such a short distance apparently requires a proton between the corresponding oxygen atoms, as illustrated in Figure 3.

We present in the Supporting Information the results of vibrational analysis using the QM/MM. Here we focus on the prominent band experimentally assigned to the C4═O vibration of flavin (Figure 1). According to ref 4, the FTIR difference spectrum of the AppA BLUF domain showed a sharp band at 1707(−)/1684(+) cm^{-1}. Consistent with the FTIR data, a change in the Raman spectrum upon formation of the signaling state is approximately 20 cm^{-1} in the C4═O stretch (observed at 1706 cm^{-1}).[16] Therefore, the experimental red shift in the C4═O frequency upon passing from the dark to the light-induced state in AppA BLUF was measured as 20−23 cm^{-1}. The computed frequencies for the dark (Figure 2) and signaling (Figure 3) states, 1700 and 1675 cm^{-1}, yield a shift of 25 cm^{-1} which is in excellent agreement with the experimental value.

Analysis of MD trajectories for the presumably dark state showed that we could distinguish two conformations of the Met106 side chain with smaller (around 53°) and larger (around 114°) values of the CB−CG−S−CE dihedral angles occurring along the MD run. We also noted oscillations occurring along MD trajectories between two conformations of the His44 side chain with respect to the flavin moiety. Therefore, we took respective four MD snapshots for QM/MM optimization of molecular clusters for which the excitation energies were calculated (Figure 2 illustrates geometry parameters corresponding to the lowest energy). We observed only a slight dependence of the computed excitation energies on the parent MD frame; the respective SOS-CIS(D)/cc-pVDZ values of the S_0-S_1 excitation energies with the oscillator strengths 0.66−0.69 were 2.89, 2.92, 2.92, and 2.91 eV. The corresponding wavelengths are 429, 425, 425, and 426 nm. For the presumably light-induced state, the MD trajectories did not show noticeable changes in the configurations in the QM subsystem (Figure 3), hence, we considered the only QM/MM optimized structure for which the excitation energy 2.81 eV (with the oscillator strength 0.63) was obtained corresponding to 441 nm for the wavelength. Therefore, the computed red shift in the absorption band maxima between the dark and light-induced states is 12÷16 nm, to be compared with the 13 nm shift measured for AppA BLUF.[4]

It is important to note that positions of the CT states corresponding to the Tyr → flavin electronic redistribution are much higher in both structures (3.7−3.9 and 3.5 eV for the dark and light-induced forms, respectively), and there were no difficulties in computing the excitation energies for the S_0-S_1 transition. It is also worth noting the magnitude of the SOS-CIS(D) corrections to the CIS excitation energies: Without these corrections, the computed S_0-S_1 energy gaps are about 1.1 eV higher.

Discussion and Conclusion

It is instructive to compare the present results with those from the previous simulations, especially with our earlier results of Domratcheva et al. for AppA BLUF[12] and those of the most recent QM/MM calculations of Sadeghian et al.[8] initiated from the crystal structure of the BLUF domain from another flavoprotein BlrB (PDB code 2BYC).

In ref 12, the CASSCF-based MCQDPT2 method for simpler molecular clusters mimicking the dark and light-induced forms of AppA BLUF was employed. The model clusters shown in Figure 3 of ref 12 illustrate the same general trend: The (Trp_{out}/Met_{in}) conformation of BLUF is characterized by a shorter wavelength and represents the dark state, while the (Trp_{in}/Met_{out}) conformation exhibits a longer wavelength and represents the light state. The results of the MCQDPT2 calculations agreed well with the experimental data. The theoretical and experimental shifts were 10 and 13 nm, respectively. These calculations involved tremendous efforts including manual inspection of the configuration state functions and the orbitals of the underlying CASSCF solutions.

The CC2 calculations at the QM/MM optimized geometries by Sadeghian et al.[8] also reproduced the red shift of 9 nm in absorption maximum when passing from the dark to light-activated form within their model illustrated in Scheme 5 (ignoring the wrong ordering of local excitation and CT states in the latter form). A very small red shift (8 cm^{-1}) in the C4═O vibrational frequency was obtained as well.

Unfortunately, it is difficult to compare the structures obtained in this work with those of Sadeghian et al.[8] since only limited information on the geometries can be extracted from their paper. Considering the dark-state model (Figure 4), we note quite reasonable agreement between the results of the two different theoretical approaches. However, for the only parameter that can be directly compared to the corre-

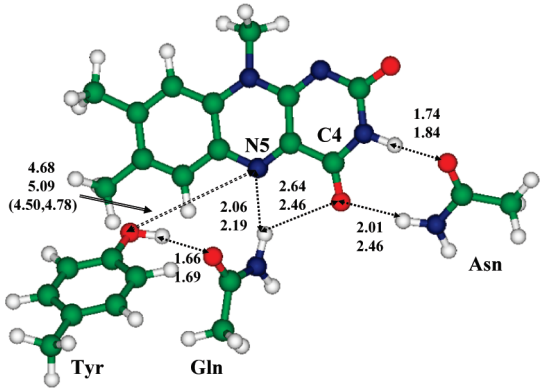


Figure 4. The comparison of selected distances in the dark-state structure obtained in this work (upper values for distances in Å) and by Sadeghian et al.[8] (lower values). For the distance between oxygen from Tyr and N5 from flavin, the bottom values in parentheses refer to the crystal structures 2IYG (AppA) and 2BYC (BlrB).

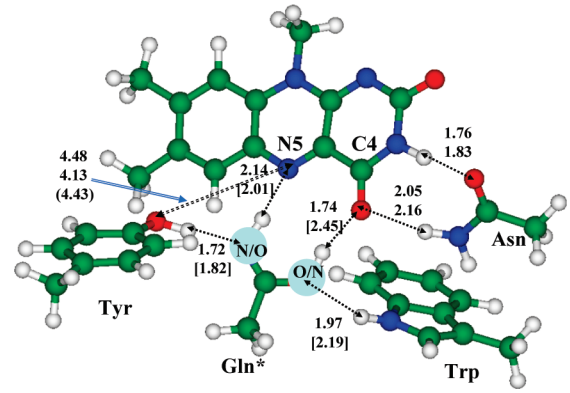


Figure 5. The comparison of selected distances in the light-induced structure obtained in this work (upper values for distances in Å) and by Sadeghian et al.[8] (lower values). For the distance between oxygen from Tyr and N5 from flavin, the bottom value in parentheses refers to the crystal structure 1YRX (AppA). We do not specify O or N atoms in the head of Gln*.

sponding values in the crystals, i.e., the distance between oxygen from Tyr and N5 from flavin, we note a large discrepancy between the computed[8] value (5.09 Å) and those measured in AppA (4.50 Å) and BlrB (4.78 Å). The agreement between our results (4.68 Å) and the crystal data is much better.

The models for the light-induced state suggested by Domratcheva et al.[12] (which is confirmed by this work) and the one suggested by Sadeghian et al.[8] agree that the tautomerized form of Gln is present in this form, however, the two models differ with respect to the orientation of the Gln functional side chain. In our model, nitrogen from Gln participates in hydrogen bonding with Tyr and (Gln*)NH with N5 of flavin, while (Gln*)OH is hydrogen bonded to O4 from flavin. In the model of Sadeghian et al.,[8] oxygen from Gln is involved in hydrogen bonding with Tyr and (Gln*)OH with N5 of flavin, while (Gln*)NH is hydrogen bonded to O4 from flavin. Similarly to the dark state (Figure 4), we note better agreement of our result with the crystal data for the distance between oxygen from Tyr and N5 from flavin; our computed distance of 4.48 Å is closer to the experimental value of 4.43 Å than the distance of 4.13 Å reported by Sadeghian et al.[8] More important is the distance between O4 from flavin to hydrogen-bound carbon from Gln* (1.74 Å in our model) or nitrogen from Gln* (2.45 Å in the model of Sadeghian et al.).[8] Since the result of light activation (production of red-shifted transient form) is believed to be a formation of a new hydrogen bond with O4 from flavin, the distance from O4 to hydrogen is a key issue. The hydrogen-heavy atom distance of 2.45 Å is apparently too long to account for a noticeable (10−15 nm) red shift in the absorption spectrum or a downshift (~20 cm^{-1}) in the vibrational spectrum. In fact, the authors reported very small computed vibrational shift of 8 cm^{-1}.[8]

A significant conclusion of ref 8 is that the results of standard TD-DFT calculations of the excited states for molecular clusters composed of the flavin chromopore and the surrounding aromatic amino acid residues are not reliable due to the contamination of the valence states by the CT ones. Interestingly, Götze and Saalfrank[19] and Obanayama et al.[7] who applied TD-DFT for such model systems did not report difficulties due to artificial CT states. Obanayama et al.[7] wrote that the obtained shifts in S_0-S_1 absorption maximum and the C4═O frequency were of the opposite sign relative to the experimental values.

In this work, we applied the modern tools of molecular modeling including molecular dynamics, QM/MM, and quantum chemistry methods to characterize the two structures of AppA BLUF corresponding to the dark and signaling states. The SOS-CIS(D) results are in excellent agreement with the experimental data. The computational cost of SOS-CIS(D) is slightly higher than that of TDDFT (which is not reliable in this case) but considerably lower than that of the CASSCF-based approaches or CC2. The errors in absolute values of the vertical excitation energy (2.89÷2.92 eV in the calculations for the molecular cluster versus 2.80 eV observed experimentally for the protein in the dark state and 2.81 versus 2.72, respectively, in the light state) are also quite small. The computed band shift of +(12÷16) nm is perfectly consistent with the experimental value of +13 nm. No preliminary work in selecting or rearranging the orbitals for configuration interaction calculations was involved in these calculations. Vibrational calculations performed at the QM/MM level for single experimentally reliable bands corresponding to the C4═O mode also show an excellent agreement with the experiment.

Comparing the present calculations with the previous work,[12] we note that despite large differences in computational protocols, the main qualitative result is the same. According to the calculations, the (Trp_{out}/Met_{in}) conformation of AppA BLUF should be associated with the dark state and the (Trp_{in}/Met_{out}) conformation with the light-induced state.

The detailed discussion of the reaction photocycle of AppA BLUF taking into consideration the results of fast spectroscopy studies and molecular modeling will be presented elsewhere. Here we concentrated mostly on quantum chemical aspects of modeling and showed that a careful selection of appropriate computational methods could provide firm validation of the mechanism of photoinduced reactions.

Acknowledgment. We thank Prof. Ilme Schlichting for generous help and valuable discussions. This work is supported by the grant from the U.S. Civilian Research and Development Foundation (Project no. RUC1-2914-MO-07) and the Russian Foundation for Basic Research (Project no. 10-03-00139). The Russian team thanks the SKIF-GRID program for providing resources at the SKIF-MSU, SKIF-SIBERIA, and UGATU computational facilities. M.K. acknowledges partial support from the grant MK-64815.2010.4. This work was conducted under the auspices of the iOpen-Shell Center for Computational Studies of Electronic Structure and Spectroscopy of Open-Shell and Electronically Excited Species http://iopenshell.usc.edu supported by the National Science Foundation through the Grant no. CRIF: CRF CHE-0625419_0624602_0625237 as well as through Grant no. CHE-0616271 grant (A.K.).

Note Added after ASAP Publication. This article was published ASAP on July 1, 2010. Initials have been added to all author names. The correct version was published on July 15, 2010.

Supporting Information Available: Cartesian coordinates of the systems shown in Figures 2 and 3. Results of vibrational analysis using the QM/MM. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Gomelsky, M.; Klug, G. *Trends Biochem. Sci.* **2002**, *27*, 497–500.

(2) van der Horst, M. A.; Hellingwerf, K. J. *Acc. Chem. Res.* **2004**, *37*, 13–20.

(3) Masuda, S.; Bauer, C. E. *Cell* **2002**, *110*, 613–623.

(4) Masuda, S.; Hasegawa, K.; Ono, T. *Biochemistry* **2005**, *44*, 1215–1224.

(5) Jung, A.; Domratcheva, T.; Tarutina, M.; Wu, Q.; Ko, W.-H.; Shoeman, R.; Gomelsky, M.; Gardner, K. H.; Schlichting, I. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 12350–12355.

(6) Anderson, S.; Dragnea, V.; Masuda, S.; Ybe, J.; Moffat, K.; Bauer, C. *Biochemistry* **2005**, *44*, 7998–8005.

(7) Obanayama, K.; Kobayashi, H.; Fukushima, K.; Sakurai, M. *Photochem. Photobiol.* **2008**, *84*, 1003–1010.

(8) Sadeghian, K.; Bocola, M.; Schütz, M. *J. Am. Chem. Soc.* **2008**, *130*, 12501–12513.

(9) Grinstead, J. S.; Avila-Perez, M.; Hellingwerf, K. J.; Boelens, R.; Kaptein, R. *J. Am. Chem. Soc.* **2006**, *128*, 15066–15067.

(10) Grinstead, J. S.; Hsu, S. T. D.; Laan, W.; Bonvin, A.; Hellingwerf, K. J.; Boelens, R.; Kaptein, R. *ChemBioChem* **2006**, *7*, 187–193.

(11) Laan, W.; Gauden, M.; Yeremenko, S.; van Grondelle, R.; Kennis, J. T. M.; Hellingwerf, K. J. *Biochemistry* **2006**, *45*, 51–60.

(12) Domratcheva, T.; Grigorenko, B. L.; Schlichting, I.; Nemukhin, A. V. *Biophys. J.* **2008**, *94*, 3872–3879.

(13) Wu, Q.; Gardner, K. H. *Biochemistry* **2009**, *48*, 2620–2629.

(14) Dragnea, V.; Waegele, M.; Balascuta, S.; Bauer, C.; Dragnea, B. *Biochemistry* **2005**, *44*, 15978–15985.

(15) Dragnea, V.; Arunkumar, A. I.; Yuan, H.; Giedroc, D. P.; Bauer, C. E. *Biochemistry* **2009**, *48*, 9969–9979.

(16) Unno, M.; Sano, R.; Masuda, S.; Ono, T.; Yamauchi, S. *J. Phys. Chem. B* **2005**, *109*, 12620–12626.

(17) Jung, A.; Reinstein, J.; Domratcheva, T.; Shoeman, R. L.; Schlichting, I. *J. Mol. Biol.* **2006**, *362*, 717–732.

(18) 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.

(19) Götze, J.; Saalfrank, P. *J. Photochem. Photobiol., B* **2009**, *94*, 87–95.

(20) Dreuw, A.; Weisman, J. L.; Head-Gordon, M. *J. Chem. Phys.* **2003**, *119*, 2943–2946.

(21) Magyar, R. J.; Tretiak, S. *Chem. Theory Comput.* **2007**, *3*, 976–987.

(22) Molina, V.; Merchán, M. *Proc. Nat. Acad. Sci. U.S.A.* **2001**, *98*, 4299–4304.

(23) Martin, M. E.; Negri, F.; Olivucci, M. *J. Am. Chem. Soc.* **2004**, *126*, 5452–5464.

(24) Coto, P. B.; Strambi, A.; Ferré, N.; Olivucci, M. *Proc. Nat.Acad. Sci. U.S.A.* **2006**, *103*, 17154–17159.

(25) Frutos, L. M.; Andruniów, T.; Santoro, F.; Ferré, N.; Olivucci, M. *Proc. Nat. Acad. Sci. U.S.A.* **2007**, *104*, 7764–7769.

(26) Strambi, A.; Coto, P. B.; Frutos, L. M.; Ferre', N.; Olivucci, M. *J. Am. Chem. Soc.* **2008**, *130*, 3382–3388.

(27) Bravaya, K.; Bochenkova, A.; Granovsky, A.; Nemukhin, A. *J. Am. Chem. Soc.* **2007**, *129*, 13035–13042.

(28) Bravaya, K. B.; Bochenkova, A. V.; Granovsky, A. A.; Savitsky, A. P.; Nemukhin, A. V. *J. Phys. Chem.* **2008**, *112*, 8804–8810.

(29) Tokmachev, A. M.; Boggio-Pasqua, M.; Bearpark, M. J.; Robb, M. A. *J. Phys. Chem. A* **2008**, *112*, 10881–10886.

(30) Filippi, C.; Zaccheddu, M.; Buda, F. *J. Chem. Theory Comput.* **2009**, *5*, 2074–2087.

(31) Valsson, O.; Filippi, C. *J. Chem. Theory Comput.* **2010**, *6*, 1275–1292.

(32) Olsen, S.; Smith, S. C. *J. Am. Chem. Soc.* **2008**, *130*, 8677–8689.

(33) Grimme, S. *J. Chem. Phys.* **2003**, *118*, 9095–9102.

(34) Rhee, Y. M.; Head-Gordon, M. *J. Phys. Chem. A* **2007**, *111*, 5314–5326.

(35) Head-Gordon, M.; Rico, R. J.; Oumi, M.; Lee, T. J. *Chem. Phys. Lett.* **1994**, *219*, 21–29.

(36) Stanton, J. F.; Bartlett, R. J. *J. Chem. Phys.* **1993**, *98*, 7029–7039.

(37) Krylov, A. I. *Annu. Rev. Phys. Chem.* **2008**, *59*, 433–462.

(38) DiStacio, R. A.; Head-Gordon, M. *Mol. Phys.* **2007**, *105*, 1073.

(39) Epifanovsky, E.; Polyakov, I.; Grigorenko, B.; Nemukhin, A.; Krylov, A. I. *J. Chem. Theory Comput.* **2009**, *5*, 1895–1906.

(40) Gauden, M.; van Stokkum, I. H.; Key, J. M.; Luhrs, D.; van Grondelle, R.; Hegemann, P.; Kennis, J. T. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 10895–10900.

(41) Gauden, M.; Grinstead, J. S.; Laan, W.; van Stokkum, I. H. M.; Avila-Perez, M.; Toh, K. C.; Boelens, R.; Kaptein, R.; van Grondelle, R.; Hellingwerf, K. J.; Kennis, J. T. M. *Biochemistry* **2007**, *46*, 7405–7415.

(42) Stelling, A. L.; Ronayne, K. L.; Nappa, J.; Tonge, P. J.; Meech, S. R. *J. Am. Chem. Soc.* **2007**, *129*, 15556–15564.

(43) Unno, M.; Masuda, S.; Ono, T. A.; Yamauchi, S. *J. Am. Chem. Soc.* **2006**, *128*, 5638–5639.

(44) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781–1802.

(45) *NAMD 2.6*; Theoretical and Computational Biophysics Group, University of Illinois at Urbana-Champaign: Urbana, IL, 2006; http://www.ks.uiuc.edu/Research/namd/ (accessed June 2010).

(46) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157–1174.

(47) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. *J. Mol. Graphics* **2006**, *25*, 247–260.

(48) *CHARMM22*; Harvard University: Cambridge, MA, 1998; http://mackerell.umaryland.edu/CHARMM_ff_params.html (accessed June 2010).

(49) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol.Graphics.* **1996**, *14*, 33–38.

(50) Sanchez-Garcia, E.; Doerr, M.; Hsiao, Y.-W.; Thiel, W. *J. Phys. Chem. B* **2009**, *113*, 16622–16631.

(51) Bakowies, D.; Thiel, W. *J. Phys. Chem.* **1996**, *100*, 10580–10594.

(52) Straatsma, T. P.; Aprà, E.; Windus, T. L.; Bylaska, E. J.; de Jong, W.; Hirata, S.; Valiev, M.; Hackler, M.; Pollack, L.; Harrison, R.; Dupuis, M.; Smith, D. M. A.; Nieplocha, J.; Tipparaju, V.; Krishnan, M.; Auer, A. A.; Brown, E.; Cisneros, G.; Fann, G.; Früchtl, H.; Garza, J.; Hirao, K.; Kendall, R.; Nichols, J.; 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.; 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 4.6; Pacific Northwest National Laboratory: Richland, WA, 2004 (accessed June, 2010).

(53) Shao, Y.; Molnar, L. F.; Jung, Y.; Kussmann, J.; Ochsenfeld, C.; Brown, S.; Gilbert, A. T. B.; Slipchenko, L. V.; Levchenko, S. V.; O'Neil, D. P.; Distasio, R. A., Jr.; 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.; Korambath, P. P.; Adamson, R. D.; Austin, B.; Baker, J.; Bird, E. F. C.; Daschel, H.; Doerksen, R. J.; Drew, A.; Dunietz, B. D.; Dutoi, A. D.; Furlani, T. R.; Gwaltney, S. R.; Heyden, A.; Hirata, S.; Hsu, C.-P.; Kedziora, G. S.; Khalliulin, R. Z.; Klunziger, P.; Lee, A. M.; Liang, W. Z.; 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 III, H. L.; Zhang, W.; Bell, A. T.; Chakraborty, A. K.; Chipman, D. M.; Keil, F. J.; Warshel, A.; Herhe, 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.

CT100179P

JCTC Journal of Chemical Theory and Computation

An Implicit Solvent Model for SCC-DFTB with Charge-Dependent Radii

Guanhua Hou, Xiao Zhu, and Qiang Cui*

Department of Chemistry and Theoretical Chemistry Institute, University of Wisconsin, Madison, 1101 University Ave., Madison, Wisconsin 53706

Received April 5, 2010

Abstract: Motivated by the need to rapidly explore the potential energy surface of chemical reactions that involve highly charged species, we have developed an implicit solvent model for approximate density functional theory, SCC-DFTB. The solvation free energy is calculated using a popular model that employs Poisson−Boltzmann for electrostatics and a surface-area term for nonpolar contributions. To balance the treatment of species with different charge distributions, we make the atomic radii that define the dielectric boundary and solute cavity depend on the solute charge distribution. Specifically, the atomic radii are assumed to be linearly dependent on the Mulliken charges and solved self-consistently together with the solute electronic structure. Benchmark calculations indicate that the model leads to solvation free energies of comparable accuracy to the SM6 model (especially for ions), which requires much more expensive DFT calculations. With analytical first derivatives and favorable computational speed, the SCC-DFTB-based solvation model can be effectively used, in conjunction with high-level QM calculations, to explore the mechanism of solution reactions. This is illustrated with a brief analysis of the hydrolysis of monomethyl monophosphate ester (MMP) and trimethyl monophosphate ester (TMP). Possible future improvements are also briefly discussed.

I. Introduction

Many chemical reactions take place in solution, so a proper description for the solvation effect is one of the most important challenges in computational chemistry. Although major progress has been made in QM/MM[1−6] and *ab initio* molecular dynamics[7] methods in which the solvent molecules are treated explicitly, the cost of such calculations is still rather high. Therefore, implicit solvent models remain an attractive choice for many studies. In the context of studying chemical reactions, the most commonly used framework for treating solvent implicitly is the dielectric continuum model,[8,9] in which the solvent is replaced by a homogeneous dielectric medium. More sophisticated treatments based on integral equations have also been developed, such as (MC)SCF-RISM,[10] although they tend to be computationally more expensive than dielectric continuum models.

Over the past few decades, many different dielectric solvent models have been developed in the quantum chemistry community, such as the self-consistent reaction field (SCRF) model,[11,12] polarized continuum model (PCM),[13−25] generalized Born (GB) model,[26−32] conductor-like screening model (COSMO),[33−38] and the Langevin dipole model.[39] For the application to chemical reactions involving large solutes, there are two practical issues. First, the computational cost of implicit solvent model calculations is still rather high, especially when used with a high-level QM method. Therefore, it is fairly common to perform gas-phase optimization for stationary points and then carry out single-point energy calculations in solution using a dielectric continuum model. This can be problematic when there is significant difference between the gas phase and solution potential energy landscape,[40] a scenario which is not uncommon when the solute is highly charged or zwitterionic. The second problem is that most implicit solvent models employ a set of fixed atomic radii to define the solvent/solute dielectric boundary, and these radii are typically preoptimized on the basis of the experimental solvation free energies of a set of small molecules[8,9,41] and are therefore limited by the diversity of

* Corresponding author e-mail: cui@chem.wisc.edu.

10.1021/ct1001818 © 2010 American Chemical Society
Published on Web 07/07/2010

the training set. The use of fixed atomic radii causes additional errors in the application to chemical reactions, as the description of transition states is rarely included during parametrization; moreover, the atom type for atoms[41] explicitly involved in the reaction is often ill-defined in a transition state. Methods have been developed in which the molecular cavity is determined on the basis of the electron isodensity surface,[42,43] although an optimal value for the electron density cutoff is not always straightforward to determine.[44]

Motivated by these considerations, we have implemented a dielectric solvent model for an approximate density functional theory, the self-consistent-charge density-functional-tight-binding (SCC-DFTB) method.[45] SCC-DFTB is an approximation to density functional theory (DFT) based on a second-order expansion of the DFT total energy around a reference electron density. With respect to computational efficiency, SCC-DFTB is comparable to widely used semiempirical methods such as AM1 and PM3, i.e., being 2−3 orders of magnitude faster than popular DFT methods. In terms of accuracy, fairly extensive benchmark calculations have indicated that it is particularly reliable for structural properties, while energetics are generally comparable to AM1 and PM3.[46−48] With recent developments of SCC-DFTB[49,50] for metal ions[51−54] and a few other elements that require d orbitals for a reliable description (e.g., phosphorus[55]), an effective implicit solvent model for SCC-DFTB will be very useful and complementary to existing models based on other semiempirical methods.[26,56,57] Our model takes advantage of the finite difference Poisson−Boltzmann approach[58,59] implemented in CHARMM[60] and has analytic first derivatives.[61] This makes it possible to perform geometry optimization, reaction path searchers, and vibrational frequency calculations (based on the numerical finite difference of first derivatives).

Our main aim is to use SCC-DFTB for quickly exploring minimum energy paths for reactions in solution and then refine selected results on the basis of higher level theories. To be able to describe transition state and stable structures on equal footing, it is desirable to determine the atomic radii in a self-consistent fashion on the basis of the electronic structure of the solute. The simple model we have adopted is to make the atomic radii depend on the Mulliken charges, which are fundamental to SCC-DFTB[45] and are solved self-consistently via an iterative procedure (see the Methods). A similar idea was explored in the context of an implicit solvent model for PM3.[62] More recently, as this work was in progress, charge-dependent radii have been developed for a DFT-based COSMO approach,[63,64] and much improved results (solvation free energies and chemical reactions) compared to fixed-radii models have been reported for small ions.

We have developed two sets of solvation radii parameters for SCC-DFTB. The first set is for the standard second-order SCC-DFTB[45] with parameters for C, H, O, and N. We recommend using this set for general applications to molecules consisting of these elements. The second set is for SCC-DFTBPR,[55] which is a specific version parametrized for the phosphate hydrolysis reaction and includes third-order on-site terms for C, H, O, and P; this set can be useful for studying phosphate hydrolysis reactions, although we caution that SCC-DFTBPR has been parametrized mainly for monoanionic phosphates and a limited set of hydrolysis reactions. Two rather large training sets for solvation free energy with an emphasis on biorelated molecules (including 103 and 57 solutes for SCC-DFTB and SCC-DFTBPR, respectively) are used to develop the solvation radii parameters. Calculations on two additional sets of test molecules show that the performance for neutral and charged species is rather well balanced, and the error is comparable to the SM6 model,[31] which is more sophisticated yet also much more expensive computationally. To illustrate the applicability of our model to chemical reactions in solution, we briefly study the hydrolysis of monomethyl monophosphate ester (MMP) and trimethyl monophosphate ester (TMP). The results from the current implicit solvent model are generally consistent with previous *ab initio* calculations in conjunction with PCM[65,66] or the Langevin dipole solvation models,[67] as well as with our explicit solvent simulations using SCC-DFTBPR/TIP3P.[55] Compared to the latter, however, the significant overstabilization of the zwitterionic intermediate is avoided, which highlights the complementary value of implicit solvent models to explicit solvent methods for studying reactions that involve highly charged species.

The paper is organized as follows: In section II, we summarize the key theoretical foundation for our implicit solvent model for SCC-DFTB; details for the parametrization and benchmark calculations are also included. In section III, we present results and discussions of the parametrization and benchmark data, including the overall performance for both the training and test sets of molecules, and results for the hydrolysis of MMP/TMP. Finally, we summarize our findings in section IV.

II. Methods

A. SCC-DFTB. Here, we briefly recall the basic elements of SCC-DFTB[45,51] that are important to the development of an implicit solvent model. The SCC-DFTB approach is based on a second-order expansion of the DFT total energy around a reference density, ρ_0,

$$E = \sum_i^{occ} \langle \Psi_i | \hat{H}^0 | \Psi_i \rangle + \frac{1}{2} \int\int \left(\frac{1}{|\vec{r} - \vec{r}'|} + \frac{\delta^2 E_{xc}}{\delta\rho\delta\rho'}|_{\rho_0} \right) \delta\rho\delta\rho' - \frac{1}{2} \int\int \frac{\rho'_0 \rho_0}{|\vec{r} - \vec{r}'|} + E_{xc}[\rho_0] - \int V_{xc}[\rho_0]\rho_0 + E_{cc} \quad (1)$$

where $\hat{H}^0 = \hat{H}[\rho_0]$ is the effective Kohn−Sham Hamiltonian evaluated at the reference density ρ_0, and the Ψ_i's are the Kohn−Sham orbitals. E_{xc} and V_{xc} are the exchange-correlation energy and potential, respectively, and E_{cc} is the core−core repulsion energy. With a minimal basis set, a monopole approximation for the second-order term and the two-center approximation to the integrals, the SCC-DFTB total energy is given in the following form:

$$E = \sum_{i\mu\nu} c_\mu^i c_\nu^i H_{\mu\nu}^0 + \frac{1}{2} \sum_{\alpha\beta} \gamma_{\alpha\beta} \Delta q_\alpha \Delta q_\beta + \frac{1}{2} \sum_{\alpha\beta} U[R_{\alpha\beta}; \rho_0^\alpha, \rho_0^\beta] \quad (2)$$

where $c^i_{\mu/\nu}$ represents orbital coefficients, $\Delta q_{\alpha/\beta}$ represents the Mulliken charges on atom α/β, and $\gamma_{\alpha\beta}$ is the approximate second-order kernel derived on the basis of two interacting spherical charges. The last pairwise summation gives the so-called repulsive potential term, which is the core−core repulsion plus double counting terms and is defined relative to infinitely separated atomic species.

As discussed in our recent work,[2,49,50] it was found that further including the third-order contribution can substantially improve calculated proton affinity; for a set of biologically relevant small molecules, significant improvements were observed even with only the on-site terms included. The corresponding expression for the SCC-DFTB total energy is

$$E = \sum_{i\mu\nu} c^i_\mu c^i_\nu H^0_{\mu\nu} + \frac{1}{2}\sum_{\alpha\beta}\gamma_{\alpha\beta}\Delta q_\alpha \Delta q_\beta + \frac{1}{2}\sum_{\alpha\beta} U[R_{\alpha\beta};\rho^\alpha_0,\rho^\beta_0] + \frac{1}{6}\sum_\alpha U^d_\alpha \Delta q^3_\alpha \quad (3)$$

where U^d_α is the derivative of the Hubbard parameter of atom α with respect to atomic charge. For the development of SCC-DFTBPR for phosphorus-containing systems,[55] we found it was useful to adopt an empirical Gaussian functional form for the Hubbard charge derivative, i.e.

$$U^d_\alpha(q) = U^d_{0\alpha} + D_0 \exp[-\Gamma_0(\Delta q_\alpha - Q_0)^2] \quad (4)$$

where the charge-independent parameter ($U^d_{0\alpha}$) is dependent on the element type, whereas the three parameters associated with the Gaussian (D_0, Γ_0, Q_0) are taken to be independent of the element type to minimize the number of parameters.

B. The Solvation Model Based on Surface Area and Poisson−Boltzmann. The implicit solvent framework that we adapt is based on the popular formulation[68] that includes a surface-area-dependent nonpolar component and an electrostatic component:

$$\Delta G_{sol} = \Delta G_{np} + \Delta G_{elec} \quad (5)$$

where

$$\Delta G_{np} = \gamma S \quad (6)$$

Here, S is the solvent accessible surface area (SASA), which is dependent on atomic radii,[69] and γ is a phenomenological surface tension coefficient.

The electrostatic solvation free energy ΔG_{elec} for a given charge distribution $\rho(\mathbf{r})$ is generally given by

$$\Delta G_{elec} = \frac{1}{2}\int\int d\mathbf{r}\, d\mathbf{r}'\, \rho(\mathbf{r})\, G(\mathbf{r},\mathbf{r}')\, \rho(\mathbf{r}') \quad (7)$$

where 1/2 reflects the linearity of the dielectric medium[70] and the reaction field Green's function $G(\mathbf{r},\mathbf{r}')$ corresponds to the reaction field potential at $\mathbf{r}$ due to a unit charge at $\mathbf{r}'$:[71]

$$\phi_{rf}(\mathbf{r}) = \int d\mathbf{r}'\, G(\mathbf{r},\mathbf{r}')\, \rho(\mathbf{r}') \quad (8)$$

For a set of point charges, $\rho(\mathbf{r}) = {}_\alpha q_\alpha \delta(\mathbf{r} - \mathbf{r}_\alpha)$, ΔG_{elec} is simplified to

$$\Delta G_{elec} = \frac{1}{2}\sum_\alpha q_\alpha \phi_{rf}(\mathbf{r}_\alpha) \quad (9)$$

The reaction-field potential $\phi_{rf}(\mathbf{r})$ is obtained by subtracting a reference electrostatic potential computed in a vacuum, $\phi_v(\mathbf{r})$, from the electrostatic potential computed in the dielectric solvent medium, $\phi_s(\mathbf{r})$. The electrostatic potentials are determined as solutions of the (linearized) Poisson−Boltzmann (PB) equation:[59,72]

$$\nabla\cdot[\varepsilon(\mathbf{r})\nabla\phi(\mathbf{r})] - \bar{\kappa}^2(\mathbf{r})\,\phi(\mathbf{r}) = -4\pi\rho(\mathbf{r}) \quad (10)$$

with the appropriate dielectric boundary ($\varepsilon(\mathbf{r})$) and charge distributions in finite difference (FD) form using iterative numerical techniques. The solution yields the electrostatic potential at every grid point, and the total electrostatic solvation free energy is given by

$$\Delta G_{elec} = \frac{1}{2}\sum_i q_i(\phi_{s,i} - \phi_{v,i}) \quad (11)$$

where q_i and ϕ_i are the charge and calculated potential at the ith gridpoint, for the cases of a vacuum (v) and a solution (s).

In SCC-DFTB, ΔG_{elec} in eq 7 is also simplified by the fact that the charge (electrons plus nuclei) density is represented by a collection of atom-centered Mulliken charges:[45,73]

$$\rho(\mathbf{r}) = \sum_\alpha \Delta q_\alpha \delta(\mathbf{r} - \mathbf{R}_\alpha) \quad (12)$$

where Δq_α is the Mulliken charge of atom α. Thus, calculating ΔG_{elec} is a straightforward extension of the classical expression

$$\begin{aligned}\Delta G_{elec} &= \frac{1}{2}\int\int d\mathbf{r}\, d\mathbf{r}'\rho(\mathbf{r})\, G(\mathbf{r},\mathbf{r}')\, \rho(\mathbf{r}') \\ &= \frac{1}{2}\int d\mathbf{r}\, \rho(\mathbf{r})\, \phi_{rf}(\mathbf{r}) \\ &= \frac{1}{2}\sum_\alpha \Delta q_\alpha\, \phi_{rf}(\mathbf{R}_\alpha)\end{aligned} \quad (13)$$

Using variational principles, the solvation contribution to the total solute energy leads to additional terms in the SCC-DFTB matrix elements during SCF iterations:

$$\frac{1}{2}S_{\mu\nu}[\phi_{rf}(\mathbf{R_C}) + \phi_{rf}(\mathbf{R_D})]\mu \in C, \nu \in D \quad (14)$$

where μ and ν run over a minimal set of localized pseudoatomic Slater orbitals located on atoms C and D, respectively, and $S_{\mu\nu}$ is the overlap integral associated with the two basis functions.

Additional analytical gradient components from the solvation are calculated on the basis of the finite difference force proposed by Im et al.[61] They used a continuous, spline-based dielectric boundary, which has been shown to give accurate and numerically stable forces for PB calculations. The total solvation force acting on atom α is given by

$$\begin{aligned}\mathbf{F}_\alpha^{\mathrm{sol}} &= -\frac{\partial \Delta G_{\mathrm{sol}}}{\partial \mathbf{R}_\alpha} \\ &= -\frac{\partial \Delta G_{\mathrm{elec}}}{\partial \mathbf{R}_\alpha} - \frac{\partial \Delta G_{\mathrm{np}}}{\partial \mathbf{R}_\alpha} \\ &= \mathbf{F}_\alpha^{\mathrm{RF}} + \mathbf{F}_\alpha^{\mathrm{DB}} + \mathbf{F}_\alpha^{\mathrm{IB}} + \mathbf{F}_\alpha^{\mathrm{NP}}\end{aligned} \tag{15}$$

This method calculated the electrostatic solvation force as a sum of individual terms:[61] the reaction field force ($\mathbf{F}_\alpha^{\mathrm{RF}}$) arising from the variation of atomic positions assuming the dielectric boundary remains constant, the dielectric boundary force ($\mathbf{F}_\alpha^{\mathrm{DB}}$) caused by the spatial variations of the dielectric function $\varepsilon(\mathbf{r})$ from the solvent to the solute interior, and the ionic boundary force ($\mathbf{F}_\alpha^{\mathrm{IB}}$) resulting from spatial variations of the modified Debye−Hückel screening factor $\bar{\kappa}(\mathbf{r})$. In SCC-DFTB/PB approach, for the atom α located at position $\mathbf{R}_\alpha$, the three terms in the limit of infinitesimal grid spacing are

$$\begin{aligned}\mathbf{F}_\alpha^{\mathrm{RF}} &= -\int_V \mathrm{d}\mathbf{r}\left[(\phi_{\mathrm{s}} - \phi_{\mathrm{v}})\frac{\partial \Delta q_\alpha}{\partial \mathbf{R}_\alpha}\right] \\ \mathbf{F}_\alpha^{\mathrm{DB}} &= -\frac{1}{8\pi}\int_V \mathrm{d}\mathbf{r}\phi_{\mathrm{s}}\nabla\cdot\left[\left(\frac{\partial \varepsilon}{\partial \mathbf{R}_\alpha} + \frac{\partial \varepsilon}{\partial \Delta q_\alpha}\frac{\partial \Delta q_\alpha}{\partial \mathbf{R}_\alpha}\right)\nabla\phi_{\mathrm{s}}\right] \\ \mathbf{F}_\alpha^{\mathrm{IB}} &= \frac{1}{8\pi}\int_V \mathrm{d}\mathbf{r}(\phi_{\mathrm{s}})^2\frac{\partial \bar{\kappa}^2}{\partial \mathbf{R}_\alpha}\end{aligned} \tag{16}$$

Calculations for the derivative of the converged Mulliken charge, dielectric function, and modified Debye−Hückel screening factor have been discussed in previous studies;[61] note that the derivatives for the Mulliken charge with respect to the orbital coefficients are not needed for the reaction field force ($\mathbf{F}_\alpha^{\mathrm{RF}}$) because the reaction field contribution to the solute energy is included in a variational fashion.[74] As preliminary tests indicate, the contribution from the second term in $\mathbf{F}_\alpha^{\mathrm{DB}}$ is rather small; therefore, we omit it to simplify calculation (i.e., to avoid solving the coupled-perturbed KS equations[74] for the derivative of the MO coefficients).

C. Charge-Dependent Radii Scheme. To establish a simple relationship between the dielectric boundary and the electronic structure of the solute, we take the atomic radius of a solute atom α to be linearly dependent on its Mulliken charge, Δq_α

$$R_\alpha = A_{i(\alpha)} + B_{i(\alpha)}\Delta q_\alpha \tag{17}$$

where $A_{i(\alpha)}$ and $B_{i(\alpha)}$ are element-type-dependent parameters that need to be determined on the basis of a training set (see below). Higher-order polynomials have also been tested, although no systematic improvement in the results is observed.

Since the atomic radii have an impact on the solvation free energy and therefore on the solute wave function and the Mulliken charges, R_α and Δq_α need to be determined self-consistently through an iterative scheme:

1. Perform a gas-phase SCC-DFTB energy calculation to obtain the initial solute wave function and Mulliken charges.

2. Substitute Mulliken charges into eq 17 to obtain the atomic radii and establish the dielectric boundary.

3. Solve the PB equation (eq 10) to obtain the reaction field, $\phi_{\mathrm{rf}}(\mathbf{R}_\alpha)$.

4. Resolve SCC-DFTB in the presence of reaction field perturbation (eq 14) to obtain a new set of Mulliken charges.

5. Check the convergence of energy (0.001 kcal/mol used for this work); if the convergence criterion is not met, return to step 2.

6. On the basis of converged atomic radii, calculate SASA,[69] the nonpolar contribution, and the total energy of the solute in solution. For most molecules tested here, it requires less than 10 iterations (typically 4−8) of atomic radii/Mulliken charges to update each geometry.

D. Parameter Optimization. The new parameters in the SCC-DFTB/PB-based solvation model are the $A_{i(\alpha)}$ and $B_{i(\alpha)}$ in eq 17, which are dependent only on the element type. Although in principle the surface tension parameter in eq 6 can also be optimized, we have not done so (see below) because, for the systems of interest, the nonpolar contribution tends to be overwelmed by the electrostatic component; the value of γ adopted is 0.005 kcal/(mol·Å^2), which is commonly used in protein simulations using implicit solvent models.[75] For optimizing $A_{i(\alpha)}$ and $B_{i(\alpha)}$, two training sets with molecules of broad chemical compositions have been constructed (see the Supporting Information), for which the experimental solvation free energies are taken from refs 26, 31, and 76. Set 1 is used for parametrizing the solvation model with the standard (second-order) SCC-DFTB method and includes 103 species that contain C, H, O, and N; the list includes alkane, alkene, alkyne, arene, alcohol, aldehyde, carboxylic acid, ketone, ester, amine, amide, and other biorelated molecules and ions. Set 2 is used for parametrizing the solvation model with SCC-DFTBPR and includes 57 species that contain C, H, O, and P; the list includes representative species from set 1 plus phosphorus-containing molecules. Both sets contain a large number of charged species (57 in set 1 and 24 in set 2), which is essential for parametrizing the charge dependence of atomic radii.

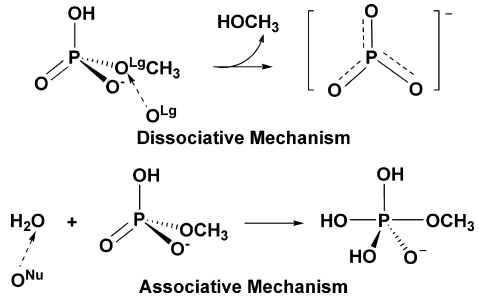
The parameters are optimized using a genetic algorithm (GA)[77] in which the "fitness" (ξ) is defined as the inverse of a weighted sum of difference between solvation free energies determined from calculation and experimentation:

$$\xi^{-1} = \frac{\sum_{i=1} w_i[\Delta G_i^{\mathrm{solv}}(\mathrm{exp}) - \Delta G_i^{\mathrm{solv}}(\mathrm{calc})]^2}{\sum_{i=1} w_i} \tag{18}$$

where i is the index of species in the training set, and the sum is over all molecules in the training set. For the weighting factors (w_i), 1.0 and 0.1 are used for the neutral molecules and ions according to the typical uncertainties in the experimental values; as analyzed by Kelly et al.,[31] the typical uncertainties in *experimental data* for neutral molecules and ions are 0.2 and 3 kcal/mol, respectively. During optimization, a micro-GA technique with a population of 10 chromosomes was allowed to operate for 500 generations with uniform crossovers; see ref 78 for detailed descriptions and recommendations for GA options.

In principle, geometry change upon solvation should be taken into consideration for a meaningful comparison to

Scheme 1



experimental results. In practice, this is very time-consuming for parameter fitting, even with the semiempirical QM method (SCC-DFTB) we employ here. Several authors discussed this point[31,63] and concluded that the change in geometry is generally small. However, in several cases, such as alcohol anions, we have observed significant structural changes upon solvation that have a substantial influence on the calculated solvation free energy. Therefore, a compromise is adopted: the gas-phase geometries are used to obtain the initial set of solvation parameters ($A_{i(\alpha)}$, $B_{i(\alpha)}$). With this set of parameters, solutes that have solvation free energy changes larger than 5 kcal/mol upon geometry optimization in solution are identified, and their geometries in solution are updated for the optimization of a new set of $A_{i(\alpha)}$, $B_{i(\alpha)}$. This cycle continues until all cases with major structural changes upon solvation have been taken into account.

Finally, it is worth noting that systematic optimization of surface tension coefficient γ (eq 6) results in negligible improvements for both neutral molecules alone and the complete training sets. A possible reason is that the nonpolar contribution to solvation free energy in the current scheme is also charge-dependent due to the use of charge-dependent atomic radii in SASA calculations. Therefore, compared with the fixed-radii scheme, the dependence of solvation free energy on γ is substantially smaller.

E. Additional Benchmark Calculations and Studies of (H)MMP/TMP Hydrolysis. To test the transferability of the optimized parameters, test sets are constructed (see the Supporting Information), which contain 32 for SCC-DFTB and 22 for SCC-DFTBPR. The calculated solvation free energies (including full geometry optimization in solution) are compared to the experimental values; similar to the training sets, the test cases contain a significant number of ionic species. As a comparison to popular and well-established solvation models, we also studied the same sets of molecules with the SM6 model of Cramer and co-workers.[31]

In addition, we have studied the mechanism[79,80] (first steps of both dissociative and associative pathways, see Scheme 1) of monomethyl monophosphate ester (MMP) hydrolysis using the SCC-DFTBPR/PB model. The potential energy surface is first explored by adiabatic mapping; the reaction coordinates include the P−$O^{Lg/Nu}$ distance (where O^{Lg} is the oxygen atom of the leaving group, methanol, and O^{Nu} is the oxygen in the nucleophilic water) and the antisymmetric stretch that describes the relevant proton transfers that involve $O^{Lg/Nu}$. The antisymmetric stretch is defined as the donor−proton distance minus the acceptor−proton distance. Each point in the 2D adiabatic map is obtained by starting the constrained optimization from several different initial structures and taking the lowest energy value. Following the adiabatic mapping calculations, the structures along the approximate reaction path are examined carefully to ensure that the change of geometry is continuous along the path; in addition, the saddle point is optimized by conjugated peak refinement (CPR).[81] Finally, frequency calculations are carried out to confirm the nature of the stationary points and to compute the vibrational entropy and zero point energies. The results are compared to previous calculations with *ab initio* QM-based implicit solvent model calculations,[65,66,82] SCC-DFTBPR/MM calculations by us,[55] and available experimental data. To correct for intrinsic errors of SCC-DFTBPR, we also explore corrections based on single-point energy calculations with B3LYP/6-311++G(d,p) at SCC-DFTBPR geometries in the gas phase; this level of theory was found to give very similar results for the reactions of interest compared to MP2 and large basis sets.[55] As discussed in the literature,[83] such a simple correction may not always improve the energetics for semiempirical methods given the errors in the geometries; however, our previous tests[55] indicated that this correction scheme appears useful for SCC-DFTBPR since the method gives fairly reliable structures, even for transition states.

Finally, we briefly compare the energetics of protonated MMP (HMMP) and trimethyl monophosphate ester (TMP) hydrolysis with OH^- as the nucleophile (see Scheme 2). This is motivated by the previous work of Warshel and Florián,[84] who discussed the roles of neutral water vs OH^- as the nucleophile in MMP hydrolysis. Since SCC-DFTBPR was developed on the basis of MMP hydrolysis with water as the nucleophile,[55] this study helps to gain initial insights into the transferability of SCC-DFTBPR and lays the groundwork for possible future developments. For better comparison to previous calculations,[67,84] we follow the same two-dimensional adiabatic mapping calculations with the bond lengths for the forming and breaking P−O bonds as the reaction coordinates. Single-point B3LYP/6-311++G(d,p) calculations in the gas phase are used as an attempt to correct for intrinsic errors of SCC-DFTBPR.

III. Results and Discussions

A. Performance for the Training and Test Sets. The trends in optimized atomic radii (see Table 1) are consistent with other implicit solvent models and chemical intuition. For example, P has the largest charge-independent radius ($A_{i(\alpha)}$), while C, O, and N have comparable values, leaving H as the smallest. The absolute values are larger than those in SM6 and also the Bondi radii.[85] Compared with the atom-type-based charge-dependent radii in CD-COSMO by Dupuis et al.,[63] comparable values are found for nitrogen and oxygen in our model and the "internal −N", "terminal oxygen", and "internal −O" in CD-COSMO. The hydrogen radius (∼1.4 Å) in our model is larger than that (polar hydrogen) in CD-COSMO (1.202 Å). In terms of the charge dependence, the typical $B_{i\alpha}$ values are around −0.10, although they are substantially larger (∼−0.2) for C in SCC-DFTB and H in

Scheme 2

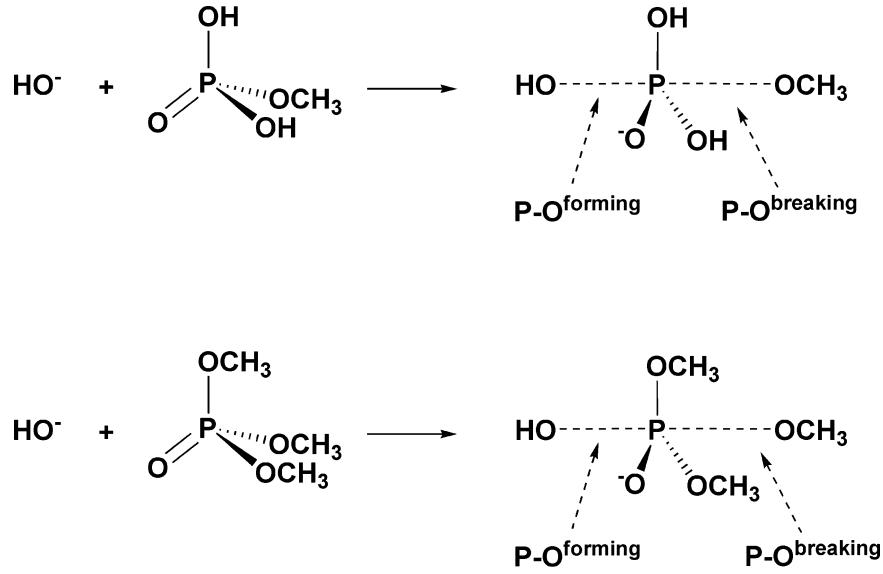


Table 1. Optimized Atomic Radii Parameters and Comparison to Other Values from the Literature[a]

	SCC-DFTB		SCC-DFTBPR			
element	$A_{i(\alpha)}$	$B_{i(\alpha)}$	$A_{i(\alpha)}$	$B_{i(\alpha)}$	SM6[31]	Bondi[85]
C	1.85	−0.24	2.07	−0.05	1.57	1.70
O	1.70	−0.11	1.87	−0.07	1.52	1.52
N	1.94	−0.01	N/A	N/A	1.61	1.55
P	N/A	N/A	2.47	−0.10	1.80	1.80
H	1.47	−0.11	1.41	−0.25	1.02	1.20

[a] $A_{i(\alpha)}$ in Å, $B_{i(\alpha)}$ in Å per charge. The values shown are fitted with solution geometry optimization (see the Methods).

SCC-DFTBPR. Even the latter are nearly half of the values in CD-COSMO, which is probably due to the use of different charges in SCC-DFTB (Mulliken) and CD-COSMO (CHELPG). It is worth emphasizing that the parameters in our model depend only on element type, rather than atom type as in CD-COSMO; therefore, CD-COSMO probably tends to be more accurate (see below for some comparison) while our scheme tends to be less problematic for studying transition states, which likely involve change in atom types.

As shown in the Supporting Information, the absolute value of solvation free energy is usually less than 10 kcal/mol for neutral molecules but larger than 60 kcal/mol for ions. Therefore, it is generally challenging to reproduce the solvation free energy of ions in a reliable fashion. Nevertheless, as shown in Table 2, the overall performance of our SCC-DFTB(PR)-based solvation model is very encouraging. For example, for ions, the mean unsigned error (MUE) for SCC-DFTB is ∼3 kcal/mol either without or with geometry optimization in solution. For SCC-DFTBPR, the error is slightly larger, with the corresponding MUE values of 5 and 4 kcal/mol. These values can be compared to results from the SM6 model,[31] which is one of the most sophasticated and well-calibrated models developed with *ab initio* DFT methods; the MUE values are 4 and 5 kcal/mol for the first (for SCC-DFTB) and second (for SCC-DFTBPR) training sets, respectively, which are even slightly larger than the values for our SCC-DFTB(PR)-based solvation model.

The level of performance deteriorates slightly for the test sets. As shown in Table 3, for example, the MUE for the ions in the first and second test sets is 3 and 5 kcal/mol, respectively, when geometry optimization in solution is carried out; without solution geometry optimization, the MUE values are 4 and 6 kcal/mol. By comparison, the SM6MUE values are 5 and 7 kcal/mol, again slightly larger than the SCC-DFTB(PR) values. These benchmark calculations indicate that the good performance of our model is fairly transferrable. This is very encouraging since the SCC-DFTB(PR)-based calculations are much faster than the DFT (MPW1PW91/6-31+G(d,p))-based SM6 calculations. Compared with CD-COSMO,[64] which is also DFT-based and involves more elaborate parametrization of charge dependence of atomic radii, it is again encouraging to see that for the three ions tested by both models, the performance is comparable. For example, for hydroxide, SCC-DFTB with or without solution geometry optimization gives an error of 2 kcal/mol, while CD-COSMO gives 3 kcal/mol. For ammonium, SCC-DFTB has an error of −3 kcal/mol, while CD-COSMO gives −2 kcal/mol. For methylamine(+1), the corresponding values are −3 and −4 kcal/mol, respectively.

We note that, relatively speaking, the performance of our model for neutral molecules is less stellar. In fact, for both the training and test cases, the SM6 model consistently outperforms the SCC-DFTB(PR) solvation model; e.g., the MUE is typically smaller by ∼1 kcal/mol with SM6 (see Tables 2 and 3). This is likely because parameters in the nonpolar component, which makes a significant (relative to ions) contribution to the total solvation free energy of neutral molecules, we have not optimized in the current model. Indeed, in the work of Xie and Liu,[76] who have implemented a GBSA model with SCC-DFTB, a root mean square error (RMSE) of 1.1 kcal/mol was obtained for 60 neutral molecules containing C, H, O, N, and S when the nonpolar parameters were optimized. On the other hand, we note that for most chemical reactions of biological relevance, the nonpolar contribution likely plays a much less significant role compared to the electrostatic component. Finally, as

Table 2. Error (in kcal/mol) Analysis of Solvation Free Energies for Training Sets 1 and 2[a]

	single point[b]			optimization[c]			SM 6[d]		
	RMSE	MUE	MSE	RMSE	MUE	MSE	RMSE	MUE	MSE
neutral	2.0	1.7	0.6	2.1	1.7	0.4	0.8	0.7	0.4
ions	4	3	2	3	3	0	4	4	2
all data	3	3	1	3	2	0	3	2	1
neutral	1.6	1.3	−0.5	2.0	1.9	−1.3	1.5	0.9	0.6
ions	4	5	5	4	4	2	4	5	5
all data	4	3	2	4	3	0	4	3	2

[a] The first three rows are for the first training set (for SCC-DFTB), and the three bottom rows are for the second training set (for SCC-DFTBPR). RMSE, root-mean-square error; MUE, mean unsigned error; MSE, mean signed error. All errors measured against experimental solvation free energies, which have typical uncertainties of 0.2 and 3 kcal/mol for neutral molecules and ions, respectively. [b] With gas-phase geometries. [c] With solution-phase geometry optimizations (see Methods). [d] Results are obtained by MPW1PW91/6-31+G(d,p).

Table 3. Error Analysis (in kcal/mol) of Solvation Free Energies for Test Sets 1 and 2[a]

	single point			optimization			SM6		
	RMSE	MUE	MSE	RMSE	MUE	MSE	RMSE	MUE	MSE
neutral	2.2	1.8	0.7	2.3	1.9	0.2	1.0	0.8	−0.2
ions	5	4	2	4	3	1	6	5	2
all data	4	3	1	3	3	0	4	2	1
neutral	1.5	1.4	−1.2	2.1	2.1	−2.0	0.9	0.7	−0.1
ions	7	6	2	7	5	0	7	7	5
all data	4	3	0	4	3	−1	5	3	2

[a] See Table 2 for format.

shown in the Supporting Information, our solvation model gives rather large errors for amine and amide molecules; for example, the error for ammonia is more than 3.2 kcal/mol with or without solution geometry optimization, which is more than 70% off the experimental value. This behavior was noted in a previous analysis of implicit solvation models,[12] and it was argued that hydrogen-bonding energies are poorly correlated with classical electrostatic interaction energies, and therefore more sophisticated treatments are needed for such short-range interactions.

B. MMP Hydrolysis Reaction with Neutral Water As Nucleophile. Experimental studies of the MMP hydrolysis reaction[86−88] determined that the reaction rate peaks at a pH of 4−5 with an activation energy of 31 kcal/mol. The reaction mechanism is traditionally regarded as dissociative, though dispute still exists.[89] Here, as a benchmark calculation for the new solvation model, we investigate the first steps of both dissociative and associative pathways (see Scheme 1) and compare the results with previous theoretical studies.[55,65,67]

For the dissociative pathway, the adiabatic map in solution with our new solvation model (Figure 1a) is qualitatively consistent with previous PMF results obtained using explicit solvent SCC-DFTBPR/MM simulations.[55] The transition state region involves largely an intramolecular proton transfer from the protonated oxygen in MMP to the oxygen in the leaving group (O^{Lg}), and the $P-O^{Lg}$ bond is only slightly stretched compared to MMP. As discussed in ref 55, the $P-O^{Lg}$ bond in the transition state decreases significantly from the gas phase (∼2.1 Å) to solution (∼1.7−1.8 Å in SCC-DFTBPR/MM PMF simulations); thus our model has captured this solvation effect adequately. Following the proton transfer, a zwitterionic intermediate is formed, which is again in qualitative agreement with both SCC-DFTBPR/MM PMF calculations[55] and the previous DFT-PCM study.[65]

More quantitatively, the fully optimized structures for MMP, the transition state (dis_ts), and the zwitterionic intermediate (dis_zt) at the SCC-DFTBPR level are in decent agreement with previous calculations; the optimized structure does not depend sensitively on the grid size in the PB calculations (for a comparison of 0.2 vs 0.4 Å grid sizes, see Figure 2, which also contains an illustration for the imaginary mode in the optimized transition state, dis_ts, with a frequency of $1742i$ cm^{-1}). Compared to the work of Vigroux et al.,[65] in which the structures were optimized at the level of B3LYP-PCM with a double-ζ-quality basis set plus diffuse and polarization functions and a pseudopotential for non-hydrogen atoms, the only major difference is that their optimized $P-O^{Lg}$ distances in dis_ts and dis_zt are longer by ∼0.1 Å and 0.25 Å, respectively. The study of Florián and Warshel[67] did not examine the zwitterionic intermediate, and the $P-O^{Lg}$ distance in their transition state is substantially longer than both values from this work and from ref 66. This is likely because the geometries of Florian et al.[67] were mainly optimized in the gas phase, and the transition state in solution was only approximately located by single-point Langevin dipole calculations along the minimum energy path from gas-phase calculations.

For the energetics, the free energy barrier estimated with the current SCC-DFTBPR-based solvation model is 34.8 kcal/mol; including a single-point B3LYP/6-311++G(d,p) gas-phase correction lowers the barrier to be 31.3 kcal/mol. As shown in Table 4, these values are consistent with previous calculations[65,67] and experimental studies,[90] which range from 30.7 to 34 kcal/mol. For the zwitterionic intermediate, which was first discussed in the work of Bianciotto et al.,[65,66] the current solvation model with SCC-DFTBPR predicts a free energy of 13.7 kcal/mol above the MMP reactant; with the B3LYP correction, the value becomes 21.1 kcal/mol. The large magnitude of the gas-phase

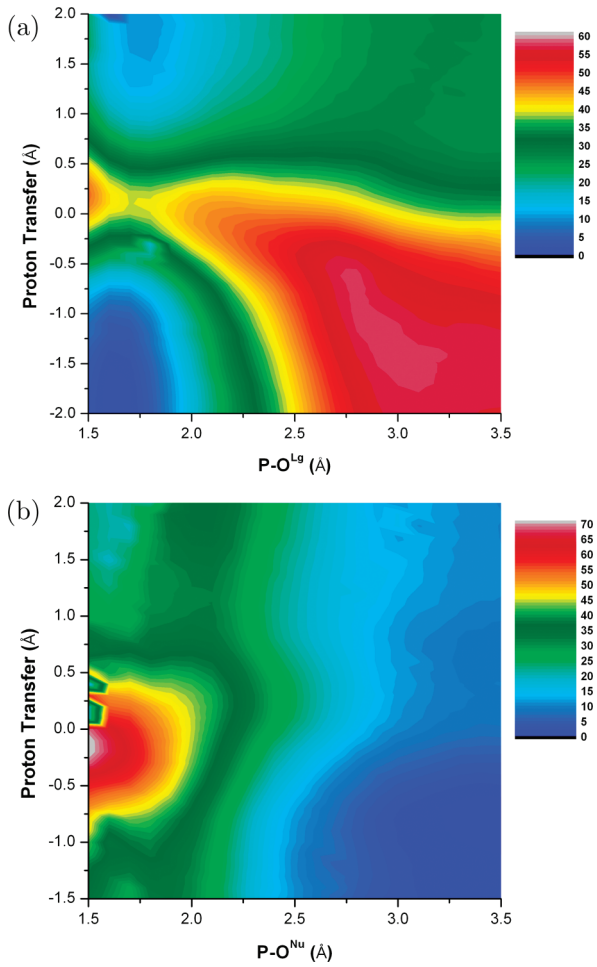


Figure 1. Adiabatic mapping results (energies in kcal/mol) for the first step of (a) the dissociative and (b) the associative pathway for the hydrolysis of monomethyl monophosphate ester (MMP). The O^{Lg} stands for the oxygen in the leaving group (see Scheme 1), which is methanol in this case; O^{Nu} stands for the oxygen in water (see Scheme 1). In a, the proton transfer coordinate is the antisymmetric stretch that describes the intramolecular proton transfer between the protonated oxygen in MMP and O^{Lg}; in b, the proton transfer coordinate is the antisymmetric stretch that describes the proton transfer between the nucleophilic water and the basic oxygen in MMP.

correction was discussed in our previous study,[55] which emphasized that the SCC-DFTBPR model was developed without any information concerning the zwitterionic region of the potential energy surface. The B3LYP corrected free energy value is in close agreement with the DFT-PCM study of Bianciotto et al.,[65] who predicted a value of 21.2 kcal/mol. Most importantly, our solvation model does not suffer from the unphysically large stabilization found in explicit solvent SCC-DFTBPR/MM simulations, which predicted that the zwitterionic intermediate is *lower* than the reactant (MMP) by ~3 kcal/mol. As discussed in ref 55, such significant overstabilization of the zwitterionic intermediate highlighted the need to improve QM/MM interactions beyond the typical form with parameters that do not reflect the electronic structure of the QM region.[91] The success of the current solvation model, on the other hand, illustrates that the charge dependence of QM/MM interactions can be effectively treated by adopting charge dependent radii in implicit solvent calculations when studying solution reactions that involve significant charge redistribution.

For the associative pathway, the adiabatic map (Figure 1b) is qualitatively similar to the PMF from explicit solvent SCC-DFTBPR/MM simulations.[55] For example, the potential energy surface is rather flat in regions with long $P-O^{Nu}$ distances but positive proton transfer coordinates, which suggests that proton transfer from the nucleophilic water to MMP can occur prior to the nucleophilic attack. Indeed, we obtained a local minimum with geometry optimization that corresponds to a molecular complex between OH^- and protonated MMP (HMMP) on the potential energy surface. Compared to the reaction complex between water and MMP (asc_pre), this complex (asc_hydro) is substantially higher in energy by ~15 kcal/mol; including the B3LYP/6-311++G(d,p) gas-phase correction further increases the value to ~26.2 − 3.6 = 22.6 kcal/mol (see Table 5). Once again, the large magnitude of the correction reflects deficiency in the current SCC-DFTBPR approach for balancing the proton affinity of phosphate and nonphosphate species, which remains an interesting challenge for future improvement.[55]

Both the adiabatic mapping and saddle point optimization point to an associative transition state in which the $P-O^{Nu}$ distance is ~2 Å and the water proton is already transferred to the phosphate oxygen (see Figure 3 for the structure of the transition state, asc_ts). Compared to the structure optimized by Florián and Warshel[67] with the Langevin dipole model, the key difference is that the proton transfer is halfway in their structure, with an $O^{Nu}-H$ distance of 1.44 Å, compared to the value of 2.18 Å in our case. Since our structure is consistent with the previous PMF results based on SCC-DFTBPR/MM simulations, we suspect that the difference is again due to the limited solution geometry optimization in the work of Florián and Warshel[67] (see the discussions above for the dis_ts). The agreement in the optimized structures for the penta-valent intermediate, asc_int, from the two sets of studies is much better, as expected (see Figure 3).

As for the energetics for the associated pathway, the SCC-DFTBPR-based solvation model gives a free energy barrier of 33.1 kcal/mol, which increases slightly to 37.5 kcal/mol when gas-phase B3LYP correction is included. These values, especially the one with B3LYP correction, are close to previous computational studies (see Table 5) but somewhat higher compared to the experimental value of 30.7 kcal/mol.[90] The pentavalent species, asc_int, is also less stable by a few kcal/mol compared to the study of Florián and Warshel.[67] We note that all calculations found that the barrier for the associative pathway is higher than that in the dissociative pathway, although the difference is fairly small (~1−2 kcal/mol) with either SCC-DFTBPR/MM or the Langevin dipole model, while the SCC-DFTBPR-based solvation model gives the largest difference (~6 kcal/mol) when B3LYP correction is included. Before more systematic analysis into the quantitative nature of B3LYP correction is performed, it remains premature to conclude that MMP hydrolysis strongly prefers a dissociative pathway.

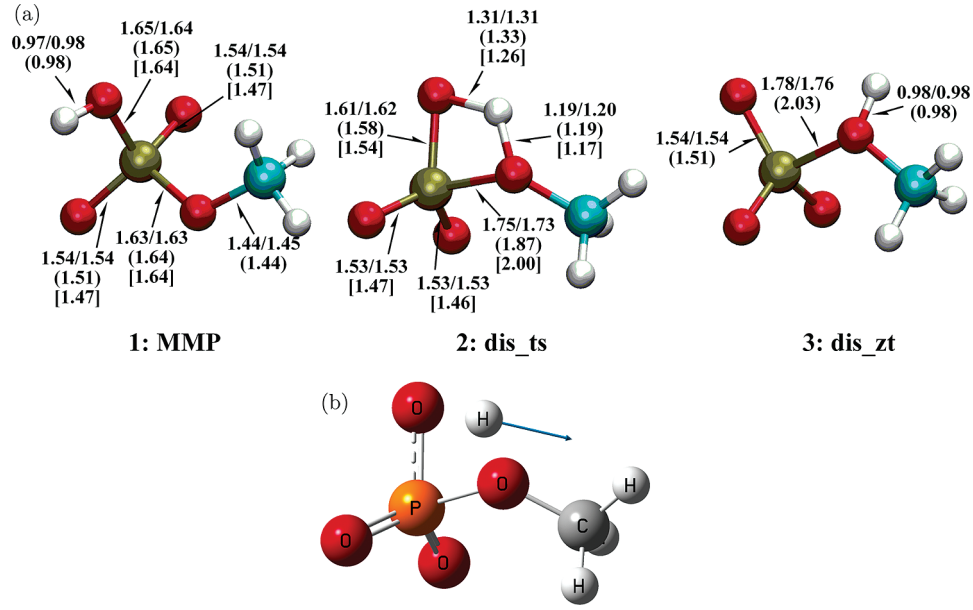


Figure 2. Geometries of the reactant, transition state, and the zwitterionic intermediate for the first step of the dissociative pathway for the hydrolysis of monomethyl monophosphate ester (MMP). (a) Values (in Å) without parentheses are from the current SCC-DFTBPR-based solvation model calculations with a grid size of 0.2/0.4 Å. Values with parentheses are from ref 65, which were obtained with B3LYP-PCM and a double-ζ-quality basis set plus diffuse and polarization functions. Values with brackets are from ref 68, which were obtained with HF/6-31G(d) in the gas phase with approximate adjustments for solvation using the Langevin dipole model. (b) An illustration of the imaginary vibrational mode in dis_ts.

Table 4. Energetics for the First Step of the Dissociative Pathway of MMP Hydrolysis from Current[a] and Previous Studies[b]

species	ΔE^c	$T\Delta S^c$	ΔZPE^c	ΔG^c	ref 55	ref 65	ref 67	exptl.[90]
MMP	*−11774.6*	*24.5*	*39.8*	*−11759.2*				
dis_ts	39.0/35.5	1.4	−2.8	34.8/31.3	32	33.5	34	30.7
dis_zt	12.6/20.0	−0.1	1.0	13.7/21.1	−3	21.2		

[a] For MMP, the total energies are given (in italics); for other species, energetics relative to MMP are given in kcal/mol. The entropic contribution ($T\Delta S$, T = 373 K in all tables, including for the experimental rate constants) and zero-point energy correction (ZPE) are calculated with the SCC-DFTBPR-based solvation model and harmonic-oscillator-rigid-rotor approximation. [b] Reference 55 employs explicit solvent SCC-DFTBPR/MM PMF simulations. Reference 65 used B3LYP-PCM and a double-ζ-quality basis set plus diffuse and polarization functions, and the pseudopotential for non-hydrogen atoms. In ref 67, geometries were obtained with HF/6-31G(d) in the gas phase with approximate adjustments for solvation using the Langevin dipole model; single-point calculations are performed at the MP2/6-31+G(d,p) level with the Langevin dipole for solvation. [c] Numbers before the slash are SCC-DFTBPR results; numbers after the slash are results after single-point gas-phase correction at the level of B3LYP/6-311++G(d,p).

Table 5. Energetics for the First Step of the Associative Pathway of MMP Hydrolysis[a]

species	ΔE	$T\Delta S$	ΔZPE	ΔG	ref 55	ref 67	exptl.[90]
MMP + H_2O	*−14345.9*	*38.3*	*53.1*	*−14331.2*			
asc_pre	−8.8/−7.0	−9.3	1.3	1.8/3.6			
asc_hydro	6.8/16.5	−9.1	0.6	16.5/26.2			
asc_ts	22.6/27.0	−9.6	0.8	33.1/37.5	34	35	30.7
asc_int	20.6/23.0	−10.4	1.4	32.5/34.9		29	

[a] Same format as in Table 4; the reference is infinitely separated MMP and H_2O.

C. HMMP and TMP Hydrolysis with OH^- as Nucleophile. A long-standing mechanistic postulate for MMP hydrolysis is that it is possible to exclude the nucleophilic attack of OH^- on the neutral phosphate. The argument was based on the high activation energy measured for the OH^- attack of trimethyl monophosphate (TMP) at high pH, which is around 25 kcal/mol (at 373K),[92] and the underlying assumption was that HMMP and TMP hydrolysis reactions have similar activation barriers. However, as pointed out by Florián and Warshel,[84] this analogy was not necessarily valid, and their calculations based on MP2 and the Langevin dipole solvation model found that the barriers for OH^- attack of HMMP and TMP differ by more than 10 kcal/mol. Moreover, the barrier of ∼12 kcal/mol found for HMMP was sufficiently low to make the OH^- attack pathway a competing mechanism of MMP hydrolysis. As an interesting benchmark of our solvation and the transferability of SCC-DFTBPR, we compare the barriers for the hydrolysis of HMMP and TMP with OH^- as the nucleophile (see Scheme 2).

As shown in Figure 4, the overall energy landscapes are quite similar for HMMP and TMP, both undergoing an associative mechanism with the new $P-O^{forming}$ bond largely formed before the $P-O^{breaking}$ broke. The transition state from the adiabatic mapping for HMMP is very consistent with the optimized saddle point asc_ts, which clearly is more

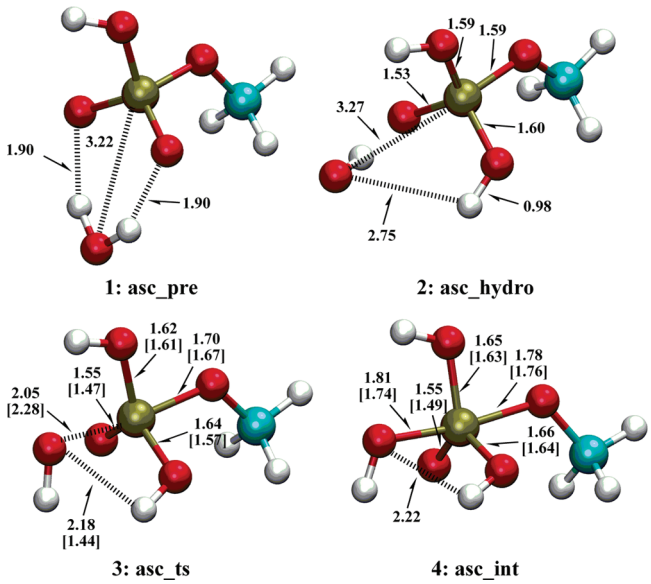


Figure 3. Similar to Figure 2, but for structures along the first step of the associative pathway for MMP hydrolysis.

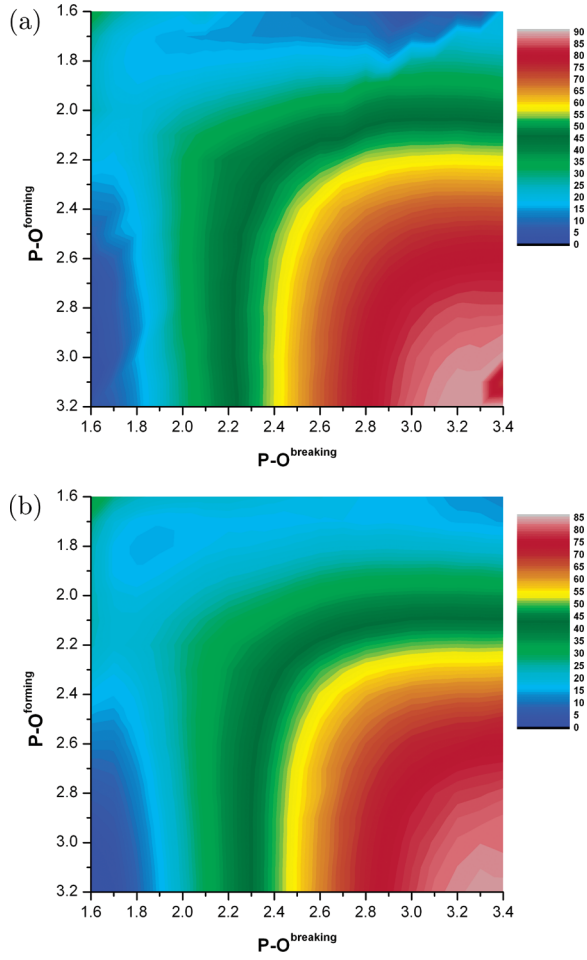


Figure 4. Adiabatic mapping results (energies in kcal/mol) for the hydrolysis of (a) hydrogen methyl monophosphate ester (HMMP) and (b) trimethyl monophosphate ester (TMP) by hydroxide. See Table 6 for the summary of the barrier heights, in which the reference is infinitely separated reactant molecules.

appropriately classified as the transition state for the OH^- attack of HMMP. According to Table 6, the corresponding energy barriers are 24.0 and 30.9 kcal/mol, with the TMP case higher by ~7 kcal/mol. Including single-point B3LYP/6-311++G(d,p) gas-phase correction further increases the gap to ~13 kcal/mol, which agrees very well with the result of Warshel and Florián.[84] This is a satisfying observation since SCC-DFTBPR was mainly parametrized on the basis of MMP and dimethyl monophosphate ester (DMP) hydrolysis; as speculated in our original work,[55] however, the parameters are likely transferrable to other phosphates that follow similar reaction mechanisms because the number of parameters is fairly small. On the absolute scale, it appears that our estimates (for both HMMP and TMP) are systematically higher, by ~4 kcal/mol, than the results of Florián and Warshel[84] and the experimental barrier for TMP.[92]

Table 6. Relative Free Energies of Key Species for the Hydrolysis of MMP and TMP along Associative Pathway with Hydroxide as the Nucleophile[a]

species	ΔE	$T\Delta S$	ΔZPE	ΔG	ref 84	Exptl.[90,92]
asc_ts	15.6/7.0	−6.9	1.4	24.0/15.4	11.7	
tmp_ts	21.9/19.4	−7.9	1.2	30.9/28.5	24.7	24.6

[a] Same format as in Table 4; the reference is infinitely separated HMMP/TMP and hydroxide.

IV. Conclusion

We report the development of an implicit solvent model for SCC-DFTB(PR) in which the solvation free energy is computed on the basis of Poisson−Boltzmann for electrostatics and a surface area term for nonpolar contributions. The unique aspect of our model is that the atomic radii that define the dielectric boundary of the solute are dependent on the solute charge distribution and are determined in a self-consistent fashion with the electronic structure of the solute. This self-consistency makes it possible to balance the solvation treatment of species with different charge distributions, such as neutral vs ionic species and structures along a chemical reaction pathway. Indeed, benchmark calculations have shown that, even for ions, our model leads to results of comparable accuracy to the much more sophisticated SM6 model; this is very encouraging since SCC-DFTB(PR) calculations are at least hundreds of times faster than the DFT calculations required in the SM6 model. We expect that the accuracy of the calculations can be further improved if more sophisticated charge schemes[9,93] are used to replace the simple Mulliken analysis.

Since our implementation has analytic first derivatives, the solvation model can be readily used to explore potential energy surfaces for solution reactions, which is the major motivation for the current work. This is demonstrated with a brief study of dissociative and associative pathways of MMP hydrolysis, as well as the hydrolysis of protonated MMP and TMP with OH^- as the nucleophile. The results (geometries and energetics) are largely in good agreement with previous computational studies using QM/MM or *ab initio*/DFT in conjunction with dielectric continuum models, as well as with available experiments. In particular, the solvation model avoids the overstabilization of the zwitterionic species along the dissociative pathway as found in explicit solvent SCC-DFTBPR/MM simulations.[55] This

highlights the complementary nature of implicit solvent model to explicit solvent approaches for studying solution reactions that involve significant charge reorganizations.

Due largely to the computational efficiency of SCC-DFTB(PR), we anticipate that the current solvation model can be effectively used in semiquantitative exploration of mechanisms for solution reactions, such as ruling out certain reaction pathways and obtaining approximate structures of key transition states and intermediates, which can be further refined with higher-level calculations. As further developments occur, it would be interesting to extend the formulation of charge-dependent radii to more approximate solvation models such as generalized Born,[75] which can be computationally more efficient than Poisson−Boltzmann; this is particularly true in molecular dynamics simulations, which can be effective for estimating the entropic contribution to reaction energetics in the framework of quasiharmonic analysis. Along this line, as extensively discussed in the literature, the first solvation shell of the solute can be treated explicitly, either at the same level of QM theory[9,12] or with a molecular mechanics model.[94,95] Since SCC-DFTB(PR) is fast, making such an extension of the molecular model for a better treatment of solvation is likely more cost-effective than with *ab initio*/DFT methods. Finally, to further improve the computational efficiency and numerical robustness of the PB aspect of the calculations, switching to a boundary element framework[96] is also of considerable interest.

Acknowledgment. This research was supported by the National Institutes of Health (R01-GM071428 and R01-GM084028) and, in part, by the National Science Foundation (CRC-CHE-0404704). Computational resources from the National Center for Supercomputing Applications at the University of Illinois and the Centre for High Throughput Computing (CHTC) at UW−Madison are greatly appreciated.

Supporting Information Available: The details of calculated and experimental solvation free energies for training and testing sets are included. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Gao, J. L.; Ma, S. H.; Major, D. T.; Nam, K.; Pu, J. Z.; Truhlar, D. G. *Chem. Rev.* **2006**, *106*, 3188–3209.

(2) Riccardi, D.; Schaefer, P.; Yang, Y.; Yu, H.; Ghosh, H.; Prat-Resina, X.; König, P.; Li, G.; Xu, D.; Guo, H.; Elstner, M.; Cui, Q. *J. Phys. Chem. B* **2006**, *110*, 6458–6469.

(3) Zhang, Y. K. *Theor. Chem. Acc.* **2006**, *116*, 43–50.

(4) Kamerlin, S. C. L.; Haranczyk, M.; Warshel, A. *J. Phys. Chem. B* **2009**, *113*, 1253–1272.

(5) Hu, H.; Yang, W. T. *Annu. Rev. Phys. Chem.* **2008**, *59*, 573–601.

(6) Senn, H. M.; Thiel, W. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198–1229.

(7) Marx, D.; Hutter, J. *Ab initio molecular dynamics: Basic theory and advanced methods*; Cambridge University Press: Cambridge, U.K., 2009; pp 9−75.

(8) Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, *99*, 2161–2200.

(9) Cramer, C. J.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 760–768.

(10) Sato, H.; Hirata, F.; Kato, S. *J. Chem. Phys.* **1996**, *105*, 1546–1551.

(11) Tannor, D. J.; Marten, B.; Murphy, R.; Friesner, R. A.; Sitkoff, D.; Nicholls, A.; RingaldaI, M.; Goddard, W. A.; Honig, B. *J. Am. Chem. Soc.* **1994**, *116*, 11875–11882.

(12) Marten, B.; Kim, K.; Cortis, C.; Friesner, R. A. *J. Phys. Chem.* **1996**, *100*, 11775–11788.

(13) Miertus, S.; Tomasi, *J. Chem. Phys.* **1982**, *65*, 239–245.

(14) Cossi, M.; Barone, V.; Cammi, R.; Tomasi, J. *Chem. Phys. Lett.* **1996**, *255*, 327–335.

(15) Barone, V.; Cossi, M.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3210–3221.

(16) Cancès, E.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3032–3041.

(17) Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *106*, 5151–5158.

(18) Amovilli, C.; Mennucci, B. *J. Phys. Chem. B* **1997**, *101*, 1051–1057.

(19) Cossi, M.; Barone, V.; Mennucci, B.; Tomasi, J. *Chem. Phys. Lett.* **1998**, *286*, 253–260.

(20) Barone, V.; Cossi, M.; Tomasi, J. *J. Comput. Chem.* **1998**, *19*, 404–417.

(21) Li, H.; Jensen, J. H. *J. Comput. Chem.* **2004**, *25*, 1449–1462.

(22) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. Chem. Phys.* **2001**, *114*, 5691–5701.

(23) Cossi, M.; Scalmani, G.; Rega, N.; Barone, V. *J. Chem. Phys.* **2002**, *117*, 43–54.

(24) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. Chem. Comput.* **2003**, *24*, 669–681.

(25) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2009**, *113*, 6378–6396.

(26) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem.* **1996**, *100*, 19824–19839.

(27) Qiu, D.; Shenkin, P. S.; Hollinger, F. P.; Still, W. C. *J. Phys. Chem. A* **1997**, *101*, 3005–3014.

(28) Ghosh, A.; Rapp, C. S.; Friesner, R. A. *J. Phys. Chem. B* **1998**, *102*, 10983–10990.

(29) Lee, M. S.; Salsbury, F. R.; Brooks, C. L. *J. Chem. Phys.* **2002**, *116*, 10606–10614.

(30) Im, W. P.; Lee, M. S.; Brooks, C. L. *J. Comput. Chem.* **2003**, *24*, 1691–1702.

(31) Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Theory Comput.* **2005**, *1*, 1133–1152.

(32) Marenich, A. V.; Olson, R. M.; Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Theory Comput.* **2007**, *3*, 2011–2033.

(33) Klamt, A.; Schuurmann, G. *J. Chem. Soc., Perkin Trans.* **1993**, *2*, 799–805.

(34) Klamt, A. *J. Phys. Chem.* **1995**, *99*, 2224–2235.

(35) klamt, A.; Jonas, V.; Burger, T.; Lohrenz, J. C. W. *J. Phys. Chem. A* **1998**, *102*, 5074–5085.

(36) Barone, V.; Cossi, M. *J. Phys. Chem. A* **1998**, *102*, 1995–2001.

(37) York, D. M.; Karplus, M. *J. Phys. Chem. A* **1999**, *103*, 11060–11079.

(38) Dolney, D. M.; Hawkins, G. D.; Winget, P.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. *J. Comput. Chem.* **2000**, *21*, 340–366.

(39) Florián, J.; Warshel, A. *J. Phys. Chem.* **1997**, *101*, 5583–5595.

(40) Wales, D. *Energy Landscapes*; Cambridge University Press: Cambridge, U.K., 2004.

(41) Barone, V.; Cossi, M.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3210–3221.

(42) Foresman, J. B.; Keith, T. A.; Wiberg, K. B.; Snoonian, J.; Frisch, M. J. *J. Phys. Chem.* **1996**, *100*, 16098–16104.

(43) Vilkas, M. J.; Zhan, C. G. *J. Chem. Phys.* **2008**, *129*, 194109.

(44) Zhan, C. G.; Chipman, D. M. *J. Chem. Phys.* **1998**, *109*, 10543–10558.

(45) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. *Phys. Rev. B* **1998**, *58*, 7260–7268.

(46) Kruger, T.; Elstner, M.; Schiffels, P.; Frauenheim, T. *J. Chem. Phys.* **2005**, *122*, 114110.

(47) Sattelmeyer, K. W.; Tirado-Rives, J.; Jorgensen, W. *J. Phys. Chem. A* **2006**, *110*, 13551–13559.

(48) Otte, N.; Scholten, M.; Thiel, W. *J. Phys. Chem. A* **2007**, *111*, 5751–5755.

(49) Elstner, M. *J. Phys. Chem. A* **2007**, *111*, 5614–5621.

(50) Yang, Y.; Yu, H.; York, D.; Cui, Q.; Elstner, M. *J. Phys. Chem. B* **2007**, *111*, 10861–10873.

(51) Elstner, M.; Cui, Q.; Munih, P.; Kaxiras, E.; Frauenheim, T.; Karplus, M. *J. Comput. Chem.* **2003**, *24*, 565–581.

(52) Cai, Z.; Lopez, P.; Reimers, J. R.; Cui, Q.; Elstner, M. *J. Phys. Chem. A* **2007**, *111*, 5743–5750.

(53) Zheng, G. S.; Witek, H. A.; Bobadova-Parvanova, P.; Irle, S.; Musaev, D. G.; Prabhakar, R.; Morokuma, K. *J. Chem. Theo. Comp.* **2007**, *3*, 1349–1367.

(54) Moreira, N. H.; Dolgonos, G.; Aradi, B.; da Roasa, A. L.; Frauenheim, T. *J. Chem. Theor. Comput.* **2009**, *5*, 605–614.

(55) Yang, Y.; Yu, H.; York, D.; Elstner, M.; Cui, Q. *J. Chem. Theory Comput.* **2008**, *4*, 2067–2084.

(56) York, D. M.; Lee, T. S.; Yang, W. T. *Chem. Phys. Lett.* **1996**, *263*, 297–304.

(57) Gogonea, V.; Merz, K. M. *J. Phys. Chem. A* **1999**, *103*, 5171–5188.

(58) Davis, M. E.; McCammon, J. A. *Chem. Rev.* **1990**, *90*, 509.

(59) Honig, B.; Nicholls, A. *Science* **1995**, *268*, 1144–1149.

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

(61) Im, W.; Beglov, D.; Roux, B. *Comput. Phys. Commun.* **1998**, *111*, 59–75.

(62) Aguilar, M. A.; del Valle, F. J. O. *Chem. Phys.* **1989**, *129*, 439–450.

(63) Ginovska, B.; Camaioni, D. M.; Dupuis, M.; Schwerdtfeger, C. A.; Gil, Q. *J. Phys. Chem. A* **2008**, *112*, 10604–10613.

(64) Ginovska, B.; Camaioni, D. M.; Dupuis, M. *J. Chem. Phys.* **2008**, *129*, 014506.

(65) Bianciotto, M.; Barthelat, J. C.; Vigroux, A. *J. Am. Chem. Soc.* **2002**, *124*, 7573–7587.

(66) Bianciotto, M.; Barthelat, J. C.; Vigroux, A. *J. Phys. Chem. A* **2002**, *106*, 6521–6526.

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

(68) Roux, B.; Simonson, T. *Bio. Chem.* **1999**, *78*, 1–20.

(69) Lee, M. S.; Feig, M.; Salsbury, F. R.; Brooks, C. L. *J. Comput. Chem.* **2003**, *24*, 1348–1356.

(70) Jackson, J. D. *Classical Electrodynamics*, 3rd ed.; John Wiley & Sons: New York, 2001; pp 145−174.

(71) Im, W.; Berneche, S.; Roux, B. *J. Chem. Phys.* **2001**, *114*, 2924–2937.

(72) McQuarrie, D. A. *Statistical Mechanics*; Harper & Row: New York, 1976; pp 326−357.

(73) Cui, Q.; Elstner, M.; Kaxiras, E.; Frauenheim, T.; Karplus, M. *J. Phys. Chem. B* **2001**, *105*, 569–585.

(74) Yamaguchi, Y.; Goddard, J. D.; Osamura, Y.; Schaefer, H. *A new dimension to quantum chemistry: Analytic derivative methods in Ab initio molecular electronic structure theory*; Oxford University Press: Oxford, U.K., 1994; pp 448−471.

(75) Feig, M.; Brooks, C. L. I. *Curr. Opin. Struct. Biol.* **2004**, *14*, 217–224.

(76) Xie, L.; Liu, H. *J. Comput. Chem.* **2002**, *23*, 1404–1415.

(77) Goldberg, D. E. *Genetic algorithms in search, optimization, and machine learning*; Addison-Wesley: Reading, MA, 1989; pp 1−25.

(78) Carroll, D. L. FORTRAN Genetic algorithm (GA) driver. http://cuaerospace.com/carroll/ga.html (accessed Jun 2010).

(79) Ba-Saif, S. A.; Davis, A. M.; Williams, A. *J. Org. Chem.* **1989**, *54*, 5483–5486.

(80) Barnes, J. A.; Wilkie, J.; Williams, I. H. *J. Chem. Soc., Faraday Trans.* **1994**, *90*, 1709–1714.

(81) S., F.; Karplus, M. *Chem. Phys. Lett.* **1992**, *194*, 511–527.

(82) Kamerlin, S. C. L.; Haranczyk, M.; Warshel, A. *ChemPhysChem* **2009**, *10*, 1125–1134.

(83) Cui, Q.; Karplus, M. *J. Phys. Chem B* **2002**, *106*, 1768–1798.

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

(85) Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441–451.

(86) Barnard, P. W. C.; Bunton, C. A.; Llewellyn, D. R.; Oldham, K. *Chem. Ind. (London)* **1955**, *760*, 2420–2423.

(87) Butcher, W. W.; Wesheimer, F. H. *J. Am. Chem. Soc.* **1955**, *77*, 2420–-.

(88) Bunton, C. A.; Llewellyn, D. R.; Oldham, K. G.; Vernon, C. A. *J. Chem. Soc* **1958**, 3574-.

(89) Åqvist, J.; Kolmodin, K.; Florián, J.; Warshel, A. *Chem. Biol.* **1999**, *6*, R71–R80.

(90) Bunton, C. A.; Llewellyn, D. R.; Oldham, K. G.; Vernon, C. A. *J. Chem. Soc.* **1958**, 3574–3587.

(91) Giese, T. J.; York, D. M. *J. Chem. Phys.* **2007**, *127*, 194101.

(92) Barnard, P. W. C.; Bunton, C. A.; Llewellyn, D. R.; Vernon, C. A.; Welch, V. A. *J. Chem. Soc.* **1961**, 2670–2676.

(93) Nicholls, A.; Wlodek, S.; Grant, J. A. *J. Phys. Chem. B* **2009**, *113*, 4521–4532.

(94) Cui, Q. *J. Chem. Phys.* **2002**, *117*, 4720–4728.

(95) Li, H.; Gordon, M. S. *J. Chem. Phys.* **2007**, *126*, 124112.

(96) Lange, A. W.; Herbert, J. M. *J. Phys. Chem. Lett.* **2010**, *1*, 556–561.

CT1001818

JCTC Journal of Chemical Theory and Computation

Dependence of Excited State Potential Energy Surfaces on the Spatial Overlap of the Kohn−Sham Orbitals and the Amount of Nonlocal Hartree−Fock Exchange in Time-Dependent Density Functional Theory

Jürgen Plötner,[†] David J. Tozer,[‡] and Andreas Dreuw*,[†]

Institute of Physical and Theoretical Chemistry, Goethe-University Frankfurt, Max von Laue-Strasse 7, 60438 Frankfurt, Germany, and Department of Chemistry, University of Durham, South Road, Durham DH1 3LE, United Kingdom

Received April 13, 2010

Abstract: Time-dependent density functional theory (TDDFT) with standard GGA or hybrid exchange-correlation functionals is not capable of describing the potential energy surface of the S_1 state of Pigment Yellow 101 correctly; an additional local minimum is observed at a twisted geometry with substantial charge transfer (CT) character. To investigate the influence of nonlocal exact orbital (Hartree−Fock) exchange on the shape of the potential energy surface of the S_1 state in detail, it has been computed along the twisting coordinate employing the standard BP86, B3LYP, and BHLYP xc-functionals as well as the long-range separated (LRS) exchange-correlation (xc)-functionals LC-BOP, ωB97X, ωPBE, and CAM-B3LYP and compared to RI-CC2 benchmark results. Additionally, a recently suggested Λ-parameter has been employed that measures the amount of CT in an excited state by calculating the spatial overlap of the occupied and virtual molecular orbitals involved in the transition. Here, the error in the calculated S_1 potential energy curves at BP86, B3LYP, and BHLYP can be clearly related to the Λ-parameter, i.e., to the extent of charge transfer. Additionally, it is demonstrated that the CT problem is largely alleviated when the BHLYP xc-functional is employed, although it still exhibits a weak tendency to underestimate the energy of CT states. The situation improves drastically when LRS-functionals are employed within TDDFT excited state calculations. All tested LRS-functionals give qualitatively the correct potential energy curves of the energetically lowest excited states of P. Y. 101 along the twisting coordinate. While LC-BOP and ωB97X overcorrect the CT problem and now tend to give too large excitation energies compared to other non-CT states, ωPBE and CAM-B3LYP are in excellent agreement with the RI-CC2 results, with respect to both the correct shape of the potential energy curve as well as the absolute values of the calculated excitation energies.

Introduction

During the last 25 years, linear-response time-dependent density functional theory (TDDFT)[1−3] has become one of the most widely used quantum chemical methods to study excited electronic states of medium sized to large molecules with say 30 up to 300 second row atoms. (See, for example, refs 4−7.) In general, locally excited electronic states that lie energetically well below the ionization potential, e.g., typical $\pi\pi^*$ or $n\pi^*$ transitions, are usually well described within TDDFT using standard exchange-correlation (xc)-functionals, very often with accuracies close to those of sophisticated wave function based methods. Typically, these $\pi\pi^*$ excited states carry most of the oscillator strength, and thus, experimental absorption spectra are very often ac-

* Corresponding author e-mail: andreas@theochem.uni-frankfurt.de.
[†] Goethe-University Frankfurt.
[‡] University of Durham.

10.1021/ct1001973 © 2010 American Chemical Society
Published on Web 07/02/2010

curately reproduced.[8,9] The excellent cost-to-accuracy relationship that TDDFT offers for the calculation of such experimental spectra is the basis for the great early success of TDDFT.

In contrast, charge transfer (CT) excited states are often described very poorly using standard local xc-functionals.[10−13] In principle, these states suffer from two related deficiencies of the standard functionals. On one hand, so-called electron transfer self-interaction is not correctly eliminated by the response of local xc-potentials leading to a physically incorrect asymptotic behavior of the potential energy curves with respect to a separation coordinate of the generated positive and negative (partial) charges in a CT excited state.[3,11,12] Second, the excitation energies tend to be significantly underestimated, which can be related to the integer discontinuity[14] in the case of asymptotic intermolecular excitations.[15] It was recognized very early that these failures are related to the locality of the employed xc-potentials and the inclusion of long-range, nonlocal exact orbital (Hartree−Fock) exchange would alleviate these problems.[11] Along these lines of thought, several long-range separated xc-potentials have been proposed, e.g., CAM-B3LYP,[16] LC-BLYP or LC-BOP,[17−19] ωPBE,[20,21] ωB97x,[22,23] LC-ωPBE,[24,25] and others,[26,27] and it has been demonstrated that these functionals indeed improve the description of CT excited states in TDDFT substantially. Also, the application of exact local Kohn−Sham exchange[28,29] instead of nonlocal Hartree−Fock exchange is a viable route to overcome the CT problem as first tests have recently shown.[30]

However, a potential problem of the inclusion of long-range Hartree−Fock (HF) exchange into TDDFT is the corruption of its well-known accuracy for locally excited states described above. This relates to the fact that Hartree−Fock exchange leads to an increase of the energy gap between the occupied and virtual orbitals owing to Koopmans' Theorem, making virtual-occupied orbital energy differences a poor zeroth order estimate for excitation energies. Therefore, it is not always useful or desirable to turn to long-range separated xc-functionals, and instead, one may wish to stay with the standard xc-functionals. In such situations, it is very important to possess a diagnostic tool to estimate the influence of the CT failure, i.e., to judge the expected quality of the computed excitation energies and as such the usefulness of the chosen xc-functional for the molecular system under consideration.

It is clear that the error in excitation energies due to the CT failure depends on the extent of charge transfer, and therefore, one should aim at measuring the degree of CT in a particular excited state. An intuitive measure of charge transfer character is given by the spatial overlap of the electron donating orbitals and the electron accepting orbitals.[31] In an excited state, the former are given by the occupied orbitals that are depopulated upon excitation while the latter correspond to the virtual orbitals that get populated. Peach et al. have recently proposed such a CT measure, Λ as the sum of the spatial overlaps O_{ia} between occupied and virtual orbitals involved in one excited state weighted by the square of their transition amplitudes κ_{ia},[31]

$$\Lambda = \frac{\sum_{ia} \kappa_{ia}^{\,2} O_{ia}}{\sum_{ia} \kappa_{ia}^{\,2}} \tag{1}$$

where the spatial overlap is given as

$$O_{ia} = \langle |\varphi_i||\varphi_a| \rangle = \int |\varphi_i(r)||\varphi_a(r)|\,dr \tag{2}$$

It was clearly demonstrated that the parameter Λ correlates with the error in excitation energies, i.e., the smaller the Λ, the larger are both the CT character and the error.[31−35]

In this contribution, we evaluate the performance of long-range separated functionals in the description of intramolecular charge transfer and, in addition, demonstrate the usefulness of the CT diagnostic Λ for the investigation of real photochemical problems. For that purpose, we reinvestigate the photochemistry of Pigment Yellow 101 (P. Y. 101, Figure 1), a commercially available fluorescent yellow pigment. Besides its fluorescence as the main excited state decay channel, P. Y. 101 exhibits a surprisingly rich photochemistry.[8,36−38] Upon photoexcitation into the first excited S_1 state, different competing processes occur: excited state intramolecular proton transfer (ESIPT), isomerization at the central N−N bond and, of course, fluorescence. We have chosen P. Y. 101, since its vertical absorption spectrum is very nicely reproduced using TDDFT with the standard local BP86 xc-functional. However, on the other hand, it has been shown that the potential energy surface of the S_1 state exhibits an artificial minimum at a twisted geometry with a CNNC (C═N−N═C) dihedral angle of the central bisazomethine group of 90° at this level of theory adulterating the excited state dynamics.[38] This artificial minimum occurs due to the CT failure of TDDFT, since the increasing amount of CT character of the S_1 state along the twisting coordinate leads to a spurious lowering of the excited state energy. The same effect has also been observed for 4-(dimethylamino)benzonitrile.[34] In summary, P. Y. 101 is one example of those molecules where one rather sticks to the common local xc-functionals due to the high quality of the vertical description of the S_1 $\pi\pi^*$ state, but still, one needs a diagnostic tool to evaluate the quality of the potential energy surface along some reaction coordinate where the amount of CT character may increase. We will show here the parameter Λ is a suitable tool for that purpose.

The paper is organized as follows. In the following section, the computational details of our theoretical investigation will be outlined. Our results will be presented and discussed, first starting with benchmark calculations at the RI-CC2 level of theory, then turning to TDDFT results employing standard local and hybrid xc-functionals and the application of the CT measure Λ, and finally giving TDDFT results obtained using long-range separated xc-functionals. The paper concludes with a brief summary of the main conclusions.

Computational Details

The minimum energy pathway for excited state intramolecular proton transfer (ESIPT) of P. Y. 101 has been computed at the TDDFT/BP86/DZP, and it has been noted

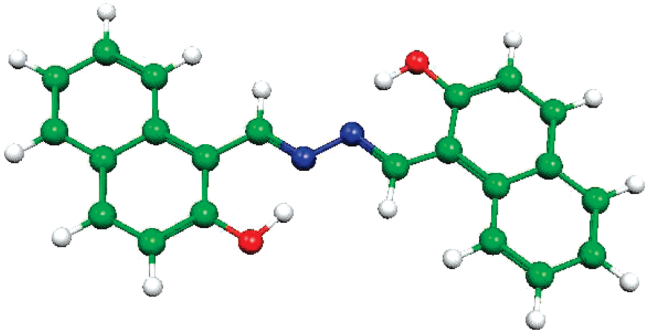


Figure 1. Molecular structure of Pigment Yellow 101.

previously that the central bisazomethine CNNC subunit supposedly twists upon proton transfer at that level of theory.[38] In fact, the minimum on the excited state potential energy surface corresponding to the twisted geometry of the keto-form of P. Y. 101 with a central CNNC dihedral angle of almost 90° is an artifact of the employed BP86 xc-functional and originates most likely from the CT failure. Therefore, we focus here on the twisting of the keto-form of P. Y. 101 only, and the C═N−N═C dihedral angle is scanned from 180° (planar keto-minimum) to 90° (twisted artificial minimum). The potential energy curves have been recomputed along this pathway at the TDDFT/BP86 optimized geometries using TDDFT with the standard B3LYP[39,40] and BHLYP[41] hybrid xc-functionals and with the newly developed long-range separated xc-functionals LC-BOP,[17,18] ωPBE,[20,21] and ωB97X[22,23] as well as CAM-B3LYP[16] using the standard 6-31G* basis set. In addition, the approximate coupled-cluster theory of second order exploiting the resolution-of-the-identity approximation (RI-CC2) has been used for benchmarking purposes.[42] For the RI-CC2 calculations, the DZP basis set has been used along with the auxiliary SVP basis set for the density fitting, and the energetically sixteen lowest occupied molecular orbitals were frozen. The equations for the computation of the Λ parameter to measure the CT character of an excited state have been implemented into a development version of the Q-Chem 3.2 program package.[43] This code has been used to compute Λ for the lowest excited S_1 state of P. Y. 101 along the twisting coordinate for all xc-functionals employed in our TDDFT calculations. All calculations have been performed using the Turbomole 5.10 program package,[44] a development version of Q-Chem 3.2,[43] and the CAM-B3LYP calculations have been done with Dalton.[45]

Twisting P. Y. 101 at the CC2 Level of Theory. To evaluate the performance of different xc-functionals in the computation of the potential energy surface of the S_1 electronic state of P. Y. 101 along the reaction coordinate representing the twisting of the central bisazomethine unit, we have first computed the five lowest excited states using RI-CC2 along the TDDFT/BP86 optimized reaction path. The obtained potential energy curves are displayed in Figure 2. The curve obtained for the S_1 state, which is responsible for the excited state dynamics of P. Y. 101, serves as benchmark for the evaluation of the quality of the TDDFT results with respect to the CT failure discussed in later sections. This is justified, since CC2 being an ab initio method does not suffer from the charge transfer failure and,

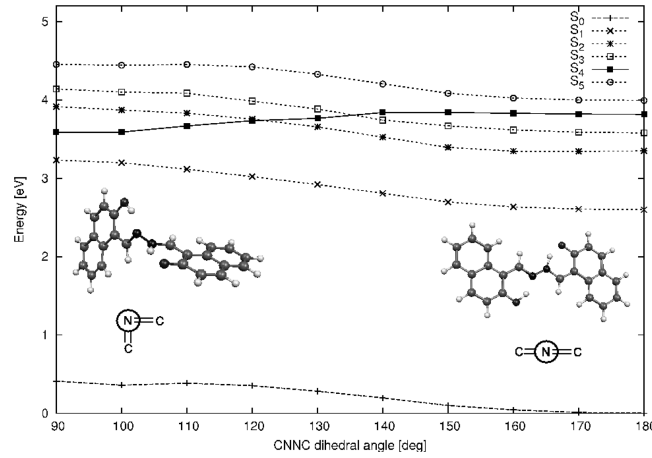


Figure 2. Potential energy curves of the ground and five lowest excited electronic states computed at the level of RI-CC2 along the twisting coordinate of the central bisazomethine subunit of the keto-form of P. Y. 101. In the insets, the planar (180°) and twisted (90°) structures as well as the corresponding Newman projections are also displayed.

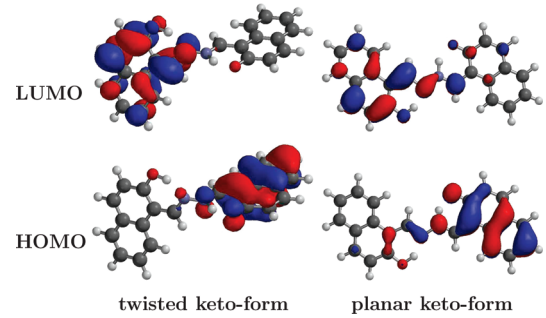


Figure 3. Highest occupied and lowest unoccupied molecular orbitals (HOMO and LUMO) at the theoretical level of Hartree−Fock for the planar and twisted keto-forms of P. Y. 101.

on the contrary, describes singly excited states with an accuracy comparable to EOM-CCSD, i.e., with an error of 0.1−0.3 eV.[42,46]

While the planar keto-form of P. Y. 101 with a central CNNC dihedral angle of 180° corresponds to a true local minimum on the S_1 potential energy surface at the theoretical level of RI-CC2, the twisted form with a dihedral angle of 90° corresponds to a transition state connecting the trans and cis keto-isomers of P. Y. 101. At the theoretical level of RI-CC2, the twisted geometry lies 0.63 eV above the trans keto-minimum. Overall, the potential energy curve of the S_1 state follows the one of the electronic ground state along the investigated CNNC twisting coordinate. The barrier for twisting, however, is with 0.63 eV slightly higher in the S_1 state than in the ground state, where the barrier exhibits a value of 0.4 eV. In the molecular orbital picture, the S_1 electronic state is mainly characterized by a single excitation of an electron out of the highest occupied molecular orbital (HOMO) into the lowest unoccupied molecular orbital (LUMO), which are both displayed in Figure 3 for the planar as well as twisted geometries of the keto-form of P. Y. 101. As is easily seen, the bonding character of the N−N bond does slightly differ in the HOMO and the LUMO. While

the HOMO exhibits weak antibonding character along the N−N bond, the LUMO is weakly bonding explaining the small increase in the twisting barrier. Moreover, the HOMO and LUMO also reveal the charge transfer character of the S_1 state already in the planar trans keto-form, since the lobes of the two orbitals are located mainly on opposite rings. Going to the twisted structure, the localization of the orbitals becomes more pronounced and, thus, the spatial overlap between these orbitals decreases even further (Figure 3).

Closer examination of Figure 2 shows that the potential energy curves of all excited states are parallel to the ground state, except for the S_4 state of the planar keto-isomer with a vertical excitation energy of 3.8 eV at RI-CC2 level of theory. This state does exhibit a minimum at the twisted geometry at a dihedral angle of practically 90°. In the molecular orbital picture, this state is best represented by an electronic transition from the HOMO to the LUMO + 1 (not shown). While at the planar geometry of the keto-isomer, the orbitals are mainly localized on the keto side of P. Y. 101; they still do extend slightly over to the other ring. However, upon twisting, the HOMO and LUMO + 1 strictly localize on the keto side. Therefore, the S_4 state of the planar keto-isomer can be classified as a local excitation on that naphthalene ring where the proton has been transferred and which now possesses the keto group. As we will see later, this state seems to disappear from the low-energy region at the TDDFT/BP86 and TDDFT/B3LYP levels of theory and is only found when the BHLYP xc-functional is employed (Figure 4).

Twisting P. Y. 101 at the TDDFT Level of Theory Using Standard xc-Functionals. Let us now turn to the detailed investigation of how time-dependent density functional theory performs in the description of the potential energy surface of the S_1 state of P. Y. 101 along the twisting coordinate when standard local and hybrid xc-functionals are used. As noticed previously,[38] employing the BP86 xc-functional TDDFT generates a spurious minimum on the S_1 surface corresponding to the twisted geometry with a CNNC dihedral angle of 90° (upper panel in Figure 4). At that level of theory, the twisted geometry is 0.25 eV lower in energy than the planar trans keto-form of P. Y. 101. Comparison of the potential energy curve of the S_1 state of P. Y. 101 computed at the TDDFT/BP86 level with those obtained at the RI-CC2 level (Figure 2) suggests that S_1 may correspond to the state previously denoted as S_4 at the RI-CC2 level, since their corresponding curves exhibit practically identical shapes. However, this is not the case. Inspection of the HOMO and the LUMO orbitals involved in the S_1 state at the TDDFT/BP86 level (Figure 5) shows that they are practically identical to the ones employed in the RI-CC2 calculation, and thus, the S_1 state is the same state at TDDFT/BP86 as well as at the RI-CC2 level. Thus, it is not the ordering of the states that changes, but rather, it is the shape of the potential energy curves.

Naturally, the question arises why the shape of the S_1 potential energy curve changes so drastically from RI-CC2 to TDDFT/BP86. As we have mentioned before, this is most likely due to the CT failure of TDDFT when standard local xc-functionals are employed. A first indication is the increase

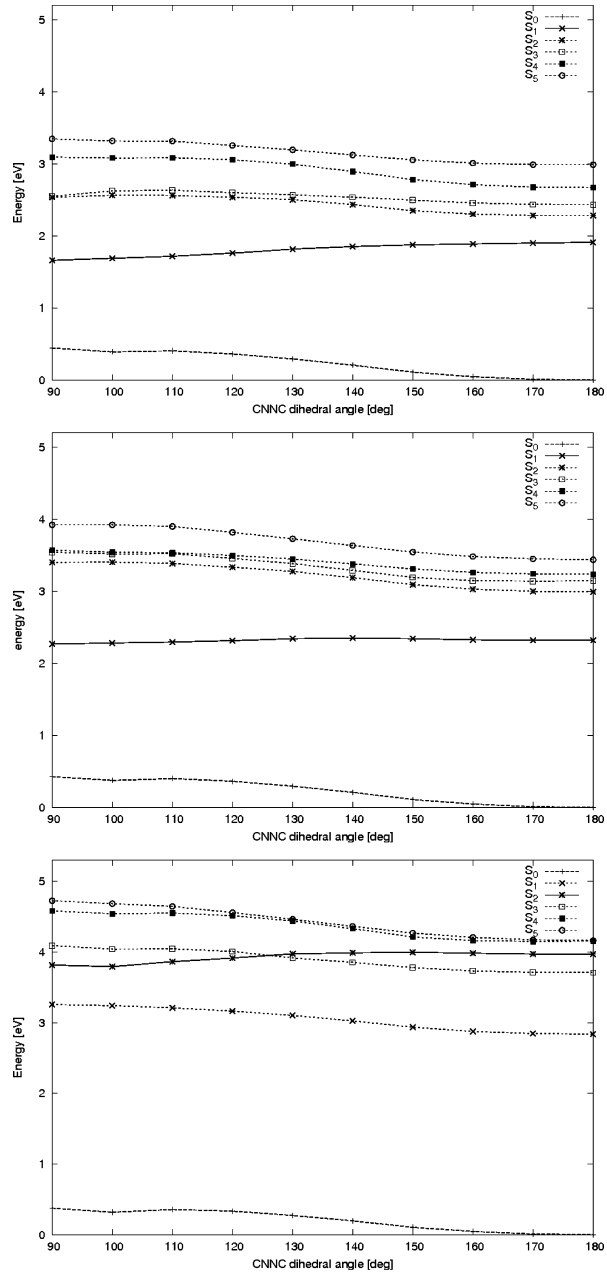


Figure 4. Potential energy curves of the ground and energetically lowest five excited singlet states calculated at the theoretical levels of TDDFT using the BP86 (top), B3LYP (middle), and BHLYP (bottom) xc-functional along the dihedral angle of the central CNNC bisazomethine subunit of P. Y. 101.

of the excited state dipole moment along the twisting coordinate from about 10 D to more than 16 D, pointing at a rise of the charge transfer character. In principle, this is not unusual, since twisted intramolecular charge transfer (TICT) states have been found to exist in a plethora of aromatic push−pull systems. (See, for example, refs 47 and 48.) However, here in P. Y. 101, this twisting is undoubtedly an artifact, since the twisted structure corresponds to a transition state and not to a local minimum at the level of RI-CC2, our benchmark method.

A general simple procedure to disguise the CT failure is to use xc-functionals with an increasing amount of nonlocal Hartree−Fock (HF) exchange, since this gradually alleviates

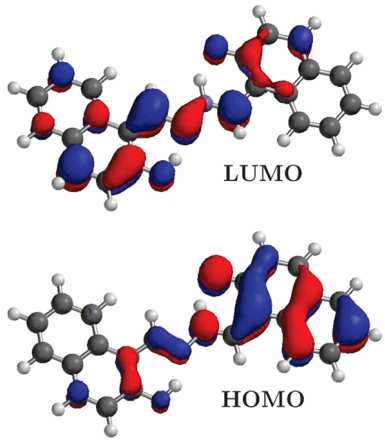


Figure 5. Highest occupied and lowest unoccupied molecular orbital of the planar keto-form of P. Y. 101 at the theoretical level of DFT/BP86.

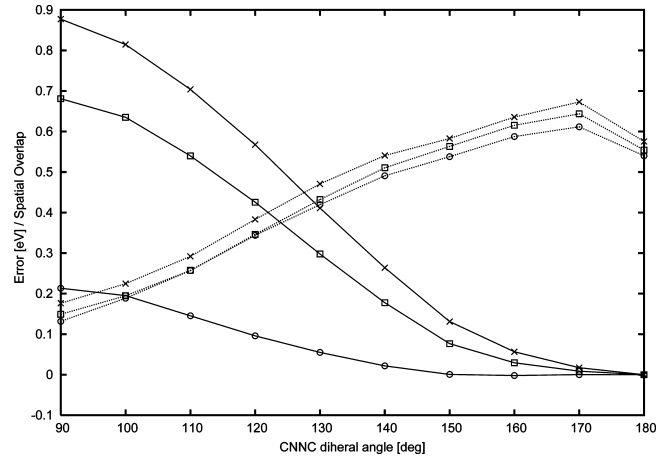


Figure 6. Comparison of the errors of the potential energy curves (solid lines) with the spatial overlap Λ of the involved orbitals at the TDDFT level employing the BP86 (crosses), B3LYP (boxes), and BHLYP (circles) xc-functionals along the twising coordinate of P. Y. 101. For the definition of the error, see text.

the problem. Therefore, we have recomputed the potential energy curve of the S_1 state along the twisting coordinate also using the B3LYP and BHLYP xc-functionals which contain a constant amount of 20% and 50% HF exchange, respectively. As can be seen in the middle of Figure 4, the shape of the S_1 curve does indeed change when the B3LYP xc-functional is employed in the calculation of the twisting path. The S_1 curve is much flatter with the twisted and the planar structure being practically degenerate at the level of TDDFT/B3LYP. The twisted structure, however, is still slightly lower in energy than the planar one by only 0.05 eV. As a consequence, an unconstrained geometry optimization at the TDDFT/B3LYP level of theory converges onto the artificially twisted geometry. It is important to emphasize that the S_1 state is the same at TDDFT/BP86 and TDDFT/B3LYP: both correspond essentially to HOMO to LUMO transitions with the orbitals being indistinguishable (not shown).

Increasing the amount of HF exchange even further to 50%, i.e., using the BHLYP xc-functional in the TDDFT calculation of the twisting path, the potential curve of the S_1 state changes again. In the lower panel of Figure 4, it is seen that at the level of TDDFT/BHLYP the twisted minimum has disappeared and is now a transition state, as in RI-CC2, lying 0.41 eV above the planar keto-form of P. Y. 101. The obtained picture of the potential energy curves of the lowest excited electronic states at the level of TDDFT/BHLYP agrees qualitatively quite nicely with the one obtained with our benchmark method RI-CC2. Again, the S_1 state is the same at all levels of theory employed, and the potential energy surfaces have been computed along the same twisting path optimized at the level of TDDFT/BP86. Owing to the gradual removal of the local minimum corresponding to the twisted geometry with increasing amounts of nonlocal HF exchange, it can be undoubtedly concluded that the artificial twisted minimum on the S_1 surface of P. Y. 101 at the TDDFT/BP86 and TDDFT/B3LYP levels originates solely from the CT failure.

Nevertheless, the experimental absorption spectrum of P. Y. 101 is reasonably well reproduced using TDDFT/BP86 when the computed excitation energies are shifted by 0.41 eV to higher energies to correct for errors in the method and solution effects.[8,36] In fact, comparison with experimental absorption spectra is very often the only possibility to validate the employed theoretical method, in particular when the molecules under investigation are too large for proper benchmark calculations. Furthermore, the computational effort is smaller for local GGA xc-functionals than for hybrid functionals, since for the latter also the HF exchange contribution needs to be evaluated. For very large molecules or molecular assemblies, this may make a significant difference. In such situations, an intrinsic diagnostic is very useful that helps to judge the influence of the CT failure and to estimate the quality of excitation energies or potential energy surfaces obtained with TDDFT. The Λ parameter as defined in the Introduction and explained in more detail in ref 31 is a useful measure of charge transfer character of an excited state and, thus, a useful diagnostic of the reliability of TDDFT calculations. Here, we have computed the value of Λ for the relevant S_1 state along the twisting coordinate, i.e., along the central CNNC dihedral angle of the bisazomethine unit for all three applied xc-functionals BP86, B3LYP, and BHLYP (Figure 6). The spatial overlap of the orbitals involved in the S_1 transition drops significantly from the planar to the twisted structure at all levels of theory in a practically identical way from a value of about 0.7 to below 0.2. The overall decrease of Λ, i.e., the reduction of spatial orbital overlap, demonstrates the increase in charge transfer character of the S_1 state along the twisting coordinate, which is in agreement with our previous observation that the static dipole moment of the S_1 state increases significantly upon twisting.

Also plotted in Figure 6 is the deviation of the TDDFT potential energy curves from the RI-CC2 curves for the different xc-functionals. For the calculation of the error, the energy of the planar keto-form is set to zero for all curves and the difference between the CC2 curve and the corresponding TDDFT curve defines the error. Despite the similar behavior of Λ for the three functionals, the error in the

Table 1. Comparison of the Absolute Excitation Energies of the Five Lowest Excited States Obtained at the RI-CC2 Level with Those Calculated by TDDFT Employing Different xc-Functionals at the Planar Structure of the Keto-Form of P.Y. 101 and at the Twisted Structure[a]

state	RI-CC2	BP86	B3LYP	BHLYP	LC-BOP	ωPBE	ωB97X	CAM-B3LYP
				planar keto-form (180°)				
S_1	2.58 (0.84)	1.90 (0.29)	2.30 (0.51)	2.82 (0.80)	2.90 (0.85)	2.81 (0.83)	2.96 (0.86)	2.77 (0.80)
		[−0.68]	[−0.28]	[+0.24]	[+0.32]	[+0.23]	[+0.38]	[+0.19]
S_2	3.32 (0.00)	2.27 (0.00)	2.97 (0.36)	3.68 (0.22)	3.75 (0.10)	3.64 (0.05)	3.80 (0.11)	3.63 (0.17)
		[−1.05]	[−0.35]	[+0.36]	[+0.43]	[+0.32]	[+0.48]	[+0.31]
S_3	3.55 (0.15)	2.42 (0.27)	3.12 (0.00)	3.94 (0.03)	3.80 (0.00)	3.67 (0.08)	3.97 (0.00)	3.76 (0.00)
		[−1.13]	[−0.33]	[0.39]	[+0.25]	[+0.12]	[+0.42]	[+0.21]
S_4	3.79 (0.01)	2.65 (0.21)	3.21 (0.03)	4.12 (0.00)	4.22 (0.33)	4.08 (0.07)	4.27 (0.31)	3.94 (0.02)
		[−1.14]	[−0.58]	[+0.34]	[+0.43]	[+0.29]	[+0.48]	[+0.15]
S_5	3.96 (0.21)	2.97 (0.60)	3.41 (0.01)	4.13 (0.16)	4.35 (0.02)	4.11 (0.22)	4.40 (0.03)	4.06 (0.24)
		[−0.99]	[−0.55]	[+0.17]	[+0.39]	[+0.15]	[+0.44]	[+0.10]
				twisted keto-form (90°)				
S_1	2.80 (0.01)	1.21 (0.00)	1.83 (0.00)	2.86 (0.00)	3.38 (0.57)	3.10 (0.00)	3.43 (0.60)	2.90 (0.00)
		[−1.59]	[−0.97]	[+0.06]	[+0.58]	[+0.30]	[+0.63]	[+0.10]
S_2	3.15 (0.55)	2.08 (0.06)	2.95 (0.00)	3.42 (0.85)	3.44 (0.07)	3.34 (0.61)	3.52 (0.06)	3.33 (0.61)
		[−1.07]	[−0.20]	[+0.27]	[+0.29]	[+0.19]	[+0.37]	[+0.18]
S_3	3.48 (0.11)	2.09 (0.01)	3.09 (0.28)	3.69 (0.04)	3.70 (0.04)	3.61 (0.05)	3.76 (0.04)	3.60 (0.05)
		[−1.39]	[−0.39]	[+0.21]	[+0.22]	[+0.13]	[+0.28]	[+0.12]
S_4	3.70 (0.00)	2.63 (0.01)	3.12 (0.29)	4.18 (0.19)	4.00 (0.00)	3.91 (0.00)	4.15 (0.00)	4.01 (0.00)
		[−1.07]	[−0.58]	[+0.48]	[+0.30]	[0.21]	[+0.45]	[+0.31]
S_5	4.01 (0.24)	2.88 (0.38)	3.47 (0.00)	4.32 (0.01)	4.21 (0.55)	4.10 (0.21)	4.25 (0.24)	4.08 (0.20)
		[−1.13]	[−0.54]	[+0.31]	[+0.20]	[+0.09]	[+0.24]	[+0.07]

[a] Excitation energies are given in eV, oscillator strengths are given in parentheses, and the difference to the RI-CC2 reference value is given in eV in square brackets.

potential energy curves is significantly different. For the BP86 xc-functional, the error increases drastically along the twisting coordinate and reaches a value of 0.88 eV at the twisted geometry. In other words, at the twisted geometry, the excitation energy of the S_1 state is strongly underestimated by that value, in contrast to the planar structure. This underestimation of the excitation energy at a particular geometry as observed here for TDDFT/BP86 introduces spurious minima on excited state potential energy surfaces. For B3LYP, the error increase is generally smaller due to its 20% nonlocal HF exchange, which partially corrects for the CT failure. The error in the potential energy surface at the twisted geometry amounts to 0.69 eV, i.e., 0.2 eV less than at the BP86 level, although the spatial overlap of the involved orbitals is in analogy to BP86 below 0.2. Using BHLYP finally improves the situation substantially. The inclusion of 50% HF exchange leads to a decrease of the error at the twisted geometry to only 0.21 eV. Again, the spatial overlap is unaffected by the inclusion of HF exchange. In this series, BHLYP is the only xc-functional that yields the qualitatively correct shape of the potential energy surface, however, still with a small bias toward twisted structures with large charge transfer character.

Let us briefly return to the electronic state which was found as S_4 at the RI-CC2 level of theory (Figure 2) which exhibits a minimum at the twisted geometry. We have proven above that the S_1 state is indeed the same state at all employed levels of TDDFT and RI-CC2 although the shapes of the curves are remarkably different. However, no other states of those displayed at the TDDFT/BP86 and TDDFT/B3LYP levels exhibit a potential energy curve similar to the one of the S_4^{CC2} state except S_1. At TDDFT/BHLYP, the S_4^{CC2} state reapears as S_3 at the planar geometry and corresponds also to the HOMO to LUMO + 1 electronic transition. Note that the orbitals do not change in the valence region, so the HOMO − 1, HOMO, LUMO, and LUMO + 1 are virtually identical at all employed levels of theory. Naturally, one may, thus, ask the question, if S_1 does not correspond to S_4^{CC2} where is this state at the TDDFT/BP86 and TDDFT/B3LYP levels? In fact, the corresponding state does exist, but it is not found among the five lowest excited states displayed in Figure 4. This has again to do with the CT failure of TDDFT and the artificial lowering of the excitation energies of CT excited states compared to locally excited states as the S_4 state. With decreasing amount of HF exchange from BHLYP to BP86, more and more CT excited states drop below the locally excited S_4^{CC2} state making this state seemingly disappear.

Comparison of the absolute excitation energies (Table 1) obtained at BP86, B3LYP and BHLYP show the typical trend to higher excitation energies. For example, the S_1 state possesses an excitation energy of 1.90 eV at the TDDFT/BP86 level and 2.30 and 2.82 eV using the B3LYP and BHLYP xc-functionals, repectively. This is a well-known phenomenon directly related to the amount of nonlocal Hartree−Fock exchange and the concomitant increase of the occupied-virtual gap. Much more interesting, however, is the absolute error of all states compared to our theoretical benchmark RI-CC2 results. If the error of all computed states is roughly the same, one can reliably compute absorption spectra and investigate excited state processes, in which more than one electronic state is involved. For the latter, precise knowledge of state crossings or conical intersection, i.e., of the relative position of the involved states, is an inevitable prerequisite. As is easily seen in Table 1, the deviations of the BP86 excitation energies of the S_1−S_5 states cover a wide range from −0.68 up to −1.23 eV at the planar geoemtry and −1.07 to −1.59 eV at the twisted geometry revealing

an unbalanced treatment of the different excited states of P.Y. 101. The situation is slightly improved for the B3LYP xc-functional; here, the absolute deviation is smaller at the planar geometry (−0.28 to −0.58 eV), and the shift is similar for all states. Although the deviation from the RI-CC2 results is still a bit large, the quality of the states is roughly the same for all resulting in a more balanced treatment. However, at the twisted geometry, the quality of the B3LYP results degrades strongly and the deviation from the RI-CC2 results becomes as large as −0.97 eV for the S_1 state due to its increased CT character leading to an unbalanced treatment of the states. The situation is greatly improved when the BHLYP xc-functional is emplyed within TDDFT. The deviation from the RI-CC2 results ranges from +0.17 to +0.48 eV, slightly overestimating the excitation energies, but most importantly, the quality of the results does not decrease at the twisted geometry of P.Y. 101. The deviation of the S_1 state is here actually the smallest with only 0.06 eV due to the tendency to underestimate excitation energies of CT states compensating the general overestimation of the excitation energies by BHLYP.

Twisting P. Y. 101 at the TDDFT Level of Theory Using Long-Range Separated xc-Functionals. In the previous section, we have seen that inclusion of at least 50% long-range Hartree−Fock exchange, as in BHLYP, largely corrects for the false shape of the potential energy surface of the S_1 state of P.Y. 101 along the twisting coordinate when local and GGA xc-functionals are employed with TDDFT. Using BHLYP, the PES exhibits correctly a transition state at a 90° dihedral angle and not a local minimum (Figure 4). However, BHLYP still has a bias to favor charge transfer (Figure 6) and, in addition, it overestimates excitation energies generally by about 0.2−0.5 eV as compared to RI-CC2, which has been observed to be generally the case. Recently, long-range separated (LRS) xc-functionals have been suggested and designed to also correct for the charge transfer failure of local and GGA xc-functionals in TDDFT. They do not contain a constant fraction of nonlocal Hartree−Fock exchange but instead use an increasing amount with growing electron−electron distance. While most LRS-functionals include at long-range 100% HF exchange, CAM-B3LYP for instance reaches only 65%. Thereby, functionals including 100% exact orbital exchange do by construction correctly describe the $1/R$ dependence of potential energy curves of long-range charge transfer excited states along a charge-separation coordinate R at the TDDFT level, and the longer-ranged the CT the better. LRS-functionals with lower fractions of HF exchange recover only the corresponding lower fraction of the $1/R$ asymptote.[11] However, it has not been investigated and compared yet, how different LRS-functionals perform in the description of intramolecular charge transfer problems. For this objective, the investigated twisting coordinate of the keto-form of P.Y. 101 is a perfect test case. On one hand, at the planar configuration with a central bisazomethine dihedral angle of 180°, only little nonlocal Hartree−Fock is needed to obtain reasonable excitation energies and reliable spectra, for example B3LYP would be sufficient (Table 1), while on the other hand increasing amounts of Hartree−Fock exchange are required

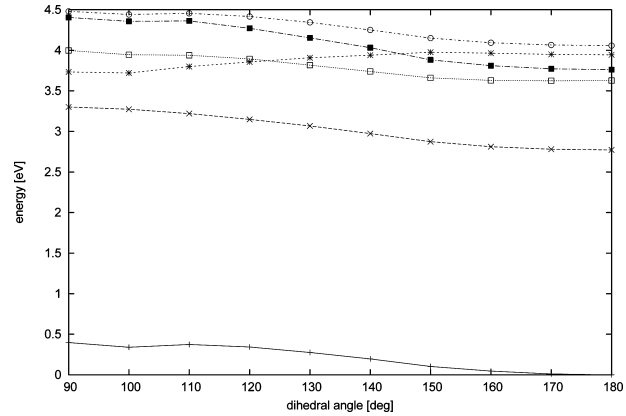


Figure 7. Potential energy curves of the ground and energetically lowest five excited singlet states calculated at the theoretical level of TDDFT using the long-range separated xc-functional CAM-B3LYP along the dihedral angle of the central CNNC bisazomethine subunit of P. Y. 101.

to describe the potential energy surface correctly, more than 50%, since BHLYP is still favoring charge transfer (Figure 6). Therefore, we have computed the potential energy surfaces of the five energetically lowest excited states along the twisting coordinate of P.Y. 101 using the LRS-functionals LC-BOP,[17−19] ωPBE,[20,21] ωB97X,[22,23] and CAM-B3LYP[16] in strict analogy to the calculations reported above. The curves obtained with all these LRS-functionals are very similar with minor deviations in the absolute energies; the overall qualitative picture, however, is practically identical. As one example, the CAM-B3LYP curves are displayed in Figure 7. Comparison with the computed curves at RI-CC2 (Figure 2) and TDDFT/BHLYP (lowest panel of Figure 4) reveals an excellent agreement. As it was the case for all tested xc-functionals so far, BP86, B3LYP, and BHLYP, the spatial overlap of the molecular orbitals involved in the S_1 state measured by the Λ-parameter decreases also for CAM-B3LYP from 0.65 to 0.15, indicating an increasing CT character of S_1 along the twisting coordinate.

For a detailed analysis of the performance of the various LRS-functionals with respect to the correct description of the shape of the potential energy surface of the critical S_1 state along the twisting coordinate, we have again compared it to the potential energy curve at the RI-CC2 level in analogy to above. The differences between the potential energy curves at the TDDFT level employing the LRS-functionals and RI-CC2 are plotted in Figure 8; a positive sign of the error corresponds to an underestimation of the energy of the twisted form, and a negative one corresponds to an overestimation. As can be seen, the deviations from the RI-CC2 curve are much smaller for all LRS-functionals than for BP86 or B3LYP and are of the same order of magnitude as the deviation of the BHLYP curve, i.e., at most 0.2 eV at the twisted structure of P. Y. 101 which exhibits pronounced charge transfer character. While CAM-B3LYP has a small tendency to underestimate the energy of the twisted charge transfer structure relative to the planar one by as little as 0.1 eV, all other LRS-functionals overestimate its energy. This is probably related to the fact that CAM-B3LYP includes only 65% or long-range HF exchange, while all other tested

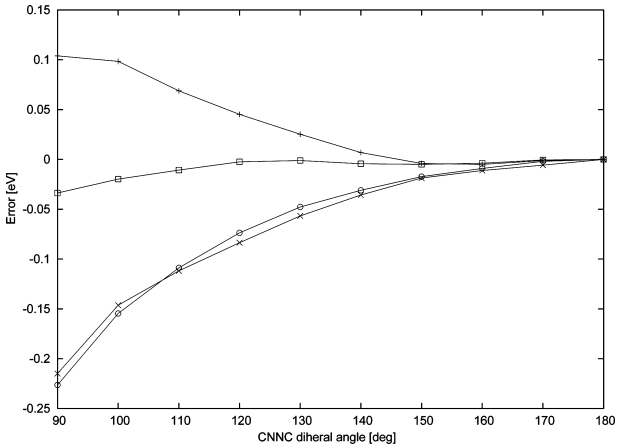


Figure 8. Errors of the potential energy curves of the S_1 state of P. Y. 101 along the twisting coordinate at the TDDFT level employing the long-range separated xc-functionals LC-BOP (circles), ωB97X (crosses), ωPBE (boxes), and CAM-B3LYP (pluses) relative to the curve calculated at RI-CC2 level. For a detailed definition of the error see text.

LRS-functionals include 100%. Both LC-BOP and ωB97X overestimate the excitation energy of the twisted charge transfer structure by about 0.22 eV relative to the planar keto-form of P. Y. 101, i.e., they overcompensate the charge transfer failure of TDDFT and shift excitation energies of charge transfer states potentially to too high values. However, the curve computed with ωPBE deviates only as little as 0.03 eV from RI-CC2 and yields a practically identical curve. However, the observed deviations of the curves are small, in particular the ones of CAM-B3LYP and ωPBE, and lie clearly within the error of the applied TDDFT method. In fact, using LRS-functionals prevents the overemphasizing charge transfer character as the conventional GGA xc-functionals or hybrids with low fractions of HF exchange do.

Let us finally turn to the absolute errors in the excitation energies of the five lowest excited states, where again the RI-CC2 results serve as a benchmark. Here, it is again important to emphasize that a consistent treatment of all excited states is of particular importance for reliable calculation of absorption spectra and excited state dynamics. Thus, it is desirable that one functional exhibits the same quality for all energetically low-lying states. The vertical excitation energies of the five energetically lowest electronic states S_1 to S_5 of the planar and twisted structure of the keto-form of P. Y. 101 computed using TDDFT and the various LRS-functionals are compiled in Table 1.

LC-BOP and ωB97X reach about the same level of accuracy, which is also the same for BHLYP, as they deviate consistently by +0.2 to +0.6 from the RI-CC2 values, i.e., the computed excitation energies at those TDDFT levels are consistently too high. The above noticed tendency of LC-BOP and ωB97X to shift CT states to too high excitation energies is also seen here, since their deviation from the RI-CC2 results increases from +0.3 eV to +0.6 eV upon twisting, i.e., with increase of CT character. The two xc-functionals that yield the best agreement with the RI-CC2 results and which describe all tested excited states at the planar as well as twisted structure of P. Y. 101 in the most balanced way are CAM-B3LYP and ωPBE. While CAM-B3LYP has a slightly better agreement in the absolute values of the excitation energies, ωPBE treats the excited states of the twisted and planar structures in a slightly more balanced way. However, these deviations are very small and are much smaller than the typical errors of TDDFT excitation energies.

Summary and Conclusions

It has been noticed previously that local, GGA, and hybrid xc-functionals cannot describe the potential energy curve of the S_1 state of Pigment Yellow 101 (P. Y. 101) along a twisting coordinate of the central bisazomethine subgroup. The increasing intramolecular CT character of the S_1 state along this coordinate leads to an underestimation of the energy of the twisted configuration relative to the planar one, resulting in an artificial twisted minimum on its potential energy surface.

Here, the potential energy curves of the five energetically lowest excited states have been recomputed along this coordinate employing TDDFT with various xc-functionals ranging from the standard hybrid-functionals B3LYP and BHLYP up to modern long-range separated (LRS) xc-functionals LC-BOP, ωB97X, ωPBE, and CAM-B3LYP. The obtained results are evaluated with respect to RI-CC2 computations, which serve as our benchmark. Furthermore, a newly introduced diagnostic Λ parameter to measure the degree of CT character of a particular state has been employed to relate the failure of the standard functionals to the spatial overlap of the molecular orbitals involved in the electronic transition.

Using a series of TDDFT calculations employing the BP86, B3LYP, and BHLYP xc-functionals with increasing amounts of Hartree−Fock exchange, we could unequivocally demonstrate that the problems in the description of the S_1 surface arise from the CT failure of TDDFT. Furthermore, the Λ parameter is a very useful instrument to identify problematic CT cases for TDDFT, since the overlap of the involved orbitals and, thus, Λ decreases smoothly with the increasing error along the twisting coordinate. It is also seen that BHLYP largely alleviates the CT problem for P. Y. 101; however, this xc-functional still has a bias toward underestimating energies of CT states.

The situation is greatly improved when LRS-functionals are employed. While ωB97X and LC-BOP tend to overcorrect for the CT failure of TDDFT and yield the twisted structure at slightly too high energies, ωPBE and CAM-B3LYP yield potential energy surfaces for the S_1 state in excellent agreement with the RI-CC2 curve. The LRS-functionals consistently overestimate the vertical excitation energies of the five energetically lowest excited states, CAM-B3LYP and ωPBE by only 0.1−0.3 eV, the latter again in excellent agreement with the RI-CC2 values. On the basis of these findings, one can undoubtedly recommend LRS-functionals, preferably CAM-B3LYP and ωPBE, to be used in the calculation of excited states of molecules with certain intramolecular charge transfer character. For P. Y. 101 and similar cases, BHLYP also

yields qualitatively correct potential energy curves and reasonably balanced values for the energetically low lying excited states but still with a small bias to underestimate the energy of CT states.

Finally, it is important to note that P. Y. 101 is just one example for which the different xc-functionals have been evaluated. However, it is particularly representative since the electronic structure of excited states usually does change along a reaction coordinate, sometimes quite drastically and very often the amount of charge transfer changes substantially. For instance, this is the case in all twisted intramolecular charge transfer (TICT) systems, for which dimethylaminobenzonitrile (DMABN) is the proto-typical example. Indeed, the findings of the present study concerning Λ and the performance of LRS-functionals are fully consistent with the DMABN study of ref 34. Therefore, the observed trends can be expected to be valid for most other systems as well, but it might well be that the relative performance of the LRS-functionals may be different. A thorough evaluation of the performance of standard xc- or LRS-functionals and of the quality of the results obtained with TDDFT is, thus, strongly recommended or even demanded.

Another critical note concerns the chosen RI-CC2 method as the benchmark. For the molecular system under investigation and the chosen problem of intramolecular charge transfer, the RI-CC2 method is certainly adequate, since all states are undoubtedly single excitations. However, it has been demonstrated recently that the overall error in excitation energies is about the same for RI-CC2 and TDDFT calculations once LRS xc-functionals are employed,[49] which is essentially the same result that we observe here for the PES of P.Y.101. It is also well-known that RI-CC2 does not systematically provide more accurate results than TDDFT, but of course the qualitative evolution of excitation energies and transition moments is better reproduced by RI-CC2.[49−51] At present, other ab initio computations with higher accuracy are not feasible for the excited states of P. Y. 101 due to its molecular size. One could have chosen a smaller model test system, but we wanted to demonstrate the applicability and the usefulness of the Λ-parameter and available LRS-functionals for the investigation of a real chemical system.

Acknowledgment. Andreas Dreuw acknowledges financial support by the Deutsche Forschungsgemeinschaft as a Heisenberg-Professor. This project is funded within the Priority Programme SPP1145 of the Deutsche Forschungsgemeinschaft. Computation time has been generously provided by the Center of Scientific Computing of the Goethe University of Frankfurt/Main.

References

(1) Runge, E.; Gross, E. K. U. *Phys. Rev. Lett.* **1984**, *52*, 997.

(2) Casida, M. E. In *Recent Advances in Density Functional Methods, Part I*; Chong, D. P., Ed.; World Scientific: Singapore, 1995; pp 155−192.

(3) Dreuw, A.; Head-Gordon, M. *Chem. Rev.* **2005**, *105*, 4009.

(4) Hay, P. J. *J. Phys. Chem. A* **2002**, *106*, 1634.

(5) Nazeeruddin, M. K.; De Angelis, F.; Fantacci, S.; Selloni, A.; Viscardi, G.; Liska, P.; Ito, S.; Takeru, B.; Grätzel, M. *J. Am. Chem. Soc.* **2005**, *127*, 16835.

(6) Dreuw, A. *Chem. Phys. Chem.* **2006**, *7*, 2259.

(7) Suramitr, S. P. W.; Meeto, W.; Hannongbua, S. *Theor. Chem. Acc.* **2010**, *125*, 35.

(8) Dreuw, A.; Plötner, J.; Lorenz, L.; Wachveitl, J.; Djanhan, J. E.; Brüning, J.; Bolte, M.; Schmidt, M. U. *Ang. Chem. Int. Ed.* **2005**, *44*, 7783.

(9) Diercksen, M.; Grimme, S. *J. Chem. Phys.* **2004**, *120*, 3544.

(10) Tozer, D. J.; Amos, R. D.; Handy, N. C.; Roos, B. J.; Serrano-Andres, L. *Mol. Phys.* **1999**, *97*, 859.

(11) Dreuw, A.; Weisman, J. L.; Head-Gordon, M. *J. Chem. Phys.* **2003**, *119*, 2943–2946.

(12) Dreuw, A.; Head-Gordon, M. *J. Am. Chem. Soc.* **2004**, *126*, 4007–4016.

(13) Sobolewski, A. L.; Domcke, W. *Chem. Phys.* **2003**, *294*, 73.

(14) Perdew, J. P.; Parr, R. G.; Levy, M.; Balduz, J. L., Jr. *Phys. Rev. Lett.* **1982**, *49*, 1691.

(15) Tozer, D. J. *J. Chem. Phys.* **2003**, *119*, 12697.

(16) Yanai, T.; Tew, D. P.; Handy, N. C. *Chem. Phys. Lett.* **2004**, *393*, 51.

(17) Tawada, Y.; Tsuneda, T.; Yanagisawa, S.; Yanai, T.; Hirao, K. *J. Chem. Phys.* **2004**, *1210*, 8425.

(18) Sato, T.; Tsuneda, T.; Hirao, K. *J. Chem. Phys.* **2007**, *126*, 234114.

(19) Chiba, M.; Tsuneda, T.; Hirao, K. *J. Chem. Phys.* **2006**, *124*, 144106.

(20) Rohrdanz, M. A.; Herbert, J. M. *J. Chem. Phys.* **2008**, *129*, 034107.

(21) Rohrdanz, M. A.; Martins, K. M.; Herbert, J. M. *J. Chem. Phys.* **2009**, *130*, 054112.

(22) Chai, J.-D.; Head-Gordon, M. *J. Chem. Phys.* **2008**, *128*, 084106.

(23) Chai, J.-D.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615.

(24) Vydrov, O. A.; Heyd, J.; Krukau, A. V.; Scuseria, G. E. *J. Chem. Phys.* **2006**, *125*, 074106.

(25) Vydrov, O. A.; Scuseria, G. E. *J. Chem. Phys.* **2006**, *125*, 234109.

(26) Baer, R.; Neuhauser, D. *Phys. Rev. Lett.* **2005**, *94*, 043002.

(27) Stein, T.; Kronik, L.; Baer, R. *J. Am. Chem. Soc.* **2009**, *131*, 2818.

(28) Görling, A. *Phys. Rev. Lett.* **1999**, *83*, 5459.

(29) Ivanov, S.; Hirata, S.; Bartlett, R. J. *Phys. Rev. Lett.* **1999**, *83*, 5455.

(30) Gimon, T.; A, I.; Heβelmann, A.; Görling, A. *J. Chem. Theory Comput.* **2009**, *5*, 781.

(31) Peach, M. J. G.; Benfield, P.; Helgaker, T.; Tozer, D. J. *J. Chem. Phys.* **2008**, *128*, 044118.

(32) Peach, M. J. G.; Le Sueur, C. R.; Ruud, K.; Guillaume, M.; Tozer, D. J. *Phys. Chem. Chem. Phys.* **2009**, *11*, 4465.

(33) Peach, M. J. G.; Tozer, D. J. *J. Mol. Struct.: Theochem.* **2009**, *914*, 110.

(34) Wiggins, P.; Williams, J. A. G.; Tozer, D. J. *J. Chem. Phys.* **2009**, *131*, 091101.

(35) Dwyer, A. D.; Tozer, D. J. *Phys. Chem. Chem. Phys.* **2010**, *12*, 2816.

(36) Plötner, J.; Dreuw, A. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1197.

(37) Lorenz, L.; Matylitsky, V.; Plötner, J.; Dreuw, A.; Wachtveitl, J. *J. Phys. Chem. A* **2007**, *111*, 10891.

(38) Plötner, J.; Dreuw, A. *Chem. Phys.* **2008**, *347*, 472.

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

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

(41) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372.

(42) Haettig, C.; Weigend, F. *J. Chem. Phys.* **2000**, *113*, 5154.

(43) Kong, J.; et al. *J. Comput. Chem.* **2000**, *21*, 1532.

(44) TURBOMOLE Version 5.7.1.

(45) DALTON, a molecular electronic structure program, Release 2.0; 2005; see http://daltonprogram.org/.

(46) Christiansen, O.; Koch, H.; Jørgensen, P.; Helgaker, T. *Chem. Phys. Lett.* **1996**, *263*, 530.

(47) Grabowski, Z. R.; Rotkiewicz, K.; Rettig, W. *Chem. Rev.* **2003**, *103*, 3899.

(48) Zakharov, M.; Krauss, O.; Nosenko, Y.; Brutschy, B.; Dreuw, A. *J. Am. Chem. Soc.* **2009**, *131*, 461.

(49) Jacquemin, D.; Wathelet, V.; Perpete, E. A.; Adamo, C. *J. Chem. Theory Comput.* **2009**, *5*, 2420.

(50) Caricato, M.; Trucks, G. W.; Frisch, M. J.; Wiberg, K. B. *J. Chem. Theory Comput.* **2010**, *6*, 370.

(51) Silva-Junior, M. R.; Schreiber, M.; Sauer, S. P. A.; Thiel, W. *J. Chem. Phys.* **2008**, *129*, 104103.

CT1001973

JCTC Journal of Chemical Theory and Computation

Efficient Structure Optimization with Second-Order Many-Body Perturbation Theory: The RIJCOSX-MP2 Method

Simone Kossmann[†] and Frank Neese*[,†,‡]

Institut für Physikalische und Theoretische Chemie, Universität Bonn, Wegelerstrasse 12, D-53115 Bonn, Germany and Max-Planck Institut für Bioanorganische Chemie, Stiftstrasse 34−36, D-45470 Mülheim an der Ruhr, Germany

Received April 14, 2010

Abstract: Efficient energy calculations and structure optimizations employing second-order Møller−Plesset perturbation theory (MP2) are presented. The application of the RIJCOSX approximation, which involves different approximations for the formation of the Coulomb- and exchange-type matrices, to MP2 theory is demonstrated. The RIJCOSX approximation incorporates the 'resolution of the identity' approximation in terms of a Split-RI-J variant for the evaluation of the Coulomb matrices and a seminumeric exchange treatment via the 'chain-of-spheres' algorithm for the formation of the exchange-type matrices. Beside the derivation of the working equations, the RIJCOSX-MP2 method is benchmarked against the original MP2 and the already highly efficient RI-MP2 method. Energies as well as gradients are computed employing various basis sets and are compared to the conventional MP2 results concerning accuracy and total wall clock times. Speedups of typically a factor of 5−7 in comparison to MP2 can be obeserved for the largest basis set employed in our study. Total energies are reproduced with an average error of ≤0.8 kcal/mol and minimum energy geometries differ by ∼0.1 pm in bond lengths and typically ∼0.2 degrees in bond angles. The RIJCOSX-MP2 gradient parallelizes with a speedup of 8.2 on 10 processors. The algorithms are implemented into the ORCA electronic structure package.

Introduction

Second-order many-body perturbation theory (MBPT2) with the Møller−Plesset partitioning of the Hamiltonian (MP2)[1] is the simplest ab initio method that accounts for dynamic electron correlation effects. MP2 theory has proven to be a great improvement over Hartree−Fock (HF) theory and is therefore widely used in computational chemistry. Besides its formal simplicity MP2 features, few elementary properties: (a) MP2 is size extensive but only size consistent when based on a size consistent reference wave function, (b) MP2 recovers 80−90% of the basis set correlation energy, (c) MP2 scales only with $\mathcal{O}(N^5)$, and (d) MP2 based on a HF reference wave function only contains double excitation amplitudes.

The shortcomings of MP2 theory are also well-known, i.e., MP2 is not variational or even stationary with respect to wave function parameters, and it does not incorporate any orbital relaxation when based on a HF reference determinant. Orbital relaxation is partially taken into account, if Brillouins' theorem[2] no longer applies, and single excitations contribute to the first-order wave function. The nonvariational character of MP2 is of minor consequence for the calculation of electron correlation effects but complicates and increases the computational effort for the computation of MP2 derivatives.

In the last two decades many attempts have been undertaken in the improvement of the quality and efficiency of MP2 calculations. On the one hand, empirical parameters have been determined to individually scale the parallel and

* Corresponding author. E-mail: theochem@thch.uni-bonn.de.
[†] Universität Bonn.
[‡] Max-Planck Institut für Bioanorganische Chemie.

10.1021/ct100199k © 2010 American Chemical Society
Published on Web 07/02/2010

antiparallel spin components in MP2 theory, e.g., Grimme's 'spin-component scaled' MP2 (SCS-MP2)[3,4] or the simplified approach of Head-Gordon et al. 'scaled opposite-spin' MP2 (SOS-MP2)[5−7] that also leads to reduced computational scaling from $\mathcal{O}(N^5)$ to $\mathcal{O}(N^4)$. On the other hand, the development of double-hybrid functionals has introduced semilocal dynamic correlation effects by adding a perturbative second-order correction in the framework of density functional theory.[8,9]

Substantial progress to improve the efficiency of MP2 calculations was made by Almlöf and Saebo by introducing an integral direct MP2 algorithm to avoid storage of $\mathcal{O}(N^4)$ intermediates.[10] Further modifications to the original algorithm have been reported by Head-Gordon and Pople[11] as well as by Ahlrichs.[12] Probably, the most efficient semidirect ansatz, without avoiding the storage of $\mathcal{O}(N^4)$ quantities on disk, was proposed by Pulay and co-workers,[13−16] who presented very large MP2 calculations with more than 2000 basis functions.[17] A linear scaling integral direct MP2 code based on the Laplace transformation technique introduced by Almlöf,[18] and discussed by various authors,[19−21] has been developed by Ochsenfeld and co-workers.[22−25] Among others Werner, Schütz, and co-workers have developed efficient approximate linear scaling approaches[26−30] which employed the correlation domain concept of Pulay and Saebo.[31−33] The most popular approximation used in combination with the MP2 method is the 'resolution of the identity' (RI) approximation, in which products of orbitals are expanded in an auxiliary basis set.[34] The RI-MP2 method was first reported by Feyereisen et al.[35] and was based on the results of Vahtras, Feyereisen, and Almlöf who showed that the RI technique performs best in the Coulomb metric.[36] The outstanding performance of the RI-MP2 method is indisputable. The speedup of one to two orders of magnitude for large basis sets in comparison to the canonical result with errors in energies usually smaller than 0.1 mE_h/atom demonstrates the impressive efficiency of the RI-MP2 method.

The first derivation and implementation of the MP2 gradient was reported by Pople at the end of the 1970s.[37] Direct and semidirect variants of the MP2 gradient implementation have been developed by Frisch, Head-Gordon, and Pople more than 10 years later.[38,39] Ahlrichs implemented a modification of the semidirect MP2 gradient with reduced disk storage requirements and the exploit of nonabelian point group symmetry in the well-know MPGRAD program that is part of the TuboMole program suite.[12]

Weigend and Häser demonstrated the efficient use of the RI approximation in MP2 gradient calculations.[40] Further refinement was made by Head-Gordon and co-workers who first presented the restricted open-shell MP2 gradient within the RI approximation[41] and later on proposed a more efficient RI-MP2 gradient algorithm by utilizing a semidirect batching approach.[42] Analytic derivatives associated with the recently developed double-hybrid functionals have been reported by Neese, Schwabe, and Grimme,[9] who also demonstrated the efficient application of the RI approximation to the incorporated MP2 correction.

Like any other wave function-based electron correlation theory, the basis set requirements for MP2 calculations are much more stringent than for self-consistent field (SCF) calculations. Hence, in order to obtain results that properly reflect the intrinsic accuracy of the MP2 method rather than basis set incompleteness artifacts, one needs to employ at least a triple-ζ basis with at least two or three sets of polarization functions.[43] However, in order to obtain truly converged results much larger basis sets are necessary. Fortunately, basis set extrapolation techniques are known that allow one to extrapolate to the MP2 basis set limit. However, the preferred level for a reliable extrapolation still involves triple- and quadruple-ζ basis sets. In this case, traditional SCF and MP2 calculations become very expensive and highly time consuming. Thus, it is desirable to search for algorithms that perform efficiently with such extended basis sets for at least medium-sized molecules that are described by 500−2000 basis functions. In this respect, the RI-MP2 method is a great achievement since it performs much better than the standard MP2 algorithms for extended basis sets. In fact, in calculations with extended basis sets the SCF calculation, despite its more favorable computational scaling, usually strongly dominates over the RI-MP2 step in terms of execution time. We have recently shown that great speedups by up to an order of magnitude can be obtained with negligible loss of accuracy, if the SCF step is performed in an approximate way that involves the RI approximation for the Coulomb integrals, while performing a semianalytic integration of the exchange term in the HF equations.[44] This concept can be easily adopted for RI-MP2 energy calculations, since the employed approximations to the HF equations do not enter the familiar calculation of the RI-MP2 energy correction. The seminumeric algorithm used to approximate the exchange term is called 'chain of spheres' (COSX), hence, this algorithm is termed RIJCOSX-MP2. COSX is closely related to Friesners pioneering pseudospectral techniques but is conceptually simpler, as discussed in detail in ref 44. In this manuscript we extend this concept to the first derivatives of the RIJCOSX-MP2 energy. Substantial savings compared to the standard MP2 and the RI-MP2 methods arise from the more efficient SCF step, the more efficient treatment of derivative integrals in the atomic orbital (AO) basis, and the accelerated solution of the z-vector equations.[45] We emphasize that the computational savings immediatly carry over to the case of double-hybrid density functional theory that inherits from RI-MP2 the significant basis set dependence.

Theory

In this section, the theory of the RI-MP2 gradient in the RIJCOSX-MP2 approximation is described in some detail. Throughout this paper, indices i, j, k, ... refer to occupied orbitals in the Hartree−Fock reference determinant, a, b, c, ... to virtual orbitals, and p, q, r, ... to general orbitals from either set. Capital letters denote auxiliary basis functions, whereas K, L, M, and N assign basis functions from the Coulomb-fitting basis sets and R, S, T, and U indicate optimized basis functions for the RI treatment of the correlation correction. Greek letters μ, ν, κ, ... refer to atomic basis functions.

Seminumeric Exchange Matrix. We briefly review the seminumerical exchange treatment that is efficiently implemented in the ORCA program package within the 'chain of spheres' algorithm (COSX).[44] The basic idea for the efficient evaluation of exchange-type matrices is to combine a numeric integration in the physical space with an analytic integration over the Coulomb singularity.

The conventional exchange-type matrix is given as

$$K_{\mu\nu} = \sum_{\kappa\tau} P_{\kappa\tau}(\mu\kappa|\nu\tau) \tag{1}$$

where $P_{\kappa\tau}$ is any density-type matrix and $(\mu\kappa|\nu\tau)$ is a two-electron repulsion integral over atomic basis functions $\{\varphi\}$.

The exchange integrals can be approximated as

$$\begin{aligned}(\mu\kappa|\nu\tau) &= \int d\mathbf{r}'\varphi_\mu(\mathbf{r}')\varphi_\kappa(\mathbf{r}')\int d\mathbf{r}\frac{\varphi_\nu(\mathbf{r})\varphi_\tau(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|} \\ &\approx \sum_g w_g\varphi_\mu(\mathbf{r}_g)\varphi_\kappa(\mathbf{r}_g)\int d\mathbf{r}\frac{\varphi_\nu(\mathbf{r})\varphi_\tau(\mathbf{r})}{|\mathbf{r}-\mathbf{r}_g|}\end{aligned} \tag{2}$$

where the first analytical integration over the coordinates $\mathbf{r}'$ is replaced by a numerical integration over grid points $\mathbf{r}_g$. The corresponding grid weights w_g are determined by Becke's weighting scheme.[46] The exchange-type matrix can then be efficiently evaluated as,

$$K_{\mu\nu} \approx \sum_g X_{\mu g}\sum_\tau A_{\nu\tau}(\mathbf{r}_g)\sum_\kappa X_{\kappa g}P_{\kappa\tau} \tag{3}$$

with

$$X_{\mu g} = w_g^{1/2}\varphi_\mu(\mathbf{r}_g) \tag{4}$$

$$A_{\nu\tau}(\mathbf{r}_g) = \int \frac{\varphi_\nu(\mathbf{r})\varphi_\tau(\mathbf{r})}{|\mathbf{r}-\mathbf{r}_g|}d\mathbf{r} \tag{5}$$

We further define

$$F_{\tau g} = \sum_\kappa P_{\kappa\tau}X_{\kappa g} \tag{6}$$

$$G_{\nu g} = \sum_\tau F_{\tau g}A_{\nu\tau}(\mathbf{r}_g) \tag{7}$$

As mentioned in the Introduction, these equations are closely related to Friesners pseudospectral method. However, there a least-squares fitting operator $Q_{\kappa g}$ is used that is simply replaced by $X_{\kappa g}$ in our formulation. The connection between the COSX and pseudospectral methods is discussed in detail in ref 47.

The exchange energy can be evaluated as

$$E_X = \sum_{\mu\nu} tr(\mathbf{FG}^+)_{\mu\nu} \tag{8}$$

Seminumeric Exchange Gradient. The exchange gradient can be written as

$$\frac{\partial K_{\mu\nu}}{\partial\lambda} = \sum_{\kappa\tau} P_{\kappa\tau}\frac{\partial(\mu\kappa|\nu\tau)}{\partial\lambda} \tag{9}$$

Due to the permutational symmetry of the exchange integrals, the derivative expression can be rearranged as follows

$$\sum_{\kappa\tau} P_{\kappa\tau}(\mu\kappa|\nu\tau^\lambda) \approx \sum_g X_{\mu g}\sum_\tau A_{\nu\tau^\lambda}(\mathbf{r}_g)\sum_\kappa X_{\kappa g}P_{\kappa\tau} \tag{10}$$

$$= \sum_{\kappa\tau} P_{\kappa\tau}(\nu\tau^\lambda|\mu\kappa) \approx \sum_g X_{\nu g}\sum_\kappa A_{\mu\kappa}(\mathbf{r}_g)\sum_\tau X_{\tau^\lambda g}P_{\kappa\tau} \tag{11}$$

where $X_{\tau^\lambda g}$ is defined by

$$X_{\tau^\lambda g} = w_g^{1/2}\frac{\partial\varphi_\tau(\mathbf{r}_g)}{\partial\lambda} \tag{12}$$

The formulation of the exchange gradient differs only slightly from the derivation of the seminumeric exchange itself. Due to the rearrangement of the two-electron exchange gradient integrals, only the derivatives of the basis functions on the grid are needed instead of the derivatives of the analytic integrals.

Density Fitting. The Coulomb matrix can be efficiently approximated by expanding products of basis functions in an extended auxiliary basis set that is usually two to four times larger than the orbital basis set.[36,48,49] The variant of this method that is used in the ORCA program is called 'Split-RI-J' and is described in ref 50.

In terms of the RI approximation, the Coulomb matrix is given as

$$J_{\mu\nu} \approx \sum_K d_K(\mu\nu|K) \tag{13}$$

With the three-index two-electron repulsion integral:

$$(\mu\nu|K) = \int \mu(\mathbf{r}_1)\nu(\mathbf{r}_1)r_{12}^{-1}K(\mathbf{r}_2)d\mathbf{r}_1d\mathbf{r}_2 \tag{14}$$

where $K(\mathbf{r_2})$ is a member of the auxiliary basis set $\{\eta\}$. The vector d_K represents the density in the auxiliary basis which is best obtained from the linear equation system $\mathbf{Vd} = \mathbf{g}$, with $V_{KL} = (K|L)$ and $g_K = \sum_{\mu\nu}P_{\mu\nu}(\mu\nu|K)$. Feyereisen et al.[35] proposed the RI approximation for MP2 energy calculations (RI-MP2) in the mid 1990s. The MP2 energy expression within the RI framework becomes

$$E_{\mathrm{MP2}} = \frac{1}{4}\sum_{ijab}[(ia|jb) - (ib|ja)]t_{ab}^{ij} \tag{15}$$

$$E_{\mathrm{RI-MP2}} = \frac{1}{2}\sum_{iaR}(ia|R)\Gamma'^R_{ai} \tag{16}$$

where we defined the three-index two-particle density as $\Gamma'^R_{ai} = \sum_{jbS}V_{RS}^{-1}(S|bj)t_{ab}^{ij}$. The RI-MP2 double excitation amplitudes t_{ab}^{ij} are given as

$$t_{ab}^{ij} = -\sum_{RS}\frac{(ia|R)V_{RS}^{-1}(S|jb) - (ib|R)V_{RS}^{-1}(S|ja)}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j} \tag{17}$$

RI-MP2 Gradient. The RI-MP2 gradient is efficiently calculated employing the RI approximation.[40] The general RI-MP2 gradient expression written in AO basis is

$$E_2^\lambda = \sum_{\mu\nu} D_{\mu\nu} h_{\mu\nu}^\lambda + \sum_{\mu\nu} W_{\mu\nu} S_{\mu\nu}^\lambda + \sum_{\mu\nu\kappa\tau} \Gamma_{\mu\nu\kappa\tau}^S (\mu\nu|\kappa\tau)^\lambda + \Gamma_{\mu\nu\kappa\tau}^{NS} (\mu\nu|\kappa\tau)_{RI}^\lambda \quad (18)$$

where $D_{\mu\nu}$ denotes the relaxed MP2 density, $W_{\mu\nu}$ is the energy weighted density matrix, and $\Gamma_{\mu\nu\kappa\tau}$ indicates the separable ('S') and nonseparable ('NS') part of the two-particle density matrix. The separable fraction of the two-particle density matrix contains products of the SCF and MP2 density matrices, whereas the nonseparable contribution consists of the MP2 double excitation amplitudes. While $h_{\mu\nu}^\lambda$ denotes the derivative of the one-electron operator and the corresponding basis functions with respect to nuclear displacements $\lambda = x$, y, and z, $S_{\mu\nu}^\lambda$ is the derivative of the overlap matrix, and $(\mu\nu|\kappa\tau)^\lambda$ represents the derivative of the two-electron repulsion integrals with or without the RI approximation.

The relaxed MP2 density matrix is defined as

$$D_{\mu\nu} = \sum_{pq} c_{\mu p} c_{\nu q} D_{pq} \quad (19)$$

$$D_{ij} = -\frac{1}{2} \sum_{kab} t_{ab}^{ik} t_{ab}^{jk} \quad (20)$$

$$D_{ab} = \frac{1}{2} \sum_{ijc} t_{ac}^{ij} t_{bc}^{ij} \quad (21)$$

$$D_{ai} = D_{ia} = \frac{1}{2} z_{ai} \quad (22)$$

where z_{ai} is the solution of the 0th order z-vector equations.[45]

The coupled perturbed self-consistent field (CP-SCF) equations to be solved are $(\varepsilon_a - \varepsilon_i) z_{ai} + R(\mathbf{z})_{ai} = -L_{ai}$, with the MP2 Lagrangian given by

$$L_{ai} = 2 \sum_{cR} (ac|R) \Gamma'^R_{ic} - 2 \sum_{kR} (ki|R) \Gamma'^R_{ka} + R(\mathbf{D})_{ai} \quad (23)$$

The response operator $R(\mathbf{P})_{pq}$ in the AO basis is defined in a general way as

$$R(\mathbf{P})_{\mu\nu} = \sum_{\kappa\tau} P_{\kappa\tau} (2(\mu\nu|\kappa\tau) - (\mu\tau|\kappa\nu) - (\nu\tau|\kappa\mu)) \quad (24)$$

where $\mathbf{P}$ denotes any density-type matrix.

The energy weighted density matrix incorporates all terms that have to be contracted with the derivative of the overlap matrix:

$$W_{\mu\nu} = \sum_{pq} c_{\mu p} c_{\nu q} W_{pq} \quad (25)$$

$$W_{ij} = -\sum_{aR} (ja|R) \Gamma'^R_{ia} - \frac{1}{2} D_{ij} (\varepsilon_i + \varepsilon_j) - \frac{1}{2} R(\mathbf{D})_{ij} \quad (26)$$

$$W_{ab} = -\sum_{iR} (ib|R) \Gamma'^R_{ia} - \frac{1}{2} D_{ab} (\varepsilon_a + \varepsilon_b) \quad (27)$$

$$W_{ai} = -2 \sum_{kR} (ki|R) \Gamma'^R_{ka} - D_{ai} \varepsilon_i \quad (28)$$

Finally, the last two terms in eq 18 contain the derivative of the non-RI and RI two-electron repulsion integrals with respect to nuclear displacements:

$$\Gamma_{\mu\nu\kappa\tau}^S (\mu\nu|\kappa\tau)^\lambda = (D_{\mu\nu} P_{\kappa\tau}^{SCF} - D_{\mu\kappa} P_{\nu\tau}^{SCF}) (\mu\nu|\kappa\tau)^\lambda \quad (29)$$

$$\Gamma_{\mu\nu\kappa\tau}^{NS} (\mu\nu|\kappa\tau)_{RI}^\lambda = \sum_{ijab} c_{\mu i} c_{\nu a} c_{\kappa j} c_{\tau b} t_{ab}^{ij} \times \left[\sum_{RS} \{ (\mu\nu|R)^\lambda V_{RS}^{-1} (S|\kappa\tau) + (\mu\nu|R) V_{RS}^{-1} (S|\kappa\tau)^\lambda \} - \sum_{RSTU} (\mu\nu|T) V_{TR}^{-1} V_{RS}^\lambda V_{SU}^{-1} (U|\kappa\tau) \right] \quad (30)$$

RIJCOSX-MP2 Gradient. The RIJCOSX algorithm combines the seminumeric exchange treatment for all exchange terms with the RI approximation for all Coulomb-type contributions. The total RIJCOSX-MP2 energy is calculated as

$$\begin{aligned} E_{\text{RIJCOSX-MP2}} &= E_{\text{RIJCOSX-HF}} + E_{\text{RI-MP2}} \\ &= \sum_{\mu\nu} P_{\mu\nu} h_{\mu\nu} + \frac{1}{2} \sum_{\mu\nu} P_{\mu\nu} \sum_{\kappa\tau} P_{\kappa\tau} \times \sum_{KL} (\mu\nu|K) V_{KL}^{-1} (L|\kappa\tau) \\ &\quad - \frac{1}{2} \sum_{\mu\nu} P_{\mu\nu} \sum_g X_{\mu g} \sum_\tau A_{\nu\tau}(\mathbf{r}_g) \sum_\kappa X_{\kappa g} P_{\kappa\tau} \\ &\quad + \frac{1}{2} \sum_{\mu\nu R} (\mu\nu|R) \Gamma'^R_{\mu\nu} \end{aligned} \quad (31)$$

Thus, the RIJCOSX-MP2 energy correction corresponds to the original RI-MP2 energy correction, and the unrelaxed MP2 difference density is calculated as in the standard RI-MP2 formulation. The COSX approximation enters into the SCF energy and then into the evaluation of the response-type operators (eq 24) required for the construction of the Lagrangian (eq 23) as well as the internal part of the energy weighted density matrix in eq 26.

For the calculation of the response operator, the Coulomb-type two-electron integral is evaluated employing the RI approximation, whereas both exchange-type integrals are calculated using the COSX algorithm. Thus, the RIJCOSX response-type operator is approximated as

$$\begin{aligned} R(\mathbf{P})_{\mu\nu} \approx & \sum_{\kappa\tau} P_{\kappa\tau} \left(2 \sum_{KL} (\mu\nu|K) V_{KL}^{-1} (L|\kappa\tau) \right) - \\ & \sum_g X_{\mu g} \sum_\kappa A_{\kappa\nu}(\mathbf{r}_g) \sum_\tau X_{\tau g} P_{\kappa\tau} - \sum_g X_{\nu g} \sum_\kappa A_{\kappa\mu}(\mathbf{r}_g) \sum_\tau X_{\tau g} P_{\kappa\tau} \end{aligned} \quad (32)$$

In this way, all exchange- and Coloumb-type two-electron integrals occurring in the Fock and response operators are evaluated in the SCF and CP-SCF procedures.

In the RIJCOSX treatment, the derivatives of the basis functions with respect to nuclear replacements are contracted with the separable part of the two-particle density matrix (eq 29) as follows

$$\begin{aligned}\Gamma^{S}_{\mu\nu\kappa\tau}(\mu\nu|\kappa\tau)^{\lambda} &= D_{\mu\nu}P^{\mathrm{SCF}}_{\kappa\tau}((\mu\nu|\kappa\tau)^{\lambda} - (\mu\kappa|\nu\tau)^{\lambda}) \\ &= D_{\mu\nu}P^{\mathrm{SCF}}_{\kappa\tau}[2\sum_{KL}(\mu\nu|K)^{\lambda}V^{-1}_{KL}(L|\kappa\tau) - \\ &\quad \sum_{KLMN}(\mu\nu|M)V^{-1}_{MK}V^{\lambda}_{KL}V^{-1}_{LN}(N|\kappa\tau)] \\ &\quad + 2\sum_{g}\sum_{\mu\nu}\frac{\partial F^{\mathrm{SCF}}_{\mu g}}{\partial\lambda}G^{\mathrm{MP2}}_{\nu g} \\ &\quad + 2\sum_{g}\sum_{\mu\nu}\frac{\partial F^{\mathrm{MP2}}_{\mu g}}{\partial\lambda}G^{\mathrm{SCF}}_{\nu g}\end{aligned} \tag{33}$$

where the superscripts SCF and MP2 indicate either the SCF density matrix or the MP2 difference density matrix with

$$\frac{\partial F_{\mu g}}{\partial\lambda} = \sum_{\kappa}P_{\mu\kappa}X_{\mu^{\lambda}g} \tag{34}$$

The seminumeric MP2 exchange gradient is implemented, as shown in Figure 1. The first step is to obtain the basis functions and its derivatives on the grid and to scale them with its corresponding weighting factor. Then the 'Batch-Density' is constructed as contraction of the density matrix elements that correspond to nonzero basis function values. $F_{\mu g}$ and the derivatives $F^{\lambda}_{\mu g}$ are calculated with the SCF density matrix as well as with the MP2 density matrix, where λ corresponds to the three Cartesian coordinates. The $F_{\mu g}$ terms are required for the calculation of the exchange energy, see eq 8. After the calculation of the intermediate $G_{\nu g}$, the three components of the gradient can be evaluated as presented in Figure 1. Since we allow for different auxiliary basis sets to approximate the Coulomb-type matrices (/J) and the MP2 exchange operators (/C), all terms that originate from the HF and coupled-perturbed HF equations are calculated with the smaller *J*-auxbasis set, indicated by labels *K*, *L*, *M*, and *N*. These terms are the Coulomb-type terms in the RIJCOSX-HF energy and in the response operator (eq 24) as well as the derivative Coulomb integrals contracted with the separable part of the two-electron density (eq 33). By contrast, the three-index integrals for the evaluation of the RI-MP2 energy correction and the derivative Coulomb integrals that are contracted with the nonseparable density have to be approximated by employing the *C*-auxbasis set, denoted by indices *R*, *S*, *T*, and *U*.

Calculations

Computational Details. All calculations were performed with a development version of the ORCA program package.[51] The split-valence ((def2-)SV),[52] triple-ζ valence ((def2-)TZV),[53] and quadruple-ζ valence ((def2-)QZV)[54] basis sets developed by the Karlsruhe group together with the appropriate polarization functions from the TurboMole library were used throughout.[55] In order to obtain optimum results, different auxiliary basis sets should be used for the approximation of the Coulomb term[56] and the MP2 energy[57−60] (/J and /C in ORCA notation, respectively). Alternatively, the larger Coulomb plus exchange[61] (/JK in ORCA notation) auxiliary basis sets can be used for the entire calculation, although the MP2 energies obtained with these fitting bases are noticeably inferior to the ones obtained with the C-bases. Nevertheless, the noncompleteness errors in the RI approximation are so smooth that the choice of auxiliary basis set does play a minor role for chemical applications, and an unbalanced or erroneous behavior was never observed with any of the tested auxiliary basis sets.

```
Calculate the basis function φ and its derivatives φ^λ on the grid
Scale φ and φ^λ with w_g^½
Loop over number of input densities
  Construct BatchDensity matrix P^SCF and P^MP2
  Calculate F^SCF_μg ⇐ P^SCF ⋆ φ and F^MP2_μg ⇐ P^MP2 ⋆ φ
  Calculate F^SCF,λ_μg ⇐ P^SCF ⋆ φ^λ and F^MP2,λ_μg ⇐ P^MP2 ⋆ φ^λ, λ = x, y, z
End loop over number of input densities
Calculate G^SCF_νg and G^MP2_νg
Calculate GRAD^λ[atom] ⇐ tr(F^SCF,λ G^MP2†_νg) + tr(F^MP2,λ G^SCF†_νg)
```

Figure 1. Pseudocode for the implementation of the separable COSX-MP2 gradient.

Calculations labeled RI-MP2 used the RI[36] approximation only for the calculation of the MP2 correction. In RIJ-DX-MP2 calculations exact analytic integration of the exchange contribution and the RI approximation for the Coulomb terms were employed. The label RIJCOSX-MP2 refers to a RI-MP2 calculation where the COSX approximation[44] in conjunction with Split-RI-J[50] (in the SCF part), and the standard RI approximation (in the MP2 part) was employed. In RIJCOSX-MP2 calculations the default SCF grids[44] were used to obtain the HF energy. For the calculation of the response operator (eq 24) and for the solution of the *z*-vector equations, a Lebedev-50 angular integration grid with an integration accuracy of 3.34 has been determined to be sufficient in extended test calculations. However, for the calculation of the derivatives of the basis functions on the grid (eq 12), a large Lebedev-302 angular grid with a radial integration accuracy of 4.34 is required to prevent the buildup of numerical noise.

Total energies were generally converged to 10^{-8} E_{h}. To benchmark the optimized structures, energy convergence tolerances of 10^{-10} E_h and tight optimization criteria were enforced. The frozen-core approximation[62] was employed throughout, and the time savings compared to the all-electron treatment are ~10% for the RIJCOSX-MP2 treatment, determined on the example of the D-glucose molecule in combination with the TZVPP or QZVP basis set.

Integration Grids. The integration grids employed for the numerical integration within the COSX algorithm were derived from standard density functional theory (DFT) grids included in the ORCA program package but are considerably smaller. The accuracy of the integration grids is controlled by the radial resolution parameter ε, which defines the number of radial shells n_r for a given atom and is calculated as $n_r = 15\varepsilon + 5r - 40$,[63] where r is the row of the periodic table to which the atom belongs. For a radial integration accuracy of $\varepsilon = \{3.34; 4.01; 4.34; 5.01\}$ the number of radial grid points for the first to third row of the periodic table is {(15,20,25); (25,30,35); (30,35,40); (40,45,50)}. The pruning algorithm of Gill et al. is used to reduce the number of points in the inner and outer regions.[64] For comparison, we have chosen the same four representative molecules with the same diverse basis sets, as for the earlier evaluation of the SCF grids.[44]

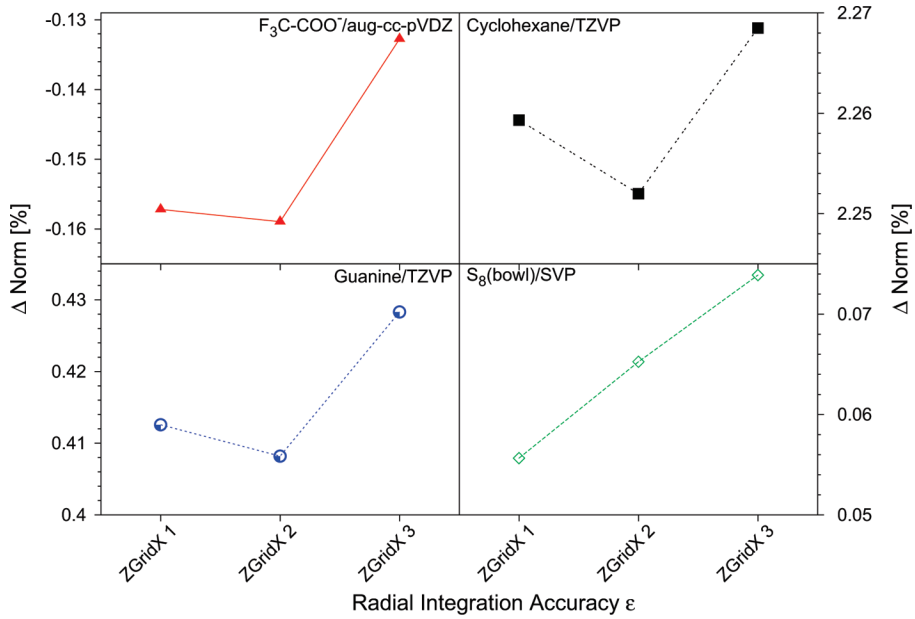


Figure 2. Deviation of the RIJCOSX-MP2 gradient norm (%) depending on the SCF grids with regard to the original MP2 method. ZGridX 1−3 refer to Lebedev-50/3.34, -50/4.01, and -194/4.34.

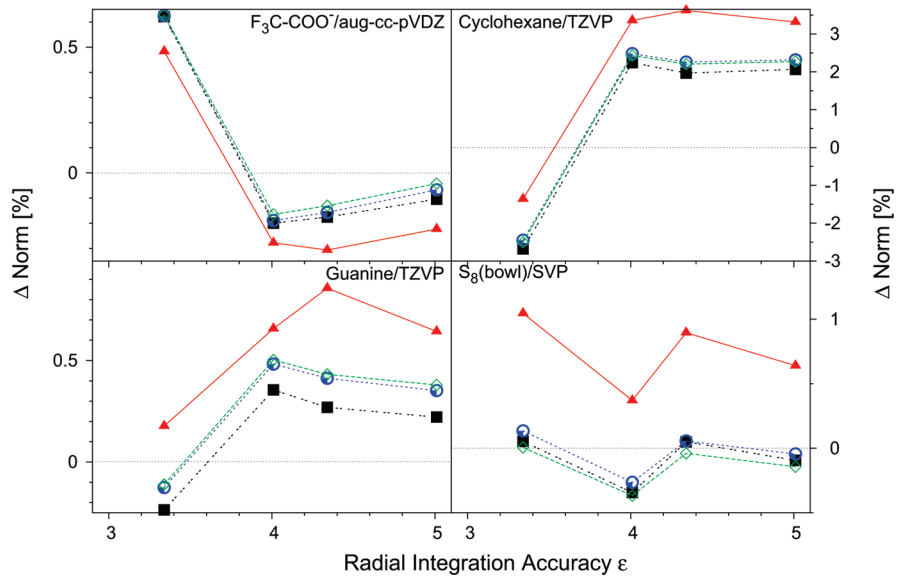


Figure 3. Deviation of the RIJCOSX-MP2 gradient norm (%) depending on the grids used for the evaluation of the basis function derivatives with regard to the original MP2 method. Triangles (▲) = 110 point, filled squares (■) = 195 point, circles (●) =302 point, and empty squares (□) = 434 point Lebedev grids.

The grids used for the calculation of the response-type operators are dependent on the chosen SCF grids, since they originate from the HF equations. Therefore, we tested the influence of the employed integration grids for the solution of the *z*-vector equations on the quality of the calculated gradient norm compared to the original MP2 gradient norm. As shown in Figure 2, the deviations of the gradient norm are very small (<3%), and the very limited deviations follow no general trend. We therefore have considered the smallest SCF grid as being appropriate for the present purpose.

The calculation of the derivative of the basis functions on the grid required for the RIJCOSX-MP2 gradient is independent of the employed SCF grids. Therefore, we tested different angular grids in combination with several radial integration accuracies. In Figure 3 the deviation of the norm calculated with the RIJCOSX approximation compared to the norm of the original MP2 method is presented.

The trend of the norm deviation changes between different molecules but is almost independent of the angular grid. The maximum deviation of the gradient norm is 3.5% for the cyclohexane molecule but is considerably smaller for the other molecules in this test suite (<1%). The curves corresponding to the large 302 and 434 point Lebedev grids nearly coincide, and even the deviations obtained with the smaller 194 point Lebedev grid are still reasonable. However, extended test calculations have shown that the best convergence in geometry optimizations is obtained with the Lebedev-302 radial grid in combination with a radial integration accuracy of 4.43. Hence, we choose this very conservative scheme as our default setting.

Table 1. Overview of the Different Approximations Employed for the SCF and MP2 Parts[a]

	approximations				
	SCF		MP2		
methods	Coulomb	exchange	correction	separable gradient	non-separable gradient
MP2 energy	exact	exact	exact	–	–
MP2 gradient	exact	exact	exact	exact	exact
RI-MP2 energy	exact	exact	RI	–	–
RI-MP2 gradient	exact	exact	RI	exact	RI
RIJ-DX-MP2 energy	RI	exact	RI	–	–
RIJ-DX-MP2 gradient	RI	exact	RI	J-RI/K-exact	RI
RIJCOSX-MP2 energy	Split-RI-J	COSX	RI	–	–
RIJCOSX-MP2 gradient	Split-RI-J	COSX	RI	J-RI/K-COSX	RI

[a] Indices J and K refer to the Coulomb and exchange contributions, respectively, RI indicates the original resolution of the identity approximation, and COSX denotes the seminumeric exchange treatment.

Table 2. Efficiency and Accuracy of the RIJ-DX and RIJCOSX Algorithms for MP2 Energy Calculations on Medium-Size Molecules with Different Basis Sets[a]

				ΔE			speedup SCF		
molecule	N_{atoms}	basis	N_{basis}	RI	RIJ-DX	RIJCOSX	RIJ-DX	RIJCOSX	speedup RI-MP2
adenine	15	TZVPP	150	0.1	−0.1	0.7	1.0	3.9	29.8
		QZVP	220	0.1	−0.1	1.1	1.0	6.9	18.7
beclometahsone	57	TZVPP	1277	0.4	−0.1	0.8	1.1	7.7	20.2
D-glucose	24	TZVPP	540	0.2	−0.2	0.0	1.1	4.5	26.3
		QZVP	1044	0.1	−0.3	0.8	1.1	7.9	21.6
dibenzo-crown-18−6	50	TZVPP	1142	0.3	−0.3	0.6	1.3	6.1	18.5
epinephrine	26	TZVPP	585	0.2	−0.2	0.5	1.1	4.4	27.2
		QZVP	1131	0.1	−0.3	1.3	1.0	7.7	24.2
menthol	31	TZVPP	621	0.2	−0.1	0.2	1.2	5.8	33.5
		QZVP	1227	0.1	−0.2	1.6	1.1	11.4	29.2
tyrosine	24	TZVPP	557	0.2	−0.1	0.6	1.1	4.9	26.0
		QZVP	1071	0.1	−0.2	1.4	1.1	7.4	22.9

[a] Speedup refers to the ratio of wall clock times of the SCF or the RI-MP2 modules (RI-MP2) required to finish one energy calculation. Speedups are given relative to the original MP2 method without any approximations. Deviation of the single point energies to the MP2 reference is denoted ΔE_{RI}, $\Delta E_{RIJ\text{-}DX}$, and $\Delta E_{RIJCOSX}$ in (kcal/mol).

Results

In this section, the RIJCOSX-MP2 method is benchmarked against the conventional MP2 as well as the RI-MP2 method in terms of accuracy and efficiency. In order to not mix up the different approximations employed, an overview of the abbreviations used is given in Table 1. Whenever the original or conventional MP2 method is mentioned, no approximations are involved. For RI-, RIJ-DX-, and RIJCOSX-MP2, the MP2 energy correction is identically calculated applying the RI approximation.

Efficiency and Accuracy of RIJCOSX-MP2 Energy Calculations. In order to demonstrate the efficiency and accuracy of the RIJCOSX approximation for MP2 energy calculations, we have calculated single point energies of seven medium-sized molecules consisting of 15−57 atoms, employing two large basis sets with up to ∼1200 basis functions. The speedups listed in Table 2 are classified in terms of speedups in the SCF calculation and in the evaluation of the MP2 correction. Due to the fact that neither the RIJ-DX nor the RIJCOSX approximation has any influence on the calculation of the RI-MP2 energy correction, the speedups displayed in the last column of Table 2 are identical for RI, RIJ-DX, and RIJCOSX. The speedups for the evaluation of the MP2 correction in terms of the RI approximation range from ∼18.5−33.5, which again demonstrates the high efficiency of the density fitting technique.

Whereas the RIJ-DX treatment represents generally no improvement over the original RI approximation, the benefit in computation time for the RIJCOSX approximation originates from the immense speedup during the SCF calculation. The HF equations in the RIJCOSX framework are solved about ∼4−11 times faster compared to that of the approximation-free RHF method. Since MP2 energy calculations are strongly dominated by the solution of the HF equations, the overall speedup that can be achieved with several approximations corresponds roughly to the speedup obtained in the SCF iterations (∼1−2 and ∼5−14 for RIJ-DX and RIJCOSX, respectively). The speedups in the SCF module are considerably lower than the speedups reported in ref 44. This is due to the fact that the integral generation package of ORCA was recently replaced with the LIBINT package designed by Valeev and Ferman that is inherently more efficient, in particular for high-angular momentum basis functions.[65] Hence, the favorable scaling of the COSX approximation with respect to higher angular momenta becomes less pronounced than previously reported. In general, the speedups obtained with the RIJCOSX approximation during the SCF iterations decrease by roughly

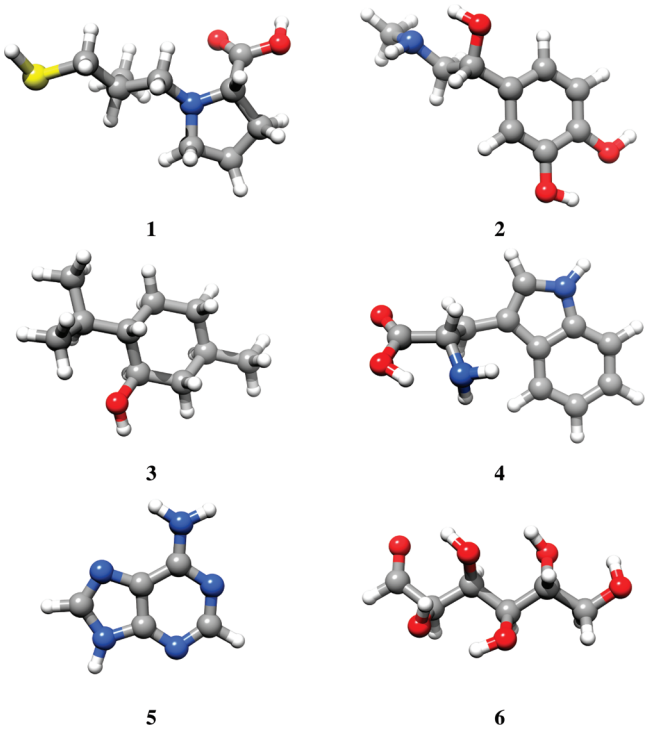


Figure 4. Systems used for testing the accuracy of geometry optimizations (**1** = captopril, **2** = epinephrine, **3** = menthol, **4** = tryptophane, **5** = adenine, and **6** = D-glucose).

a factor of 4 for the TZVPP basis set and ~6 for the QZVP basis set (cf. speedup for the menthol molecule in QZVP basis: with LIBINT package −11.6, without LIBINT package −61.2).

The speedups listed in Table 2 concerning the MP2 energy correction obtained with the RI approximation compared to the conventional MP2 correction generally decrease with larger basis sets. This is caused by the higher angular momenta in the auxiliary basis set and the associated generation of the three-index integrals in the atomic orbital basis. Replacing the QZVP/C auxbasis by the smaller TZVPP/C auxbasis set, i.e., reducing the highest angular momentum from *h* to *g*, yields roughly a speedup of ~35 in the RI-MP2 module compared to that of the conventional MP2 for the adenine molecule in a QZVP orbital basis.

The deviation in total single point energies introduced by the different approximations is on the same order of magnitude with absolute errors amounting to 0.1−0.4 kcal/mol with RI and RIJ-DX. The error for the RIJ-DX approximation is overall negative, which results from the well-known variational nature of Coulomb energy by the RI approximation.[66,67,48] The error in total single point energies introduced by the RIJCOSX approximation is about one order of magnitude larger than what is obtained with RIJ-DX and is overall positive. The errors are larger for the extended QZVP basis set, which has also been reported for the RIJCOSX-HF energy in ref 44. These errors can be reduced by applying a larger Lebedev final grid.

Accuracy of the RIJCOSX-MP2 Gradient. In order to prove the accuracy of the geometries predicted with the RI, RIJ-DX, and RIJCOSX approximations in restricted MP2 (RMP2) calculations, we performed structure optimizations on some representative medium-sized molecules (Figure 4).

In Table 3, deviations of the structural parameters relative to the original MP2 method are collected. The RIJ-DX approximation introduces no noticeable error compared to the original RI-MP2 method. The deviations introduced by the RI approximation are on average ≤0.01 pm in bond distances and 0.005° in bond angles. The maximum absolute errors are 0.03 pm in bond lengths and 0.02° in bond angles. The RIJCOSX approximation introduces deviations, which are about one order of magnitude larger. Bond distances deviate by about 0.1 pm from their parent values, whereas bond angles differ typically by less than 0.2°. The largest

Table 3. Statistical Analysis of Errors in the Optimized Geometries of Molecules 1−6 from the RI, RIJ-DX, and RIJCOSX Approximations Compared to the Parent RMP2 Method[a]

		|Δ mean distances|	|Δ mean absolute distances|	|Δ max distances|	|Δ mean angles|	|Δ mean absolute angles|	|Δ max angles|
1	RI-MP2	−0.011	0.011	0.020	−0.001	0.003	0.020
	RIJ-DX-MP2	−0.013	0.013	0.020	−0.001	0.004	0.020
	RIJCOSX-MP2	−0.001	0.102	0.400	−0.005	0.145	0.400
2	RI-MP2	−0.007	0.007	0.020	0.001	0.002	0.020
	RIJ-DX-MP2	−0.008	0.008	0.030	0.000	0.002	0.010
	RIJCOSX-MP2	−0.031	0.075	0.170	−0.014	0.139	0.380
3	RI-MP2	−0.012	0.012	0.020	0.000	0.001	0.010
	RIJ-DX-MP2	−0.013	0.013	0.030	0.000	0.002	0.010
	RIJCOSX-MP2	−0.036	0.077	0.240	−0.001	0.133	0.400
4	RI-MP2	−0.004	0.004	0.020	0.000	0.002	0.010
	RIJ-DX-MP2	−0.005	0.006	0.020	0.000	0.005	0.010
	RIJCOSX-MP2	−0.001	0.075	0.220	0.001	0.164	0.730
5	RI-MP2	−0.004	0.004	0.020	0.003	0.004	0.020
	RIJ-DX-MP2	−0.004	0.004	0.020	0.003	0.004	0.020
	RIJCOSX-MP2	−0.031	0.044	0.130	0.015	0.073	0.280
6	RI-MP2	−0.006	0.006	0.020	0.000	0.002	0.010
	RIJ-DX-MP2	−0.007	0.007	0.020	0.000	0.002	0.010
	RIJCOSX-MP2	−0.011	0.086	0.310	0.004	0.106	0.360

[a] Errors are given relative to the parent MP2 method. All calculations were done with the TZVP basis set (distances in pm and angles in degrees). *N*(distances) for molecules 1−6 = 30, 26, 31, 38, 16, and 23, *N*(angles) for molecules 1−6 = 56, 42, 61, 46, 24, and 28. Errors in final total energies are for RI = 0.7, 0.3, 0.9, 0.3, 0.1, and 0.2 kcal/mol, RIJ-DX = 0.5, 0.2, 0.7, 0.1, − 0.1, and 0.1 kcal/mol, and RIJCOSX = 0.6, 0.9, 1.6, 0.8, 0.4, and 0.4 kcal/mol.

Table 4. Efficiency of the RIJ-DX and RIJCOSX Algorithms for MP2 Calculations on Medium-Sized Molecules with Different Basis Sets[a]

						speedup		
molecule	N_{atoms}	basis	N_{basis}	$\Delta_{RIJ\text{-}DX}$	$\Delta_{RIJCOSX}$	RI	RIJ-DX	RIJCOSX
adenine	15	SV(P)	150	−0.0001	−0.0009	1.3	1.3	1.3
		TZVP	220	−0.0001	−0.0004	1.4	1.4	2.3
		TZVPP	380	0.0001	−0.0009	1.3	1.3	4.1
cysteine	14	SV(P)	116	−0.0002	−0.0010	1.2	1.2	0.8
		TZVP	178	−0.0002	−0.0013	1.2	1.4	1.7
		TZVPP	318	0.0000	−0.0011	1.4	1.4	3.7
cytosine	13	SV(P)	122	−0.0001	−0.0017	1.2	1.2	1.1
		TZVP	182	−0.0002	−0.0015	1.2	1.4	1.8
		TZVPP	318	0.0001	−0.0016	1.2	1.3	3.6
D-glucose	24	SV(P)	192	0.0004	−0.0001	1.5	1.5	1.2
		TZVP	300	0.0004	0.0002	1.6	1.8	2.8
		TZVPP	540	−0.0002	−0.0002	1.6	2.0	5.6
epinephrine	26	SV(P)	208	0.0000	0.0000	1.4	1.4	1.3
		TZVP	325	0.0002	0.0014	1.6	1.9	3.0
		TZVPP	585	0.0000	0.0011	1.7	1.9	5.6
ferrocene	21	SV(P)	184	−0.0002	−0.0008	1.2	1.3	1.6
		TZVP	283	−0.0005	−0.0012	1.3	1.3	3.0
		TZVPP	490	0.0002	−0.0009	1.3	1.7	5.3
glycine	10	SV(P)	80	−0.0002	−0.0007	1.2	1.0	0.7
		TZVP	125	−0.0003	−0.0002	1.2	1.1	1.1
		TZVPP	225	0.0000	0.0000	1.3	1.2	2.8
guanine	16	SV(P)	164	0.0000	−0.0015	1.4	1.3	1.5
		TZVP	239	0.0000	−0.0017	1.4	1.5	2.5
		TZVPP	411	0.0001	−0.0014	1.3	1.4	4.4
histidine	20	SV(P)	172	−0.0003	−0.0003	1.4	1.3	1.2
		TZVP	263	−0.0003	−0.0002	1.5	1.7	2.6
		TZVPP	467	0.0000	0.0001	1.6	1.8	5.1
menthol	31	SV(P)	194	0.0002	0.0003	1.3	1.6	1.0
		TZVP	329	0.0003	0.0000	1.6	1.8	2.9
		TZVPP	621	0.0002	−0.0001	1.9	2.4	7.5
nitroglycerine	17	SV(P)	178	0.0001	−0.0005	1.4	1.4	1.3
		TZVP	258	0.0001	−0.0002	1.5	1.7	2.9
		TZVPP	442	0.0001	−0.0004	1.5	1.7	5.2
thymine	15	SV(P)	138	0.0000	−0.0008	1.2	1.1	1.0
		TZVP	207	0.0000	−0.0011	1.4	1.4	2.0
		TZVPP	363	0.0001	−0.0012	1.4	1.5	4.3
tyrosine	24	SV(P)	204	−0.0003	−0.0006	1.6	1.6	1.3
		TZVP	313	−0.0002	−0.0006	1.5	1.7	3.0
		TZVPP	557	0.0000	−0.0005	1.6	1.9	5.8

[a] Speedup refers to the ratio of total wall clock times required to finish one gradient calculation. The speedups are given relative to the original MP2 method without any approximations. The deviation of the gradient norm w.r.t. the MP2 reference is denoted $\Delta_{RIJ\text{-}DX}$ and $\Delta_{RIJCOSX}$ in (E_h /bohr).

error in bond lengths is up to 0.4 pm, the corresponding error in bond angles is 0.7°. The mean absolute deviations for the RIJCOSX-MP2 method are in the same range observed previously for the RIJCOSX-HF method.[44] The deviations of the final total energies from the canonical values remain below 2 kcal/mol and are significantly smaller when compared to the RI-MP2 method (<1 kcal/mol).

Efficiency of the RIJCOSX-MP2 Method. In Tables 4 and 5 the efficiency of the RIJ-DX-MP2 and RIJCOSX-MP2 methods is demonstrated compared to the original MP2 as well as to the already very efficient RI-MP2 method.

The speedups listed in Table 4 show the efficiency of the RI, RIJ-DX, and RIJCOSX approximations for 13 medium-sized molecules for 3 different basis sets with up to ∼620 basis functions. The RIJ-DX approximation shows speedups up to a factor of 2.4 in the gradient calculations but represents no major improvement over the original RI approximation. In general, the speedup becomes noticeable for ∼200 basis functions. The speedups tend to become more remarkable with larger basis sets, but the enhancements are almost not observable.

The RIJCOSX approximation shows only minor speedups for the smallest employed basis set SV(P). For the smallest molecules in our test suite, the RIJCOSX approximation even slows down the MP2 gradient calculations. For the next larger basis set TZVP, the RIJCOSX approximation accelerates the calculations by a factor of ∼3 for molecules with around 300 basis functions. The speedup increases noticeably with basis set and system size. In the case of the menthol molecule, a factor of ∼7.5 in computation time is gained, if the RIJCOSX approximation is applied. For comparison, the speedup reduces to ∼4 when RI-MP2 serves as reference.

The deviation of the gradient norm shows a deviation for RIJCOSX-MP2, which is naturally somewhat larger than for the RIJ-DX approximation but as has been shown in the previous subsection, this has almost no influence on the resulting structures.

Table 5. Efficiency of the RIJ-DX and RIJCOSX Algorithms for RI-MP2 Calculations on Medium-Sized Molecules with Different Basis Sets.[a]

						speedup (RIJ-DX)		speedup (RIJCOSX)	
molecule	N_{atoms}	basis	N_{basis}	Δ_{RIJ-DX}	$\Delta_{RIJCOSX}$	RI-MP2	total	RI-MP2	total
adenine	15	SV(P)	150	0.0000	−0.0009	1.0	1.0	1.0	1.0
		TZVP	220	0.0000	−0.0004	1.1	1.0	1.5	1.7
		TZVPP	380	0.0001	−0.0009	1.0	1.0	2.9	3.1
		QZVP	720	0.0001	−0.0010	1.1	1.1	5.9	6.3
beclometahsone	57	SV(P)	454	0.0004	0.0007	1.2	1.2	1.1	1.3
		TZVP	709	0.0003	0.0004	1.3	1.4	2.0	2.5
		TZVPP	1277	0.0000	0.0001	1.2	1.3	2.4	3.3
captopril	30	SV(P)	220	0.0002	0.0008	1.1	1.1	0.8	0.9
		TZVP	352	0.0001	0.0005	1.3	1.3	1.6	2.0
		TZVPP	644	0.0002	0.0006	1.2	1.2	2.9	3.4
		QZVP	1264	0.0002	0.0005	1.4	1.4	4.9	6.2
cholesterole	74	SV(P)	484	0.0004	0.0001	1.1	1.1	1.0	1.1
		TZVP	808	0.0003	−0.0004	1.4	1.4	1.7	2.2
		TZVPP	1512	0.0002	−0.0001	1.4	1.4	2.1	2.9
CO-heme	86	SV(P)	782	0.0000	−0.0001	1.2	1.2	1.1	1.3
		TZVP	1180	0.0000	−0.0002	1.2	1.3	1.6	2.2
cysteine	14	SV(P)	116	−0.0002	−0.0009	1.0	1.0	0.6	0.7
		TZVP	178	−0.0002	−0.0013	1.2	1.2	1.2	1.3
		TZVPP	318	0.0000	−0.0011	1.0	1.0	2.6	2.7
		QZVP	622	0.0000	−0.0010	1.2	1.2	5.1	5.5
cytosine	13	SV(P)	122	−0.0001	−0.0016	1.0	1.0	0.8	0.9
		TZVP	182	−0.0002	−0.0014	1.1	1.2	1.4	1.5
		TZVPP	318	0.0001	−0.0016	1.0	1.0	2.7	2.9
		QZVP	606	0.0001	−0.0016	1.1	1.1	5.6	6.0
D-glucose	24	SV(P)	192	0.0003	−0.0002	1.0	1.0	0.8	0.8
		TZVP	300	0.0003	0.0001	1.1	1.1	1.4	1.7
		TZVPP	540	−0.0002	−0.0002	1.3	1.3	3.1	3.6
		QZVP	1044	0.0000	−0.0002	1.4	1.3	4.5	5.5
dibenzo-crown-18-6	50	SV(P)	412	−0.0001	0.0000	1.2	1.2	1.0	1.1
		TZVP	638	0.0003	0.0001	1.3	1.3	1.6	2.0
		TZVPP	1142	0.0002	0.0000	1.3	1.3	2.1	2.7
epinephrine	26	SV(P)	208	−0.0001	0.0000	1.1	1.0	0.9	0.9
		TZVP	325	0.0002	0.0014	1.2	1.2	1.5	1.8
		TZVPP	585	0.0000	0.0011	1.1	1.1	2.7	3.4
		QZVP	1131	0.0000	0.0008	1.4	1.4	5.2	6.7
ferrocene	21	SV(P)	184	−0.0001	−0.0007	1.1	1.1	1.4	1.3
		TZVP	283	−0.0005	−0.0012	1.1	1.0	2.0	2.4
		TZVPP	490	0.0003	−0.0009	1.0	1.3	3.5	4.0
		QZVP	954	0.0003	−0.0020	1.0	1.0	4.5	5.7
flutamide	30	SV(P)	288	0.0004	0.0016	1.1	1.1	1.0	1.1
		TZVP	427	0.0001	0.0014	1.2	1.2	1.6	2.0
		TZVPP	743	0.0000	0.0013	1.1	1.1	2.5	3.0
		QZVP	1413	0.0000	0.0013	1.4	1.4	4.5	5.8
glycine	10	SV(P)	80	−0.0002	−0.0007	0.8	0.9	0.5	0.6
		TZVP	125	−0.0002	−0.0002	1.0	0.9	0.9	0.9
		TZVPP	225	0.0000	0.0000	1.0	0.9	2.2	2.2
		QZVP	435	0.0000	−0.0003	1.1	1.0	4.6	4.7
guanine	16	SV(P)	164	0.0001	−0.0013	0.9	0.9	1.0	1.1
		TZVP	239	0.0001	−0.0015	1.1	1.1	1.6	1.8
		TZVPP	411	0.0001	−0.0013	1.0	1.0	3.1	3.3
		QZVP	777	0.0001	−0.0017	1.2	1.2	5.9	6.6
histidine	20	SV(P)	172	−0.0002	−0.0002	0.9	0.9	0.8	0.9
		TZVP	263	−0.0002	−0.0001	1.1	1.1	1.4	1.7
		TZVPP	467	0.0001	0.0001	1.2	1.1	2.8	3.2
		QZVP	897	0.0000	0.0001	1.3	1.2	5.3	6.2
menthol	31	SV(P)	194	0.0001	0.0002	1.2	1.2	0.6	0.7
		TZVP	329	0.0003	0.0000	1.3	1.1	1.5	1.8
		TZVPP	621	0.0001	−0.0002	1.3	1.2	3.2	3.9
		QZVP	1227	0.0000	0.0001	1.4	1.4	5.3	7.3
morphin	40	SV(P)	332	0.0001	0.0002	1.2	1.2	1.3	1.4
		TZVP	513	0.0001	0.0005	1.3	1.2	2.3	2.9
		TZVPP	917	0.0001	0.0008	1.3	1.3	3.1	4.2
		QZVP	1767	0.0001	0.0004	1.3	1.3	5.1	7.0
NiTrenNCS2	35	SV(P)	292	0.0000	0.0006	1.1	1.1	1.3	1.5
		TZVP	451	0.0001	0.0005	1.1	1.4	1.7	2.6
		TZVPP	794	0.0000	0.0002	1.3	1.3	2.9	4.0
		QZVP	1562	0.0000	0.0005	1.4	1.4	5.1	9.2
nitroglycerine	17	SV(P)	178	0.0001	−0.0005	1.1	1.0	0.9	0.9
		TZVP	258	0.0001	−0.0002	1.1	1.1	1.7	1.9

Table 5 Continued

molecule	N_{atoms}	basis	N_{basis}	Δ_{RIJ-DX}	$\Delta_{RIJCOSX}$	speedup (RIJ-DX) RI-MP2	speedup (RIJ-DX) total	speedup (RIJCOSX) RI-MP2	speedup (RIJCOSX) total
		TZVPP	442	0.0001	−0.0004	1.1	1.1	3.0	3.4
		QZVP	834	0.0001	−0.0003	1.3	1.2	4.8	5.6
penicilin	42	SV(P)	376	0.0000	0.0007	1.1	1.1	1.1	1.2
		TZVP	567	0.0001	−0.0002	1.2	1.2	1.6	1.9
		TZVPP	999	0.0001	0.0010	1.3	1.3	2.3	2.9
		QZVP	1921	0.0001	0.0007	1.5	1.5	3.6	4.7
tetracycline	56	SV(P)	496	0.0001	0.0019	1.2	1.2	1.2	1.4
		TZVP	752	0.0001	0.0022	1.2	1.3	1.8	2.2
		TZVPP	1328	0.0000	0.0019	1.2	1.3	2.2	2.9
thymine	15	SV(P)	138	0.0000	−0.0007	0.9	0.9	0.8	0.8
		TZVP	207	0.0001	−0.0011	1.0	1.0	1.3	1.5
		TZVPP	363	0.0001	−0.0012	1.1	1.1	2.2	2.9
		QZVP	693	0.0001	−0.0014	1.1	1.1	5.3	5.8
tryptophane	27	SV(P)	234	−0.0002	0.0000	1.1	1.1	1.0	1.1
		TZVP	357	−0.0002	−0.0001	1.1	1.1	1.6	2.0
		TZVPP	633	0.0001	0.0001	1.2	1.2	2.9	3.6
		QZVP	1215	0.0001	0.0000	1.4	1.4	5.2	6.7
tyrosine	24	SV(P)	204	−0.0003	−0.0006	1.0	1.0	0.8	0.9
		TZVP	313	−0.0002	−0.0006	1.1	1.1	1.7	1.9
		TZVPP	557	0.0000	−0.0004	1.1	1.1	3.1	3.6
		QZVP	1071	0.0000	−0.0003	1.3	1.3	4.8	5.9

[a] Speedup refers to the ratio of wall clock times of either the RI-MP2 module (RI-MP2) or the total wall clock time (total) required to finish one gradient calculation. The deviation of the gradient norm w.r.t. the RI-MP2 reference is denoted Δ_{RIJ-DX} and $\Delta_{RIJCOSX}$ in (E_h /bohr).

In Table 5 the speedup of the RIJ-DX and RIJCOSX approximations to the MP2 gradient is related to the very efficient RI-MP2 gradient. Therefore, we studied 24 molecules in the range of 10−86 atoms with ∼100−1900 basis functions. The speedup obtained with the RIJ-DX-MP2 method compared to RI-MP2 is about a factor of ∼1.2. The inclusion of the SCF wall clock time has almost no consequence on the overall speedup.

For the smallest basis set in our test set SV(P), the speedup for the RIJCOSX approximation in the MP2 module is almost negligible, and the inclusion of the SCF wall clock times has no influence on the entire performance. The speedup of the RIJCOSX approximation becomes more observable for the triple-ζ basis set TZVP and for molecules with ≥300 basis functions (∼1.5−2.3).

For the TZVPP basis set, the speedup varies from a factor of 2−3 for the MP2 module, and even reaches a factor of 4.2 for the morphin molecule when the SCF calculation is taken into account. Excellent speedups are obtained with the more extended QZVP basis set. Speedups of a factor ∼4.5−6 are obtained for the MP2 module, and even a factor of 9.2 is achieved for the NiTrenNCS2 molecule when the SCF wall clock time is included in the comparison. This behavior has to be expected due to the favorable scaling of the RIJCOSX approximation with respect to higher angular momenta.

In order to verify the quality of the calculated gradients, we also listed the errors in the norm of the MP2 gradient in Table 5.

Timing Analysis of the RIJCOSX-MP2 Method. A detailed analysis of the timings for the different parts of the MP2 gradient is presented in this subsection. For the comparison of the original MP2 with the RIJCOSX-MP2 method, the menthol molecule in a TZVPP basis serves as an example. The calculation of one MP2 gradient takes about ∼18 h. The time spent in the MP2 module is ∼14 h (80% of the total wall clock time).

Table 6. Timings in (s) of the Individual Contributions to the MP2 Gradient[a]

gradient components	MP2	RI-MP2	RIJ-DX-MP2	RIJCOSX-MP2
integral transformation	18 061	133	129	136
K^{ij}(a,b)	65	145	144	148
T^{ij}(a,b)	27	17	16	17
D(virtual)	60	61	60	60
D(internal)	1537	124	117	122
W(virtual)	120	121	119	121
W(internal)	191	217	194	201
L(3-ext)	1510	260	250	257
L(3-int)	16	6	6	6
R(*D*)	1170	1209	949	146
R(Z)	1209	1213	969	148
Z-vector solution	5410	5497	4062	773
separable gradient (S)		10918	8937	3733
nonseparable gradient (NS)		582	325	327
gradient (S + NS)	20 290	11 500	9262	4060
total time	49 689	20 732	16 533	6460

[a] Depending on the employed approximation obtained for the menthol molecule (31 atoms) in a TZVPP basis set.

A detailed analysis of the different components of the MP2 gradient is given in Table 6. The integral transformation in the original MP2 method corresponds to a three-fourths transformation of AO integrals to the molecular orbital (MO) basis, and the calculation of separable and nonseparable gradient contributions is the most time-consuming step. The evaluation of the internal part of the unrelaxed density matrix **D**(internal) is more expensive than in the RI-MP2 variants due to the generation of the required integrals containing one frozen-core index on the fly. The formation of the three-external part of the Lagrangian L(3-ext), eq 23 is also more costly in terms of computation time due to the large AO↔MO transformations. The calculation of the nonseparable and separable gradient, eq 18, is done simultaneously

in the original MP2 code, and the time-determining step is the calculation of the derivative of the two-electron repulsion integrals.

The situation changes drastically when going to the RI-MP2 variants. The time required for the integral transformation is negligible, since in RI-MP2 methods only three-index integrals are needed, which can be calculated very efficiently. The formation of the exchange integrals becomes somewhat more expensive due to the larger matrix multiplications, but the effect on the overall timing is insignificant. The evaluation of **D**(internal) is more effcient due to prestored frozen-core integrals, and the three-external part of the Lagrangian only requires the transformation of one virtual index. The calculation of the separable gradient is done analogously to the parent MP2 method, but the nonseparable part is very efficiently calculated employing three- and two-index derivative integrals, eq 30.

The speedup achieved with the RIJ-DX approximation results from the more efficient calculation of the response-type operators and, as a consequence, the accelerated solution of the *z*-vector equations. The separable gradient profits from the application of the RI approximation to the Coulomb contribution, whereas the nonseparable gradient saves computation time through storage of the three-index/two-particle density and the use of large BLAS level two operations.

The RIJCOSX approximation introduces a highly efficient treatment of the exchange-type contributions to the MP2 gradient. The calculation of the response-type operators, eq 32, and the solution of the *z*-vector equations is highly dominated by the formation of the exchange-type contributions. The employment of the RIJCOSX approximation speeds the calculation of these contributions up by a factor of ∼7−8. A speedup of ∼2 is gained by exploiting the efficient calculation of the basis function derivatives on the grid, as ingredient to the separable gradient.

Parallelization. The separable COSX-MP2 gradient, eq 29, is parallelized over batches of grid points. A reasonable load balancing is achieved by distributing the batches of grid points over the processors such that all processors work on all parts of the molecular system to nearly the same extent. A single gather operation is then performed at the end of the COSX integration loop, in order to keep the communication overhead negligible. The same strategy has been followed for the parallel exchange treatment in the SCF module.

In Figure 5 the efficiency of the parallel implementation is demonstrated for a medium-sized molecule (dibenzo-crown-18-6/QZVP, 50 atoms, 2202 basis functions). The scaling behavior is excellent for up to 16 processors, although the example is not really large. However, the evaluation of the basis function derivatives on the grid requires the large Lebedev-302 angular grid with an integration accuracy of 4.34, which yields about 185 000 grid points that are divided among the processors. Some of the plotted speedups in Figure 5 seem to be better than the linear speedup, which is mainly due to the inaccuracy in the time measurement, since no statistics have been performed.

The parallelization of the entire MP2 gradient module, presented in Figure 6, is less efficient than that for the separable COSX-MP2 gradient contribution. However, for up to at least 10 processors reasonable speedups are achieved. The reason for the apparent stagnation of the speedup with 12 processors is the poor scaling of the three-external part to the MP2 Lagrangian and of the separable gradient calculated in the RI approximation with the number of processors employed. The speedup of the evaluation of the three-external Lagrangian contribution achieves a factor 4.4 for 10 processors and decreases to a factor of 3.8 for 12 and 16 processors. A similar situation is found for the separable gradient calculated within the RI approximation. Although not rate determining, the speedup amounts to only a factor of ∼3. The scaling of the internal part of the energy weighted density matrix is also reduced to a factor of 10 for 16 processors. All three contributions together determine 30% of the total wall clock time needed for the calculation on 16 processors. Furthermore, some smaller parts of the parallel MP2 gradient do not scale ideally with the number of used processors. Taken together, these factors are responsible for the worsened scaling starting with 10−12 employed processors.

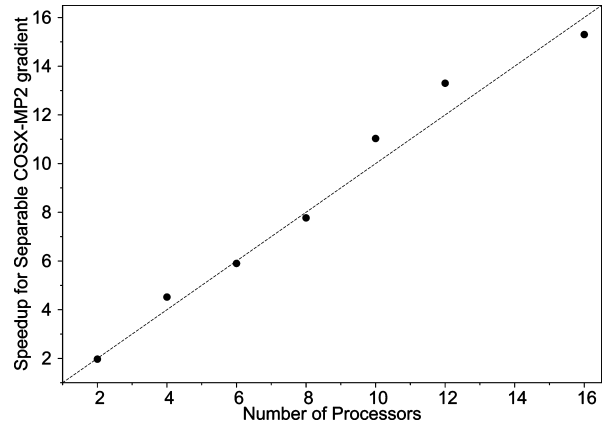


Figure 5. Parallelization efficiency of the COSX gradient algorithm with dibenzo-crown-18-6/QZVP as an example. Plotted is the wall clock time required for formation of the entire separable COSX-MP2 gradient relative to the time taken for the same operation by a single process.

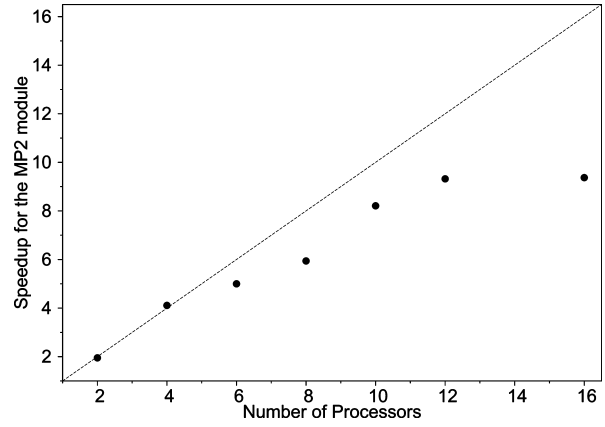


Figure 6. Parallelization efficiency of the MP2 gradient algorithm with dibenzo-crown-18-6/QZVP as an example. Plotted is the wall clock time required for formation of the entire COSX-MP2 gradient relative to the time taken for the same operation by a single process.

Discussion

In the present work, the applicability of the seminumeric exchange treatment via the COSX approximation to the exchange-type contributions of the MP2 gradient has been reported. The entire RIJCOSX-MP2 gradient algorithm fulfills all important requirements required by a reliable approximate method. The optimized structures show only negligible deviations from the MP2 structures. The RIJ-COSX-MP2 gradient is efficient with observed speedups in wall clock times of ∼7−7.5 for the TZVPP basis set compared to conventional MP2 regarding one gradient calculation (including the SCF time). RIJCOSX-MP2 versus RI-MP2 yields also speedups of ∼5−9 for one entire energy and gradient calculation with the QZVP basis.

The RIJCOSX-MP2 method is generally applicable and does neither depend on the molecular system nor on the chosen basis set. The program is parallelized and scales reasonably well with the number of used processors, but some more work is required before efficient large-scale parallelization on many processors is achieved. Overall, we believe that the present development adds to the applicability of MP2-type methods, such as MP2, SCS-MP2, or double hybrid density functional theory in large-scale chemical applications. The method is implemented in the freely available ORCA program.

References

(1) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618–622.

(2) Brillouin, L. *Actual. Sci. Ind.* **1934**, *71*, 159.

(3) Grimme, S. *J. Chem. Phys.* **2003**, *118*, 9095–9102.

(4) Gerenkamp, M.; Grimme, S. *Chem. Phys. Lett.* **2004**, *392*, 229–235.

(5) Jung, Y. S.; Lochan, R. C.; Dutoi, A. D.; Head-Gordon, M. *J. Chem. Phys.* **2004**, *121*, 9793–9802.

(6) Lochan, R. C.; Head-Gordon, M. *J. Chem. Phys.* **2007**, *126*, 164101.

(7) Jung, Y. S.; Shao, Y.; Head-Gordon, M. *J. Comput. Chem.* **2007**, *28*, 1953–1964.

(8) Grimme, S. *J. Chem. Phys.* **2006**, *124*, 034108.

(9) Neese, F.; Schwabe, T.; Grimme, S. *J. Chem. Phys.* **2007**, *126*, 124115.

(10) Saebø, S.; Almlöf, J. *Chem. Phys. Lett.* **1989**, *154*, 83–89.

(11) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. *Chem. Phys. Lett.* **1988**, *153*, 503–506.

(12) Haase, F.; Ahlrichs, R. *J. Comput. Chem.* **1993**, *14*, 907–912.

(13) Saebø, S.; Pulay, P. *J. Chem. Phys.* **2001**, *115*, 3975–3983.

(14) Baker, J.; Pulay, P. *J. Comput. Chem.* **2002**, *23*, 1150–1156.

(15) Saebø, S.; Baker, J.; Wolinski, K.; Pulay, P. *J. Chem. Phys.* **2004**, *120*, 11423–11431.

(16) Ishimura, K.; Pulay, P.; Nagase, S. *J. Comput. Chem.* **2006**, *27*, 407–413.

(17) Baker, J.; Wolinski, K.; Malagoli, M.; Kinghorn, D.; Wolinski, P.; Magyarfalvi, G.; Saebø, S.; Janowski, T.; Pulay, P. *J. Comput. Chem.* **2008**, *30*, 317–335.

(18) Almlöf, J. *Chem. Phys. Lett.* **1991**, *181*, 319–320.

(19) Häser, M.; Almlöf, J. *J. Chem. Phys.* **1992**, *96*, 489–494.

(20) Häser, M. *Theor. Chim. Acta* **1993**, *87*, 147–173.

(21) Ayala, P. Y.; Kudin, K. N.; Scuseria, G. E. *J. Chem. Phys.* **2001**, *115*, 9698–9707.

(22) Lambrecht, D. S.; Doser, B.; Ochsenfeld, C. *J. Chem. Phys.* **2005**, *123*, 184102.

(23) Doser, B.; Lambrecht, D. S.; Ochsenfeld, C. *Phys. Chem. Chem. Phys.* **2008**, *10*, 3335–3344.

(24) Schweizer, S.; Doser, B.; Ochsenfeld, C. *J. Chem. Phys.* **2008**, *128*, 154101.

(25) Doser, B.; Lambrecht, D. S.; Kussmann, J.; Ochsenfeld, C. *J. Chem. Phys.* **2009**, *130*, 064107.

(26) Rauhut, G.; Pulay, P.; Werner, H. J. *J. Comput. Chem.* **1998**, *19*, 1241–1254.

(27) El Azhary, A.; Rauhut, G.; Pulay, P.; Werner, H. J. *J. Chem. Phys.* **1998**, *108*, 5185–5193.

(28) Hetzer, G.; Schütz, M.; Stoll, H.; Werner, H. J. *J. Chem. Phys.* **2000**, *113*, 9443–9455.

(29) Werner, H. J.; Manby, F. R.; Knowles, P. J. *J. Chem. Phys.* **2003**, *118*, 8149–8160.

(30) Schütz, M.; Werner, H. J.; Lindh, R.; Manby, F. R. *J. Chem. Phys.* **2004**, *121*, 737–750.

(31) Saebø, S.; Pulay, P. *Chem. Phys. Lett.* **1985**, *113*, 13–18.

(32) Saebø, S.; Pulay, P. *Annu. Rev. Phys. Chem.* **1993**, *44*, 213–236.

(33) Pulay, P. *Chem. Phys. Lett.* **1983**, *100*, 151–154.

(34) Kendall, R. A.; Früchtl, H. A. *Theor. Chim. Acta* **1997**, *97*, 158–163.

(35) Feyereisen, M.; Fitzgerald, G.; Komornicki, A. *Chem. Phys. Lett.* **1993**, *208*, 359–363.

(36) Vahtras, O.; Almlöf, J.; Feyereisen, M. W. *Chem. Phys. Lett.* **1993**, *213*, 514–518.

(37) Pople, J. A.; Krishnan, R.; Schlegel, H. B.; Binkley, J. S. *Int. J. Quantum Chem.* **1979**, *16*, 225–241.

(38) Frisch, M. J.; Head-Gordon, M.; Pople, J. A. *Chem. Phys. Lett.* **1990**, *166*, 275–280.

(39) Frisch, M. J.; Head-Gordon, M.; Pople, J. A. *Chem. Phys. Lett.* **1990**, *166*, 281–289.

(40) Weigend, F.; Häser, M. *Theor. Chem. Acc.* **1997**, *97*, 331–340.

(41) Rhee, Y. M.; DiStasio, R. A.; Lochan, R. C.; Head-Gordon, M. *Chem. Phys. Lett.* **2006**, *426*, 197–203.

(42) DiStasio, R. A.; Steele, R. P.; Rhee, Y. M.; Shao, Y.; Head-Gordon, M. *J. Comput. Chem.* **2007**, *28*, 839–856.

(43) Simandiras, E. D.; Handy, N. C.; Amos, R. D. *Chem. Phys. Lett.* **1987**, *133*, 324–330.

(44) Neese, F.; Wennmohs, F.; Hansen, A.; Becker, U. *Chem. Phys.* **2009**, *356*, 98–109.

(45) Handy, N. C.; Schaefer, H. F., III *J. Chem. Phys.* **1984**, *81*, 5031–5033.

(46) Becke, A. D. *J. Chem. Phys.* **1988**, *88*, 2547–2553.

(47) Ko, C.; Malick, D. K.; Braden, D. A.; Friesner, R. A.; Martinez, T. J. *J. Chem. Phys.* **2008**, *128*, 104103.

(48) Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. *Chem. Phys. Lett.* **1995**, *242*, 652–660.

(49) Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. *Theor. Chem. Acc.* **1997**, *97*, 119–124.

(50) Neese, F. *J. Comput. Chem.* **2003**, *24*, 1740–1747.

(51) Neese, F.; Becker, U.; Ganyushin, D.; Koβmann, S.; Hansen, A.; Liakos, D.; Petrenko, T.; Riplinger, C.; Wennmohs, F. *ORCA*, version 2.7.0; University of Bonn: Bonn, Germany, 2009.

(52) Schäfer, A.; Horn, H.; Ahlrichs, R. *J. Chem. Phys.* **1992**, *97*, 2571–2577.

(53) Schäfer, A.; Huber, C.; Ahlrichs, R. *J. Chem. Phys.* **1994**, *100*, 5829–5835.

(54) Weigend, F.; Furche, F.; Ahlrichs, R. *J. Chem. Phys.* **2003**, *119*, 12753–12762.

(55) Ahlrichs, R.; Furche, F.; Hättig,C.; Klopper, W. M.; Sierka, M.; Weigend, F. *TurboMole basis set library*; TURBOMOLE GmbH: Karlsruhe, Germany; ftp.chemie.uni-karlsruhe.de/pub/BASES (accessed June 24, 2010).

(56) Weigend, F. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057–1065.

(57) Weigend, F.; Häser, M.; Patzelt, H.; Ahlrichs, R. *Chem. Phys. Lett.* **1998**, *294*, 143–152.

(58) Weigend, F.; Köhn, A.; Hättig, C. *J. Chem. Phys.* **2002**, *116*, 3175–3183.

(59) Hättig, C. *Phys. Chem. Chem. Phys.* **2005**, *7*, 59–66.

(60) Hellweg, A.; Hättig, C.; Hoefener, S.; Klopper, W. *Theor. Chim. Acta* **2007**, *117*, 587–597.

(61) Weigend, F. *J. Comput. Chem.* **2008**, *29*, 167–175.

(62) Aikens, C. M.; Webb, S. P.; Bell, R. L.; Fletcher, G. D.; Schmidt, M. W.; Gordon, M. S. *Theor. Chim. Acta* **2003**, *110*, 233–253.

(63) Krack, M.; Köster, A. M. *J. Chem. Phys.* **1998**, *108*, 3226–3234.

(64) Gill, P. M.-W.; Johnson, B. G.; Pople, J. A. *Chem. Phys. Lett.* **1993**, *209*, 506–512.

(65) Valeev E. ; Fermann J. T. LIBINT integral library; Virginia Tech University: Blacksburg, VA; http://www.files.chem.vt.edu/chem-dept/valeev/software/libint/download.html (accessed June 24, 2010).

(66) Whitten, J. L. *J. Chem. Phys.* **1973**, *58*, 4496–4501.

(67) Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R. *J. Chem. Phys.* **1979**, *71*, 3396–3402.

CT100199K

JCTC Journal of Chemical Theory and Computation

Parallel Calculation of CCSDT and Mk-MRCCSDT Energies

Eric Prochnow,* Michael E. Harding,† and Jürgen Gauss

Institut für Physikalische Chemie, Universität Mainz, D-55099 Mainz, Germany

Received April 14, 2010

Abstract: A scheme for the parallel calculation of energies at the coupled-cluster singles, doubles, and triples (CCSDT) level of theory, several approximate iterative CCSDT schemes (CCSDT-1a, CCSDT-1b, CCSDT-2, CCSDT-3, and CC3), and for the state-specific multireference coupled-cluster ansatz suggested by Mukherjee with a full treatment of triple excitations (Mk-MRCCSDT) is presented. The proposed scheme is based on the adaptation of a highly efficient serial coupled-cluster code leading to a communication-minimized implementation by parallelizing the time-determining steps. The parallel algorithm is tailored for affordable cluster architectures connected by standard communication networks such as Gigabit Ethernet. In this way, CCSDT and Mk-MRCCSDT computations become feasible even for larger molecular systems and basis sets. An analysis of the time-determining steps for CCSDT and Mk-MRCCSDT, namely the computation of the triple-excitation amplitudes and their individual contributions, is carried out. Benchmark calculations are presented for the N_2O, ozone, and benzene molecules, proving that the parallelization of these steps is sufficient to obtain an efficient parallel scheme. A first application to the case of 2,6-pyridyne using a triple-ζ quality basis (222 basis functions) is presented demonstrating the efficiency of the current implementation.

I. Introduction

Coupled-cluster (CC) theory represents one of the most powerful approaches to solve the electronic Schrödinger equation for atoms and molecules.[1−4] Based on the CC ansatz for the wave function a hierarchy of methods can be derived which provides increasing accuracy as well as a fast and systematic convergence toward the full configuration-interaction (FCI) limit. Despite the success of these CC methods, there are a large number of problems such as, for example biradicals, transition-metal compounds, and bond breaking processes, for which the underlying assumption that the wave function is dominated by one reference determinant breaks down. To overcome these limitations, several attempts to generalize single-reference CC theory to a multireference CC (MRCC) ansatz have been reported in the literature (for some examples of genuine MRCC approaches, see refs 5−10). At the moment, the most widely discussed MRCC approaches are based on a state-specific ansatz. Such a strategy enhances the reliability and naturally avoids convergence problems due to intruder states by focusing on one state at a time.[3] However, to obtain a unique state-specific theory, supplementary specifications, i.e., the so-called sufficiency conditions, need to be imposed. Accordingly, different formulations of state-specific MRCC theory are possible.[11] A size-extensive ansatz is the state-specific MRCC method suggested by Mukherjee and co-workers (Mk-MRCC).[5] Mk-MRCC shows faster convergence toward the FCI limit and also yields more accurate energies than other MRCC approaches.[12] In light of these factors efforts in our group toward the development of a state-specific MRCC theory have recently focused on the Mk-MRCC approach.[12−15]

When aiming for quantitative accuracy it is essential to consider in the CC treatment triple-excitation effects. While in most calculations triple excitations are often treated using approximate schemes such as the CC singles and doubles (CCSD) approach with a perturbative treatment of triples (CCSD(T)),[16] full inclusion of triple excitations within the

* Corresponding author e-mail: Eric.Prochnow@uni-mainz.de.
† Present address: Department of Chemistry and Biochemistry, The University of Texas at Austin, Austin, TX 78712.

10.1021/ct1002016 © 2010 American Chemical Society
Published on Web 07/08/2010

CC singles, doubles, and triples (CCSDT)[17,18] model has turned out to be of importance in areas where quantitative accuracy is needed.[19−25] Furthermore, explicit expressions for the implementation of full triples have not only been presented for the single-reference case[26] but also within the Mk-MRCC framework in order to tackle multireference cases.[14] However, the application of CC methods including a full treatment of triple excitations to larger chemical problems is hampered by the considerable increase of computational effort with increasing number of electrons and basis functions. If N represents a measure of the system size the operation count and therefore the execution time scales as N^8 for CCSDT and $d \cdot N^8$ for Mk-MRCCSDT with d denoting the number of chosen reference determinants. In comparison to that the storage requirements for intermediates and CC amplitudes only scales as N^6 and $d \cdot N^6$, respectively. Therefore, the limiting factor of CCSDT calculations is the execution time and not the storage requirement.

To circumvent limitations due to the execution time and motivated by the cheap computing power available through computer clusters, several parallel implementations of single-reference CC methods have been recently presented.[27−31] For most of these parallelization schemes additional tools have been developed. The parallelization in ref 28, for example, is based on the use of Global Arrays (GA)[32] as a parallelization tool. GA simulates shared memory programming on distributed memory computer clusters by striping large arrays across nodes. In view of the large memory and disk space available nowadays, the striping of individual arrays may no longer be necessary. Therefore, an alternative approach, namely Array Files (AF),[33] has been developed where the whole array is stored on a given node. The concept of AF has been used for a parallel implementation of CCSD energy calculations by Janowski et al.[31] A highly sophisticated parallel algorithm for the computation of CCSD(T) energies has been furthermore reported by Olson et al. thereby combining distributed and shared memory techniques by using the Distributed Data Interface (DDI/3).[34,35] This implementation is tailored to multiprocessor and multicore nodes connected via a dedicated communication network. An ansatz which works without an additional layer of complexity provided by libraries such as GA, AF, or DDI/3 has been reported recently for the parallel calculation of CCSD and CCSD(T) energies as well as analytic first and second derivatives.[29] This scheme is based on the message passing interface (MPI)[36] and the adaptation of an efficient serial algorithm to parallel environments where all nonparallel steps run redundantly on all available processors at the same time.

In the field of MRCC theory, parallel algorithms are hardly found as MRCC calculations are so far not considered a routine application. Only Piecuch and Landman reported a parallelization scheme for state-universal CC calculations with single and double excitations.[37] In this scheme the underlying serial algorithm has been modified by parallelizing the most time-determining steps using OpenMP.[38]

To the best of our knowledge, parallel CC algorithms are so far only capable to compute perturbative treatments of triple excitations such as in CCSD(T), but no parallel implementation for CCSDT or a corresponding multireference ansatz has been reported, with the sole exception of the parallel implementation of CCSDT within the general CC programs of Kállay[39] and the NWChem program package.[40] In this paper we describe an approach based on the adaptation of an efficient serial algorithm and a detailed analysis of the time-determining steps for a full treatment of triple excitations as well as the resulting parallel algorithm in the quantum-chemical program package CFOUR[41] is presented. The applicability of our parallelization strategy is demonstrated by computations for N_2O, ozone (O_3), benzene (C_6H_6), and 2,6-pyridyne (C_5NH_3).

II. Parallelization of CCSDT and Mk-MRCCSDT Energy Computations

In our efforts to parallelize CCSDT and Mk-MRCCSDT energy calculations, we follow the strategy already described in ref 29 for the parallel computation of CCSD and CCSD(T) energies as well as analytic first and second derivatives. To avoid the communication of intermediate quantities, most of the quantities needed within the CC iterations are stored completely on every node. Thus, the algorithm presented here is tailored to cluster architectures with moderate hardware specifications as well as inexpensive and rather slow standard interconnect structures such as, for example, Gigabit Ethernet. Furthermore, it is assumed that enough fast memory and disk space are available locally on every node to store the full set of t-amplitudes. In this way communication is minimized as only the CC amplitudes need to be communicated.

In the actual algorithm, parts of the intermediates are contracted with the proper t_3-amplitudes to yield parts of the resulting quantities which, at the end of the parallel operation, are broadcasted to all other nodes. The costs associated with the communication of the triples amplitudes thus scales as occ^3vrt^3 for CCSDT and $d \cdot occ^3vrt^3$ for Mk-MRCCSDT, where occ (vrt) denotes the number of occupied (virtual) orbitals. This is two orders of magnitudes smaller than the required time for the computation of the amplitudes which has a scaling of occ^3vrt^5 and $d \cdot occ^3vrt^5$, respectively. Disregarding communication latencies, the dominating steps have been implemented therefore in a way that the total amount of communication between the parallel processes does not depend on the number of involved processors due to the MPI routine used. Thus, the distribution of the time-determining steps to different nodes reduces the overall walltime significantly, particularly when larger examples are considered.

As the underlying theory and the general implementation of CCSDT and Mk-MRCCSDT have been described in the literature[13,14,26,42] we will report here only the essential equations and in this way avoid a reiteration of the detailed theory.

As for the parallelization of CCSD and CCSD(T) calculations,[29] it is important to identify the time-determining steps. CCSDT energy calculations are dominated by the computation of the t_3-amplitudes yielding a scaling of occ^3vrt^5. Furthermore, contributions of the t_3-amplitudes to the t_2-amplitude equations having a formal scaling of N^7 need to be considered. Thus, the time spent for the calculation of the triples amplitudes and

Table 1. Timings (Walltime in s) for the N_2O Molecule for One CC Iteration and the Triples Part within This Iteration[a]

no. of nodes	CC iter., outside	CC iter., inside	triples part, outside	triples part, inside
1	1065	1062	1046 (98%)	1043
2	547	545	528 (97%)	526
4	288	289	269 (93%)	269
8	158	157	139 (88%)	138
16	93	94	74 (80%)	75

[a] The code has been parallelized outside or inside the loop over occupied orbitals. The computation has been carried out at the CCSDT/cc-pCVTZ level (129 basis functions). For the communication outside the loop the ratio of the triples part to the whole CC iteration is given in parentheses. On each node 1 core has been used.

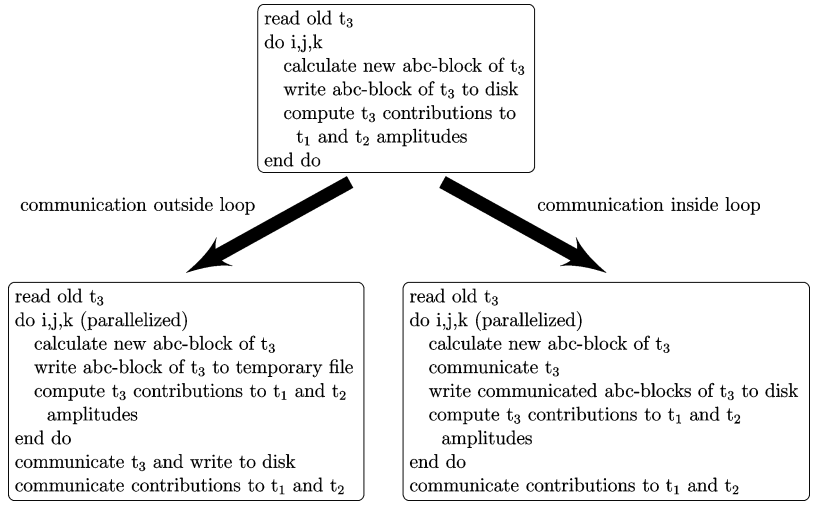


Figure 1. Schematic representation of the two different parallelization strategies. The serial algorithm (top) has been parallelized by communicating the triples amplitudes outside (left) or inside (right) the loop over occupied orbitals i,j,k.

the resulting contributions to the t_2-amplitude equations increases more rapidly with the number of basis functions than in the pure CCSD part (the scaling of CCSD is occ^2vrt^4) which renders the triples part by far the time-determining step in CCSDT calculations. This becomes obvious when considering the timings for a CCSDT energy calculation, e.g., for the N_2O molecule employing the cc-pCVTZ basis set[43] (Table 1). For this example the time spent in the triples part for a serial calculation is about 98% of a CC iteration.

Our approach to parallelize the computation of the triple amplitudes starts with the t_3-amplitude equations given as

$$t_{ijk}^{abc} D_{ijk}^{abc} = \langle \Phi_{ijk}^{abc} | e^{-\hat{T}} (\hat{H}_N - \hat{H}_0) e^{\hat{T}} | 0 \rangle \tag{1}$$

with

$$\hat{H}_0 = \sum_p f_{pp} \{ \hat{a}_p^\dagger \hat{a}_p \} \tag{2}$$

The denominator array D_{ijk}^{abc} is defined as

$$D_{ijk}^{abc} = f_{ii} + f_{jj} + f_{kk} - f_{aa} - f_{bb} - f_{cc} \tag{3}$$

where $i, j, k, \ldots$ denotes occupied and $a, b, c, \ldots$ virtual spin orbitals. The normal-ordered Hamiltonian $\hat{H}_N$ is given as

$$\hat{H}_N = \sum_{pq} f_{pq} \{ \hat{a}_p^\dagger \hat{a}_q \} + \frac{1}{4} \sum_{pqrs} \langle pq || rs \rangle \{ \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \} \tag{4}$$

with the Fock-matrix elements

$$f_{pq} = h_{pq} + \sum_k^{occ} \langle pk || qk \rangle \tag{5}$$

as well as the one-electron integrals h_{pq} and the antisymmetrized two-electron integrals $\langle pk||qk \rangle$. The strings $\{\hat{a}_p^\dagger \hat{a}_p\}$, $\{\hat{a}_p^\dagger \hat{a}_q\}$, and $\{\hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r\}$ denote normal-ordered sequences of creation ($\hat{a}_p^\dagger$) and annihilation ($\hat{a}_q$) operators.

The basic feature in the serial CFOUR algorithm for the formation of the t_3-amplitudes in a CCSDT energy calculation is an outer loop over an index triple i, j, k of the t_{ijk}^{abc} amplitudes (for a detailed description of the implementation see refs 26 and 42). Blocks of a, b, c index triples are computed within the loop one at a time and stored on disk for the next iteration. In addition, these blocks are used immediately to calculate the contributions to the singles and doubles amplitude equations. A schematic representation of this algorithm is depicted in the upper box of Figure 1.

The Mk-MRCCSDT ansatz is implemented using the same loop structure over indices i, j, k. The triples-amplitude equations of the Mk-MRCC ansatz are given as

$$\langle \Phi_{ijk}^{abc}(\mu) | e^{-\hat{T}_\mu} \hat{H} e^{\hat{T}_\mu} | \Phi_\mu \rangle c_\mu^\alpha + \sum_{\nu(\neq\mu)} \langle \Phi_{ijk}^{abc}(\mu) | e^{-\hat{T}_\mu} e^{\hat{T}_\nu} | \Phi_\mu \rangle H_{\mu\nu}^{\mathrm{eff}} c_\nu^\alpha = 0 \tag{6}$$

where c_μ^α denotes the weighting coefficients of the reference determinants in the wave function. These are obtained by diagonalizing the effective Hamiltonian $H_{\mu\nu}^{\mathrm{eff}}$ with its matrix elements defined as

$$H_{\mu\nu}^{\text{eff}} = \langle \Phi_\mu | \hat{H} e^{\hat{T}_\nu} | \Phi_\nu \rangle \overset{\text{CMS}}{=} \langle \Phi_\mu | e^{-\hat{T}_\nu} \hat{H} e^{\hat{T}_\nu} | \Phi_\nu \rangle \quad (7)$$

The second equality in eq 7 holds for a complete model space (CMS), i.e., a model space that contains all determinants generated by distributing m electrons in n_{act} active orbitals (for a detailed description of the Mk-MRCC ansatz, see refs 5 and 15). The advantage of the Mk-MRCC method is the fact that the amplitude equations (eq 6) can be separated into a single-reference and a coupling part.[44] The first part is very similar to eq 1 with the only difference that all quantities are now specific for one reference determinant μ. Consequently, the Mk-MRCC ansatz can be implemented within a single-reference CC (SRCC) code as, for example, provided by the CFOUR package by looping over the reference determinants and adding the coupling terms.[13] This does not affect the general structure of the algorithm for the calculation of the t_3-amplitudes as mentioned before. The only difference is an extra loop over reference determinants around the loop over the index triples i, j, k and the additional coupling terms which are added inside the i, j, k loop.

For a parallelization of the loop over the index triples i, j, k the t_3-amplitudes and the triples contributions to the t_1- and t_2-amplitudes computed on separate nodes need to be communicated. For the communication of the triples amplitudes two different strategies are possible as shown in Figure 1. The t_3-amplitudes are either written to a temporary file and communicated outside the loop or communicated directly after their computation inside the loop. The contributions to the t_1- and t_2-amplitudes are in both cases computed within the same i, j, k loop and communicated outside the loop. Using the strategy described above all triple terms and contributions are parallelized simultaneously. Each parallel process has to calculate a similar amount of i, j, k index triples which is then broadcasted to all other processors.

The performance of the two possible schemes is demonstrated via computations for the N_2O molecule[45] at the CCSDT level of theory using the cc-pCVTZ basis (129 basis functions).[46] The timings for one CC iteration and the respective triples part are given in Table 1. It can be seen that the walltime for one CC iteration is reduced considerably with increasing number of processors. Thus, our implementation will make calculations at the CCSDT level feasible within days or weeks that would take otherwise months to calculate, provided that an appropriate number of nodes is used. In addition, with a larger number of nodes the triples part becomes less dominant in a CCSDT calculation. While for two nodes 97% of the time are spent computing triples contributions, it is reduced to 80% on 16 nodes. The corresponding speedup[47] of the parallel algorithm is depicted in Figure 2. The speedup observed with increasing number of nodes is quite close to the optimal one. With 16 nodes the speedup of one CCSDT iteration reaches values between 11 and 12, while the triples part alone yields a speedup of 14. It is interesting to note that the two outlined communication strategies for the t_3-amplitudes (outside and inside the i, j, k loop) show very similar performance. However, a careful investigation of Figure 2 indicates some preferences to the "outside communication". Therefore, this communication strategy is used in the following.

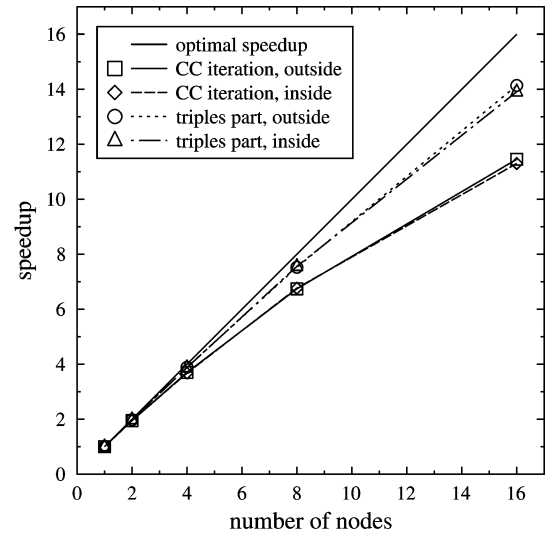


Figure 2. Parallel scaling of energy calculation at the CCSDT/cc-pCVTZ level of theory for N_2O. Depicted is the speedup for one CCSDT iteration and for the triples part within this iteration. The t_3-amplitudes are communicated either outside or inside the loop over *i, j, k* indices. On each node 1 core has been used.

Table 2. Timings (Walltime in s) for One CC Iteration of the Computation for the N_2O Molecule Using Approximate Schemes of CCSDT[a]

no. of nodes	CCSDT-1a	CCSDT-1b	CCSDT-2	CCSDT-3	CC3
1	596	605	650	660	667
2	311 (1.9)	318 (1.9)	356 (1.8)	373 (1.8)	373 (1.8)
4	170 (3.5)	175 (3.5)	214 (3.0)	231 (2.9)	229 (2.9)
8	99 (6.0)	102 (5.9)	142 (4.6)	158 (4.2)	157 (4.2)
16	64 (9.3)	67 (9.0)	106 (6.1)	123 (5.4)	123 (5.4)

[a] The computations have been carried out using the cc-pV5Z basis set (273 basis functions). The speedups for the corresponding computations are given in parentheses. On each node 1 cpu has been used.

As the full inclusion of triple excitations in CC calculation is computational demanding due to its scaling behavior (N^8), several approximate iterative schemes of CCSDT have been developed, namely CCSDT-n (n = 1a, 1b, 2, 3)[48,49] and CC3.[50] In these schemes certain terms of the CCSDT amplitude equations are skipped so that the computational scaling is reduced to N^7. As these methods are implemented within the same i, j, k loop in CFOUR their parallelization is straightforward but should be mentioned at this point. The timings for these methods obtained for the computation on N_2O using the cc-pV5Z basis[51] (273 basis functions)[52] are reported in Table 2. As for CCSDT the walltime is reduced considerably when increasing the number of nodes used, e.g., the calculation time for one iteration of the CCSDT-1a method is about 10 min on 1 node in comparison to about 1 min on 16 nodes yielding a speedup of 9.3. The speedup for 16 nodes observed here is smaller than for CCSDT which is explained by the fact that these computations are less dominated by the triples part. While the time spend calculating the triples contributions for the CCSDT-1a method is 95% of the whole CC iteration on 1 node (compared to 98%

Table 3. Timings (Walltime in s) for the O_3 Molecule for One CC Iteration and the Triples Part within This Iteration[a]

no. of nodes	CC iter.	triples part
1	5864	5811 (99%)
2	3065	3013 (98%)
4	1597	1544 (97%)
8	830	778 (94%)
16	436	384 (88%)

[a] The communication has been performed outside the loop. The computation has been carried out at the Mk-MRCCSDT/cc-pCVTZ level (129 basis functions). The ratio of the triples part to the whole CC iteration is given in parentheses. On each node 1 core has been used.

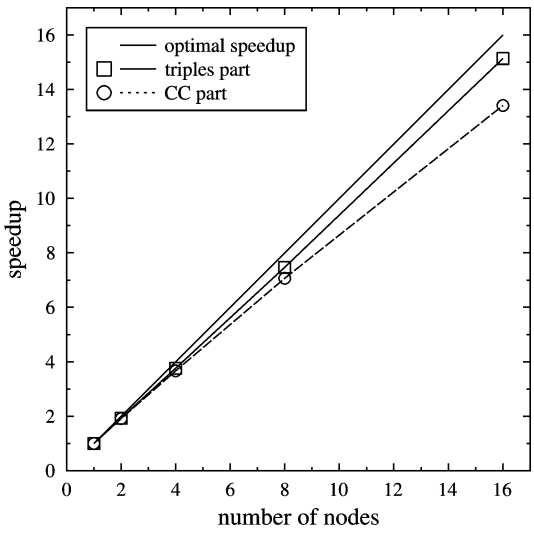


Figure 3. Parallel scaling of energy calculation at the Mk-MRCCSDT/cc-pCVTZ level of theory for O_3. Depicted is the speedup for one Mk-MRCCSDT iteration and for the triples part within this iteration. The t_3-amplitudes are communicated outside the loop over *i, j, k* indices. On each node 1 core has been used.

for CCSDT using a smaller basis) it is reduced to 58% on 16 nodes (80% for CCSDT).

The performance of the parallel computation of Mk-MRCCSDT energies is shown for the ozone molecule. The computation[53] has been performed using the cc-pCVTZ basis set and two-configurational SCF orbitals (with the active-space orbitals having b_1 and a_2 symmetry) using the experimental geometry ($R = 1.2717$ Å, $\angle = 116.8°$).[54] The timings for one CC iteration and the corresponding triples part are given in Table 3. The time for one CC iteration is considerably reduced with an increasing number of processors yielding a computation time of about 7 min on 16 nodes in comparison to 98 min on 1 node. The triples part of this calculation is even more dominant than for a single-reference CCSDT case. For Mk-MRCCSDT the time spent in the triples part is 99% on 1 node and 88% on 16 nodes (compared to 98% and 80% for CCSDT). The speedup obtained for the Mk-MRCCSDT computations is displayed in Figure 3. With 16 nodes the speedup observed lies between 13 and 14, while the triples part alone yields a speedup of over 15. These speedups are even closer to the optimal speedup as the ones obtained for the CCSDT calculation of N_2O, thus, demonstrating the efficiency of the parallelization scheme presented here for Mk-MRCCSDT computations.

Table 4. Timings (Walltime in s) for One CC Iteration for the Benzene Molecule and the Triples Part within This Iteration[a]

no. of cores	CC iteration	triples part
1	21669	21494
2	11036	10861
4	5866	5685
8	2986	2805
16	1542	1371
20	1302	1131
22	1209	1037
24	1327	1126
28	1160	963
32	1081	886
64	809	594

[a] The computation has been carried out at the fc-CCSDT/cc-pVTZ level (264 basis functions). The parallel processes have been distributed in a round-robin fashion.

III. Results and Discussion

In this section we focus on the overall performance and applicability of our parallelization scheme. Results of two applications are presented that involve typical problems in quantum chemistry for which the inclusion of triple excitation is desirable but so far extremely time-consuming.

A. The CCSDT Energy of Benzene. For the investigation of the energetics of molecules, such as the atomization energy, several schemes have been reported in the literature to achieve high-accuracy results.[22−24] From these studies it is obvious that triple excitations are of importance and that contributions beyond CCSD(T) due to a full treatment of triple excitations at the CCSDT level of theory are non-negligible: the contribution from nonperturbative triple excitations, approximated by the difference of CCSDT and CCSD(T) extrapolated from triple- and quadruple-ζ quality basis sets is found for example to be crucial for obtaining chemical accuracy (about 1 kcal/mol) for the molecules N_2, C_2H_2, CO_2, HCN, and O_2 (see ref 22) as well as for vinyl chloride,[55] cyclopropenylidene, propadienylidene,[56] and many others. However, these computations are often very time-consuming, and the applicability of these schemes thus is limited to rather small molecules. Using the parallelization scheme presented here, high-level computations on larger molecules as benzene or substituted benzenes become feasible.

To demonstrate the applicability of our parallelization scheme for medium-sized molecules we have chosen benzene (with geometrical parameters R(CC) = 1.3911 Å, R(CH) = 1.0800 Å)[57] as a typical example. The timings for the frozen-core (fc) energy calculation at the CCSDT level of theory using the cc-pVTZ basis set[51] are given in Table 4. Again, as for N_2O, the walltime is reduced considerably with increasing number of nodes. The walltime for one CC iteration on one node is about six hours, whereas the same computation on 16 nodes only requires about 25 min. Thus, the whole calculation is carried out within hours instead of days. In Figure 4 the corresponding speedup is visualized. In comparison to the results obtained for N_2O (Figure 2) the

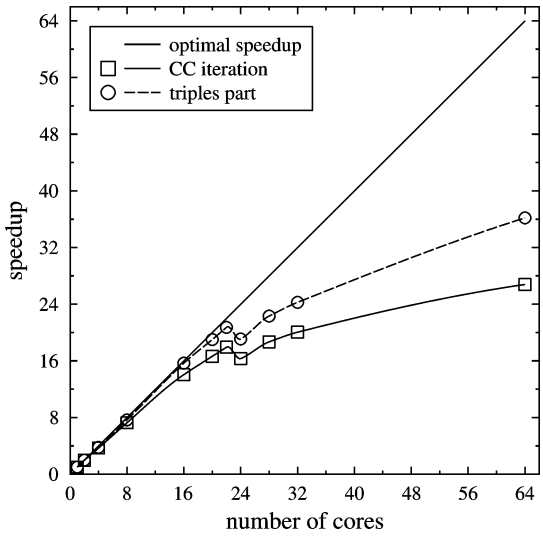


Figure 4. Parallel scaling of frozen-core energy calculation at the CCSDT/cc-pVTZ level for benzene. Depicted is the speedup for one CCSDT iteration and for the triples part within this iteration. The parallel processes have been distributed in a round-robin fashion.

Table 5. Timings (Walltime in s) and Corresponding Speedup for One Iteration in the Calculation of Benzene at the fc-CCSDT/cc-pVTZ Level (264 Basis Functions) with Increasing Number of cores per Node[a]

no. of cores per node	timing	speedup
1	1542	14
2	1812	12
3	2360	9

[a] The computation has been performed using 16 cores.

speedup for benzene is closer to the optimal speedup. Using 16 nodes the speedup observed is 14 for one CC iteration and larger than 15 for the triples part (compare to 11−12 and 14 in the case of N_2O). Thus, the scaling of the total time with the number of processors is improving for increasing system size (129 basis functions and 16 correlated electrons for N_2O, 264 basis functions and 30 correlated electrons for benzene) as the importance of the parallelized, time-determining steps is even larger. It is worthwhile to note the drop in the speedup when using 24 instead of 22 cores. All the computations presented in this section have been carried out on 22 nodes with four cores each. The processes are distributed in a round-robin fashion, i.e., every node first obtains only one job. After that all remaining jobs are distributed in the same way. That means that for computations with up to 22 parallel processes only one core per node has been used. After that point at least some of the nodes need to handle more than one process at a time. The influence of this on the computation time is shown in Table 5 for a computation using 16 cores. The same calculation has been carried out using a different number of cores per node. While using one core per node the walltime observed is 25 min, the computation of one iteration takes about five minutes longer when the number of cores is increased from one to two yielding a speedup of 12 instead of 14. Increasing the number of cores per node again by one, the computational effort grows by additional 9 min, lowering the speedup to nine. Thus, when increasing the number of cores per node used in a calculation a loss of hardware efficiency is observed which worsens with a higher number of cores per node.

B. The 2,6-Pyridyne (C_5NH_3) Molecule. We further demonstrate the applicability of our parallelization scheme by investigating the 2,6-isomer of the didehydropyridine (pyridyne). This choice of example has been motivated by the fact that the chemically interesting didehydroarenes (arynes)[58−61] exhibit a strong biradical character and that their theoretical treatment remains a challenge until today. Multireference approaches are the natural choice for the theoretical treatment of arynes. The focus in the following will be on biradicals derived from pyridine, namely pyridyne. These systems have been recently investigated using density-functional theory (DFT), multiconfigurational SCF (MCSCF), the reduced multireference CC approach, and Mk-MRCC.[13,62−66]

A particularly intriguing aspect of *meta*-arynes is the possibility that they may exist as monocyclic or bicyclic compounds. While for *meta*-benzyne the bicyclic structure is understood to be an artifact displayed at various levels of theory,[67] the issue has not been fully settled in the case of the didehydro compounds of heteroarenes. An earlier study of the geometry of 2,6-pyridyne at the CCSD, CCSD(T), and Mk-MRCCSD level did not provide a final answer for this class of molecules.[13] While CCSD and Mk-MRCCSD predicts both forms to be a minimum (with the monocyclic form to be more stable at the Mk-MRCCSD level and the bicyclic form to be more stable for CCSD) CCSD(T) only yields a minimum for the monocyclic structure. In order to narrow down the problem high-level multireference ab initio calculations are necessary. The recent development of the Mk-MRCCSDT approach[14] allows the inclusion of full triple excitations. However, calculations at this level of theory are demanding and time-consuming when proper basis sets are to be used.

To extend the previous study of 2,6-pyridyne single-point energy calculations at the Mk-MRCCSDT/cc-pVTZ level have been carried out. The geometries of the bicyclic and the monocyclic form as well as the transition state are obtained at the Mk-MRCCSD/cc-pCVTZ level (see ref 13). Two closed-shell determinants are used as reference with the orbitals taken from restricted-Hartree-Fock (RHF) calculations. The first, Φ_1, corresponds to the HF solution, while the second, Φ_2, is obtained by replacing the highest-occupied MO in the monocyclic form (a_1 symmetry) by the lowest unoccupied MO (b_2 symmetry).

The resulting energies are depicted in Figure 5. For comparison, the single-point energies computed at the fc-CCSD and fc-Mk-MRCCSD levels have been added to the figure. The CCSD energy for the bicyclic form is lower (2.9 kcal mol^{-1}) than that for the monocyclic form. The energy of the transistion state lies in between. Mk-MRCCSD yields a lower energy for the monocyclic isomer (3.9 kcal mol^{-1}) and a barrier of 0.2 kcal mol^{-1} from the bicyclic form. When the effect of triple excitations is included at the CCSDT and Mk-MRCCSDT levels of theory the monocyclic form is found to be lower in energy than the bicyclic one. The energy computed at the Mk-MRCCSD/cc-pCVTZ transition-state geometry lies in between. The inclusion of triple excitations

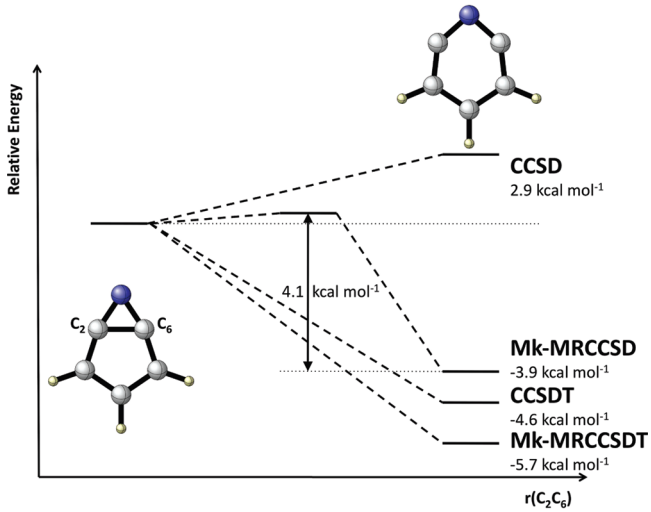


Figure 5. Energetic ordering of the monocyclic and bicyclic forms of 2,6-pyridyne as obtained in frozen-core computations at the CCSD, Mk-MRCCSD, CCSDT, and Mk-MRCCSDT levels of theory using the cc-pVTZ basis set and the RHF-Mk-MRCCSD/cc-pCVTZ geometries from ref 13. The energies are given relative to the bicyclic form.

Table 6. Timings (Walltime in s) for One CC Iteration for the Energy Computations of 2,6-Pyridyne at Geometries for the Bicyclic and Monocyclic Form As Well As the Transition State[a]

	bicyclic	transition state	monocyclic
Mk-MRCCSDT	10707	10718	10699
CCSDT	2529	2548	2526

[a] The frozen-core calculations have been carried out on 8 nodes (using 1 CPU each) employing the cc-pVTZ basis set (222 basis functions) and the Mk-MRCCSD/cc-pCVTZ geometry.

at the CCSDT level leads to a considerable stabilization of the monocyclic isomer (7.5 kcal mol^{-1}). The additional stabilization at the Mk-MRCCSDT level is significantly smaller (1.8 kcal mol^{-1}).

These results indicate that the energy minimum of the bicyclic isomer of 2,6-pyridyne may be an artifact of theory due to the lack of triple excitations. Interestingly, computations at the Mk-MRCCSD(T) level, where triple excitations are treated in an approximate manner, yield comparable findings supporting our conclusion.[64]

As the calculations at the Mk-MRCCSDT level are computationally demanding they have only been carried out using 8 and 16 nodes. The walltimes observed for the calculations are given in Table 6. The computation time for the monocyclic isomer is about 178 min on 8 nodes and about 102 min on 16 nodes yielding a speedup of 1.8. For the corresponding computation at the CCSDT level a speedup of 1.6 is observed.

IV. Conclusions

A detailed analysis of CCSDT and Mk-MRCCSDT energy computations indicates that these calculations are completely dominated by the determination of the triples amplitudes and their contributions to the singles and doubles amplitude equations. Therefore, a parallelization scheme is presented that starts from the efficient serial algorithm available within the CFOUR program package and parallelizes the most time-consuming parts in the calculation, namely steps related to the triple amplitudes. The central aspect of our parallel implementation is the distributed computation of the triple amplitudes as well as their contributions to the single and double amplitudes. In this way an algorithm is obtained for which sufficient local memory and disk space are needed but which does not depend on high-speed network connections. The parallel implementation presented here will be available in a future public release of the CFOUR program package.

Benchmark calculations demonstrate the applicability of our parallel algorithm. Computations at the CCSDT/cc-pCVTZ level of theory for the N_2O molecule (129 basis functions) exhibit a good scaling with the number of processors. In this case a speedup of 11 to 12 is observed on 16 processors. For larger systems the behavior is even better. For a frozen-core calculation on benzene at the CCSDT/cc-pVTZ level (264 basis functions) a speedup of 14 is obtained when 16 nodes are used. The excellent scaling behavior may be observed as long as the time-determining steps in the serial calculation dominate the overall time in the parallel computation. However, this might only be a problem when a larger number of nodes is used, as the computation on benzene still spends about 73% of the whole time in the triples part even when 64 processors are used.

We further present parallel Mk-MRCCSDT computations for ozone for which benchmark calculations at the Mk-MRCCSDT/cc-pCVTZ level (129 basis functions) have been carried out yielding speedups even closer to the optimal speedup than single-reference CCSDT. The speedup observed is 13 to 14 when 16 nodes are used. The applicability of the parallelization of the multireference algorithm is demonstrated by investigating the 2,6-pyridyne molecule. This compound requires a multireference treatment due to a strong biradical character and the inclusion of at least triple excitations for a qualitative correct description of the potential energy surface. Frozen-core computations at the Mk-MRCCSDT/cc-pVTZ level (222 basis functions) in combination with the parallelization scheme presented here become feasible in a resonable amount of time.

In this work we have demonstrated that by adapting an efficient serial algorithm to a parallel environment, calculations at the CCSDT and Mk-MRCCSDT level of theory are feasible within days or weeks instead of months or years even when larger basis sets are used or larger systems are investigated. Thus, this kind of implementation opens the field of application for CCSDT and Mk-MRCCSDT where high-accuracy is decisive.

V. Technical Details

All calculations were carried out on a 22 node cluster. The cluster nodes were equipped with two dual core Intel Xeon 5160 processors running at 2.93 GHz, 32 GB FB-DIMM RAM, and 8 striped SATA disks. For the network communication the onboard Gigabit Ethernet controler was used.

For the parallel implementation the message passing interface (MPI)[36] is used. The results presented here are

obtained by using LAM/MPI.[68,69] The communication of the t_3-amplitudes is done by the MPI_BCAST subroutine. The triples contributions to the t_1- and t_2-amplitudes are communicated using the MPI_ALLREDUCE routine.

Note Added after ASAP Publication. This article was published ASAP on July 8, 2010. Equation 1 has been modified. The correct version was published on July 14, 2010.

Acknowledgment. This work has been supported by the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie.

References

(1) Gauss, J. In *The Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., Schreiner, P. R., Eds.; Wiley: Chichester, 1998; p 615.

(2) Crawford, T. D.; Schaefer, H. F. *Rev. Comput. Chem.* **2000**, *14*, 33.

(3) Bartlett, R. J.; Musiał, M. *Rev. Mod. Phys.* **2007**, *79*, 291.

(4) Shavitt, I.; Bartlett, R. J. *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory*; Cambridge University Press: Cambridge, 2009; p 1.

(5) Mahapatra, U. S.; Datta, B.; Mukherjee, D. *Mol. Phys.* **1998**, *94*, 157.

(6) Mukherjee, D.; Moitra, R. K.; Mukhopadhyay, A. *Mol. Phys.* **1977**, *33*, 955.

(7) Jeziorski, B.; Monkhorst, H. J. *Phys. Rev. A* **1981**, *24*, 1668.

(8) Paldus, J.; Piecuch, P.; Pylypow, L.; Jeziorski, B. *Phys. Rev. A* **1993**, *47*, 2738.

(9) Hubač, I.; Pittner, J.; Čársky, P. *J. Chem. Phys.* **2000**, *112*, 8779.

(10) Hanrath, M. *J. Chem. Phys.* **2005**, *123*, 084102.

(11) Kong, L. *Int. J. Quantum Chem.* **2009**, *109*, 441.

(12) Evangelista, F. A.; Allen, W. D.; Schaefer, H. F. *J. Chem. Phys.* **2006**, *125*, 154113.

(13) Prochnow, E.; Evangelista, F. A.; Schaefer, H. F.; Allen, W. D.; Gauss, J. *J. Chem. Phys.* **2009**, *131*, 064109.

(14) Evangelista, F. A.; Simmonett, A. C.; Allen, W. D.; Schaefer, H. F.; Gauss, J. *J. Chem. Phys.* **2008**, *128*, 124104.

(15) Evangelista, F. A.; Allen, W. D.; Schaefer, H. F. *J. Chem. Phys.* **2007**, *127*, 024102.

(16) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479.

(17) Noga, J.; Bartlett, R. J. *J. Chem. Phys.* **1987**, *86*, 7041.

(18) Scuseria, G. E.; Schaefer, H. F. *Chem. Phys. Lett.* **1988**, *152*, 382.

(19) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic Structure Theory*; Wiley: Chichester, 2001.

(20) Halkier, A.; Larsen, H.; Olsen, J.; Jørgensen, P.; Gauss, J. *J. Chem. Phys.* **1999**, *110*, 734.

(21) Sordo, J. A. *J. Chem. Phys.* **2001**, *114*, 1974–1980.

(22) Harding, M. E.; Vázquez, J.; Ruscic, B.; Wilson, A. K.; Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **2008**, *128*, 114111.

(23) Császár, A. G.; Allen, W. D.; Schaefer, H. F. *J. Chem. Phys.* **1998**, *108*, 9751.

(24) Karton, A.; Rabinovich, E.; Martin, J. M. L.; Ruscic, B. *J. Chem. Phys.* **2006**, *125*, 144108.

(25) Feller, D.; Peterson, K. A. *J. Chem. Phys.* **2009**, *131*, 154306.

(26) Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **2002**, *116*, 1773.

(27) Rendell, A. P.; Lee, T. J.; Lindh, R. *Chem. Phys. Lett.* **1992**, *194*, 84.

(28) Hirata, S. *J. Phys. Chem. A* **2003**, *107*, 9887.

(29) Harding, M. E.; Metzroth, T.; Gauss, J.; Auer, A. A. *J. Chem. Theory. Comput.* **2008**, *4*, 64.

(30) Lotrich, V.; Flocke, N.; Ponton, M.; Yau, A. D.; Perera, A.; Deumens, E.; Bartlett, R. J. *J. Chem. Phys.* **2008**, *128*, 194104.

(31) Janowski, T.; Ford, A. R.; Pulay, P. *J. Chem. Theory. Comput.* **2007**, *3*, 1368.

(32) Nieplocha, J.; Harrison, R. J.; Littlefield, R. In *Proceedings of Supercomputing*; IEEE Computer Society Press: Washington, DC, 1994; p 340.

(33) Ford, A. R.; Janowski, T.; Pulay, P. *J. Comput. Chem.* **2007**, *28*, 1215.

(34) Olson, R. M.; Bentz, J. L.; Kendall, R. A.; Schmidt, M. W.; Gordon, M. S. *J. Chem. Theory. Comput.* **2007**, *3*, 1312.

(35) Olson, R. M.; Schmidt, M. W.; Gordon, M. S.; Rendell, A. P. *Enabling the Efficient Use of SMP Clusters: The GAMESS/DDI Approach; In Supercomputing*; ACM/IEEE Conference: Phoenix, AZ, 2003; p 41.

(36) The MPIForum. MPI: a message passing interface. In *Proceedings of the 1993 ACM/IEEE Conference on Supercomputing*; ACM Press: Portland, OR, U.S.A., 1993.

(37) Piecuch, P.; Landman, J. I. *Parallel Comput.* **2000**, *26*, 913.

(38) OpenMP Fortran Application Programming Interface. http://www.openmp.org. Accessed March 12, 2010.

(39) Kállay, M.; Harding, M. E. Parallel version of the string-based general coupled-cluster program MRCC, 2006. See http://www.mrcc.hu. Accessed March 18, 2010.

(40) Bylaska, E. J. *NWChem, A Computational Chemistry Package for Parallel Computers, Version 5.1*; Pacific Northwest National Laboratory: Richland, Washington 99352-0999, U.S.A., 2007.

(41) CFOUR, a quantum chemical program package written by Stanton, J. F.; Gauss, J.; Harding, M. E. Szalay, P. G. with contributions from Auer, A. A.; Bartlett, R. J.; Benedikt, U.; Berger, C.; Bernholdt, D. E.; Bomble, Y. J.; Christiansen, O.; Heckert, M.; Heun, O.; Huber, C.; Jagau, T.-C.; Jonsson, D.; Jusélius, J.; Klein, K.; Lauderdale, W. J.; Matthews, D. A.; Metzroth, T.; O'Neill, D. P.; Price, D. R.; Prochnow, E.; Ruud, K.; Schiffmann, F.; Stopkowicz, S.; Vázquez, J.; Wang, F.; Watts, J. D. and the integral packages MOLECULE (Almlöf, J.; Taylor, P. R.),PROPS (Taylor, P. R.), ABACUS (Helgaker, T.; Jensen, H. J. Aa.; Jørgensen, P.; Olsen, J.),and ECP routines by Mitin, A. V.; van Wüllen, C. For the current version, see http://www.cfour.de. Accessed February 18, 2010.

(42) Gauss, J.; Stanton, J. F. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2047.

(43) Woon, D. E.; Dunning, Jr, T. H. *J. Chem. Phys.* **1995**, *103*, 4572.

(44) Mahapatra, U. S.; Datta, B.; Mukherjee, D. *J. Chem. Phys.* **1999**, *110*, 6171.

(45) R(NN) = 1.1290 Å, R(NO) = 1.1849 Å, ∠(NNO) = 180.0°.

(46) The absolute energy obtained is−184.572494 hartree.

(47) The speedup is defined as the ratio of the serial execution time and the execution time on *n* cores.

(48) Urban, M.; Noga, J.; Cole, S. J.; Bartlett, R. J. *J. Chem. Phys.* **1985**, *83*, 4041.

(49) Noga, J.; Bartlett, R. J.; Urban, M. *Chem. Phys. Lett.* **1987**, *134*, 126.

(50) Koch, H.; Christiansen, O.; Jørgensen, P.; Sanchez de Merás, A. M.; Helgaker, T. *J. Chem. Phys.* **1997**, *106*, 1808.

(51) Dunning, Jr, T. H. *J. Chem. Phys.* **1989**, *90*, 1007.

(52) The absolute energies obtained are (in hartree): CC3:−184.595095, CCSDT-1a:−184.594467, CCSDT-1b:−184.594525, CCSDT-2:−184.588903, CCSDT-3:−184.589449.

(53) The absolute energy obtained is−225.306818 hartree.

(54) Tanaka, T.; Morino, Y. *J. Mol. Spectrosc.* **1970**, *33*, 538.

(55) Harding, M. E.; Gauss, J.; Pflüger, K.; Werner, H.-J. *J. Phys. Chem. A* **2007**, *111*, 13623.

(56) Vázquez, J.; Harding, M. E.; Gauss, J.; Stanton, J. F. *J. Phys. Chem. A* **2009**, *113*, 12447.

(57) Gauss, J.; Stanton, J. F. *J. Phys. Chem. A* **2000**, *104*, 2865.

(58) Myers, A. G.; Cohen, S. B.; Kwon, B. M. *J. Am. Chem. Soc.* **1994**, *116*, 1670.

(59) Schottelius, M. J.; Chen, P. *J. Am. Chem. Soc.* **1996**, *118*, 4896.

(60) Hoffner, J.; Schottelius, M. J.; Feichtinger, D.; Chen, P. *J. Am. Chem. Soc.* **1998**, *120*, 376.

(61) Winkler, M.; Cakir, B.; Sander, W. *J. Am. Chem. Soc.* **2004**, *126*, 6135.

(62) Debbert, S. L.; Cramer, C. J. *Int. J. Mass. Spectrom.* **2000**, *201*, 1.

(63) Li, X.; Paldus, J. *J. Chem. Phys.* **2008**, *129*, 174101.

(64) Evangelista, F. A.; Prochnow, E.; Gauss, J.; Schaefer, H. F. *J. Chem. Phys.* **2010**, *132*, 074107.

(65) Jagau, T.-C.; Prochnow, E.; Evangelista, F. A.; Gauss, J. *J. Chem. Phys.* **2010**, *132*, 144110.

(66) Manohar, P. U.; Koziol, L.; Krylov, A. I. *J. Phys. Chem. A* **2009**, *113*, 2591.

(67) Smith, C. E.; Crawford, T. D.; Cremer, D. *J. Chem. Phys.* **2005**, *122*, 174309.

(68) Burns, G.; Daoud, R.; Vaigl, J. LAM: An Open Cluster Environment for MPI. *Proceedings of Supercomputing Symposium*; 1994; p 379.

(69) Squyres, J. M.; Lumsdaine, A. A Component Architecture for LAM/MPI. *Proceedings, 10th European PVM/MPI Users' Group Meeting*, Venice, Italy, 2003; p 379.

CT1002016

JCTC Journal of Chemical Theory and Computation

Auxiliary Density Matrix Methods for Hartree−Fock Exchange Calculations

Manuel Guidon, Jürg Hutter, and Joost VandeVondele*

Physical Chemistry Institute, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

Received April 27, 2010

Abstract: The calculation of Hartree−Fock exchange (HFX) is computationally demanding for large systems described with high-quality basis sets. In this work, we show that excellent performance and good accuracy can nevertheless be obtained if an auxiliary density matrix is employed for the HFX calculation. Several schemes to derive an auxiliary density matrix from a high-quality density matrix are discussed. Key to the accuracy of the auxiliary density matrix methods (ADMM) is the use of a correction based on standard generalized gradient approximations for HFX. ADMM integrates seamlessly in existing HFX codes and, in particular, can be employed in linear scaling implementations. Demonstrating the performance of the method, the effect of HFX on the structure of liquid water is investigated in detail using Born−Oppenheimer molecular dynamics simulations (300 ps) of a system of 64 molecules. Representative for large systems are calculations on a solvated protein (Rubredoxin), for which ADMM outperforms the corresponding standard HFX implementation by approximately a factor 20.

1. Introduction

The success of density functional theory (DFT) can be attributed to the fact that it can provide an accurate description of the electronic structure at a moderate computational cost. DFT has become a unique tool to describe systems containing hundreds to thousands of atoms. Not only is it possible to describe molecules in the gas phase, properties of condensed phase systems such as liquids and solids can also be computed. For these systems, using contemporary computer resources, it has become possible to go beyond a static description of matter, and finite temperature effects can be included directly through ab initio molecular dynamics (MD) simulations. Large-scale, condensed phase and dynamical simulations have mostly adopted a relatively simple form for the exchange and correlation functional, namely, the semilocal generalized gradient approximation (GGA). However, it becomes increasingly clear that an improved description of the electronic structure, and thus more accurate results, can only be obtained by functionals that go beyond the GGA form and incorporate a nonlocal term such as Hartree−Fock exchange (HFX). The computational cost of these nonlocal terms is typically much larger than that of the local terms. Consequently, there is significant interest in finding efficient approaches to deal with these nonlocal forms.

The efficiency of a HFX calculation depends strongly on the algorithm employed. A straightforward implementation based on localized basis sets scales with the fourth power of the system size. However, integral screening[1] reduces the scaling with system size to quadratic and, for short-range operators, such as screened[2] or truncated exchange,[3−6] to linear scaling. Nonmetallic systems furthermore allow for a screening on the density matrix,[7] which leads to linear scaling also for long-range operators. Using these techniques, HFX can be evaluated also for condensed phase systems containing a few thousand atoms[3] and can be used to perform ab initio molecular dynamics simulations.[8] Despite the favorable scaling with system size, HFX calculations scale very poorly with basis set quality. This is an important issue, since high-quality results not only require accurate functionals but also good basis sets. There are several reasons why the cost of HFX depends very strongly on the basis employed. Indeed, even in a linear scaling code, the cost increases with the fourth power of the number of (primitive) basis functions

* Corresponding author e-mail: Joost.VandeVondele@pci.uzh.ch.

10.1021/ct1002225 © 2010 American Chemical Society
Published on Web 07/01/2010

Table 1. Impact of the Basis Set Quality for the Wavefunction Optimization of a 20 Water Cluster[a]

basis	$\kappa(S)$	threshold	cost [ERIs]	cost [s]
3-21G*	4.9×10^1	1.0×10^{-4}	2.3×10^7	0.06
6-31G**	2.1×10^2	1.0×10^{-5}	5.2×10^8	0.35
6-311G++G**	1.2×10^5	1.0×10^{-7}	1.1×10^{10}	11.71
pc-0	5.2×10^1	1.0×10^{-4}	1.7×10^7	0.07
pc-1	4.5×10^3	1.0×10^{-5}	4.4×10^8	0.50
pc-2	5.7×10^5	1.0×10^{-7}	2.0×10^{10}	11.21
aug-pc-1	1.4×10^6	1.0×10^{-8}	5.0×10^{10}	53.23
aug-pc-2	3.9×10^8	1.0×10^{-9}	1.5×10^{12}	766.92
def2-QZVP	7.1×10^4	1.0×10^{-8}	3.2×10^{11}	127.16
aug-def2-QZVP	8.5×10^5	1.0×10^{-8}	6.2×10^{11}	331.61

[a] The condition number $\kappa(S)$ of the overlap matrix determines the maximal possible screening threshold. The latter needs to be chosen more tightly, if $\kappa(S)$ gets large. This is reflected in the cost of a calculation, which is given once by the number of Cartesian four-center electron repulsion integrals (ERIs) that need to be evaluated and the time in seconds that is spent in building the Fock matrix in the first self-consistent field (SCF) step. 3-21G*, 6-31G**, and 6-311G++G** refer to basis sets by Pople and co-workers.[10−13] The polarization consistent (pc) basis sets have been developed by Jenssen,[14−16] the def2-QZVP basis by Ahlrichs and Weigend.[17] Timings are obtained on 128 cores of a CRAY-XT5.

per atom. Basis sets with a high l-quantum number (polarization functions) are therefore costly, as the number of basis functions per atom grows quadratically with l. Heavily contracted basis functions, such as the molecularly optimized basis sets proposed in ref 9, are expensive since for each quartet of basis functions a very larger number of primitive integrals needs to be considered. Very flexible basis sets, or basis sets with diffuse primitives, are costly for several reasons. First, diffuse primitives are nonzero in a larger part of space, and thus screening becomes less efficient. This is particularly important in condensed phase systems, where periodic boundary conditions provide a potentially unlimited number of interacting atomic sites. Second, uncontracted diffuse primitives influence the condition number of the overlap matrix (S) strongly, and a poor condition number in turn implies that a tighter screening threshold has to be employed[3] to obtain a stable self-consistent calculation. Third, the sparsity of the matrix representation of the density matrix (P) also depends strongly on the condition number of S, making density matrix screening less efficient for poorly conditioned basis sets. Table 1 illustrates this problem by providing costs and maximal thresholds needed in order to get converged results for a water cluster containing 20 water molecules.

Clearly, a technique which reduces the impact of the basis set on the computational cost is a significant progress.

During the past decade, much effort has been invested in solving this problem, and many different techniques have been proposed. Among them are methods that apply an approximate resolution of identity, for example, RI[18] or Cholesky decomposition.[19] These schemes rely on the introduction of auxiliary basis functions in terms of which the four center integrals can be approximated by corresponding two- and three-center terms. In order to improve efficiency, Sodt. and Head-Gordon[20] developed a local variant of RI, atomic resolution of identity (ARI). A slightly different post-Hartree−Fock approach in a dual basis was introduced in ref 21, where a reference calculation in a small basis set is perturbatively corrected to a large basis set. A different approximation for the two-electron integrals has been proposed by Friesner[22] and has been termed the "pseudo-spectral" method. Recently, Neese et al.[23] presented an algorithm called COSX that is a combination of semi-numerical methods and RI. Furthermore, there exist several schemes to achieve linear scaling in the context of plane wave basis sets such as the multiwavelet-based ansatz of Harrison et al.[24] or FFT-based algorithms as presented in refs 25−27.

In this work, we propose to employ an auxiliary density matrix to evaluate the expensive nonlocal part of the functional, while all other energy components are computed with the primary (original) density matrix. The auxiliary density matrix will be constructed in a way that allows for a rapid evaluation of the HFX energy, using any algorithm, including traditional or linear scaling approaches. In order to ensure that the quality of the calculation is influenced as little as possible by the quality of the auxiliary density matrix, a correction term is added to the exchange and correlation functional. On the basis of a GGA for exchange, this correction takes the difference between auxiliary and primary density matrices into account. All terms of the resulting density functional are straightforward to compute, but there is considerable freedom in how to obtain from a given primary density matrix a suitable auxiliary density matrix. In this paper, various procedures are discussed and tested. Tests are presented in section 3 and include gas phase thermochemistry; basis set superposition error; and electronic structure including band gaps, large systems, and liquid water. The theory is introduced in the following section, but for mathematical derivations and technical details, we refer to the Appendix.

2. Theory

2.1. Basic Concepts. In Kohn−Sham DFT, the total energy of a system consisting of N_e electrons can be written in terms of the electron density

$$\rho(\mathbf{r}) = \sum_{i=1}^{N_e} |\psi_i(\mathbf{r})|^2 \tag{1}$$

where ψ_i denotes the single particle wave functions, which are assumed to be real-valued. The total energy is then expressed in terms of a functional of the electron density as

$$E[\rho] = T_s[\rho] + J[\rho] + E_{xc}[\rho] + \int v(\mathbf{r})\,\rho(\mathbf{r})\,d\mathbf{r} \tag{2}$$

with the standard abbreviations for kinetic, Hartree, and exchange-correlation energy and the part due to the external potential. In hybrid DFT, the exchange-correlation functional is augmented by a certain fraction of Hartree−Fock exchange based on the wave functions $\{\psi_i\}$

$$E_{xc}[\rho] = \alpha E_x^{HFX}[\{\psi_i\}] + (1 - \alpha)E_x^{DFT}[\rho] + E_c^{DFT}[\rho] \tag{3}$$

where α denotes the fraction of HFX and E_x and E_c are the density functionals for exchange and correlation, respectively. In the presence of an atomic centered basis set $\{\phi_\mu(\mathbf{r})\}$

$$\psi_i(\mathbf{r}) = \sum_\mu C^{\mu i}\,\phi_\mu(\mathbf{r}) \tag{4}$$

the Hartree−Fock exchange energy can be expressed in terms of a density matrix and two-electron integrals (ERIs)

$$E_x^{\mathrm{HFX}}[P] = -\frac{1}{2}\sum_{\lambda\sigma\mu\nu} P^{\mu\sigma}P^{\nu\lambda}(\mu\nu|\lambda\sigma) \tag{5}$$

where the density matrix elements $P^{\mu\nu}$ are obtained from the molecular (MO) coefficients as

$$P^{\mu\nu} = \sum_i C^{\mu i}C^{\nu i} \Leftrightarrow P = CC^T \tag{6}$$

and the ERIs are defined as

$$(\mu\nu|\lambda\sigma) = \int\int \phi_\mu(\mathbf{r}_1)\,\phi_\nu(\mathbf{r}_1)\,g(|\mathbf{r}_2 - \mathbf{r}_1|)\,\phi_\lambda(\mathbf{r}_2)\,\phi_\sigma(\mathbf{r}_2)\,\mathrm{d}\mathbf{r}_1\mathbf{r}_2 \tag{7}$$

with the interaction potential $g(r)$ that is Coulombic $(1/r)$ in standard Hartree−Fock theory. The fourth order scaling of HFX with basis set size can be directly inferred from eq 5.

By introducing an auxiliary density matrix $\hat{P} \approx P$ that is either smaller in size or more rapidly decaying than the original one, the evaluation of HFX can be sped up significantly. The HFX energy can be written as

$$\begin{aligned} E_x^{\mathrm{HFX}}[P] = &\; E_x^{\mathrm{HFX}}[\hat{P}] + (E_x^{\mathrm{HFX}}[P] - E_x^{\mathrm{HFX}}[\hat{P}]) \\ \approx &\; E_x^{\mathrm{HFX}}[\hat{P}] + (E_x^{\mathrm{DFT}}[P] - E_x^{\mathrm{DFT}}[\hat{P}]) \end{aligned} \tag{8}$$

The assumption behind this approximation is that *the difference* in the exchange energy between primary and auxiliary density matrices is well captured by a GGA, even in those cases where GGA exchange and HFX might be qualitatively different. Equation 8 amounts to computing the HFX energy with an auxiliary density matrix, while a GGA correction is introduced which takes the difference between auxiliary and primary density matrices into account. As shown in section 3, applying this correction indeed improves upon uncorrected results. Clearly, our approach yields the original HFX energy as the quality of either the auxiliary density matrix or the correcting functional improves. In this work, we have based the GGA correction on PBE exchange[28,29] and have not explored other parametrizations or other functionals. The introduction of eq 8 in hybrid density functionals is natural and straightforward and, usually, because only a fraction of exchange is needed, will introduce a smaller error. If hybrid functionals employ a non-Coulombic operator, the exchange functional needs to be chosen consistently with the shape of the interaction potential ($g(r)$) in the ERI calculation. Currently, the GGA correction for the standard Coulomb potential, the short-range (erfc) and the truncated Coulomb potential have been implemented and tested.

2.2. Auxiliary Density Matrices. The performance and accuracy of the ADMM scheme depends on how the auxiliary density matrix is constructed, and various approaches seem possible. In this section, we present methods that either rely on the use of an auxiliary basis set or directly manipulate the sparsity of the density matrix.

The size of P obtained from a high-quality primary basis set (PBS) $\{\phi_\mu(\mathbf{r})\}$ can be reduced by introducing an auxiliary basis set (ABS) $\{\hat{\phi}_\mu(\mathbf{r})\}$ for the description of the underlying wave function

$$\hat{\psi}_i(\mathbf{r}) = \sum_\mu \hat{C}^{\mu i}\hat{\phi}_\mu(\mathbf{r}) \tag{9}$$

i.e.

$$\hat{P}^{\mu\nu} = \sum_i \hat{C}^{\mu i}\hat{C}^{\nu i} \Leftrightarrow \hat{C}\hat{C}^T \tag{10}$$

An optimal value for the MO coefficients can be obtained by requiring that the square difference for the occupied wave functions in ABS and PBS representation is minimized

$$\min_{\hat{C}} = \sum_i \int (\psi_j(\mathbf{r}) - \hat{\psi}_j(\mathbf{r}))^2\,\mathrm{d}\mathbf{r} \tag{11}$$

This yields the following expression for the auxiliary MO coefficients

$$\hat{C} = AC \tag{12}$$

where A is defined as the projector between the two basis sets

$$A = \hat{S}^{-1}Q \tag{13}$$

with the overlap matrices

$$\hat{S}_{nn'} = \int \hat{\phi}_n(\mathbf{r})\,\hat{\phi}_{n'}(\mathbf{r})\,\mathrm{d}\mathbf{r} \text{ and } Q_{nm} = \int \hat{\phi}_n(\mathbf{r})\,\phi_m(\mathbf{r})\,\mathrm{d}\mathbf{r} \tag{14}$$

A slightly more complicated formula is obtained when the auxiliary wave functions are required to minimize eq 11 subject to the constraint that they remain orthonormal. This constraint can be enforced introducing Lagrangian multipliers (Λ_{kl}) in eq 11 as

$$\min_{\tilde{C}}[\sum_j \int (\psi_j(\mathbf{r}) - \tilde{\psi}_j(\mathbf{r}))^2\,\mathrm{d}\mathbf{r} + \sum_{k,l}\Lambda_{kl}(\int \tilde{\psi}_k(\mathbf{r})\,\tilde{\psi}_l(\mathbf{r})\,\mathrm{d}\mathbf{r} - \delta_{kl})] \tag{15}$$

The coefficients $\tilde{C}$ that minimize this expression can be obtained as

$$\tilde{C} = \hat{C}\Lambda^{-1/2} \text{ with } \Lambda = \hat{C}^T\hat{S}\hat{C} \tag{16}$$

where $\hat{C}$ is defined by eq 12.

Of course, there is significant freedom in selecting the auxiliary basis set, and the choice need not to be homogeneous in space. For example, for large systems with a chemically active region, such as enzymes, it is natural to retain the high-quality primary basis where exchange matters most, while a lower-quality auxiliary basis can be used for the bulk. Furthermore, note that the explicit shape of the basis functions (Gaussian functions, Slater functions, etc.) is not important and indeed need not to be the same in the auxiliary and primary basis sets. The method thus provides an interesting approach for computing exchange contributions

in programs that do not employ Gaussian basis functions and for which the calculation of exchange is relatively difficult.

The two different sets of MO coefficients, $\hat{C}$ and $\tilde{C}$, correspond to two different density matrices:

$$\tilde{P} = \tilde{C}\tilde{C}^T = \hat{C}\Lambda^{-1}\hat{C}^T \tag{17}$$

and

$$\hat{P} = \hat{C}\hat{C}^T = APA^T \tag{18}$$

that can be used as an auxiliary density matrix. We will refer to the first as purified wave function fitting or ADMM1 and to the second as nonpurified wave function fitting or ADMM2 (see section 2.3 for an explanation of the nomenclature).

A strategy directly aimed at obtaining a sparse auxiliary density matrix relies on a blocking of the primary density matrix. This strategy is applicable if the system of interest can be divided into subsystems that have no important exchange interactions beyond what is captured with a GGA. In this case, the nonrelevant blocks in the auxiliary density matrix can just be zeroed, and to some extend this method can be considered a subsystem-based neglect of diatomic differential overlap (NDDO). If all intersubsystem blocks are zeroed, the approximate density matrix will be positive definite; however, we employ the slightly generalized form of the auxiliary density matrix as

$$\hat{P} = P \otimes B \tag{19}$$

where B is a blocking matrix with $B_{ij} \in \{1,0\}$ and $\otimes$ denotes the Hadamard product of two matrices. In this case, B can reflect the molecular topology and allow for connections between subsystems. In the general case, the resulting $\hat{P}$ need not to be positive definite. This method of obtaining an auxiliary density matrix will be referred to as blocking or ADMM3 in the following.

2.3. Density Matrix Purification. As already mentioned, an approximate density matrix might not fulfill the properties of a pure density matrix:

$$P = P^T \tag{20}$$

$$PSPS = PS \tag{21}$$

$$\mathrm{tr}(PS) = N_\mathrm{e} \tag{22}$$

that is, symmetry, idempotency, and particle conservation. For the three approximations mentioned in the previous section, all three conditions are only fulfilled by ADMM1, i.e., the purified wave function fitting scheme. ADMM2 and the block diagonal version of ADMM3 fulfill a property of ensemble averaged (finite temperature) density matrices, i.e., that the eigenvalues of $\hat{P}$ are bounded by 0 and 1, which is a relaxed version of the idempotency condition. In order to compute a GGA correction for exchange, it is essential that the approximate density matrix is at least positive semidefinite. Fortunately, there exist purification algorithms that can restore the idempotency of an approximate density matrix.

Well known is the McWeeny purification algorithm,[30] which, in the presence of an overlap matrix, is defined as follows

$$\bar{P}_{n+1} = f(\bar{P}_n) = 3\bar{P}_n S\bar{P}_n - 2\bar{P}_n S\bar{P}_n S\bar{P}_n \tag{23}$$

for an initial guess $\bar{P}_0 = \hat{P}$. The pure density matrix is than given as

$$\tilde{P} = \lim_{n\to\infty} \bar{P}_n \tag{24}$$

An interesting property of this algorithm is that it can be implemented in a linear scaling fashion.[31,32] In the current context, we prefer an extension of the McWeeny procedure based on a Cauchy integral representation[33]

$$\tilde{P} = S^{-1}\left[\frac{1}{2\pi i}\oint \frac{\Theta(z-0.5)}{S^{-1}z - \hat{P}}\,\mathrm{d}z\right]S^{-1} \tag{25}$$

where $\Theta(z)$ denotes the Heaviside function. This scheme yields a pure density matrix for all input matrices, is noniterative, but is not easily incorporated in a linear scaling procedure. Through eq 25, a purified $\tilde{P}$ can be interpreted as a matrix functional of a nonpure $\hat{P}$. This is an important property, which will be used to derive an expression for the Kohn−Sham matrix in the following section. Equation 25 can be easily computed using basic linear algebra techniques as

$$\tilde{P} = S^{-1}RLR^TS^{-1} \tag{26}$$

where R is the matrix of eigenvectors of the generalized eigenvalue problem

$$\hat{P}R = S^{-1}R\lambda \tag{27}$$

and L is the diagonal matrix $L_{ii} = \Theta(\lambda_i - 0.5)$ with the corresponding eigenvalues λ_i. At this point, and as shown in the Appendix, we remark that purification by eq 25 of the density matrix obtained from nonpurified wave function fitting (ADMM2) exactly yields the density matrix derived from the purified wave function fitting (ADMM1).

2.4. Kohn−Sham Matrix and the SCF Procedure. In a standard SCF procedure, an improved density matrix is obtained from a diagonalization of the Kohn−Sham matrix. The Kohn−Sham matrix itself is obtained as the derivative of the total energy with respect to the density matrix. In ADMM, the total energy can be considered to consist of two parts, one part depending explicitly on the primary density matrix ($E[P]$) and one part depending explicitly on the auxiliary matrix ($\tilde{E}[\tilde{P}]$):

$$E_\mathrm{total} = E[P] + \tilde{E}[\tilde{P}] \tag{28}$$

The Kohn−Sham matrix associated with this expression

$$K_\mathrm{total} = \frac{dE[P]}{dP} + \frac{d\tilde{E}[\tilde{P}]}{dP} = K + \frac{d\tilde{E}[\tilde{P}]}{dP} \tag{29}$$

contains one nontrivial term

$$\frac{d\tilde{E}[\tilde{P}]}{dP} = \frac{d\tilde{E}}{dP} = \frac{d\tilde{E}}{d\tilde{P}}\frac{d\tilde{P}}{d\hat{P}}\frac{d\hat{P}}{dP} = \tilde{K}\frac{d\tilde{P}}{d\hat{P}}\frac{d\hat{P}}{dP} \tag{30}$$

where $\tilde{K}$ is the Kohn−Sham matrix constructed from the purified density matrix. $(d\hat{P})/(dP)$ is readily evaluated for wave function fitting and blocking, while $(d\tilde{P})/(d\hat{P})$ can be obtained through the Cauchy integral eq 25. We obtain in the case of purified wave function fitting (for details, see Appendix part B)

$$\frac{d\tilde{E}}{dP} = A^T R[(R^T \tilde{S}^{-1} \tilde{K} \tilde{S}^{-1} R) \otimes M] R^T A \tag{31}$$

with R as defined above, and

$$M_{kj} = \begin{cases} \dfrac{\Theta(\lambda_k - 0.5) - \Theta(\lambda_j - 0.5)}{\lambda_k - \lambda_j} & k \neq j \\ \delta(\lambda_k - 0.5) & k = j \end{cases} \tag{32}$$

In the Appendix, computationally more efficient expressions are presented for optimization schemes that only require the derivative of the energy with respect to the MO coefficients (dE/dC) or that exploit the special structure of $\hat{P}$.

At this point, it is important to point out that the eigenvalues of the Kohn−Sham matrix in ADMM might be very different from the eigenvalues of the Kohn−Sham matrix in the primary basis. This is not an indication of the inaccuracy of the scheme, nor is it a problem for the SCF procedure, but it is related to the fact that purification as part of the energy functional partially accounts for the orthonormality constraint of the wave function (see also Appendix part G). In order to use the eigenvalues of the ADMM Kohn−Sham matrix directly as orbital energies, e.g., to calculate the band gaps of a system, an ADMM scheme without purification needs to be employed. For the nonpurified wave function fitting (ADMM2), the corresponding Kohn−Sham matrix is given by

$$K_{\text{total}} = K[P] + A^T \hat{K} A \tag{33}$$

where $\hat{K}$ is built from $\hat{P}$. This simple expression suggests an expression for use with purified wave function fitting (ADMM1); i.e., orbital energies can be obtained from eigenvalues of

$$K_{\text{total}} = K[P] + A^T \tilde{K} A \tag{34}$$

where $\tilde{K}$ is constructed from $\tilde{P}$. We will show in section 3.6 that this expression can be accurate.

3. Assessment and Validation of the Method

3.1. Computational Details. All algorithms have been implemented in CP2K,[34] a freely available molecular simulation package. CP2K is well suited for these calculations as the density functional module Quickstep,[35] implements a linear scaling and fast scheme for calculations on the basis of local functionals. Indeed, the Gaussian and plane waves (GPW) scheme[36] and its augmented (GAPW) variant[37] provide an efficient method to evaluate the Coulomb energy for pseudopotential and all-electron calculations, respectively. These approaches use Fourier transform based techniques, i.e., a plane wave auxiliary basis, and scale favorably with basis set size. Recently, an efficient, massively parallel and linear scaling implementation of Hartree−Fock exchange has been incorporated into the CP2K code.[3,8] Despite this efficiency, calculations including HFX and employing high-quality basis sets are at least one order of magnitude more expensive than calculations based on GGAs. ADMM aims at resolving this issue. Currently, the GGA correction term required for ADMM has only been implemented for use with the GPW method, and consequently all calculations are based on Goedecker, Teter, Hutter (GTH) pseudopotentials.[38] Pseudopotentials[39] constructed for the PBE functional have been used throughout. This is an approximation that poorly describes core−valence exchange and that is known to introduce errors in excess of 0.1 eV in the computation of band gaps[40,41] but appears to give reasonable results for ground state properties (see, e.g., section 3.2). In this work, both ADMM and the standard HFX implementation employ the same pseudopotential approximation, so that a meaningful comparison can be made. The all-electron implementation of ADMM and the development of pseudopotentials for hybrid functionals are beyond the scope of the current work.

Calculations based on pseudopotentials use split valence Gaussian basis sets as discussed in ref 35, the fully contracted molecularly optimized (MOLOPT) basis sets discussed in ref 9, or a reference basis (GTH-def2-QZVP), which combines the pseudoatomic orbitals of the MOLOPT basis, with uncontracted valence, and polarization exponents of the Ahlrichs quadruple-ζ (aug-)def2-QZVP[17] basis set. The latter basis can be considered close to the basis set limit. The choice of auxiliary basis for the ADMM method will in general be dictated by accuracy and performance requirements of a particular calculation. Indeed, the gain in performance for the hybrid part of the calculation might allow for better primary basis sets, or large systems can be simulated by more aggressively using a smaller auxiliary basis. Here, we are interested in exploring the accuracy of relatively small auxiliary basis sets, of which a library of eight different basis sets per atom have been constructed. This basis employs three Gaussian exponents for the valence orbitals, optimized in atomic calculations. We will refer to this uncontracted basis, without polarization functions, as FIT3, while a contraction of this basis (to double-ζ quality) is referred to as cFIT3. In order to improve accuracy, polarization functions from the standard 6-31G** basis sets have been added, yielding pFIT3 and cpFIT3 basis sets. Finally, an augmented version has been constructed by adding a "diffuse" function (typical exponents are 0.03 for hydrogen and 0.09 for oxygen), yielding aug-FIT3, aug-cFIT3, aug-pFIT3, and aug-cpFIT3 (see Table 2).

3.2. GMTKN24 Database. The GMTKN24 database is a compilation of 24 different chemically relevant benchmarks collected and established by Goerigk and Grimme.[42,43] It is based on 1049 atomic and molecular single point energies that are combined to yield 731 relative energies. These energies can be compared to available benchmark data, derived from either theory or experiment. In order to judge the quality of a computational method using a single number, the authors defined a weighted total mean absolute deviation (WTMAD) that combines all mean absolute deviations (MADs). This convenient measure is adopted here to judge

Table 2. Cost for Using the FIT3 Basis Sets on a Cluster of 20 Water Molecules[a]

basis	$\kappa(S)$	threshold	cost [ERIs]	cost [s]
cFIT3	1.3×10^2	1.0×10^{-4}	1.8×10^7	0.08
FIT3	1.5×10^2	1.0×10^{-4}	1.6×10^7	0.10
cpFIT3	1.7×10^2	1.0×10^{-4}	1.0×10^8	0.13
pFIT3	2.3×10^2	1.0×10^{-4}	9.0×10^7	0.15
aug-cFIT3	5.5×10^4	1.0×10^{-7}	1.1×10^9	3.38
aug-FIT3	6.1×10^4	1.0×10^{-7}	1.1×10^9	4.14
aug-cpFIT3	5.9×10^4	1.0×10^{-7}	3.2×10^9	6.09
aug-pFIT3	6.4×10^4	1.0×10^{-7}	3.0×10^9	6.78

[a] For comparison and details, see Table 1.

Table 3. WTMAD and $WTMAD_{ref}$ in kcal/mol for the GMTKN24 Database and the PBE0 Functional[a]

method	PBS	ABS	WTMAD	$WTMAD_{ref}$
STD	GTH-def2-QZVP		5.0	0.0
	FIT3		15.3	10.8
	pFIT3		7.1	4.2
ADMM1	GTH-def2-QZVP	cFIT3	5.3	1.0
	GTH-def2-QZVP	FIT3	5.3 (6.1)	0.7 (1.8)
	GTH-def2-QZVP	cpFIT3	5.0	0.7
	GTH-def2-QZVP	pFIT3	4.9 (5.5)	0.5 (1.2)
ADMM2	GTH-def2-QZVP	cFIT3	5.3	1.1
	GTH-def2-QZVP	FIT3	5.3	0.8
	GTH-def2-QZVP	cpFIT3	4.9	0.7
	GTH-def2-QZVP	pFIT3	4.9	0.5

[a] Whereas WTMAD refers to the weighted mean absolute deviation with respect to experimental and theoretical benchmark results, $WTMAD_{ref}$ refers to deviations with respect to PBE0 reference results obtained using a standard HFX implementation and the high-quality GTH-def2-QZVP basis. Standard (STD) HFX calculations with the GTH-def2-QZVP, FIT3, and pFIT3 basis sets are employed to establish the quality of these basis sets as a primary basis set (PBS). Wavefunction fitting results with purification (ADMM1) and without purification (ADMM2) are provided using four different auxiliary basis sets (ABS), while the GTH-def2-QZVP has been employed as a primary basis in all of these cases. The results in parentheses have been obtained using ADMM, but ignoring the GGA correction.

the quality of the wave function fitting methods ADMM1 and ADMM2 for various basis sets. Results, summarized in Table 3, are based on the hybrid PBE0 functional[44−46] without empirical dispersion correction.[47] In a first step, reference results using a standard HFX implementation have been generated for the GTH-def2-QZVP basis. As in ref 42, an augmented basis set has been used for two of the subsets. ADMM results can be directly compared to these reference results, and deviations with respect to this data is referred to as $WTMAD_{ref}$. WTMAD without a subscript is used to refer to the deviations with respect to the experimental and theoretical benchmark results. Second, to quantify the expected poor quality of the FIT3 family as a primary basis, these basis sets have been used with a standard HFX implementation. These calculations yield a $WTMAD_{ref}$ in the range 4−11 kcal/mol and WTMADs in the range 7−15 kcal/mol, far worse than the typical performance of local functionals, with a good basis set, on this database.[42] Third, ADMM calculations have been performed using the FIT3 family as auxiliary basis sets. Whenever the primary basis is augmented, an augmented auxiliary basis has been used as well. The results obtained with ADMM are in very close agreement with the reference calculations. In particular, both ADMM1 and ADMM2 using the better auxiliary basis set (pFIT3 or cpFIT3) are basically indistinguishable in terms of error with respect to the benchmark data (WTMAD) and have an error of less than 1 kcal/mol compared to the reference run ($WTMAD_{ref}$). In the case of FIT3, ADMM results improve by 10 kcal/mol as compared to standard HFX calculations with the same basis. In Table 3, it is also shown that including the GGA correction term in ADMM more than halves the $WTMAD_{ref}$, thus emphasizing the benefit of the correction term. Finally, we observe that ADMM1 and ADMM2 perform equally well, suggesting that in this case the purification is not essential. In the cases we have verified, $\hat{P}$ had eigenvalues close to 0 and 1, even for the small cFIT3 basis. These data show that results of def2-QZVP quality can be obtained at a cost similar to 6-31G**. The relatively modest cost of computing the full database with ADMM has been exploited to benchmark the quality of the PBE0-TC-LRC functional proposed in ref 3. This functional uses a truncated operator for the calculation of exchange but, like HSE,[48,49] corrects for the long-range part using a density functional. As shown in ref 3, the PBE0-TC-LRC is useful in the condensed phase but can also reduce the computational cost for (large) molecules. In Table 4, the effect of varying the range of exchange has been studied systematically, using ADMM1 with the pFIT3 basis, for PBE0-TC-LRC functionals including 20% and 25% of nonlocal exchange.

These WTMADs clearly show that the range of the truncated operator can be reduced to 2 Å without affecting the quality of the results. The lowest WTMAD, slightly smaller than the WTMAD for PBE0, is found for 20% nonlocal exchange and a range of 2.5 Å.

3.3. Basis Set Superposition Error. In this section, the impact of ADMM on the basis set superposition error (BSSE) for the water dimer is investigated. Indeed, the BSSE is a concern as soon as small, lower-quality basis sets are employed. Here, it is shown that small auxiliary basis sets introduce only a moderate BSSE, especially if compared to the BSSE in standard HFX calculations with the same basis. In order to quantify the BSSE, the counterpoise correction[50] has been computed for a water dimer at fixed equilibrium geometry, for various methods. These results are summarized in Table 5. As expected, using the nonaugmented FIT3 basis sets as primary bases leads to errors of approximately 3 kcal/mol. This error is large when compared to a basis using

Table 4. WTMADs in kcal/mol for the GTMKN24 Database and the PBE0_TC_LRC Functional for Several Different Cutoff Radii in the Range of 0.5−6.0 Å[a]

α	0.50	0.75	1.00	1.25	1.50	1.75	2.00	2.25	2.50	3.00	4.50	6.00	∞
0.20	5.5	5.3	5.6	5.6	5.3	5.0	4.9	4.8	4.7	4.7	4.8	4.8	N/A
0.25	5.5	5.5	6.1	6.3	5.9	5.4	5.2	5.0	4.9	4.9	4.9	4.9	5.0

[a] The column labeled with ∞ refers to the standard PBE0 hybrid functional. All calculations have been performed twice for different fractions of Hartree−Fock exchange, $\alpha = 0.2$ and $\alpha = 0.25$.

Table 5. Counterpoise Corrections in kcal/mol to the PBE0 Binding Energy of a Water Dimer[a]

basis set	STD	ADMM1	ADMM2
cFIT3	−3.112	0.771	0.223
FIT3	−3.128	0.520	−0.006
cpFIT3	−3.468	0.882	0.248
pFIT3	−3.448	0.604	0.004
aug-cFIT3	−1.889	−0.193	−0.346
aug-FIT3	−1.744	−0.095	−0.253
aug-cpFIT3	−1.023	−0.246	−0.325
aug-pFIT3	−1.005	−0.162	−0.247
TZV2P-MOLOPT	−0.123	−0.123	−0.123

[a] STD refers to traditional hybrid calculations, using the shown basis set as the primary basis set. ADMM1 and ADMM2 refer to the wavefunction fitting methods, using the TZV2P-MOLOPT basis set as the primary basis and the shown basis as the auxiliary basis set.

diffuse primitives, such as the TZV2P-MOLOPT basis, which has a BSSE of only 0.1 kcal/mol. However, within the ADMM scheme, the error reduces to 0.8 kcal/mol, approximately a 4-fold reduction. Using the augmented auxiliary basis sets reduces the error to approximately 0.3 kcal/mol, similar to, but not quite as good as, the quality of the primary basis set. Note that, since an auxiliary basis set method is not necessarily variational in the auxiliary basis, the counterpoise corrections can be of both signs. This can lead to an error cancellation, which is presumably the reason why ADMM2 performs surprisingly well with the lower-quality auxiliary basis sets. It can thus be concluded that both wave function fitting methods do not suffer from the large BSSE associated with the inferior quality of the auxiliary basis even though the BSSE does not reduce to the extent of the primary basis in all cases.

3.4. H_2^+ Dissociation Curve. ADMM calculations that are GGA corrected might be biased from deficiencies of the underlying GGA functional. In order to investigate this effect, dissociation curves for H_2^+ at different levels of theory have been calculated. As is well-known, GGA functionals, such as PBE exchange, describe the dissociation of this system incorrectly.[51] Figure 1 compares results obtained from a Hartree−Fock reference calculation, which is exact for this system, with results obtained from ADMM1. The primary basis was chosen to be the same as in the reference calculation (TZV2P-MOLOPT) while several different ABSs have been applied. The results clearly show that the wave function fitting is not biased by the GGA correction. Furthermore, as shown in the inset, better quality ABSs consistently improve the description of the potential around the minimum. It can thus be concluded that the qualitatively important effects of HFX are properly retained and that the GGA correction does not introduce artifacts of the underlying functionals.

3.5. The Cationic Hole in Liquid Water. In order to probe the effect of the dual basis set approach on the electronic structure directly, the spin density distribution of the cationic hole in bulk liquid water has been computed. The poor performance of local functionals for the radical cation water dimer was discussed in detail by Sodupe et al. in ref 52 and attributed to the self-interaction error, which favors configurations with a delocalized spin density distribution. Hybrid functionals with a relatively large fraction

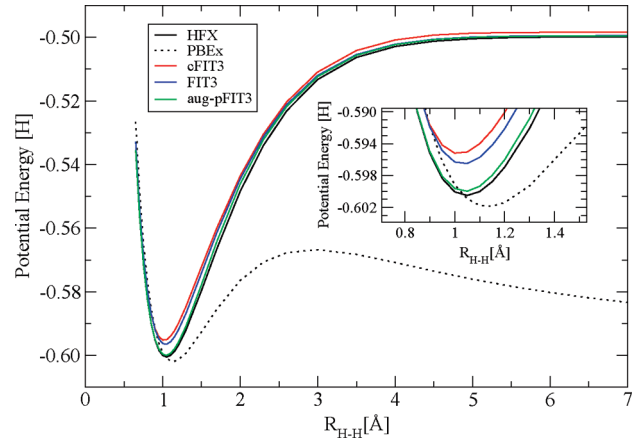


Figure 1. Dissociation curves for H_2^+ obtained from different calculations. The black solid line depicts the reference Hartree−Fock run with the TZV2P-MOLOPT basis set. Red, blue, and green lines represent results for ADMM1 for auxiliary basis sets of increasing quality, cFIT3, FIT3, and aug-pFIT3, respectively. The dotted black line shows the dissociation curve obtained from a pure GGA exchange calculation (PBEx). In the inset, a magnification of the potential energy around the minimum is presented.

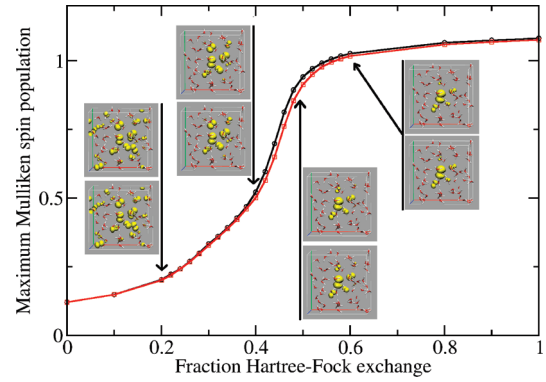


Figure 2. Localization of the spin density distribution after ionization of bulk liquid water as a function of the fraction of Hartree−Fock exchange employed in the density functional. The Mulliken spin populations of the oxygen atom on which the hole localizes is shown with a solid line, while the insets show a contour plot at 0.001 a.u. of the spin density for selected fractions (0.2, 0.4, 0.5, and 0.6) of exchange. Results obtained with the auxiliary FIT3 basis (black line and upper panels of the inserts) are almost indistinguishable from the results obtained with the primary basis only, despite the pronounced sensitivity of this system toward the use of Hartree−Fock exchange.

of exchange, for example, BH&HLYP,[53,54] perform significantly better. In ref 55, ionization of bulk liquid water was probed, and the difficulty of DFT to properly describe the electronic structure was discussed. In particular, it has been found that the electron hole, or similarly the spin density, is delocalized over the full simulation cell with local functionals, whereas it localizes on a single water molecule with Hartree−Fock exchange. Hybrid functionals with varying amounts of exchange yield intermediate degrees of localization. This is illustrated in Figure 2 for a bulk sample of liquid water (64 molecules), where the localization of the spin

Table 6. Band Gaps of Diamond As Obtained from Different Methods[a]

method	number of integrals	gap [eV]
PBE (PBS)		4.17
PBE (ABS)		4.37
PBE0 (PBS)	40 787 850 778 591	6.07
PBE0 (ABS)	23 561 509 497	6.25
PBE0 ADMM1	24 816 897 009	6.03
PBE0 ADMM2	24 795 460 638	6.02

[a] All calculations have been performed using the 3 × 3 × 3 repetition of the basic unit cell in the Γ-point approximation. For the hybrid PBE0 calculations, also the number of Cartesian integrals is shown. See the text for details on the primary and auxiliary basis set ((PBS) and (ABS)). ADMM1 is purified wavefunction fitting and ADMM2 is non-purified wavefunction fitting.

density is shown as a function of the amount of HFX in the PBE0 functional.

As a quantitative measure, the maximum value of the Mulliken spin population is reported, ranging from approximately 0.1 in the local functional to more than 1.0 in a functional containing 100% HFX. Contour plots of the spin density distribution emphasize this radical change in the electronic structure. Given this very strong dependence on the amount of Hartree−Fock exchange, this is a very stringent test for the auxiliary basis method presented in this work. Furthermore, this calculation has been performed with the relatively small FIT3 basis, i.e., without polarization functions. The results shown in Figure 2 are therefore very reassuring, since the spin distribution obtained with the auxiliary basis set approach essentially reproduces the reference density in all details for all fractions of exchange.

3.6. Diamond Band Gap. Both wave function fitting methods (ADMM1 and ADMM2) have been benchmarked with respect to their capability of predicting the band gap in diamond. The basic cubic unit cell with lattice parameter $a = 3.576$ Å containing eight carbon atoms has been extended to a large supercell in order to apply the Γ-point approximation. In a first step, PBE band gaps for supercell sizes ranging from 1 × 1 × 1 to 6 × 6 × 6 repetitions of the basic unit cell have been determined with a high-quality basis set. The band gap calculation was found to be converged for the 3 × 3 × 3 repetition, yielding a band gap of 4.17 eV in agreement with the literature.[41] This supercell has been used to calculate the PBE0 reference band gap of this system, applying the same high-quality basis set. Since the condition number of the overlap matrix with the FIT3 basis is unfavorable in the case of bulk C (1.7×10^5), an optimized FIT3 (optFIT3) basis has been constructed that served as the ABS for the two wave function fitting methods. optFIT3 was obtained by minimization of the total energy of the PBE 2 × 2 × 2 supercell with respect to the constraint of a well behaved overlap matrix (the final condition number is on the order of 10^2). This allows for rather loose screening thresholds (10^{-6}) and thus significantly reduces the amount of work in the Fock matrix construction. Results are summarized in Table 6. Both wave function fitting methods are in good agreement with the reference band gaps of the PBE0 run in the high-quality basis. In order to illustrate the cost savings, the total number of Cartesian integrals that needs to be calculated has been added to the table. The ADMM calculations are by 3 orders of magnitude more efficient than the reference PBE0 run. Both methods give very similar results, suggesting that the approximate Kohn−Sham matrix (eq 34) is a valid approximation.

3.7. Performance and Embedding for Large Systems. In order to illustrate the impact of ADMM for large systems, the electronic structure of Rubredoxin has been computed. Rubredoxin is a relatively small iron−sulfur protein that is an excellent benchmark system for electronic structure calculations, since it features an interesting active site. A realistic model including solvent and using periodic boundary conditions is comprised of only 2825 atoms and fits in a unit cell with edges 31.1 × 28.1 × 30.5 Å^3. This system has been used extensively in our earlier work. In ref 56, ab initio simulations of the full system have been combined with statistical sampling to quantify the effect of mutations on the redox potential of the active site. In ref 9, the feasibility of computing the electronic structure with accurate, molecularly optimized, basis sets has been demonstrated. In ref 3, hybrid density functional calculations using an all-electron description and a polarized triple-ζ valence basis set[57] have been performed. Molecularly optimized basis sets[9] have been employed as a primary basis for hybrid calculations (B3LYP[54,58,59]) of the same system, and the performance and accuracy of the ADMM scheme have been evaluated. Using a traditional HFX implementation with the MOLOPT basis set requires significant computational effort, despite the fact that the DZVP-MOLOPT-SR-GTH basis has been employed (22 910 basis functions), which has fewer and less diffuse primitives than the basis sets originally presented in ref 9. Indeed, the reference calculation has been run using 48 000 cores on a Cray XT5. The first SCF step required 45 min to compute 3.7×10^{14} primitive Cartesian integrals after screening with a threshold of 10^{-6}. Successive SCF steps spent only 25 s in the Hartree−Fock routines, since these calculations could be run in-core using integral compression[8] and 6.8 Tb of RAM. Due to the contracted and diffuse nature of the basis sets, this calculation is significantly more expensive than the calculations performed in ref 3. The difference in spin density between the B3LYP and a BLYP calculations is shown in the left panel of Figure 3. ADMM1 calculations using the cFIT3 basis (12 311 basis functions) require far fewer resources and have been run on 1152 cores only. The Hartree−Fock routines used 75 and 25 s in the first and successive SCF steps, respectively, and in-core operation only required 5.2 Gb of RAM. The time spent in dense linear algebra for the wave function fitting (15s, eq 16) and corresponding derivative calculation (15s, eq 101) is similar to the time spent in the HFX, suggesting that this system might benefit from linear scaling techniques for this part of the calculation. For this system, ADMM thus improves the efficiency of the calculation by a factor 20 to 1000, depending on the measure. As shown in Figure 3, the obtained spin density reproduces the reference calculation very well, even though some small differences near the Fe−S bond can be observed. To improve the accuracy, we have employed the simple embedding strategy in which the auxiliary basis for the five central atoms (Fe and S) was set

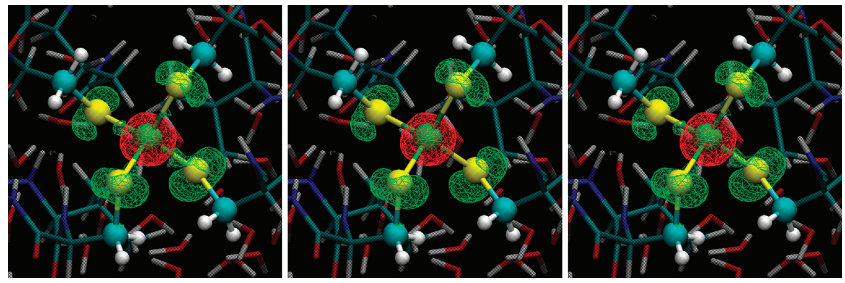


Figure 3. Isosurfaces at ±0.001 a.u. of the difference between the spin density as computed with BLYP and B3LYP for the iron−sulfur protein Rubredoxin. Left panel: Traditional calculation using only a primary basis. Middle panel: ADMM1 calculation using the cFIT3 auxiliary basis. Right panel: ADMM1 calculations using an embedding-like strategy, where the bulk of the system is described using the cFIT3 basis, but Fe and S use the primary basis as the auxiliary basis. Both ADMM calculations clearly capture the effect of Hartree−Fock exchange, a reduced delocalization of the spin density, at a small fraction of the cost of the traditional approach. The embedding strategy faithfully reproduces all details, including the change in spin density along the Fe−S bonds.

equal to the primary basis. These calculations can be performed without any significant increase in computational cost, and the right panel of Figure 3 shows that full quantitative agreement can be obtained in this way.

3.8. The Effect of Hartree−Fock Exchange on the Structure of Liquid Water. In this section, ADMM is employed to study the effect of changing the fraction of Hartree−Fock exchange in the PBE0 functional on the structure of liquid water. First, the accuracy of ADMM for describing bulk water is investigated. Second, we perform ab initio molecular dynamics simulations based on ADMM for various values of the fraction of exchange. The model system is a sample of 64 water molecules in a cubic box with edges of 12.42 Å that has previously been equilibrated using PBE0.[8] The primary basis is in all cases a TZV2P basis (2560 basis functions in total).

In order to investigate the accuracy of ADMM, the following procedure has been adopted. In a first step, a reference molecular dynamics trajectory of 2 ps starting from an equilibrated configuration has been produced. In a second step, 400 equipartitioned configurations have been chosen, and for all of these, single point ADMM calculations have been performed. The error has been quantified by computing the distribution of the difference between the reference energy and the ADMM energy. The important quantity is the variance of this difference, i.e., the energy fluctuations between the two potential energy surfaces. ADMM1 and ADMM2 have been benchmarked for various basis sets. ADMM3, which starts from a blocked density matrix, has been employed with blocked purification (eq 70) or full purification (eq 69). Nonpurified ADMM3 was found to be unstable. The subsystems have been defined as containing exactly one water molecule per block; i.e., the whole system consists of 64 diagonal sub-blocks. As shown in Figure 4, the fluctuations have approximately a Gaussian distribution. For ADMM1 and ADMM2, the associated variance gets consistently smaller when improving the quality of the auxiliary basis set. The variance for the purified wave function fitting (ADMM1) is slightly lower than the variance from nonpurified wave function fitting (ADMM2). The variance of the energy fluctuations per water molecule is below 30 micro-Hartree for all auxiliary basis sets. This variance is significantly below the variance obtained applying the same procedure with the pure density functional PBE; i.e., the difference between PBE0 and PBE is captured correctly with the ADMM1 and ADMM2 procedures. ADMM3 shows a relatively large variance, similar to direct use of the PBE functional, and its accuracy is not competitive.

With the aim of studying the effect of the fraction of exchange on the structure of the liquid, simulations employing the following functionals have been performed: PBE0 with various amounts of Hartree−Fock exchange ($\alpha \in \{0.12, 0.25, 0.37, 0.5, 0.62, 0.75, 1.00\}$), PBE, pure Hartree−Fock, PBE exchange (PBEx), and a revised parametrization[61] of PBEx (revPBEx). With these settings, trajectories longer than 30 ps have been obtained for all cases at a rate of 7 and 20 s per MD step (0.5 fs) for the pure and the hybrid functionals, respectively, on 64 cores of a Nehalem-based cluster. Compared to a standard hybrid functional calculation in the PBS without ADMM and multiple time-step MD,[8] this is a speed-up of a factor 16 per MD step. All MD simulations have been done within the isokinetic ensemble[62] at a temperature of 330 K, using ADMM1 and the FIT3 auxiliary basis. The structure has been analyzed using the oxygen−oxygen pair correlation function using the last 28 ps for each run, binning with a width of 0.03 Å. As shown in Figure 5, PBE and all variants of PBE0 yield very similar pair correlation functions. Compared to experimental results,[60] the location of the peak is correct, but the liquid is overstructured. In order to quantify the structure, the maximum value of the pair correlation function is shown in Figure 6. For PBE and all variants of PBE0, the height of the first peak falls in the range 3.45−3.75. There is no systematic trend with respect to the fraction of exchange, and the differences between the peak heights must be attributed to the limited statistics that can be collected within 30 ps for a structured liquid. Within these statistical uncertainties, these ADMM results agree with the PBE0 results obtained using traditional HFX and the same basis in ref 8, where a maximum height of 3.4 was found for PBE and PBE0 ($\alpha = 0.25$) using 7.5 ps of data. On the other hand, the liquid is significantly understructured for the pure Hartree−Fock, PBEx, and revPBEx runs. The maximum pair correlation height obtained from the Hartree−Fock simulation is in agreement with the results in ref 26, 2.34 and 2.35, respectively, where a plane wave basis set has been

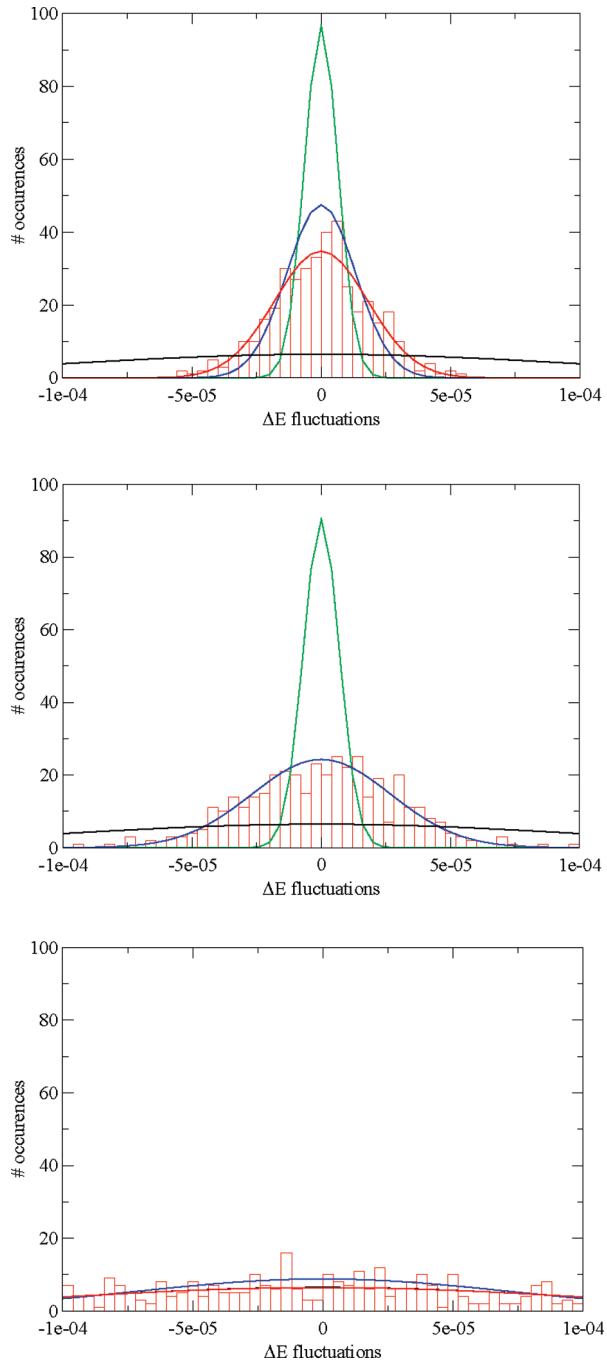


Figure 4. Centered distributions of energy differences (a.u.) between a standard PBE0 reference run and an ADMM method for bulk water. The top, middle, and bottom panels were obtained with ADMM1, ADMM2, and ADMM3, respectively. ADMM1 and ADMM2 results have been computed using cFIT3 (red), FIT3 (blue), and aug-pFIT3 (green) auxiliary basis sets. ADMM3 employs blocking on a molecular level, with blocked (red) and full purification (blue). For clarity, Gaussian distributions are shown instead of binned data, except for one data set per panel.

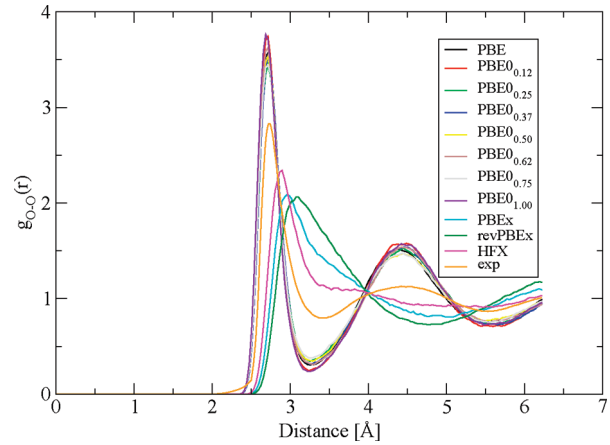


Figure 5. Oxygen−oxygen pair correlation functions as obtained for bulk liquid water, based on a sample of 64 water molecules. Seven variants of PBE0, using various amounts of Hartree−Fock exchange, and PBE almost superimpose but are overstructured as compared to the experimental result from ref 60. PBE exchange (PBEx) only, revised PBE exchange (revPBEx), and pure Hartree−Fock yield pair correlations that are similar and understructured as compared to experimental results. A detailed comparison of the maximum values of the pair correlation functions is shown in Figure 6.

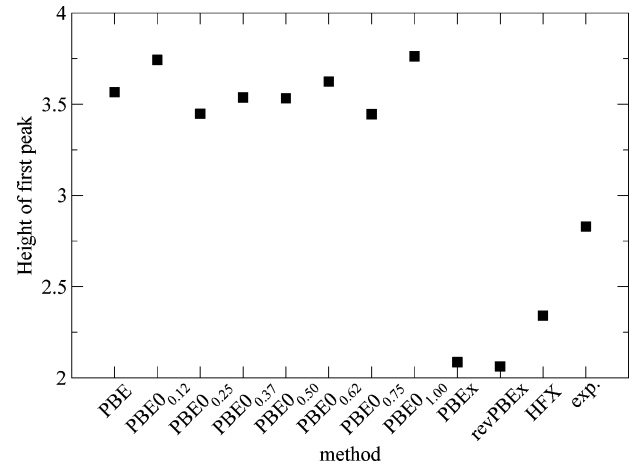


Figure 6. Height of the first peak for the oxygen−oxygen pair correlations shown in Figure 5. For the variants of PBE0, no trend in peak height with respect to the fraction of exchange can be observed.

employed. The large difference between a Hartree−Fock simulation and a PBE0 ($\alpha = 1.00$) simulation can only be attributed to the correlation functional, since all other terms in the Kohn−Sham equations are the same. Consistent with this and the observations made above, we find that the PBEx and revPBEx simulations, which do not include a correlation functional and employ a density functional to model exchange, qualitatively reproduce the Hartree−Fock simulations. The deviation between the Hartree−Fock and the revPBEx pair correlation function is somewhat larger than the deviation between the Hartree−Fock and the PBEx results.

Finally, it is important to emphasize three limitations of our simulations. First, due to the fact that the stress tensor is currently not implemented for hybrid functionals, these simulations have been performed at constant volume and not at constant pressure. In recent work, see, e.g., refs 63 and 64, it has been shown that constant volume simulations might differ significantly from constant pressure simulations for this system. Indeed, the density of the liquid, and several other macroscopic quantities,[65] most of them challenging to compute ab initio, might be more revealing about the quality

of the underlying density functional than the pair correlation function. Second, whereas the structure of the liquid for the PBE0 functional does not depend strongly on the fraction of exchange, this dependence might be different for other hybrid functionals. Third, the fact that the structure of the liquid is effectively unchanged as the fraction of exchange is varied does not imply that the properties as a solvent, i.e., the interaction of the liquid with solutes, is unchanged. In the future, the efficiency of ADMM might contribute to addressing some of these important issues.

4. Summary

We presented auxiliary density matrix methods that aim at reducing the cost of simulations based on hybrid density functionals. By constructing an approximate density matrix, which allows for a fast calculation of exchange, and by correcting the error introduced using a density functional, significant speedups have been achieved while accuracy has been retained. Wave function fitting methods that employ a small auxiliary basis to reduce the size of the density matrix appear to be a simple yet successful way to obtain an approximate density matrix. The accuracy of this approach has been investigated using a variety of tests. Calculations on the GMTKN24 database suggest that the predictivity of calculations based on wave function fitting essentially equals that of the more expensive traditional approach. Test calculations specifically aimed at difficult systems, such as BSSE calculations for the water dimer and the dissociation profile of H_2^+, have demonstrated that neither the deficiencies of the small auxiliary basis nor the correcting functional impact the quality of the results significantly. Two variants of wave function fitting, either with purification (ADMM1) or without purification (ADMM2), have been tested, and no significant differences in accuracy have been found so far. Whereas ADMM1 has the advantage of yielding a pure auxiliary density matrix, ADMM2 is particularly simple to implement and is directly suitable for a linear scaling code. ADMM3, which relies on a blocking of the density matrix, has not been tested thoroughly yet but might find its application in cases where clear subsystems, such as a solute in solution, can be easily defined. Exploiting the efficiency of the ADMM scheme, the effect of the range of exchange has been investigated for the PBE0-TC-LRC functional. The performance of this functional on the GMTKN24 database is optimal for 20% of exchange, and a range of 2.5 Å. Furthermore, ADMM has been used to perform extensive simulations of bulk water, showing that for PBE0-like functionals the amount of Hartree−Fock exchange does not directly influence the structure of the liquid. In this case, the role of correlation is more significant. Finally, a calculation on a solvated protein has been used to demonstrate that speedups in excess of a factor of 20 can be observed in actual applications.

Acknowledgment. The authors acknowledge Goerigk and Grimme for making the GMTKN24 database publically available in a convenient format. Calculations were enabled by a 2008−2010 INCITE award on the CRAY XT5 using resources of the National Center for Computational Sciences at Oak Ridge National Laboratory (ORNL), which is supported by the Office of Science of the U.S. DOE under Contract No. DE-AC05-00OR22725, and by the Swiss National Supercomputer Centre (CSCS). This work has been funded by the Swiss University Conference through the High Performance and High Productivity Computing (HP2C) Programme.

Appendix

A. Wave Function Fitting. The one-particle wave functions represented with the high-quality primary basis set (PBS) $\{\phi_\mu\}$ can be written in terms of molecular coefficients $C^{\mu i}$

$$\psi_i(\mathbf{r}) = \sum_\mu C^{\mu i}\phi_\mu(\mathbf{r}) \tag{35}$$

These wave functions are assumed to be orthonormal, i.e.

$$\int \psi_i(\mathbf{r})\,\psi_j(\mathbf{r})\,\mathrm{d}\mathbf{r} = \delta_{ij} \tag{36}$$

For the wave function fitting, a lower-quality auxiliary basis set (ABS) $\{\hat{\phi}_\mu\}$ is introduced which yields a second set of molecular coefficients $\hat{C}^{\mu i}$ and auxiliary one-particle wave functions in the following form

$$\hat{\psi}_i(\mathbf{r}) = \sum_\mu \hat{C}^{\mu i}\hat{\phi}_\mu(\mathbf{r}) \tag{37}$$

The molecular coefficients $\hat{C}^{\mu i}$ are a priori unknown but can be determined by requiring that the corresponding occupied wave functions resemble as well as possible the original ones by minimizing their square difference over all space

$$\sum_j \int (\psi_j(\mathbf{r}) - \hat{\psi}_j(\mathbf{r}))^2\,\mathrm{d}\mathbf{r} \tag{38}$$

Optionally, the auxiliary wave functions can be restricted to obey the orthonormality constraint

$$\int \hat{\psi}_i(\mathbf{r})\,\hat{\psi}_j(\mathbf{r})\,\mathrm{d}\mathbf{r} = \delta_{ij} \tag{39}$$

These two possibilities give raise to two slightly different minimization problems:

$$\min_{\hat{C}}[\sum_j \int (\psi_j(\mathbf{r}) - \hat{\psi}_j(\mathbf{r}))^2\,\mathrm{d}\mathbf{r}] \tag{40}$$

and

$$\min_{\tilde{C}}[\sum_j \int (\psi_j(\mathbf{r}) - \tilde{\psi}_j(\mathbf{r}))^2\,\mathrm{d}\mathbf{r} + \sum_{k,l}\Lambda_{kl}(\int \tilde{\psi}_k\tilde{\psi}_l - \delta_{kl})] \tag{41}$$

where in the latter case, the Lagrangian multipliers Λ_{kl} enforce condition eq 39 and the notation $\tilde{C}$ has been introduced in order to distinguish the two different sets of molecular coefficients. The overlap matrices associated with the two basis set representations are given as

$$S_{mm'} = \int \phi_m(\mathbf{r})\,\phi_{m'}(\mathbf{r})\,\mathrm{d}\mathbf{r} \text{ and } \hat{S}_{nn'} = \tilde{S}_{nn'} = \int \hat{\phi}_n(\mathbf{r})\,\hat{\phi}_{n'}(\mathbf{r})\,\mathrm{d}\mathbf{r} \quad (42)$$

In order to retain a consistent notation, $\hat{S}$ and $\tilde{S}$ have been introduced, even though both matrices are identical. Furthermore, a mixed overlap matrix Q needs to be defined that takes the overlap of both sets of basis functions into account:

$$Q_{nm} = \int \hat{\phi}_n(\mathbf{r})\,\phi_m(\mathbf{r})\,\mathrm{d}\mathbf{r} \quad (43)$$

Within this notation, the Lagrange functions associated with the two minimization problems eqs 40 and 41 can conveniently be expressed as

$$\hat{L} = \sum_j \Big(\sum_{m,m'} C_{mj} C_{m'} S_{mm'} + \sum_{n,n'} \hat{C}_{nj} \hat{C}_{n'j} \hat{S}_{nn'} - 2 \sum_{m,n} C_{mj} \hat{C}_{nj} Q_{nm} \Big) \quad (44)$$

and

$$\tilde{L} = \sum_j \Big(\sum_{m,m'} C_{mj} C_{m'j} S_{mm'} + \sum_{n,n'} \tilde{C}_{nj} \tilde{C}_{n'j} \tilde{S}_{nn'} - 2 \sum_{m,n} C_{mj} \tilde{C}_{nj} Q_{nm} + \sum_{m,n} \sum_{k,l} \Lambda_{kl} (\tilde{C}_{nj} \tilde{C}_{mk} \tilde{S}_{nm} - \delta_{kl}) \Big) \quad (45)$$

Because of eq 36, or, equivalently, $C^T S C = 1$, and due to eq 39 or $\tilde{C}^T \tilde{S} \tilde{C} = 1$ in the second case, this simplifies to

$$\hat{L} = \sum_j \Big(\sum_{n,n'} \hat{C}_{nj} \hat{C}_{n'j} \hat{S}_{nn'} - 2 \sum_{m,n} C_{mj} \hat{C}_{nj} Q_{nm} \Big) \quad (46)$$

and

$$\tilde{L} = -2 \sum_j \sum_{m,n} C_{mj} \tilde{C}_{nj} Q_{nm} + \sum_{k,l} \Lambda_{kl} (\tilde{C}_{nk} \tilde{C}_{ml} \tilde{S}_{nm} - \delta_{kl}) \quad (47)$$

respectively. From that, the unknown auxiliary molecular coefficients can be determined by taking the partial derivatives and equating them to zero. This yields

$$\frac{\partial \hat{L}}{d\hat{C}_{pq}} = -2(QC)_{pq} + 2(\hat{S}\hat{C})_{pq} \doteq 0 \quad (48)$$

and

$$\frac{\partial \tilde{L}}{d\tilde{C}_{pq}} = -2(QC)_{pq} + 2(\tilde{S}\tilde{C}\Lambda)_{pq} \doteq 0 \quad (49)$$

Thus, the final results for the MO coefficients are given by

$$\hat{C} = \hat{S}^{-1} Q C \text{ and } \tilde{C} = \tilde{S}^{-1} Q C \Lambda^{-1} \quad (50)$$

with the matrix of the Lagrangian multipliers

$$\Lambda = [(QC)^T \tilde{S}^{-1} Q C]^{1/2} \quad (51)$$

Defining $A := \hat{S}^{-1} Q = \tilde{S}^{-1} Q$ to be the projector between the PBS and ABS directly yields eqs 12 and 16 presented in section 2.2.

B. Purification. Most of the calculations that follow take advantage of the Cauchy integral theorem for matrix functions.[33] For an arbitrary matrix F, it states

$$f(F) = \frac{1}{2\pi i} \oint f(z) \frac{1}{zI - F}\,\mathrm{d}z \quad (52)$$

which, since

$$\frac{d}{dx} F^{-1} = -F^{-1} \frac{dF}{dx} F^{-1} \quad (53)$$

transforms into an explicit formula for the derivative of a matrix function

$$\frac{df(F)}{dx} = \frac{1}{2\pi i} \oint f(z) \frac{1}{F - zI} \frac{dF}{dx} \frac{1}{F - zI}\,\mathrm{d}z \quad (54)$$

Applying this formula, matrix function derivatives can be calculated through residues of its eigenvalues. Applying this to $f(x) = \Theta(z)$, where $\Theta(z)$ denotes the Heaviside function, the expression for the purified density matrix becomes

$$\tilde{P} = \hat{S}^{-1/2} \Big[\frac{1}{2\pi i} \oint \frac{\Theta(z - 0.5)}{zI - \hat{S}^{1/2} \hat{P} \hat{S}^{1/2}}\,\mathrm{d}z \Big] \hat{S}^{-1/2} \quad (55)$$

or, after some rearrangements

$$\tilde{P} = \hat{S}^{-1} \Big[\frac{1}{2\pi i} \oint \frac{\Theta(z - 0.5)}{\hat{S}^{-1} z - \hat{P}}\,\mathrm{d}z \Big] \hat{S}^{-1} \quad (56)$$

The evaluation of the contour integral can easily be performed via diagonalization. For that purpose, the following generalized eigenvalue problem needs to be solved:

$$\hat{P} R = \hat{S}^{-1} R \lambda \quad (57)$$

where R defines the matrix containing the generalized eigenvectors of $\hat{P}$. Indeed, inserting $RR^{-1} = 1$ and $(\hat{S}^{-1}R)(\hat{S}^{-1}R)^{-1}$ from left and right into eq 56 gives

$$\begin{aligned} \tilde{P} &= \hat{S}^{-1} \Big[\frac{1}{2\pi i} \oint R R^{-1} \frac{\Theta(z - 0.5)}{\hat{S}^{-1} z - \hat{P}} (\hat{S}^{-1}R)(\hat{S}^{-1}R)^{-1}\,\mathrm{d}z \Big] \hat{S}^{-1} \\ &= \hat{S}^{-1} R \Big[\frac{1}{2\pi i} \oint \frac{\Theta(z - 0.5)}{zI - D}\,\mathrm{d}z \Big] R^{-1} \\ &= \hat{S}^{-1} R \Big[\frac{1}{2\pi i} \oint \frac{\Theta(z - 0.5)}{zI - D}\,\mathrm{d}z \Big] R^T \hat{S}^{-1} \end{aligned} \quad (58)$$

where in the last step, the relation $R^T \hat{S}^{-1} R = 1$, which is valid under the assumption that $\hat{P}$ is a symmetric matrix. The integral in brackets is evaluated using the Cauchy residue theorem and can be written in terms of a diagonal matrix L. Component-wise, this yields

$$L_{ii} = \frac{1}{2\pi i} \oint \frac{\Theta(z - 0.5)}{z - \lambda_i}\,\mathrm{d}z = \mathrm{Res}\Big(\frac{\Theta(z - 0.5)}{z - \lambda_i}, z = \lambda_i \Big) = \Theta(\lambda_i - 0.5) \quad (59)$$

The final expression for $\tilde{P}$ is therefore

$$\tilde{P} = \hat{S}^{-1}RLR^T\hat{S}^{-1} \tag{60}$$

In a similar fashion, the derivative of $\tilde{P}$ with respect to $\hat{P}$, which is needed in the expression for the Kohn−Sham matrix, can be evaluated. After diagonalization of $\hat{P}$, this derivative reads

$$\frac{d\tilde{P}}{d\hat{P}} = \hat{S}^{-1}R\left[\frac{1}{2\pi i}\oint \Theta(z-0.5)\left(\frac{1}{D-zI}\right)R^{-1}\hat{S}\frac{d\hat{P}}{d\hat{P}}\times R\left(\frac{1}{D-zI}\right)\mathrm{d}z\right]R^{-1} \tag{61}$$

Again, the contour integral in brackets is computed via the Cauchy residue theorem. Since, in this case, the diagonal matrix D with the eigenvalues appears twice, the result is now a matrix M that also contains off-diagonal elements

$$\begin{aligned} M_{kj} &= \frac{1}{2\pi i}\oint \frac{\Theta(z-0.5)}{(\lambda_k - z)(\lambda_j - z)}\,\mathrm{d}z = \frac{1}{2\pi i}\oint g(z)\,\mathrm{d}z \\ &= \mathrm{Res}(g,\lambda_k) + \mathrm{Res}(g,\lambda_j) \\ &= \begin{cases} \dfrac{\Theta(\lambda_k - 0.5) - \Theta(\lambda_j - 0.5)}{\lambda_k - \lambda_j} & k \neq j \\ \delta(\lambda_k - 0.5) & k = j \end{cases} \end{aligned} \tag{62}$$

The derivative eq 61 thus becomes

$$\frac{d\tilde{P}_{cd}}{d\hat{P}_{ef}} = [\hat{S}^{-1}R(M \otimes G_{ef})R^{-1}]_{cd} \tag{63}$$

with

$$G_{ef} = R^{-1}\hat{S}\frac{d\hat{P}}{d\hat{P}_{ef}}R \tag{64}$$

This result can now be applied to the purified wave function fitting in order to obtain an expression for the Kohn−Sham matrix. In this case, $\hat{P}$ is a function of P; i.e., it holds that $\hat{P} = APA^T$, and the Kohn−Sham matrix written in terms of a derivative of the energy with respect to P is given as

$$\frac{d\tilde{E}[\tilde{P}]}{dP_{ab}} = \frac{d\tilde{E}}{d\tilde{P}_{cd}}\frac{d\tilde{P}_{cd}}{d\hat{P}_{ef}}\frac{d\hat{P}_{ef}}{dP_{ab}} = \tilde{K}_{cd}\frac{d\tilde{P}_{cd}}{d\hat{P}_{ef}}\frac{d\hat{P}_{ef}}{dP_{ab}} \tag{65}$$

where summation over same indices is assumed. The last derivative trivially amounts to

$$\frac{d\hat{P}_{ef}}{dP_{ab}} = \frac{d}{dP_{ab}}[APA^T]_{ef} = A_{e\alpha}A_{fb} \tag{66}$$

Under utilization of eq 63, this term simplifies to

$$\frac{d\tilde{E}[\tilde{P}]}{dP_{ab}} = [(A^T\tilde{S}R^{-1})[(R^T\tilde{S}^{-1}\tilde{K}\tilde{S}^{-1}R^{-T}) \otimes M]R^TA]_{ab} \tag{67}$$

Since $R^T\tilde{S}^{-1}R = 1$, this can be rewritten as

$$\frac{d\tilde{E}[\tilde{P}]}{dP} = A^TR[(R^T\tilde{S}^{-1}\tilde{K}\tilde{S}^{-1}R) \otimes M]R^TA \tag{68}$$

If $\hat{P}$ is obtained from a blocking procedure, the above expression needs to be filtered through B

$$\frac{d\tilde{E}[\tilde{P}]}{dP} = [A^TR[(R^T\hat{S}^{-1}\tilde{K}\hat{S}^{-1}R) \otimes M]R^TA] \otimes B \tag{69}$$

For the purification of the blocked density matrix, the McWeeny procedure based on the overlap matrix S can be replaced by a blocked McWeeny procedure where the overlap matrix is replaced by its blocked counterpart $S^\dagger = S \otimes B$. Equation 23 thus becomes

$$\bar{P}_{n+1} = f(\bar{P}_n) = 3\bar{P}_nS^\dagger\bar{P}_n - 2\bar{P}_nS^\dagger\bar{P}_nS^\dagger\bar{P}_n \tag{70}$$

If the matrix B is chosen to be block diagonal, the eigenvalue problem eq 27 can thus be solved within the smaller diagonal subspaces, which significantly reduces the computational workload.

C. Wave Function Fitting with and without Purification. As mentioned in section 2.3, applying the purification scheme eq 23 to the density matrix $\hat{P}$ obtained from wave function fitting without the orthonormality constraint yields exactly the density matrix $\tilde{P}$ obtained through the fitting procedure including the constraint. This can easily be seen by plugging $\hat{P}$ into the McWeeny purification algorithm. The first two iterations amount to

$$\begin{aligned} \bar{P}_1 &= 3\hat{P}\hat{S}\hat{P} - 2\hat{P}\hat{S}\hat{P}\hat{S}\hat{P} \\ &= 3\hat{C}\hat{C}^T\hat{S}\hat{C}\hat{C}^T - 2\hat{C}\hat{C}^T\hat{S}\hat{C}\hat{C}^T\hat{S}\hat{C}\hat{C}^T \\ &= 3\hat{C}\Lambda\hat{C}^T - 2\hat{C}\Lambda^2\hat{C}^T \\ &= \hat{C}(3\Lambda - 2\Lambda^2)\hat{C}^T =: \hat{C}g_1\hat{C}^T \end{aligned} \tag{71}$$

and

$$\begin{aligned} \bar{P}_2 &= 3\hat{P}_1\hat{S}\hat{P}_1 - 2\hat{P}_1\hat{S}\hat{P}_1\hat{S}\hat{P}_1 \\ &= 3\hat{C}(3\Lambda - 2\Lambda^2)\hat{C}^T\hat{S}\hat{C}(3\Lambda - 2\Lambda^2)\hat{C}^T \\ &\quad - 2\hat{C}(3\Lambda - 2\Lambda^2)\hat{C}^T\hat{S}\hat{C}(3\Lambda - 2\Lambda^2)\hat{C}^T\hat{S}\hat{C}(3\Lambda - 2\Lambda^2)\hat{C}^T \\ &= 3\hat{C}g_1^2\Lambda\hat{C}^T - 2\hat{C}g_1^3\Lambda^2\hat{C}^T \end{aligned} \tag{72}$$

Recursively, that yields

$$\bar{P}_{n+1} = \hat{C}g_{n+1}\hat{C}^T \tag{73}$$

with

$$g_{n+1} = 3g_n^2\Lambda - 2g_n^3\Lambda^2 \tag{74}$$

If the McWeeny procedure converges in the limit for $n \to \infty$, $g_n(\Lambda)$ is required to become a fixed point. Obviously this condition is fulfilled in the case of $g_n \to \Lambda^{-1}$, i.e.

$$\bar{P} = \hat{C}\Lambda^{-1}\hat{C}^T = \tilde{P} \tag{75}$$

which matches exactly the corresponding equation for the purified wave function fitting.

D. Diagonalization in the Occupied Subspace for Wave Function Fitting. As shown in Appendix part B,

in order to obtain an expression for the Kohn−Sham matrix, a general eigenvalue problem of size N_{ABS}

$$\hat{P}R = \tilde{S}^{-1}R\lambda \tag{76}$$

needs to be solved. This might become the bottleneck in an ADMM calculation because N_{ABS} is not necessarily a small quantity. However, if $\hat{P}$ can be expressed in terms of molecular coefficients, as is the case in wave function fitting, it is sufficient to diagonalize the occupied subspace only, which is typically much smaller than N_{ABS}. This can be achieved by introducing the following substitution $R \rightarrow R_\Lambda$:

$$R = \tilde{S}\hat{C}\Lambda^{-1/2}R_\Lambda = \tilde{S}\tilde{C}R_\Lambda \tag{77}$$

Equation 76 thus transforms into

$$\begin{aligned}
\hat{P}R &= \tilde{S}^{-1}R\lambda \\
\hat{C}\hat{C}^T R &= \tilde{S}^{-1}R\lambda \\
\hat{C}\hat{C}^T\tilde{S}\tilde{C}R_\Lambda &= \tilde{S}^{-1}\tilde{S}\tilde{C}R_\Lambda\lambda \\
\tilde{C}\Lambda^{1/2}\Lambda^{1/2}\tilde{C}^T\tilde{S}\tilde{C}R_\Lambda &= \tilde{C}R_\Lambda\lambda \\
(\tilde{S}\tilde{C})^T\tilde{C}\Lambda\tilde{C}^T\tilde{S}\tilde{C}R_\Lambda &= (\tilde{S}\tilde{C})^T\tilde{C}R_\Lambda\lambda \\
\Lambda R_\Lambda &= R_\Lambda\lambda
\end{aligned} \tag{78}$$

where, in the last step, the fact that $\tilde{C}\tilde{S}\tilde{C}^T = 1$ has been used. This eigenvalue problem is of the size $N_{\mathrm{mo}} \times N_{\mathrm{mo}}$, with N_{mo} being the number of occupied orbitals in the system, and therefore significantly smaller in size than the general one. Unfortunately, its solution will only provide the eigenvectors of the occupied subspace

$$R_o = \tilde{S}\hat{C}\Lambda^{-1/2}R_\Lambda \tag{79}$$

and the eigenvectors R_n for the null space are unknown. However, since $RR^T = \tilde{S}$, it follows for the decomposition into occupied and unoccupied subspaces that

$$R_nR_n^T + R_oR_o^T = \tilde{S} \tag{80}$$

which motivates the notation $R = (R_n\ R_o)$. Furthermore, the matrix M has a very characteristic structure

$$M = \begin{pmatrix} M_n & M_{no} \\ M_{on} & M_o \end{pmatrix} \tag{81}$$

with

$$M_o = M_n = \begin{pmatrix} 0 & 0 & \dots \\ 0 & 0 & \dots \\ \vdots & \vdots & \end{pmatrix},\ M_{no} = M_{on}^T = (M_{no}\ \ M_{no}\ \ M_{no}\dots) \tag{82}$$

which directly follows from its definition in eq 62. Using this decomposition, eq 68 can be rewritten in terms of occupied and unoccupied parts:

$$\begin{aligned}
\frac{d\tilde{E}[\tilde{P}]}{dP} &= A^T(R_nR_o)\left[\begin{pmatrix} R_n^T\tilde{S}^{-1}\tilde{K}\tilde{S}^{-1}R_n & R_n^T\tilde{S}^{-1}\tilde{K}\tilde{S}^{-1}R_o \\ R_o^T\tilde{S}^{-1}\tilde{K}\tilde{S}^{-1}R_n & R_o^T\tilde{S}^{-1}\tilde{K}\tilde{S}^{-1}R_o \end{pmatrix} \otimes M\right](R_n\ \ R_o)^TA \\
&= A^T(R_nR_o)\begin{pmatrix} 0 & R_n^T\tilde{S}^{-1}\tilde{K}\tilde{S}^{-1}\tilde{R}_o \\ (R_n^T\tilde{S}^{-1}\tilde{K}\tilde{S}^{-1}\tilde{R}_o)^T & 0 \end{pmatrix}(R_n\ \ R_o)^TA
\end{aligned} \tag{83}$$

with $\tilde{R}_o$ denoting the eigenvectors of the occupied subspace with the columns scaled by the vectors M_{no}. This can further be simplified, yielding an expression that only depends on the eigenvectors R_o of the occupied subsystem:

$$\begin{aligned}
\frac{d\tilde{E}[\tilde{P}]}{dP} &= A^T(R_o(R_n^T\tilde{S}^{-1}\tilde{K}\tilde{S}^{-1}\tilde{R}_o)^T\ \ R_nR_n^T\tilde{S}^{-1}\tilde{K}\tilde{S}^{-1}\tilde{R}_o)(R_n\ \ R_o)^TA \\
&= A^T(R_o\tilde{R}_o^T\tilde{S}^{-1}\tilde{K}\tilde{S}^{-1}R_nR_n^T + R_nR_n^T\tilde{S}^{-1}\tilde{K}\tilde{S}^{-1}\tilde{R}_oR_o^T)A \\
&= A^T(R_o\tilde{R}_o^T\tilde{S}^{-1}\tilde{K}\tilde{S}^{-1}(\tilde{S} - R_oR_o^T) \\
&\quad + (\tilde{S} - R_oR_o^T)\tilde{S}^{-1}\tilde{K}\tilde{S}^{-1}\tilde{R}_oR_o^T)A
\end{aligned} \tag{84}$$

An explicit expression for the products $R_o\tilde{R}_o^T$ is still required. Again, due to the special structure of M, the matrix $\tilde{R}_o$ resulting from a column scaling of R_o with M_{no} can conveniently be obtained from the eigenvalues of Λ

$$\tilde{R}_o = R_oD_\Lambda^{-1} \tag{85}$$

where D_Λ contains the eigenvalues of the matrix Λ and fulfills $R_\Lambda D_\Lambda R_\Lambda^T = \Lambda$ and similarly $R_\Lambda D_\Lambda^{-1}R_\Lambda^T = \Lambda^{-1}$, yielding

$$\begin{aligned}
R_o\tilde{R}_o^T &= \tilde{S}\hat{C}\Lambda^{-1/2}R_\Lambda(\tilde{S}\hat{C}\Lambda^{-1/2}R_\Lambda D_\Lambda^{-1})^T \\
&= \tilde{S}\hat{C}\Lambda^{-1/2}R_\Lambda D_\Lambda^{-1}R_\Lambda^T\Lambda^{-1/2}\hat{C}^T\tilde{S} \\
&= \tilde{S}\hat{C}\Lambda^{-1/2}\Lambda^{-1}\Lambda^{-1/2}\hat{C}^T\tilde{S} \\
&= \tilde{S}\hat{C}\Lambda^{-2}\hat{C}^T\tilde{S}
\end{aligned} \tag{86}$$

Inserting this result into eq 84 yields

$$\begin{aligned}
\frac{d\tilde{E}[\tilde{P}]}{dP} &= A^T(R_o\tilde{R}_o^T\tilde{S}^{-1}\tilde{K}\tilde{S}^{-1}(\tilde{S} - R_oR_o^T) \\
&\quad + (\tilde{S} - R_oR_o^T)\tilde{S}^{-1}\tilde{K}\tilde{S}^{-1}\tilde{R}_oR_o^T)A \\
&= A^T(\tilde{S}\hat{C}\Lambda^{-2}\hat{C}^T\tilde{S}\tilde{S}^{-1}\tilde{K}\tilde{S}^{-1}(\tilde{S} - R_oR_o^T) \\
&\quad + (\tilde{S} - R_oR_o^T)\tilde{S}^{-1}\tilde{K}\tilde{S}^{-1}\tilde{S}\hat{C}\Lambda^{-2}\hat{C}^T\tilde{S})A
\end{aligned} \tag{87}$$

For the last step, this equation needs to be back-transformed applying the substitution from eq 77. Since Λ is symmetric, its eigenvectors are orthonormal, i.e., $R_\Lambda R_\Lambda^T = 1$, and one finds

$$\begin{aligned}
R_oR_o^T &= \tilde{S}\hat{C}\Lambda^{-1/2}R_\Lambda R_\Lambda^T\Lambda^{-1/2}\hat{C}^T\tilde{S} \\
&= \tilde{S}\hat{C}\Lambda^{-1}\hat{C}^T\tilde{S} = \tilde{S}\tilde{P}\tilde{S}
\end{aligned} \tag{88}$$

Thus, the final expression for the Kohn−Sham matrix is given as

$$\begin{aligned}
\frac{d\tilde{E}[\tilde{P}]}{dP} &= A^T(\tilde{S}\hat{C}\Lambda^{-2}\hat{C}^T\tilde{K}(1 - \tilde{P}\tilde{S}) \\
&\quad + (1 - \tilde{S}\tilde{P})\tilde{K}\hat{C}\Lambda^{-2}\hat{C}^T\tilde{S})A
\end{aligned} \tag{89}$$

which indeed depends only on the inverse square of Λ, which has the size $N_{\mathrm{mo}} \times N_{\mathrm{mo}}$, which can be evaluated

through efficient Cholesky decomposition and does not require a diagonalization.

E. MO Derivatives. For wave function optimization algorithms that do not rely on the existence of a Kohn−Sham matrix but rather utilize the MO derivatives

$$U = \frac{dE}{dC} \tag{90}$$

such as the orbital transformation (OT) method[66] in Quickstep,[35] the explicit construction of a Kohn−Sham matrix can be omitted. This is certainly the case for the purified wave function fitting because in that case the corresponding auxiliary density matrix can be obtained from molecular coefficients. Instead of calculating the derivative of the energy with respect to the density matrix, it is thus sufficient to compute the MO derivatives

$$U_{\text{total}} = \frac{dE[P]}{dC} + \frac{d\tilde{E}[\tilde{P}]}{dC} \tag{91}$$

where only the second term is of interest here. Notice that the auxiliary density matrix $\tilde{P}$ can be expressed either in terms of purified molecular coefficients

$$\tilde{P} = \tilde{C}\tilde{C}^T \text{ with } \tilde{C} = \hat{C}\Lambda^{-1} = AC\Lambda^{-1} \tag{92}$$

or in terms of nonpurified molecular coefficients

$$\tilde{P} = \hat{C}\Lambda^{-1}\hat{C}^T \tag{93}$$

As a consequence, there exist two different approaches for calculating the desired MO derivative. The first method involves the auxiliary Kohn−Sham matrix

$$\frac{d\tilde{E}[\tilde{P}]}{dC} = \frac{d\tilde{E}}{dC} = \frac{d\tilde{E}}{d\tilde{P}}\frac{d\tilde{P}}{d\hat{C}}\frac{d\hat{C}}{dC} = \tilde{K}\frac{d\tilde{P}}{d\hat{C}}\frac{d\hat{C}}{dC} \tag{94}$$

and the second method directly takes the derivative of the energy with respect to the purified MO coefficients into account

$$\frac{d\tilde{E}[\tilde{P}]}{dC} = \frac{d\tilde{E}}{dC} = \frac{d\tilde{E}}{\tilde{d}C}\frac{d\tilde{C}}{dC} = \tilde{U}\frac{d\tilde{C}}{dC} \tag{95}$$

with

$$\tilde{U} = \frac{d\tilde{E}[\tilde{P}]}{d\tilde{C}} \tag{96}$$

The first case is algebraically straightforward, leading to

$$\frac{dE[\tilde{P}]}{dC} = 2(A^T\tilde{H}\hat{C}\Lambda^{-1}) - 2(A^T\tilde{S}\hat{C}\Lambda^{-1}\hat{C}^T\tilde{H}\hat{C}\Lambda^{-1}) \tag{97}$$

The second case is slightly more involved and requires the usage of the Cauchy integral formalism. The required derivative $(d\tilde{C})/(dC)$ involves terms such as

$$\frac{d\Lambda^{-1/2}}{dC} \tag{98}$$

which can conveniently by expressed by

$$\frac{d\Lambda^{-1/2}}{dC} = \frac{1}{2\pi i}\oint f(z)\frac{1}{\Lambda - zI}\frac{d\Lambda}{dC}\frac{1}{\Lambda - zI}\,\mathrm{d}z \tag{99}$$

with $f(z) = z^{-1/2}$. For the evaluation of the contour integral, the matrix Λ needs to be diagonalized. Using the same notation as in Appendix part D, i.e., $R_\Lambda D_\Lambda R_\Lambda^T = \Lambda$ and D_Λ defined through the eigenvalues μ_i of Λ, the resulting matrix reads

$$\begin{aligned} N_{kj} &= \frac{1}{2\pi i}\oint \frac{z^{-1/2}}{(\mu_k - z)(\mu_j - z)}\,\mathrm{d}z = \frac{1}{2\pi i}\oint g(z)\,\mathrm{d}z \\ &= \mathrm{Res}(g,\mu_k) + \mathrm{Res}(g,\mu_j) \\ &= \begin{cases} \dfrac{\mu_k^{-1/2} - \mu_j^{-1/2}}{\mu_k - \mu_j} & k \neq j \\ -\dfrac{1}{2}\mu_k^{-3/2} & k = j \end{cases} \end{aligned} \tag{100}$$

The final result for the MO derivatives is thus given by

$$\frac{d\tilde{E}[\tilde{P}]}{dC} = A^T\tilde{U}\Lambda^{-1/2} + Q^TAC(Y + Y^T) \tag{101}$$

with

$$Y = R_\Lambda([R_\Lambda^T C^T A^T \tilde{U} R_\Lambda] \otimes N)R_\Lambda^T \tag{102}$$

The first method, eq 97, has the advantage that only the inverse of Λ is needed, whereas the second method, eq 101, requires the diagonalization of Λ.

F. Analytical Ionic Forces. The derivative of the energy with respect to the atomic positions $\mathbf{R}$

$$\frac{dE}{d\mathbf{R}} = \frac{dE[P]}{d\mathbf{R}} + \frac{d\tilde{E}[\tilde{P}]}{d\mathbf{R}} \tag{103}$$

can be calculated via the MO derivatives given in eq 101. That is, for the interesting term,

$$\frac{d\tilde{E}[\tilde{P}]}{d\mathbf{R}} = \frac{d\tilde{E}[\tilde{P}]}{d\tilde{C}}\frac{d\tilde{C}}{d\mathbf{R}} = \tilde{U}\frac{d\tilde{C}}{d\mathbf{R}} \tag{104}$$

Component-wise, this yields

$$\frac{d\tilde{E}[\tilde{P}]}{d\mathbf{R}} = \tilde{U}_{ab}\left[\frac{dA}{d\mathbf{R}}C\Lambda^{-1/2}\right]_{ab} + \tilde{U}_{ab}\left[AC\frac{d\Lambda^{-1/2}}{d\mathbf{R}}\right]_{ab} \tag{105}$$

with implicit summation over repeated indices. In order to evaluate the second term, it is possible to apply the same mathematical formalism as in eq 99, and the final result reads

$$\begin{aligned}\frac{d\tilde{E}[\tilde{P}]}{d\mathbf{R}} = & -\tilde{U}_{ab}\left[\tilde{S}^{-1}\frac{d\tilde{S}}{d\mathbf{R}}\tilde{S}^{-1}QC\Lambda^{-1/2}\right]_{ab} \\ & + \tilde{U}_{ab}\left[\tilde{S}^{-1}\frac{dQ}{d\mathbf{R}}C\Lambda^{-1/2}\right]_{ab} \\ & + Y_{ab}\left[C^T\frac{dQ^T}{d\mathbf{R}}AC\right]_{ab} \\ & - Y_{ab}\left[C^TQ^T\tilde{S}^{-1}\frac{d\tilde{S}}{d\mathbf{R}}\tilde{S}^{-1}QC\right]_{ab} \\ & + Y_{ab}\left[C^TQ^T\tilde{S}^{-1}\frac{dQ}{d\mathbf{R}}C\right]_{ab}\end{aligned} \tag{106}$$

with Y_{ab} as in eq 102. This can further be simplified in terms of traces and becomes

$$\frac{d\tilde{E}}{d\mathbf{R}} = -\mathrm{tr}\left(W_{\tilde{S}}^T\frac{d\tilde{S}}{d\mathbf{R}}\right) + \mathrm{tr}\left(W_Q^T\frac{dQ}{d\mathbf{R}}\right) \tag{107}$$

with the weighted density matrices

$$W_{\tilde{S}} = \tilde{S}^{-1}\tilde{U}\Lambda^{-T/2}C^TA^T + ACYC^TA^T \tag{108}$$

and

$$W_Q = \tilde{S}^{-1}\tilde{U}\Lambda^{-T/2}C^T + ACY^TC^T + ACYC^T \tag{109}$$

G. Eigenvalues. For purified wave function fitting, the Kohn−Sham matrix obtained through the McWeeny procedure or the Cauchy integral is not suitable for the calculation of orbital energies. This problem can be illustrated by evaluating eq 89 for identical primary and auxiliary basis sets, i.e., $A = 1$, $\Lambda = 1$, $\hat{C} = C$, $\tilde{S} = S$, and $\tilde{P} = CC^T$. In that case, the eigenvalues are given as

$$\begin{aligned}C^T\frac{d\tilde{E}[\tilde{P}]}{dP}C &= C^T[A^T(\tilde{S}\hat{C}\Lambda^{-2}\hat{C}^T\tilde{K}(1-\tilde{P}\tilde{S}) + \\ &\qquad (1-\tilde{S}\tilde{P})\tilde{K}\hat{C}\Lambda^{-2}\hat{C}^T\tilde{S})A]C \\ &= C^T[SCC^T\tilde{K} - SCC^T\tilde{K}PS + \tilde{K}CC^TS - SP\tilde{K}CC^TS]C \\ &= C^T\tilde{K}C - C^T\tilde{K}PS + C^T\tilde{K}C - SP\tilde{K}C \\ &= C^T\tilde{K}C - C^T\tilde{K}C + C^T\tilde{K}C - C^T\tilde{K}C\end{aligned} \tag{110}$$

which is identically zero. Thus, eq 89 is not a suitable candidate for the calculation of the orbital energies, and a different approach needs to be taken into account. The obvious choice is to derive a similar expression as in eq 33 for the nonpurified wave function fitting, i.e., omitting the purification procedure. This assumption leads to the approximated Kohn−Sham matrix given in eq 34.

References

(1) Strout, D. L.; Scuseria, G. E. *J. Chem. Phys.* **1995**, *102*, 8448–8452.

(2) Izmaylov, A. F.; Scuseria, G. E.; Frisch, M. J. *J. Chem. Phys.* **2006**, *125*, 104103.

(3) Guidon, M.; Hutter, J.; VandeVondele, J. *J. Chem. Theory Comput.* **2009**, *5*, 3010–3021.

(4) Spencer, J.; Alavi, A. *Phys. Rev. B* **2008**, *77*, 193110.

(5) Williamson, A. J.; Rajagopal, G.; Needs, R. J.; Fraser, L. M.; Foulkes, W. M. C.; Wang, Y.; Chou, M.-Y. *Phys. Rev. B* **1997**, *55*, R4851−R4854.

(6) Kent, P. R. C.; Hood, R. Q.; Williamson, A. J.; Needs, R. J.; Foulkes, W. M. C.; Rajagopal, G. *Phys. Rev. B* **1999**, *59*, 1917–1929.

(7) Ochsenfeld, C.; White, C. A.; Head-Gordon, M. *J. Chem. Phys.* **1998**, *109*, 1663–1669.

(8) Guidon, M.; Schiffmann, F.; Hutter, J.; VandeVondele, J. *J. Chem. Phys.* **2008**, *128*, 214104.

(9) VandeVondele, J.; Hutter, J. *J. Chem. Phys.* **2007**, *127*, 114105.

(10) Binkley, J. S.; Pople, J. A.; Hehre, W. J. *J. Am. Chem. Soc.* **1980**, *102*, 939–947.

(11) Pietro, W. J.; Francl, M. M.; Hehre, W. J.; Defrees, D. J.; Pople, J. A.; Binkley, J. S. *J. Am. Chem. Soc.* **1982**, *104*, 5039–5048.

(12) Harihara, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213–222.

(13) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650–654.

(14) Jensen, F. *J. Chem. Phys.* **2001**, *115*, 9113.

(15) Jensen, F. *J. Chem. Phys.* **2002**, *116*, 7372.

(16) Jensen, F. *J. Phys. Chem. A* **2007**, *111*, 11198–11204.

(17) Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.

(18) Weigend, F. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4285–4291.

(19) Boman, L.; Koch, H.; de Merás, A. S. *J. Chem. Phys.* **2008**, *129*, 134107.

(20) Sodt, A.; Head-Gordon, M. *J. Chem. Phys.* **2008**, *128*, 104106.

(21) Liang, W.; Head-Gordon, M. *J. Phys. Chem. A* **2004**, *108*, 3206–3210.

(22) Friesner, R. A. *Chem. Phys. Lett.* **1985**, *116*, 39–43.

(23) Neese, F.; Wennmohs, F.; Hansen, A.; Becker, U. *Chem. Phys.* **2009**, *356*, 98–109.

(24) Yanai, T.; Fann, G. I.; Gan, Z.; Harrison, R. J.; Beylkin, G. *J. Chem. Phys.* **2004**, *121*, 6680–6688.

(25) Gygi, F.; Baldereschi, A. *Phys. Rev. B* **1986**, *34*, 4405–4408.

(26) Todorova, T.; Seitsonen, A. P.; Hutter, J.; Kuo, I.-F. W.; Mundy, C. J. *J. Phys. Chem. B* **2006**, *110*, 3685–3691.

(27) Sorouri, A.; Foulkes, W. M. C.; Hine, N. D. M. *J. Chem. Phys.* **2006**, *124*, 064105.

(28) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

(29) Ernzerhof, M.; Perdew, J. P. *J. Chem. Phys.* **1998**, *109*, 3313–3320.

(30) McWeeny, R. *Rev. Mod. Phys.* **1960**, *32*, 335–369.

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

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

(33) Rinehart, R. F. *Proc. Amer. Math. Soc* **1956**, *7*, 2–5.

(34) The CP2K developers group. http://cp2k.berlios.de/ (accessed June 3, 2010).

(35) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. *Comput. Phys. Commun.* **2005**, *167*, 103.

(36) Lippert, G.; Hutter, J.; Parrinello, M. *Theor. Chem. Acc.* **1999**, *103*, 124.

(37) Krack, M.; Parrinello, M. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2105–2112.

(38) Goedecker, S.; Teter, M.; Hutter, J. *Phys. Rev. B* **1996**, *54*, 1703–1710.

(39) Krack, M. *Theor. Chem. Acc.* **2005**, *114*, 145–152.

(40) Gómez-Abal, R.; Li, X.; Scheffler, M.; Ambrosch-Draxl, C. *Phys. Rev. Lett.* **2008**, *101*, 106404.

(41) Duchemin, I.; Gygi, F. *Comput. Phys. Commun.* **2010**, *181*, 855–860.

(42) Goerigk, L.; Grimme, S. *J. Chem. Theory Comput.* **2010**, *6*, 107–126.

(43) Goerigk, L.; Grimme, S. GMTKN24. http://toc.uni-muenster.de/GMTKN/GMTKNmain.html (accessed June 3, 2010).

(44) Perdew, J. P.; Ernzerhof, M.; Burke, K. *J. Chem. Phys.* **1996**, *105*, 9982–9985.

(45) Perdew, J. P.; Ernzerhof, M.; Burke, K. *Int. J. Quantum Chem.* **1997**, *64*, 285–295.

(46) Ernzerhof, M.; Scuseria, G. E. *J. Chem. Phys.* **1999**, *110*, 5029–5036.

(47) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787–1799.

(48) Heyd, J.; Scuseria, G. E.; Ernzerhof, M. *J. Chem. Phys.* **2003**, *118*, 8207–8215.

(49) Heyd, J.; Scuseria, G. E.; Ernzerhof, M. *J. Chem. Phys.* **2006**, *124*, 219906; Erratum.

(50) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553–&.

(51) Ruzsinszky, A.; Perdew, J. P.; Csonka, G. I. *J. Phys. Chem. A* **2005**, *109*, 11006–11014.

(52) Sodupe, M.; Bertran, J.; Rodriguez-Santiago, L.; Baerends, E. *J. Phys. Chem. A* **1999**, *103*, 166–170.

(53) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372–1377.

(54) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.

(55) Maršálek, O.; Elles, C. G.; Pieniazek, P. A.; VandeVondele, J.; Bradforth, S. E.; Jungwirth, P. Manuscript in preparation.

(56) Sulpizi, M.; Raugei, S.; VandeVondele, J.; Carloni, P.; Sprik, M. *J. Phys. Chem. B* **2007**, *111*, 3669.

(57) Schäfer, A.; Huber, C.; Ahlrichs, R. *J. Chem. Phys.* **1994**, *100*, 5829.

(58) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5852.

(59) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200–1211.

(60) Hura, G.; Sorenson, J. M.; Glaeser, R. M.; Head-Gordon, T. *J. Chem. Phys.* **2000**, *113*, 9140–9148.

(61) Zhang, Y.; Yang, W. *Phys. Rev. Lett.* **1998**, *80*, 890.

(62) McGrath, M. J.; Siepmann, J. I.; Kuo, I. F. W.; Mundy, C. J.; VandeVondele, J.; Hutter, J.; Mohamed, F.; Krack, M. *ChemPhysChem* **2005**, *6*, 1894–1901.

(63) 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.

(64) McGrath, M. J.; Siepmann, J. I.; Kuo, I. F. W.; Mundy, C. J.; VandeVondele, J.; Hutter, J.; Mohamed, F.; Krack, M. *J. Phys. Chem. A* **2006**, *110*, 640–646.

(65) Vega, C.; Abascal, J. L. F.; Conde, M. M.; Aragones, J. L. *Faraday Discuss.* **2009**, *141*, 251–276.

(66) VandeVondele, J.; Hutter, J. *J. Chem. Phys.* **2002**, *118*, 4365–4369.

CT1002225

JCTC Journal of Chemical Theory and Computation

Comparative Study of Selected Wave Function and Density Functional Methods for Noncovalent Interaction Energy Calculations Using the Extended S22 Data Set

Lucie Gráfová,[†] Michal Pitoňák,*[,‡,¶] Jan Řezáč,[‡] and Pavel Hobza*[,‡,§]

Department of Computing and Control Engineering, Faculty of Chemical Engineering, Institute of Chemical Technology, Technická 5, CZ-166 28 Prague, Czech Republic, Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, v. v. i. and Center for Biomolecules and Complex Molecular Systems, Flemingovo nám. 2, 166 10 Prague 6, Czech Republic, Department of Physical and Theoretical Chemistry, Faculty of Natural Sciences, Comenius University, Mlynská Dolina, 842 15 Bratislava 4, Slovak Republic, and Department of Physical Chemistry, Palacký University, 771 46 Olomouc, Czech Republic

Received April 30, 2010

Abstract: In this paper, an extension of the S22 data set of Jurecka et al. (Jurečka, P.; Šponer, J.; Černý, J.; Hobza, P. *Phys. Chem. Chem. Phys.* 2006, *8*, 1985.), the data set of benchmark CCSD(T)/CBS interaction energies of twenty-two noncovalent complexes in equilibrium geometries, is presented. The S22 data set has been extended by including the stretched (one shortened and three elongated) complex geometries of the S22 data set along the main noncovalent interaction coordinate. The goal of this work is to assess the accuracy of the popular wave function methods (MP2-, MP3- and, CCSD-based) and density functional methods (with and without empirical correction for the dispersion energy) for noncovalent complexes based on a statistical evaluation not only in equilibrium, but also in nonequilibrium geometries. The results obtained in this work provide information on whether an accurate and balanced description of the different interaction types and complex geometry distortions can be expected from the tested methods. This information has an important implication in the calculation of large molecular complexes, where the number of distant interacting molecular fragments, often in far from equilibrium geometries, increases rapidly with the system size. The best performing WFT methods were found to be the SCS-CCSD (spin-component scaled CCSD, according to Takatani, T.; Hohenstein, E. G.; Sherrill, C. D. *J. Chem. Phys.* 2008, *128*, 124111), MP2C (dispersion-corrected MP2, according to Hesselmann, A. *J. Chem. Phys.* 2008, *128*, 144112), and MP2.5 (scaled MP3, according to Pitoňák, M.; Neogrády, P.; Černý, J.; Grimme, S.; Hobza, P. *ChemPhysChem* 2009, *10*, 282.). Since none of the DFT methods fulfilled the required statistical criteria proposed in this work, they cannot be generally recommended for large-scale calculations. The DFT methods still have the potential to deliver accurate results for large molecules, but most likely on the basis of an error cancellation.

1. Introduction

The benchmark interaction energies of noncovalent complexes are extremely valuable and even indispensable data for the development and testing of new approximate methods designed to treat these applications. Of the several established

* To whom correspondence should be addressed: E-mail: pitonak@fns.uniba.sk (M.P.); pavel.hobza@uochb.cas.cz (P.H.).

[†] Department of Computing and Control Engineering, Faculty of Chemical Engineering, Institute of Chemical Technology, Technická 5.

[‡] Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, v. v. i. and Center for Biomolecules and Complex Molecular Systems.

[¶] Department of Physical and Theoretical Chemistry, Faculty of Natural Sciences, Comenius University.

[§] Department of Physical Chemistry, Palacký University.

10.1021/ct1002253 © 2010 American Chemical Society
Published on Web 07/01/2010

series of high-quality interaction energy databases, the S22 data set[1] of the Hobza group seems to be the most extensively used. In the vast majority of cases, wave function theory methods (WFT) (SCS(MI)-MP2,[2] MP2.5,[3] MP2C,[4,5] etc.), density functional theory (DFT) methods (DFT-D,[6,7] M06−2X,[8] etc.), or semiempirical methods[9−11] designed for the calculation of noncovalent interaction were either optimized toward the best performance on this data set or their performance was at least tested on the S22.[12]

Since the S22[1] was introduced in 2006, extensive research in the field of noncovalent interactions has revealed several limitations of this data set. The first of these drawbacks, the resolution of which is not the goal of this paper, originates from the selection of the benchmark method. In S22, it is the coupled-cluster singles, doubles, and perturbative triples, close to the complete basis set limit (CCSD(T)/CBS). The inaccuracies of the CCSD(T), taken as an approximation to the full configuration interaction (FCI), discussed in several papers,[13−16] are clearly not the major source of the errors. It is the incompleteness of the basis set that leads to deviations up to 0.5−0.6 kcal/mol (S22 CCSD(T)/CBS interaction energies range from −20.65 to −0.53 kcal/mol), but they are typically smaller, 0.1−0.3 kcal/mol, according to estimates from refs 16−18, confirmed by a complete reinvestigation of the S22 conducted by Sherrill.[19]

The limitation, which is of the main concern in this work, is related to the geometry of complexes in the S22. Without a single exception, the geometries of the S22 complexes correspond to the geometries of the global minima of a particular complex, in almost all cases obtained by the Møller−Plesset second-order (MP2) counterpoise-corrected gradient optimization in Dunning's cc-pVTZ basis set.[20] On the one hand, it is natural to insist on the best performance of a particular method on equilibrium complex geometries, but for real-life applications, this is certainly not sufficient. The structure of biomacromolecules is determined by the interaction of their building blocks. In the case of DNA and RNA, it is by nucleic acid bases, whereas in the case of proteins by amino acids. Only the neighboring pairs are roughly in an equilibrium geometry, while all of the other pairs are far from being in equilibrium. The so-called "clashes" in DNA,[21] or the tremendous variety of mutual amino acid residue orientations found in proteins[22] could serve as rather apparent examples. The nearest pairs obviously play the most important role, but all the other pairs contribute non-negligibly as well. The thermal effects can also cause temporary deviations of the geometry away from the equilibrium even for close contacts. Generally speaking, most of the noncovalent interactions prevalent in biological systems, in terms of distinct interaction types, are rather indirectional, or at least, have a very flat potential energy surface. The only exceptions might be the electrostatic-energy-dominated interactions, like the hydrogen or halogen bond, which have fairly strict geometrical constraints.

Interactions between nonequilibrium structures inherently take place in the calculation of large (e.g., of about hundreds of atoms and more) molecules. Presuming that the local interacting neighborhood is arranged in the "close-to-equilibrium" geometry, the middle- to long-range interaction of distant molecular fragments happens between significantly more disordered structures. Despite the fact that the contribution to the inter- or intramolecular stabilization from such interactions is much smaller in absolute values, their number increases quickly. Thus, the eventual accumulation or cancellation of errors because of such interactions certainly has to be taken into consideration. So far, these effects have not been explored enough for us to be able to estimate their magnitude accurately. A good example of such a case is, for instance, the PBE-D[23] calculations of the interaction of the hydrogen molecule with π-systems of increasing size, starting from benzene through naphthalene to the graphene surface;[24] the accuracy is enhanced upon π-system growth because of error cancellation between the inaccuracies in the equilibrium (the hydrogen molecule with the closest sp^2 carbon π-system) and the accumulated asymptotic interaction from the distant graphene surface. In even larger molecules or other large (complex) molecular systems, the mutual orientation of distant interacting fragments is expected to be almost random. Thus, an ill asymptotic behavior of a method can be expected to cancel itself out. This may not, however, be true for highly organized molecular systems (molecular crystals for instance), where the asymptotic failure of a method may accumulate and eventually propagate into an artificial (de)stabilization in a particular direction.

Finally, quickly evolving and nowadays quite feasible "on-the-fly" ab initio molecular dynamics sample large areas of the potential energy surface, far from the equilibrium. By virtue of all the arguments and examples above, it is thus clear that a database of equilibrium structures cannot properly facilitate the parametrization and testing of any new, approximate method suitable for such applications. The inclusion of nonequilibrium structures is clearly important, but certain effects, like error accumulation or cancellation in large-scale molecular complexes, cannot be simply extrapolated from their performance on small dimers. Nevertheless, the unphysical asymptotic performance indicates that the results obtained for large molecular systems should be accepted with caution, if possible validated with respect to another reliable theoretical or experimental source.

The last issue concerning the S22 database, which in our opinion deserves reinvestigation, is related to the type of interactions represented in the S22. The complexes in the S22 are divided into three groups, that is, hydrogen-bonded, dispersion-dominated and "mixed-character" interactions. This division, carried out "intuitively" by the authors of ref 1, is clear in about half of the cases (i.e., the hydrogen-bonded complexes, obviously), but in some cases it is not so clear-cut (e.g., the ethene dimer being "dispersion-dominated", whereas the ethene···ethine complexes are categorized as "mixed-character" complexes). Any statistical evaluation of performance on a method that is based on this categorization may thus lead to dubious, "general" conclusions. In this work, we will propose a more physical division, based exclusively on the highly accurate interaction energy decomposition obtained by the DFT-SAPT method.[25] In so doing, we propose a, still rough, division into the electrostatics-dominated, dispersion-dominated and "mixed-character" complexes. As discussed below, this division is still not

refined enough to distinguish even qualitatively different interaction types within the S22 (like the dispersion-dominated $\pi-\pi$ stacking interaction in the benzene dimer and the "aliphatic" dispersion in the methane dimer), but it will serve the purpose of this work better than the former intuitive division. Besides these issues, we are aware of other drawbacks of the S22 database, like its having only a few representatives of dispersion-dominated interaction other than $\pi-\pi$ stacking, electrostatics-dominated complexes being represented solely by the H-bonded complexes, the large span of the complex interactions energies (about 20 kcal/mol) leading to the eventual favoring of complexes with large interaction energies over those with small interaction energies (if the absolute, rather than relative, interaction energies are used for parametrization and testing purposes), and last of all, the fact that certain interaction types (like the cation$\cdots\pi$, halogen-bond, etc.) are not covered by the S22 at all. Such improvements are clearly beyond the scope of this article and are the subject of ongoing research in our laboratory.

In this work, we will propose the so-called "S22 × 5" database, which is essentially an extension of the S22 from ref 1. for complexes in nonequilibrium geometries. For each complex from the original S22, the S22 × 5 database contains four additional stretched structures (i.e., ×5), one (by 10%) shortened, and three elongated (by 20%, 50%, and 100%) structures along the "interaction coordinate", generated in a systematic way to respect the character of the interaction. For details, please see section 2.1. For each such structure, we have calculated an estimate of the CCSD(T)/CBS interaction energy by following the identical methodology as used in the original S22[1] (for details, again, please see section 2.1). Despite the fact that, since the date of the publication of the original S22 paper, the computational possibilities have allowed for the use of more extended basis sets (needed mostly for the CCSD(T)), we have decided to follow the original methodology for several reasons. First of all, even though the CCSD(T) calculations (being the bottleneck) are feasible nowadays in larger basis sets, calculating five different distortions for each of the twenty-two complexes would be quite time-consuming. Second, the error bars of the present methodology (stabilization energies being overestimated by about 0.5 kcal/mol, ranging in absolute values from −20.65 to −0.53 kcal/mol), are fairly accurately established either from the series of our benchmark papers on the benzene dimer,[16] uracil dimer,[17] or methyl-adenine$\cdots$methyl-thymine,[18] or by other authors.[19,26]

In parallel with our research, Molnar et al.[27] has published a similar study on the potential-energy curves obtained by a variation of the distances of the centers of mass of the monomers for twenty of the complexes from the S22. The authors of that study, however, focused more on the performance of the methods applicable to truly large-scale calculations, like the MP2 and local MP2 (with a special focus on the effect of the basis set superposition error (BSSE)) in small basis sets, DFT methods without non-local exchange and semiempirical methods. Contrary to their aims, the focus of this work is the testing of a broader range of methods, scaling as N^3-N^6 with the system's size, applied close to the complete basis set limit. The conclusions presented here should thus draw a broader picture about the "limit" accuracy of the methods, perhaps suggesting the best candidate in terms of computational demands vs accuracy ratio, without providing too much space for error cancellation arising from the basis set unsaturation or BSSE. In a sense, the paper of Molnar et al. thus can be considered complementary rather than competing.

2. Methodology

2.1. Methods and Geometries. The nonequilibrium geometries were prepared by shifting one of the monomers from the equilibrium complex geometry, without modifying the intramolecular coordinates. Both the axis along which the molecule is shifted (i.e., the interaction coordinate) and the distance from which the relative displacements 0.9, 1.0 (i.e., the original S22 geometry), 1.2, 1.5, and 2.0 were calculated are defined by selected atoms or the centers of mass of groups of atoms in each of the monomers. For instance, in hydrogen-bonded complexes, these are only the hydrogen atom involved in (one of) the hydrogen bonds of a monomer and its acceptor in the other monomer. This choice of the interaction coordinate ensures the consistent treatment of the different types of interactions.

As already mentioned, the estimates of the CCSD(T)/CBS interaction energy were calculated analogously to ref 1, that is, MP2/CBS being a two-point "$1/X^3$ " extrapolation according to Halkier et al.[28] utilizing the MP2 interaction energies in the(aug)-cc-pVXZ and (aug)-cc-pV(X+1)Z (X = T, Q, 5, depending on the complex size) basis sets, augmented by the ΔCCSD (T) higher-order correlation term from (aug)-cc-pVXZ (X = D, T, T′, Q; T′ being the cc-pVTZ basis set with f-functions removed and all but the most diffuse d-function removed). The only exception was the stacked indole$\cdots$benzene complex, in the case of which the aug-cc-pVDZ was used for calculation of the ΔCCSD(T) term because the originally used cc-pVDZ basis set was found to be strongly unsaturated.[5] The MP2C, MP2.5, and SCS-CCSD calculations were conducted using the same methodology, that is, the ΔX (X being MP2C, etc.), calculated in the same basis set as the ΔCCSD(T), combined with the MP2/CBS. The spin-component scaled MP2 methods, SCS-MP2,[29] SOS-MP2,[30,31] SOS(MI)-MP2,[2] and SCS(MI)-MP2,[2] were calculated in cc-pVTZ, except for the smallest complexes (the ammonia dimer, water dimer, methane dimer, and ethene dimer), in the case of which the cc-pVQZ basis set was used. The basis sets used for the DFT calculations (TZVP, TZVPP, and QZVP, taken from TURBOMOLE library;[32] LP = 6-311++G(3df,3pd)[33] and aDZ = aug-cc-pVDZ) are explicitly shown in the relevant tables. The frozen core approximation was applied. PBE0AC exchange-correlation functional with density fitting was used for the DFT-SAPT calculations. The basis sets used for calculation of the DFT-SAPT interaction energies and the differences (shift) between the vertical ionization potential (IP) and the highest occupied molecular orbital (HOMO) energy are specified in Table 2.

The MP2 and the DFT calculations (except for the M06-2X) were carried out using the TURBOMOLE code;[34] the

Table 1. Definition of the (Relative) Errors Used for the Statistical Evaluation of the Interaction Energies[a]

abreviation	error	equation
RMS(R)E	root-mean-square (relative)	$\sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i)^2}$
M(R)E	mean (relative)	$\frac{1}{N}\sum_{i=1}^{N}x_i$
MA(R)E	mean absolute (relative)	$\frac{1}{N}\sum_{i=1}^{N}\lvert x_i\rvert$
MAX(R)	maximum absolute (relative)	$\max_i\lvert x_i\rvert$

[a] In the case of "relative" quantities, $x_i = 100[(E_{CCSD(T)} - E_{method})/E_{CCSD(T)}]$, the error is expressed as a percentage; otherwise, $x_i = E_{CCSD(T)} - E_{method}$, where index "*i*" indicates the molecular complex and "method" stands for the name of the tested method (MP2, SCS-MP2, etc.).

M06-2X calculations were done in Gaussian;[35] and the DFT-SAPT, SCS-MP2 (and all SOS and (MI) variants), MP2C, MP2.5, and CCSD(T) calculations were carried out using the MOLPRO[36] package. The SCS-CCSD and part of the CCSD(T) calculations were run using the MOLCAS[37] suite of programs.

2.2. Statistics. For the purpose of the statistical evaluation of the method's performance, we used root-mean-square error (RMSE), mean error (ME), mean absolute error (MAE), and the maximum absolute error (MAX), applied on the difference of the interaction energy obtained by a particular method and the CCSD(T)/CBS interaction energy. Analogously, we have used the corresponding "relative" quantities, root-mean-square relative error (RMSRE), mean relative error (MRE), mean absolute relative error (MARE), and the maximum absolute relative error (MAXR), that is, applied to the percentual error of the interaction energy, as obtained by a particular method and the CCSD(T)/CBS. For more details on the definition of these quantities, please see Table 1. The reason for using the relative errors in statistics is straightforward. The S22 itself, as already mentioned, has a large span of the absolute values of the interaction energies, and in addition, the interaction energies change significantly upon the stretching of the complex geometries. In so doing, all of the complexes in all of the geometries, regardless of their absolute interaction energy, contribute equally to the statistics.

To analyze several, qualitatively different, aspects of the method performance, we have proposed five criteria. These criteria will be further referred to as "the requirements" (i.e., RQ), RQ_1−RQ_5, because the method will be considered to "score" for a particular requirement, whether the underlying criteria are below the predefined threshold or fall into the predefined error bars.

- RQ_1: The MARE for all of the interaction types and geometry distortions is below a certain threshold. This is the simplest measure of the general accuracy of a method for all the database structures on a relative scale. It is important to point out that the accuracy on a "relative scale" may seem to be stricter than accuracy on an absolute scale because of the contribution from all of the complexes (even with the smallest interaction energy). One can argue, however, that a 50% relative error on an interaction energy of 0.05 kcal/mol is irrelevant, but, in our opinion, the relative scale is the only transferable measure for complexes of different sizes and interaction strengths.
- RQ_2: The MARE of each category of interaction type (i.e., electrostatics-dominated, dispersion-dominated and mixed-character) for a particular geometry distortion should be similar. This criterion measures the uniform accuracy of a method for different interaction types for a particular geometry distortion. "Uniform accuracy" means that it should provide similarly large mean absolute relative errors for all of the interaction types.
- RQ_3: The MARE of each geometry distortion (i.e., 1.0, 1.2, 1.5, and 2.0) of a particular interaction type should be similar. This criterion measures the uniform accuracy of a method for different geometrical distortions for a particular interaction type.
- RQ_4: Close RMSRE and MARE values for a particular interaction type and geometry distortion. This criterion measures the uniform accuracy for a particular interaction type and geometry distortion. The difference between the RMSRE and MARE is very sensitive to the presence of a value that is clearly outlying from the average MARE. Please note that this requirement can be fulfilled not only if the method performs equally well but also if it performs equally poorly for the whole subgroup.
- RQ_5: The MRE for all the interaction types and geometry distortions should be close to zero. This criterion measures the overall balance of the description, thus also a tendency to over- or underestimate the interaction energy.

For each requirement, we have introduced two ad hoc thresholds, that is, strict and lenient, which in our opinion reflect the typical expectations from an accurate method for noncovalent interaction calculation. For RQ_1, it is a MARE below 10% for a strict and 20% for a lenient threshold; for RQ_2 and RQ_3, it is a MARE within ±3% (±6%) in a strict (lenient) threshold; for RQ_4, the RMSRE being less than 3% (6%) larger than the MARE in the strict (lenient) threshold; and finally, for the RQ_5, it is the MRE being within ±3% (±6%) off from zero for the strict (lenient) threshold.

Since it is more natural to think in absolute rather than relative errors, we can demonstrate the impact of these thresholds for a model complex with an interaction energy of, for instance, −20 kcal/mol. In this case, applying the strict thresholds, the fulfillment of RQ_1 leads to an mean absolute (unsigned) error of 2 kcal/mol, the fulfillment of RQ_3 means that the mean absolute error on each geometry distortion (including all or the individual interaction types) differs from the others by less than 0.6 kcal/mol, etc.

Table 2. Distinction between the Complexes As Interaction Types, E, Electrostatics-Dominated, D, Dispersion-Dominated, M, Mixed-Character, According to the DFT-SAPT Analysis[a]

			geometry distortion				
	DFT-SAPT	complex	0.9	1.0	1.2	1.5	2.0
H-bonded	aQZ (QZVP)	ammonia dimer	E	E	E	E	E
	aQZ (QZVP)	water dimer	E	E	E	E	E
	aQZ (QZVP)	formic acid dimer	E	E	E	E	E
	aQZ (QZVP)	formamide dimer	E	E	E	E	E
	aTZ (QZVP)	2-PO···0.2-PY	E	E	E	E	E
	aDZ (TZVP)	uracil dimer H	E	E	E	E	E
	aDZ (TZVP)	adenine···thymine WC	E	E	E	E	E
dispersion dominated	aQZ (QZVP)	methane dimer	D	D	D	D	D
	aQZ (QZVP)	ethene dimer	D	D	D	D	M
	aTZ (QZVP)	benzene···methane	D	D	D	D	D
	aTZ (QZVP)	benzene dimer S	D	D	D	D	M
	aTZ (QZVP)	pyrazine dimer	D	D	D	M	M
	aTZ (QZVP)	uracil dimer S	E	M	M	M	M
	aDZ (TZVP)	indole···benzene S	D	D	D	D	D
	aDZ (TZVP)	adenine···thymine S	M	M	M	M	M
Mixed character	aQZ (QZVP)	ethene···ethyne	M	M	M	M	M
	aQZ (QZVP)	benzene···water	E	M	M	M	M
	aTZ (QZVP)	benzene···ammonia	M	M	M	M	M
	aTZ (QZVP)	benzene···HCN	E	E	M	M	M
	aTZ (QZVP)	benzene dimer T	M	D	D	D	M
	aTZ (TZVP)	indole···benzene T	M	M	M	M	M
	aTZ (QZVP)	phenole dimer	M	M	M	M	M
	count	electrostatics	10	8	7	7	7
		dispersion	6	7	7	6	3
		mixed	6	7	8	9	12

[a] The DFT-SAPT column shows the basis sets used for the DFT-SAPT and HOMO energy calculations and for calculations of the vertical IP (in parentheses). aXZ stands for Dunning's aug-cc-pVXZ basis sets, and TZVP[40] and QZVP[41] stand for basis sets of Ahlrichs and coworkers.

3. Results and Discussion

3.1. Interaction Types. DFT-SAPT interaction energy decomposition was used to evaluate the character of the interaction for each complex and geometry distortion. As already mentioned, in this work we considered three distinct interaction types: electrostatic-energy-dominated, dispersion-dominated and the mixed-character. A complex in a particular geometry is assigned to, for instance, the electrostatics-dominated type, if (the absolute value of) the electrostatic component (i.e., $E^{(1)}$) of the DFT-SAPT interaction energy is at least twice as large as the dispersion component ($E^{(2)}_{\mathrm{disp}} + E^{(2)}_{\mathrm{disp-exchange}}$), and vise versa for the dispersion-dominated category. The complexes for which neither of these is true were assigned to the mixed-character interaction type group. The result of the SAPT-DFT analysis, along with the specification of the basis sets used for the DFT-SAPT and the monomer DFT shifts employed, are shown in Table 2. The arrangement of rows in the table is the same as in the original S22 paper. As clearly seen from the column corresponding to the geometry distortion 1.0, that is, the equilibrium S22 geometries, the number of the complexes in each category, 8/7/7, for electrostatics (E)/dispersion (D)/mixed (M), is almost identical to the count of hydrogen-bonded/dispersion/mixed in the original paper, that is, 7/8/7, but they are organized slightly differently (e.g., stacked uracil dimer and adenine···thymine dimer being "mixed" now, but T-shaped benzene dimer being, on the contrary, dispersion-dominated). However, upon complex dissociation, this ratio changes dramatically and in the distance elongated by 100%, that is, 2.0, the ratio is 7/3/12. This rather anticipated consequence of the different asymptotic scaling of various interaction energy components, however, strongly influences the statistical evaluation of the method's accuracy because of the reorganization of the complexes between the different interaction type categories.

As will be analyzed below in more detail, the performance of almost all of the WFT/DFT methods for the 0.9 distortion is much worse when compared to the equilibrium and elongated structures. Including this distortion would thus completely contaminate the simultaneous statistics over all of the geometry distortions. That is the reason that in all of the following tables, except for those where it is clearly stated, the 0.9 geometry distortion has been excluded from the statistics. This finding is surprising and at the same time alarming. Since these distorted geometries may occur in biomacromolecules (e.g., tight clashes in DNA), we are going to investigate this feature in more detail in the near future.

Because of space limitation, we have selected only a few statistical markers, some of which are presented only in a qualitative way (e.g., "yes/no" instead of an exact value, etc.). Exact numerical values and several other statistical evaluations can be found in the Supporting Information.

3.2. WFT Methods. In Tables 3−8, we present the statistical analysis of the most frequently used WFT methods for noncovalent interaction energy calculations (ordered according to scaling of the computation requirements with the system size): SOS-MP2, SOS(MI)-MP2 (scaling N^4), MP2, SCS-MP2, SCS(MI)-MP2, MP2C (scaling N^5), MP2.5 (scaling N^6, noniterative), and SCS-CCSD (scaling N^6, iterative). More methods can be found in the Supporting

Table 3. WFT Methods: A Statistical Evaluation of All of the S22 × 5 Complexes

	RQ_1[a]		RQ_2		RQ_3		RQ_4		RQ_5			
method/basis	10%	20%	±3%	±6%	±3%	±6%	3%	6%	±3%	±6%		score[b] (%)
MP2C	yes	yes	yes	yes	yes	yes	yes	yes	no	yes		90
SCS-CCSD	yes	yes	yes	yes	yes	yes	yes	yes	no	yes	(−)[c]	90
MP2.5	yes	yes	no	yes	yes	yes	yes	yes	no	no	(+)	70
SCS(MI)-MP2	no	yes	no	no	no	yes	no	yes	no	no		30
SCS-MP2	no	no	no	no	no	no	yes	yes	no	no	(−)	20
SOS-MP2	no	no	no	no	no	no	no	yes	no	no	(−)	10
MP2	no	no	no	no	no	no	no	no	no	no	(+)	0
SOS(MI)-MP2	no	no	no	no	no	no	no	no	no	no	(−)	0

[a] RQ_1, a small (<10% or <20%) mean absolute relative error (MARE) for all complexes and geometry distortions; RQ_2, a small (<±3% or <±6%) MARE variation for all of the investigated complexes for a particular geometry distortion; RQ_3, a small (<±3% or <±6%) MARE variation within a group of complexes of a particular interaction type for all of the geometry distortions; RQ_4, a balanced description within a group of complexes of a particular interaction type and geometry distortion evaluated by the small (<±3% or <±6%) difference between the MARE and root-mean-square relative error (RMSRE); RQ_5, a small (<±3% or <±6%) deviation of the mean relative error (MRE) from zero, indicating systematic over- or understabilization of all of the complexes and geometry distortions. For a more detailed definition and reasoning of these restrictions, see section 2. [b] The score is defined as the number of positive (yes) matches of a particular (sub)requirement, divided by the total number of requirements applied (ten in this case). [c] (+) means overstabilization; (−) means understabilization. If no symbol is provided, the method could not be unambiguously assigned to any category.

Information. Allow us first to analyze the results shown in Tables 3−6. In these tables, we present the "score" of each method and the requirement with the strict and the lenient threshold for the whole of S22 × 5 (Table 3) and for each interaction type separately (Tables 4−6). For easier orientation in each table, a horizontal line has been drawn separating the methods that reached a score above 70% (if any), which in our opinion represent methods capable of proper description of noncovalent complexes. The score is calculated as the number of fulfilled requirements, both strict and lenient, divided by the total number of requirement/threshold columns, converted to percents. Methods having the same score are ordered alphabetically.

In Table 3, the statistics for all of the categories of complexes are shown. The best performing methods, which in the spirit of the motivation of this work are expected to perform well for the calculation of biomacromolecules, are the MP2C, MP2.5. and SCS-CCSD, scoring at least to 70%. The methods performing the best were MP2C and SCS-CCSD with a score of 90%, which fully satisfies all of the requirements (four requirements with the strict and one, RQ_5, with the lenient threshold).

The most demanding requirement (not only for the entire S22 × 5 but also for different interaction types separately) was clearly RQ_5, that is, the criterion measuring the overall balance of the description across all of the interaction types and geometrical distortions (in other words, the tendency to over- or underestimate the interaction energy). None of the methods fulfilled the strict threshold for this requirement; among the best scoring methods, the MP2.5 did not even meet the lenient one. In Table 3, in the RQ_5 column, there is a sign indicating whether the method has a tendency to overstabilize (+) or understabilize (−) the complexes systematically. The best performing methods did not show the same tendency of these characteristics. The SCS-CCSD had a tendency to underestimate the interaction energies systematically, unlike the MP2.5, which had a tendency to overestimate the interaction energies systematically. The MP2C provided unbalanced description without an equable error orientation.

What is slightly surprising is the performance of the SCS(MI)-MP2 method, which was parametrized for the S22 test set and did not fulfill any of the requirements with the strict threshold. The performance of the SOS-MP2 and especially SOS(MI)-MP2 (parametrized for the S22 test set as well), confirmed the inability of the sole opposite spin contribution to describe the noncovalent interactions correctly. The zero score of the MP2 method is no surprise because of its known strong overestimations of the $\pi-\pi$ stacked complexes, which clearly is in conflict with four of our five requirements. The SCS-MP2 provided a poor score (20%), but on the other hand, surprisingly fulfilled the RQ_4 with the strict threshold, which means that the method is capable of delivering balanced accuracy for all of the interaction types and geometrical distortions.

We should now analyze the performance of the WFT methods for the different interaction types separately. Please notice that the score is now evaluated according to four/eight requirements/thresholds, since RQ_2 is not applicable here.

Exceptionally good performance was found for the electrostatics-dominated complexes (see Table 4). Except for the SOS/SCS-MP2 and SOS(MI)-MP2, all of the methods reached a score clearly above 70% (four methods even had a score of 100%) and thus can be safely recommended for the calculation of electrostatics-dominated noncovalent interactions (e.g., H-bonded complexes). This is, however, not so surprising, since already the MP2 method is known to describe the electric properties of molecules (dipole moment, quadrupole moment, etc.) fairly accurately. All of the methods fulfill requirement RQ_4 even with the strict threshold (i.e., the criteria measuring the uniform accuracy for a particular interaction type and geometry distortion), which might be a consequence of the fact that the electrostatics-dominated complexes are represented solely (except for the benzene···HCN) by the same "chemical" type of interaction, that is, hydrogen bond.

The performance of the WFT methods on the dispersion-dominated complexes, Table 5, is the worst. The reason is that the complexes in this category are mostly of the $\pi-\pi$

Table 4. WFT Methods: A Statistical Evaluation of the Electrostatics-Dominated Complexes of the S22 × 5 Set

	RQ_1[a]		RQ_3		RQ_4		RQ_5			
method/basis	10%	20%	±3%	±6%	3%	6%	±3%	±6%		score[b] (%)
MP2	yes	yes	yes	yes	yes	yes	yes	yes		100
MP2.5	yes	yes	yes	yes	yes	yes	yes	yes		100
MP2C	yes	yes	yes	yes	yes	yes	yes	yes		100
SCS(MI)-MP2	yes	yes	yes	yes	yes	yes	yes	yes		100
SCS-CCSD	yes	yes	yes	yes	yes	yes	no	yes	(−)[c]	88
SCS-MP2	no	yes	no	yes	yes	yes	no	no	(−)	50
SOS(MI)-MP2	no	yes	no	yes	yes	yes	no	no	(−)	50
SOS-MP2	no	no	no	no	yes	yes	no	no	(−)	25

[a] For a brief definition of requests, RQ_1, RQ_3, ..., RQ_5, see the footnote of Table 3. For a more detailed definition and reasoning, see section 2. [b] The score is defined as the number of positive (yes) matches of a particular (sub)requirement, divided by the total number of requirements applied (eight in this case). [c] (+) means overstabilization; (−) means understabilization. If no symbol is provided, the method could not be unambiguously assigned to any category.

Table 5. WFT Methods: A Statistical Evaluation of the Dispersion-Dominated Complexes of the S22 × 5 Set

	RQ_1[a]		RQ_3		RQ_4		RQ_5			
method/basis	10%	20%	±3%	±6%	3%	6%	±3%	±6%		score[b] [%]
SCS-CCSD	yes	yes	yes	yes	yes	yes	yes	yes		100
MP2C	yes	yes	yes	yes	yes	yes	no	yes		88
MP2.5	yes	yes	yes	yes	yes	yes	no	no	(+)[c]	75
SCS(MI)-MP2	no	yes	no	yes	no	yes	no	no	(−)	38
SCS-MP2	no	no	no	no	yes	yes	no	no	(−)	25
SOS-MP2	no	no	no	no	no	yes	no	no		13
MP2	no	no	no	no	no	no	no	no	(+)	0
SOS(MI)-MP2	no	no	no	no	no	no	no	no		0

[a] For a brief definition of requests RQ_1, RQ_3, ..., RQ_5, see the footnote of Table 3. For a more detailed definition and reasoning, see section 2. [b] The score is defined as the number of positive (yes) matches of a particular (sub)requirement, divided by the total number of requirements applied (eight in this case). [c] (+) means overstabilization; (−) means understabilization. If no symbol is provided, the method could not be unambiguously assigned to any category.

Table 6. WFT Methods: Statistical Evaluation of the Mixed-Character Complexes of the S22 × 5 Set

	RQ_1[a]		RQ_3		RQ_4		RQ_5			
method/basis	10%	20%	±3%	±6%	3%	6%	±3%	±6%		score[b] (%)
MP2C	yes	yes	yes	yes	yes	yes	yes	yes	(+)[c]	100
SCS-CCSD	yes	yes	yes	yes	yes	yes	yes	yes	(−)	100
MP2.5	yes	yes	yes	yes	yes	yes	no	yes	(+)	88
SCS(MI)-MP2	yes	yes	yes	yes	yes	yes	no	yes		88
SOS(MI)-MP2	no	yes	yes	yes	yes	yes	no	no		63
MP2	no	yes	yes	yes	no	yes	no	no	(+)	50
SCS-MP2	no	no	no	no	yes	yes	no	no	(−)	25
SOS-MP2	no	no	no	no	yes	yes	no	no	(−)	25

[a] For a brief definition of requests RQ_1, RQ_3, ..., RQ_5, see the footnote of Table 3. For a more detailed definition and reasoning, see section 2. [b] The score is defined as the number of positive (yes) matches of a particular (sub)requirement, divided by the total number of requirements applied (eight in this case). [c] (+) means overstabilization; (−) means understabilization. If no symbol is provided, the method could not be unambiguously assigned to any category.

stacking type, which are known to be the most challenging to describe in terms of the level of the WFT sophistication. The performance of the methods essentially followed the performance for the whole S22 × 5 test set, perhaps with only minor differences. The methods capable of the proper description of dispersion-dominated complexes are the SCS-CCSD, MP2C, and MP2.5, scoring above 70%. The best-performing method was the SCS-CCSD, which fulfilled all of the requirements with the strict thresholds.

The description of mixed-character complexes (see Table 6) was almost as good as that of the electrostatics-dominated ones. The methods that scored above 70% and can thus be safely used for the proper description of this class of complexes were the MP2C, SCS-CCSD (both with a score of 100%), MP2.5, and SCS(MI)-MP2. The biggest difference was in the performance of the MP2 method, which satisfied all the requirements with strict thresholds for the electrostatics-dominated complexes but now reached only a 50% score. Its most significant drawback was in its slightly unbalanced description upon the interaction-coordinate elongation (RQ_4) with a tendency to overstabilize these complexes, as indicated by requirement RQ_5.

Table 7 summarizes the absolute and relative ME, MAE, and MAX values for all of the interaction types and individual geometry distortions. In this table, the values of these statistical parameters are explicitly shown also for distortion 0.9. It is clear that the values for the 0.9 distortion do not follow the same trends as for other distortions, being almost all above 100% (or below −100%) on the relative scale. It is only the SCS-CCSD method that performs reasonably well for this geometry distortion. From the data in the MAE and MAX columns, it can be observed that these

Table 7. WFT Methods[a]

method/basis	0.9			1.0			1.2			1.5			2.0		
MP2	1.34	1.35	5.46	0.81	0.82	3.46	0.32	0.37	1.48	0.10	0.14	0.48	0.03	0.04	0.12
	(>−100)[b]	(>100)	(>100)	(−17.89)	(18.33)	(79.20)	(−10.61)	(11.09)	(44.55)	(−9.94)	(10.39)	(48.99)	(−13.48)	(13.91)	(>100)
SCS-MP2	−2.25	2.25	5.88	−1.63	1.63	3.78	−0.85	0.85	2.30	−0.33	0.33	1.11	−0.11	0.11	0.47
	(>100)	(>100)	(>100)	(27.00)	(27.00)	(52.90)	(17.59)	(17.59)	(35.49)	(14.10)	(14.1)	(30.38)	(12.77)	(12.77)	(36.59)
SOS-MP2	−3.29	3.29	8.97	−2.32	2.32	5.69	−1.17	1.17	2.77	−0.45	0.45	1.29	−0.14	0.14	0.52
	(>100)	(>100)	(>100)	(40.22)	(40.22)	(69.37)	(25.64)	(25.64)	(46.65)	(20.72)	(20.72)	(41.25)	(19.85)	(19.85)	(78.37)
SOS(MI)-MP2	−1.02	1.78	5.59	−0.92	1.34	4.53	−0.63	0.77	3.11	−0.32	0.36	1.71	−0.14	0.15	0.85
	(−67.94)	(>100)	(>100)	(7.86)	(18.68)	(45.82)	(7.83)	(12.89)	(23.8)	(7.05)	(11.39)	(23.25)	(3.60)	(16.00)	(90.35)
SCS(MI)-MP2	0.05	0.43	1.12	0.00	0.30	0.78	−0.01	0.13	0.45	0.03	0.06	0.25	0.02	0.03	0.11
	(−0.15)	(27.14)	(>100)	(2.57)	(7.98)	(41.66)	(1.03)	(4.88)	(25.63)	(−0.34)	(4.66)	(20.33)	(−1.29)	(5.68)	(18.83)
MP2C	0.04	0.30	1.31	0.06	0.19	0.83	0.02	0.10	0.40	−0.01	0.04	0.21	−0.01	0.02	0.13
	(4.80)	(20.46)	(>100)	(−2.82)	(4.11)	(11.19)	(−1.68)	(2.56)	(5.98)	(−0.88)	(1.92)	(4.36)	(0.09)	(1.77)	(4.89)
MP2.5	0.26	0.29	0.85	0.16	0.18	0.55	0.06	0.09	0.26	0.02	0.04	0.13	0.01	0.02	0.06
	(−30.29)	(32.12)	(>100)	(−3.72)	(4.45)	(16.21)	(−2.40)	(2.93)	(9.43)	(−2.56)	(3.00)	(9.4)	(−3.45)	(3.77)	(16.71)
SCS-CCSD	−0.25	0.27	1.05	−0.20	0.21	0.79	−0.12	0.12	0.47	−0.06	0.06	0.23	−0.02	0.02	0.09
	(1.52)	(4.71)	(10.84)	(1.71)	(2.42)	(4.31)	(1.5)	(1.86)	(3.61)	(1.36)	(1.73)	(3.77)	(1.20)	(1.75)	(5.22)

[a] The values in rows: in the first line of each row, ME, MAE, and MAX (in kcal/mol), in the second line, the relative quantities, MRE, MARE, and MAXR (in %) for all of the interaction types and particular geometry distortion. The absolute and relative MAX(R) in a particular table cell do not necessarily correspond to the same complex. [b] The values above 100% (or below −100%) are for simplification shown as ">100". The actual values can be found in the Supporting Information.

Table 8. WFT Methods[a]

method/CBS	$(CH_4)_2$	$(C_6H_6)_2$	Δ^b
SCS-CCSD	1.95	1.34	0.61
SOS-MP2	48.88	47.52	1.36
MP2C	4.32	1.71	2.61
MP2.5	2.93	11.38	−8.45
SOS(MI)-MP2	20.00	28.67	−8.68
SCS-MP2	36.66	24.93	11.73
SCS(MI)-MP2	26.61	3.87	22.75
MP2	1.70	54.91	−53.22

[a] The averaged absolute relative errors (AARE, in %) over all (except 0.9) geometry distortions for the methane and benzene dimers. [b] AARE(methane dimer)−AARE(benzene dimer).

"absolute-value-based" errors are far less important as compared to our statistics based on RQ_1 to RQ_5. According to the results in Table 3, the SCS-CCSD is a much better performing method than the SCS(MI)-MP2, despite the fact that both have almost the same MAE and MAX values for almost all of the geometrical distortions.

The last note in this section concerns the incapability of our statistical model to distinguish between certain interaction types, namely the $\pi-\pi$ stacking and aliphatic dispersion-dominated complexes. In Table 8, the absolute relative errors (ARE) for the methane and benzene dimers, averaged over the distortions 1.0−2.0 are shown. The last column in the table shows the difference between these two numbers, thus reflecting the balance of the simultaneous description of these two complexes. The first three methods, that is, SCS-CCSD, SOS-MP2, and MP2C, deliver exceptionally balanced descriptions, with the difference being below 3%. In the case of the SOS-MP2 method, it is only a consequence of error compensation, because the average ARE for both complexes are almost 50% off. This method thus cannot be expected to deliver a both accurate and balanced description of the dispersion-dominated complexes. Nevertheless, the most curious performance is found for the MP2 method. It is notoriously known to overestimate the stability of the $\pi-\pi$ stacked complexes (by about 55% are for benzene dimer), but its excellent performance for the methane dimer (1.7% ARE, that is, the most accurate of all the tested methods) is indeed surprising.

3.3. DFT Methods. A statistical assessment of the performance of the DFT methods is shown in Tables 9−14. We considered the following functionals (ordered according to scaling of the computation requirements with the system size): BLYP, PBE, TPSS (scaling N^3), M06-2X (scaling N^4), B2PLYP (scaling N^5), with and without empirical correction for the dispersion energy (if applicable). The M06-2X interaction energies were calculated both in "Fine" and "UltraFine" integration grid, as according to the GAUSSIAN options, with no noteworthy improvement, as could be hoped for according to ref 38.

When comparing the DFT results for all of S22 × 5, as is shown in Table 9, with the results for the WFT methods, the difference is striking. Only three methods had nonzero score, the best scoring method (PBE-D_G) achieved only 40%. None of the DFT methods can thus, according to our statistics, be clearly recommended for "black-box" application on the calculation of biomacromolecules. This does not necessarily mean that the use of the DFT methods is meaningless, but the possible results of the satisfactory accuracy obtained for large molecules are quite likely to be attributable to error cancellation.

In contrast with the generally poor performance of the DFT methods, the description of the electrostatics-dominated complexes (Table 10) is much better. According to our statistics, almost all tested DFT methods are capable of proper description of these complexes. Only three methods (B2PLYP-D, PBE- D_G, TPSS-D_G) have not reached the "critical" score of 70%. These results support the generally accepted fact that the DFT methods are suitable for the description of (at least small) H-bonded complexes. Clearly the most demanding requirement (not only for this category) was RQ_5 (like for the WFT methods), that is, the balanced treatment of all the geometry distortions. None of the methods fulfilled the strict threshold for this requirement. The best-performing methods did not have a similar tendency to over- or underestimate the stability of complexes (TPSS-D_J and PBE- D_J overestimated, B2PLYP and PBE underestimated and BLYP- D_J, BLYP- D_G and M06−2X provided unbalanced description without an equable error orientation).

Table 9. DFT Methods: A Statistical Evaluation of All the S22 × 5 Complexes

	RQ_1[a]		RQ_2		RQ_3		RQ_4		RQ_5			
method/basis	10%	20%	±3%	±6%	±3%	±6%	3%	6%	±3%	±6%		score[b] (%)
PBE-D_G/LP	no	yes	no	yes	no	yes	no	yes	no	no		40
BLYP-D_J/LP	no	no	no	no	no	yes	yes	yes	no	no	(−)	30
B2PLYP-D/TZVPP	no	no	no	no	no	no	no	yes	no	no		10
BLYP-D_G/LP	no	no	no	no	no	no	no	no	no	no		0
B2PLYP/TZVPP	no	no	no	no	no	no	no	no	no	no	(−)[c]	0
M06-2X/aDZ	no	no	no	no	no	no	no	no	no	no		0
PBE-D_J/LP	no	no	no	no	no	no	no	no	no	no	(+)	0
PBE/LP	no	no	no	no	no	no	no	no	no	no	(−)	0
TPSS-D_G/LP	no	no	no	no	no	no	no	no	no	no	(+)	0
TPSS-D_J/LP	no	no	no	no	no	no	no	no	no	no	(+)	0

[a] RQ_1, a small (<10% or <20%) mean absolute relative error (MARE) for all complexes and geometry distortions; RQ_2, a small (<±3% or <±6%) MARE variation for all of the investigated complexes for a particular geometry distortion; RQ_3, a small (< ± 3% or < ± 6%) MARE variation within a group of complexes of a particular interaction type for all of the geometry distortions; RQ_4, a balanced description within a group of complexes of a particular interaction type and geometry distortion evaluated by the small (<±3% or <±6%) difference between the MARE and root-mean-square relative error (RMSRE); RQ_5, a small (<±3% or <±6%) deviation of the mean relative error (MRE) from zero, indicating systematic over- or understabilization of all of the complexes and geometry distortions. For a more detailed definition and reasoning of these restrictions, see section 2. [b] The score is defined as the number of positive (yes) matches of a particular (sub)requirement, divided by the total number of requirements applied (ten in this case). [c] (+) means overstabilization; (−) means understabilization. If no symbol is provided, the method could not be unambiguously assigned to any category.

Table 10. DFT Methods: A Statistical Evaluation of the Electrostatics-Dominated Complexes of the S22 × 5 Set

method/basis	RQ_1[a]		RQ_3		RQ_4		RQ_5			score[b]
BLYP-D_G/LP	yes	yes	yes	yes	yes	yes	no	yes		88
M06-2X/aDZ	yes	yes	yes	yes	yes	yes	no	yes		88
TPSS-D_J/LP	yes	yes	yes	yes	yes	yes	no	yes	(+)[c]	88
BLYP-D_J/LP	yes	yes	yes	yes	yes	yes	no	no		75
B2PLYP/TZVPP	yes	yes	yes	yes	no	yes	no	yes	(−)	75
PBE/LP	yes	yes	yes	yes	yes	yes	no	no	(−)	75
PBE-D_J/LP	yes	yes	yes	yes	yes	yes	no	no	(+)	75
B2PLYP-D/TZVPP	no	yes	yes	yes	yes	yes	no	no	(−)	63
PBE-D_G/LP	no	yes	no	yes	yes	yes	no	no	(+)	50
TPSS-D_G/LP	no	yes	no	yes	yes	yes	no	no	(+)	50

[a] For a brief definition of requests RQ_1, RQ_3, ..., RQ_5, see the footnote of Table 9. For a more detailed definition and reasoning, see section 2. [b] The score is defined as the number of positive (yes) matches of a particular (sub)requirement, divided by the total number of requirements applied (eight in this case). [c] (+) means overstabilization; (−) means understabilization. If no symbol is provided, the method could not be unambiguously assigned to any category.

Table 11. DFT Methods: A Statistical Evaluation of the Dispersion-Dominated Complexes of the S22 × 5 Set

method/basis	RQ_1[a]		RQ_3		RQ_4		RQ_5			score[b]
BLYP-D_J/LP	no	no	no	yes	yes	yes	no	no		38
PBE-D_G/LP	no	yes	no	yes	no	no	no	no	(+)	25
B2PLYP-D/TZVPP	no	no	no	no	no	yes	no	no	(−)	13
BLYP-D_G/LP	no	no	no	no	no	no	no	no		0
B2PLYP/TZVPP	no	no	no	no	no	no	no	no	(−)[c]	0
M06-2X/aDZ	no	no	no	no	no	no	no	no		0
PBE/LP	no	no	no	no	no	no	no	no	(−)	0
PBE-D_J/LP	no	no	no	no	no	no	no	no	(+)	0
TPSS-D_G/LP	no	no	no	no	no	no	no	no	(+)	0
TPSS-D_J/LP	no	no	no	no	no	no	no	no	(+)	0

[a] For a brief definition of requests RQ_1, RQ_3, ..., RQ_5, see the footnote of Table 9. For a more detailed definition and reasoning, see section 2. [b] The score is defined as the number of positive (yes) matches of a particular (sub)requirement, divided by the total number of requirements applied (eight in this case). [c] (+) means overstabilization; (−) means understabilization. If no symbol is provided, the method could not be unambiguously assigned to any category.

According to the result shown in Table 11, it is clearly the dispersion-dominated complex category that dictates the overall substandard performance of the DFT methods. Only three DFT methods have a nonzero score. The most likely reason for the failure of, at least, the empirically dispersion-corrected DFT methods in delivering a balanced and accurate description of noncovalent complexes upon the interaction coordinate stretching originates from the fundamental limitation of the damping function to compensate equally well for the doublecounting of the dispersion energy at various geometry distortions.

Table 12. DFT Methods: A Statistical Evaluation of the Mixed-Character Complexes of the S22 × 5 Set

method/basis	RQ_1[a]		RQ_3		RQ_4		RQ_5			score[b]
BLYP-D_J/LP	no	yes	no	yes	yes	yes	no	no	(−)[c]	50
B2PLYP-D/TZVPP	no	yes	no	yes	yes	yes	no	no		50
PBE-D_J/LP	no	yes	no	yes	yes	yes	no	no	(+)	50
BLYP-D_G/LP	no	yes	no	yes	no	yes	no	no		38
PBE-D_G/LP	no	yes	no	no	no	yes	no	no	(+)	25
TPSS-D_J/LP	no	yes	no	no	no	yes	no	no	(+)	25
B2PLYP/TZVPP	no	no	no	yes	no	no	no	no	(−)	13
TPSS-D_G/LP	no	no	no	no	no	yes	no	no	(+)	13
M06-2X/aDZ	no	no	no	no	no	no	no	no		0
PBE/LP	no	no	no	no	no	no	no	no	(−)	0

[a] For a brief definition of requests RQ_1, RQ_3, ..., RQ_5, see the footnote of Table 9. For a more detailed definition and reasoning, see section 2. [b] The score is defined as the number of positive (yes) matches of a particular (sub)requirement, divided by the total number of requirements applied (eight in this case). [c] (+) means overstabilization; (−) means understabilization. If no symbol is provided, the method could not be unambiguously assigned to any category.

The accuracy of description of the mixed-character complexes (see Table 12) is, again, significantly better than that of all the S22 × 5 and the dispersion-dominated complexes but is still clearly inferior to the accuracy of the WFT methods. However, none of the methods reached the score of 70%.

Table 13 presents, analogously to Table 7 for the WFT methods, the absolute and relative ME, MAE, and MAX values for all of the interaction types and individual geometry distortions. It is again clear that the WFT methods provide much better accuracy than the DFT methods. For instance, unlike the WFT methods, only one of the DFT methods

Table 13. DFT Methods[a]

method/basis	0.9			1.0			1.2			1.5			2.0		
TPSS-D_G/LP	1.78	1.78	4.2	1.04	1.04	2.21	0.43	0.43	1.18	0.28	0.28	0.75	0.08	0.08	0.26
	(>−100)[b]	(>100)	(>100)	(−21.49)	(21.49)	(40.84)	(−12.67)	(12.67)	(48.86)	(−18.69)	(18.69)	(58.33)	(−19.34)	(19.34)	(>100)
TPSS-D_J/LP	−0.09	0.63	2.67	0.19	0.32	1.17	0.25	0.27	0.68	0.22	0.22	0.53	0.08	0.08	0.26
	(−6.26)	(32.29)	(>100)	(−5.90)	(8.61)	(34.70)	(−10.69)	(10.98)	(54.61)	(−18.05)	(18.05)	(64.10)	(−18.93)	(18.93)	(>100)
M06-2X/aDZ	0.82	0.85	2.59	0.55	0.6	1.95	0.00	0.20	0.75	−0.09	0.17	0.56	−0.01	0.05	0.15
	(−80.58)	(80.73)	(>100)	(−15.86)	(16.18)	(61.03)	(−3.64)	(7.38)	(60.32)	(0.94)	(12.29)	(36.51)	(16.5)	(21.00)	(>100)
BLYP-D_G/LP	0.38	0.42	1.05	0.25	0.32	0.86	−0.03	0.20	0.53	0.00	0.13	0.51	−0.02	0.07	0.16
	(−31.91)	(33.31)	(>100)	(−2.84)	(7.17)	(31.96)	(7.12)	(8.81)	(50.18)	(6.33)	(8.93)	(31.92)	(−9.50)	(17.28)	(>100)
BLYP-D_J/LP	−0.25	0.50	1.54	−0.47	0.61	2.29	−0.37	0.48	1.57	−0.17	0.21	0.51	−0.07	0.09	0.27
	(44.74)	(53.94)	(>100)	(12.80)	(15.13)	(54.64)	(17.46)	(18.73)	(66.25)	(16.37)	(16.78)	(45.73)	(2.51)	(13.23)	(50.76)
PBE/LP	−2.92	2.99	13.45	−2.21	2.24	9.58	−1.13	1.15	4.21	−0.42	0.43	1.20	−0.14	0.14	0.40
	(>100)	(>100)	(>100)	(46.78)	(47.32)	(>100)	(31.04)	(31.48)	(81.64)	(27.38)	(27.82)	(>100)	(39.80)	(39.80)	(>100)
PBE- D_G /LP	1.23	1.23	2.66	0.81	0.81	1.97	0.38	0.39	1.28	0.21	0.21	0.90	0.03	0.04	0.25
	(−81.83)	(81.83)	(>100)	(−18.32)	(18.34)	(39.11)	(−10.55)	(10.89)	(38.21)	(−10.04)	(10.04)	(31.23)	(−3.42)	(4.54)	(20.38)
PBE-D_J/LP	−0.63	1.19	6.32	0.06	0.44	2.05	0.47	0.47	1.00	0.32	0.32	0.68	0.08	0.08	0.34
	(74.14)	(>100)	(>100)	(−5.30)	(10.16)	(66.90)	(−17.56)	(17.56)	(71.15)	(−21.10)	(21.10)	(65.25)	(−17.45)	(17.45)	(95.17)
B2PLYP/TZVPP	−2.31	2.36	9.24	−1.67	1.7	6.27	−0.88	0.92	2.92	−0.37	0.39	1.03	−0.09	0.1	0.26
	(>100)	(>100)	(>100)	(40.31)	(41.07)	(>100)	(30.24)	(31.12)	(95.18)	(31.83)	(33.01)	(>100)	(33.84)	(34.81)	(>100)
B2PLYP-D /TZVPP	0.73	0.73	1.67	0.55	0.56	1.32	0.22	0.31	0.94	0.09	0.17	0.66	0.03	0.05	0.22
	(−27.96)	(28.64)	(>100)	(−7.43)	(9.42)	(23.67)	(−0.26)	(8.11)	(35.32)	(4.39)	(10.00)	(40.06)	(1.87)	(5.92)	(18.08)

[a] The values in rows: in the first line of each row, ME, MAE, and MAX (in kcal/mol); in the second line, the relative quantities MRE, MARE, and MAXR (in %) for all of the interaction types and particular geometry distortions. The absolute and relative MAX(R) in a particular table cell do not necessarily correspond to the same complex. [b] The values above 100% (or below −100%) are for simplification shown as ">100". The actual values can be found in the Supporting Information.

Table 14. DFT Methods[a]

method/basis	$(CH_4)_2$	$(C_6H_6)_2$	Δ^b
B2PLYP/TZVPP	95.78	94.55	1.23
M06-2X/aDZ	47.75	39.67	8.08
BLYP-D_J/LP	48.79	33.04	15.75
B2PLYP-D/TZVPP	26.39	10.47	15.91
TPSS-D_G/LP	18.33	34.91	−16.59
PBE-D_G/LP	29.77	9.5	20.27
BLYP-D_G/LP	39.67	18.33	21.33
TPSS-D_J/LP	51.37	29.81	21.56
PBE-D_J/LP	63.51	30.47	33.04
PBE/LP	63.69	105.92	−42.23

[a] The averaged absolute relative errors (AARE, in %) over all (except "0.9") geometry distortions for the methane and benzene dimers. [b] AARE(methane dimer)−AARE(benzene dimer).

Table 15. DFT Methods: A Comparison of the Overall Score of the Counterpoise-Corrected (CP) and Uncorrected (NCP) Calculations for the Whole S22 × 5 Set

	electrostatic		dispersion		mixed		all	
method/basis	CP	NCP	CP	NCP	CP	NCP	CP	NCP
BLYP-D_G/LP	63	38	0	0	25	38	0	0
BLYP-D_J/LP	75	50	25	25	50	50	10	10
B2PLYP/TZVPP	63	75	0	0	0	13	0	0
B2PLYP-D/TZVPP	88	63	25	0	63	50	20	10
M06-2X/aDZ	88	88	0	0	0	0	0	0
PBE/LP	50	50	0	0	0	0	0	0
PBE-D_G/LP	38	38	13	25	50	13	10	0
PBE-D_J/LP	50	25	25	0	50	38	10	0
TPSS-D_G/LP	63	50	13	0	63	13	10	0
TPSS-D_J/LP	88	88	13	0	88	25	10	0

(BLYP-D_G) had a MAX value below 1 kcal/mol (0.86 kcal/mol) at the equilibrium distance (1.0), and even in the 100% stretched geometry the mean absolute error is about 0.3 kcal/mol. On the other hand, the MAE error of, for example, TPSS-D_J or B2PLYP-D at the equilibrium distance is only 0.32 and 0.56 kcal/mol, which again demonstrates the contrast between the conclusions one can draw from the statistics used in this work and a "simplified", "MAE in equilibrium geometry" assessment.

As mentioned in the WFT section, the accuracy for qualitatively different types of dispersion-dominated complexes cannot be extracted from the statistics used throughout this work. The performance of the DFT methods on the methane and benzene dimers is shown in Table 14. The most balanced treatment is, because of error cancellation, provided by the B2PLYP method. If only those methods that describe at least one of the complexes with an error of below 20% are considered, it is the B2PLYP-D method that delivers the most balanced results, Δ being about 16%.

The final note in this section is dedicated to the importance of the counterpoise corrections for the BSSE. For the WFT methods, it is generally accepted that at least in a small (or medium-sized) basis set, counterpoise correction for the BSSE leads to interaction energies closer to the CBS limit than the use of plain ones. Because of the inherent local nature of the electron-correlation description in the DFT methods, convergence toward the CBS is much faster as compared to the WFT methods. Despite the undoubted validity of this statement, it is often overrated and the DFT methods are frequently referred to as the "BSSE-free". Table 15 shows the score of the DFT methods in counterpoise-corrected (CP) and uncorrected (NCP) calculations for the whole of S22 × 5 and each interaction type separately. Please note, that BLYP and PBE are now calculated in smaller, TZVP basis set, as often recommend in the literature. It is important to say that for the relevant cases, that is, when the method can be used for the description of noncovalent interactions at all (according to our statistics, that is, the score being above 70%), the CP enhances the accuracy (in two cases the performance was unaffected, that is, M06-2X and TPSS-D_J, for electrostatics-dominated complexes, in one case it is even worse, that is, B2PLYP, again for the electrostatics-dominated complexes) when compared to the NCP results. Furthermore, if the CP is used with the TPSS-D_J method in the calculation of mixed complexes, its score exceeds 70%, thus making it the only DFT method suitable for the

description of complexes in this category (again, according to our statistics). When dispersion corrected PBE and BLYP CP/NCP results in larger, for example LP, basis set are compared, less profound and less systematic trends are observed (see the Supporting Information). Benefits from using CP with larger basis sets in DFT methods with empirical dispersion correction are less clear because of the parametric nature of the correction.

4. Conclusions

A comparison of the performance of several WFT and DFT methods with respect to the CCSD(T)/CBS benchmark interaction energies not only in the equilibrium but also in stretched complex geometries (one shortened and three elongated in the direction of interaction coordinate) was carried out to investigate the general applicability of the approximate methods for large-scale systems. Spatially separated interacting molecular fragments, typically in non-equilibrium geometries, coexist in abundance with the "equilibrium" close contacts in biomolecules. With respect to this fact, any computational method potentially used for the calculation of large molecules must be capable of a balanced simultaneous description of noncovalent interactions in all of the purposeful conformations.

The statistical measures, proposed and used in this paper, were designed to assess the method's general accuracy, balanced treatment within the categories of interaction types (electrostatic- or dispersion-energy-dominated or the "mixed-character" complexes), as well as the balanced treatment of different complex-geometry distortions. The distinction between the interaction types is based on a DFT-SAPT analysis of the interaction energy, carried out for each complex and geometry distortion separately, thus being completely independent of the interaction type's definition on the basis of "chemical intuition".

Of the family of the WFT methods, the most potent approaches in this respect proved to be the MP2C, MP2.5, and SCS-CCSD. Of the DFT family, none of the methods tested in this work fulfilled the, rather strict, statistical criteria applied, thus being significantly inferior (at least) to the three WFT methods mentioned above. This does not necessary mean that the use of the DFT methods (obviously only those covering the dispersion energy) would be meaningless, but the possible satisfactorily accurate results obtained for the large molecules are quite likely to be obtained on the basis of error cancellation.

The importance of the counterpoise corrections in the DFT calculations was shown. For the relevant cases, that is, when the method is, according to our statistics, suitable for the description of noncovalent interactions, the counterpoise correction enhances both the balance and accuracy of the description. With increase of the basis set size, the role of the counterpoise correction becomes less important and furthermore, in the case of the DFT methods with empirical dispersion correction even less systematic.

Refined statistics on the dispersion-dominated complexes revealed that the balanced treatment of the aliphatic and $\pi-\pi$ stacked complexes is a challenging problem and that some methods capable of delivering excellent accuracy for one of these interaction types, for example, MP2 for aliphatic or SCS(MI)-MP2 for $\pi-\pi$ stacked complexes, may fail for the other interaction type.

The S22 $\times$ 5 geometries and interaction energies will be made available at www.begdb.com[39] at the time of the publication of this work.

Acknowledgment. Suggestions from one of the reviewers of this work, related to the basis set saturation of the DFT calculations, are highly appreciated. This work was a part of Research Project No. Z40550506 of the Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, and was supported by Grants No. LC512 and MSM6198959216 from the Ministry of Education, Youth and Sports of the Czech Republic. The support of Praemium Academiae, Academy of Sciences of the Czech Republic, awarded to P.H. in 2007 is also acknowledged. This work was also supported by the Korea Science and Engineering Foundation (World Class University program R32-2008-000-10180-0) and by the Slovak Grant Agency VEGA under Contracts No. 1/0428/09 and 1/0520/10. L.G. gratefully acknowledges support from Research Grant No. MSM 6046137306. A portion of the research was performed using 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.

Supporting Information Available: Complete statistics for the tested WFT and DFT methods (extended of those not shown in the paper), S22×5 CCSD(T)/CBS interaction energies and detailed information on the basis sets used. Geometries of all S22×5 complexes. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Jurečka, P.; Šponer, J.; Černý, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985.

(2) DiStasio, R. A.; Head-Gordon, M. *Mol. Phys.* **2007**, *105*, 1073.

(3) Pitoňák, M.; Neogrady, P.; Černý, J.; Grimme, S.; Hobza, P. *ChemPhysChem* **2009**, *10*, 282.

(4) Hesselmann, A. *J. Chem. Phys.* **2008**, *128*, 144112.

(5) Pitoňák, M.; Heβelmann, A. *J. Chem. Theory Comput.* **2010**, *6*, 168.

(6) Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463.

(7) Jurečka, P.; Černý, J.; Hobza, P.; Salahub, D. R. *J. Comput. Chem.* **2007**, *28*, 555.

(8) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215.

(9) McNamara, J. P.; Hillier, I. H. *Phys. Chem. Chem. Phys.* **2007**, *9*, 2362.

(10) Řezáč, J.; Fanfrłik, J.; Salahub, D. R.; Hobza, P. *J. Chem. Theory Comput.* **2009**, *5*, 1749.

(11) Korth, M.; Pitoňák, M.; Řezáč, J.; Hobza, P. *J. Chem. Theory Comput.* **2010**, *6*, 344.

(12) Riley, K. E.; Pitoňák, M.; Jurečka, P.; Hobza, P. *Chem. Rev.* [Online early access]. DOI: 10.1021/cr1000173. Published Online: May 20, 2010.

(13) Pittner, J.; Hobza, P. *Chem. Phys. Lett.* **2004**, *390*, 496.

(14) Hopkin, B. W.; Tschumper, G. S. *J. Phys. Chem. A* **2004**, *108*, 2941.

(15) Crittenden, D. L. *J. Phys. Chem. A* **2009**, *113*, 1663.

(16) Pitoňák, M.; Neogrády, P.; Řezáč, J.; Jurečka, P.; Urban, M.; Hobza, P. *J. Chem. Theory Comput.* **2008**, *4*, 1829.

(17) Pitoňák, M.; Riley, K. E.; Neogrády, P.; Hobza, P. *ChemPhysChem* **2008**, *9*, 1636.

(18) Pitoňák, M.; Janowski, T.; Neogrády, P.; Pulay, P.; Hobza, P. *J. Chem. Theory Comput.* **2009**, *5*, 1761.

(19) Takatani, T.; Hohenstein, E. G.; Malagoli, M.; Marshall, M. S.; Sherrill, C. D. *J. Chem. Phys.* **2010**, *132*, 144104.

(20) Dunning, T. H. J. *J. Chem. Phys.* **1989**, *90*, 1007.

(21) Morgado, C. A.; Jurečka, P.; Svozil, D.; Hobza, P.; Sponer, J. *J. Chem. Theory Comput.* **2009**, *5*, 1524.

(22) Berka, K.; Laskowski, R.; Riley, K. E.; Hobza, P.; Vondrášek, J. *J. Chem. Theory Comput.* **2009**, *5*, 982.

(23) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787.

(24) Rubeš, M.; Bludský, O. *ChemPhysChem* **2009**, *10*, 1868.

(25) Heßelmann, A.; Jansen, G.; Schütz, M. *J. Chem. Phys.* **2005**, *122*, 014103.

(26) Marchetti, O.; Werner, H. *J. Phys. Chem. A* **2009**, *113*, 11580.

(27) Molnar, L. F.; He, X.; Wang, B.; Merz, K. M. *J. Chem. Phys.* **2009**, *131*, 065102.

(28) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. *J. Chem. Phys.* **1997**, *106*, 9639.

(29) Grimme, S. *J. Chem. Phys.* **2003**, *118*, 9095.

(30) Jung, Y.; Lochan, R. C.; Dutoi, A. D.; Head-Gordon, M. *J. Chem. Phys.* **2004**, *121*, 9793.

(31) Aquilante, F.; Pedersen, T. B. *Chem. Phys. Lett.* **2007**, *449*, 354.

(32) The basis sets are available from the TURBOMOLE homepage via the FTP server button (in the subdirectories basen, jbasen, and cbasen). See http://www.turbomole.com (accessed Jun 15, 2010).

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

(34) *TURBOMOLE*, V5.9 2009; University of Karlsruhe and Forschungszentrum Karlsruhe GmbH: Karlsruhe, Germany, 1989−2007; TURBOMOLE GmbH: Karlsruhe, Germany, 2007−2010; available from http://www.turbomole.com (accessed Jun 15, 2010).

(35) 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.

(36) Werner, H.-J.; Knowles, P. J.; Lindh, R.; Schütz, M.; Celani, P.; Korona, T.; Manby, F. R.; Rauhut, G.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Hampel, C.; Hetzer, G.; Lloyd, A. W.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklass, A.; Palmieri, P.; Pitzer, R.; Schumann, U.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T. MOLPRO, a package of ab initio programs, version 2006.1; Birmingham, UK, 2006; see http://www.molpro.net (accessed Jun 15, 2010).

(37) Aquilante, F.; Vico, L. D.; Ferre, N.; Malmqvist, P.-Å.; Neogrady, P.; Pedersen, T.; Pitoňák, M.; Reiner, M.; Roos, B. O.; Serrano-Andres, L.; Urban, M.; Veryazov, V.; Lindh, R. *J. Comput. Chem.* **2010**, *31*, 224.

(38) Johnson, E. R.; Becke, A. D.; Sherrill, C. D.; DiLabio, G. A. *J. Chem. Phys.* **2009**, *131*, 03411.

(39) Řezáč, J.; Jurečka, P.; Riley, K. E.; Černý, J.; Valdes, H.; Pluháčkova, K.; Berka, K.; Řezáč, T.; Pitoňák, M.; Vondrášek, J.; Hobza, P. *Collect. Czech. Chem. Commun.* **2008**, *73*, 1261.

(40) Schaefer, A.; Huber, C.; Ahlrichs, R. *J. Chem. Phys.* **1994**, *100*, 5829.

(41) Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297.

CT1002253

JCTC Journal of Chemical Theory and Computation

Potential Energy Landscape of the Electronic States of the GFP Chromophore in Different Protonation Forms: Electronic Transition Energies and Conical Intersections

I. V. Polyakov,*,† B. L. Grigorenko,† E. M. Epifanovsky,‡ A. I. Krylov,‡ and A. V. Nemukhin†,§

Department of Chemistry, M.V. Lomonosov Moscow State University, Moscow 119991, Russian Federation, Department of Chemistry, University of Southern California, Los Angeles, California 90089, and Institute of Biochemical Physics, Russian Academy of Sciences, Moscow 119334, Russian Federation

Received May 3, 2010

Abstract: We present the results of quantum chemical calculations of the transition energies and conical intersection points for the two lowest singlet electronic states of the green fluorescent protein chromophore, 4′-hydroxybenzylidene-2,3-dimethylimidazolinone, in the vicinity of its cis conformation in the gas phase. Four protonation states of the chromophore, i.e., anionic, neutral, cationic, and zwitterionic, were considered. Energy differences were computed by the perturbatively corrected complete active space self-consistent field (CASSCF)-based approaches at the corresponding potential energy minima optimized by density functional theory and CASSCF (for the ground and excited states, respectively). We also report the EOM-CCSD and SOS-CIS(D) results for the excitation energies. The minimum energy S_0/S_1 conical intersection points were located using analytic state-specific CASSCF gradients. The results reproduce essential features of previous *ab initio* calculations of the anionic form of the chromophore and provide an extension for the neutral, cationic, and zwitterionic forms, which are important in the protein environment. The S_1 PES of the anion is fairly flat, and the barrier separating the planar bright conformation from the dark twisted one as well as the conical intersection point with the S_0 surface is very small (less than 2 kcal/mol). On the cationic surface, the barrier is considerably higher (∼13 kcal/mol). The PES of the S_1 state of the zwitterionic form does not have a planar minimum in the Franck−Condon region. The S_1 surface of the neutral form possesses a bright planar minimum; the energy barrier of about 9 kcal/mol separates it from the dark twisted conformation as well as from the conical intersection point leading to the cis−trans chromophore isomerization.

Introduction

The fascinating photochemical properties of the green fluorescent protein (GFP), which is employed in many areas of biotechnology and medicine as a biomarker in living cells,[1−3] have inspired numerous experimental and theoretical studies.[4−6] An important practical goal is to formulate concrete suggestions to guide the design of novel biomarkers by modifying either the structure of the chromophore or its immediate environment (specific amino acid residues) in fluorescent proteins of the GFP series. This requires a mechanistic understanding of transformations occurring with the chromophore upon photoexcitation, which includes characterization of absorption and emission bands in optical spectra as well as evolution of the system in the excited electronic states in various environments. From the theoreti-

* Corresponding author e-mail: polyakoviv@gmail.com.
† M.V. Lomonosov Moscow State University.
‡ University of Southern California.
§ Russian Academy of Sciences.

10.1021/ct100227k © 2010 American Chemical Society
Published on Web 07/20/2010

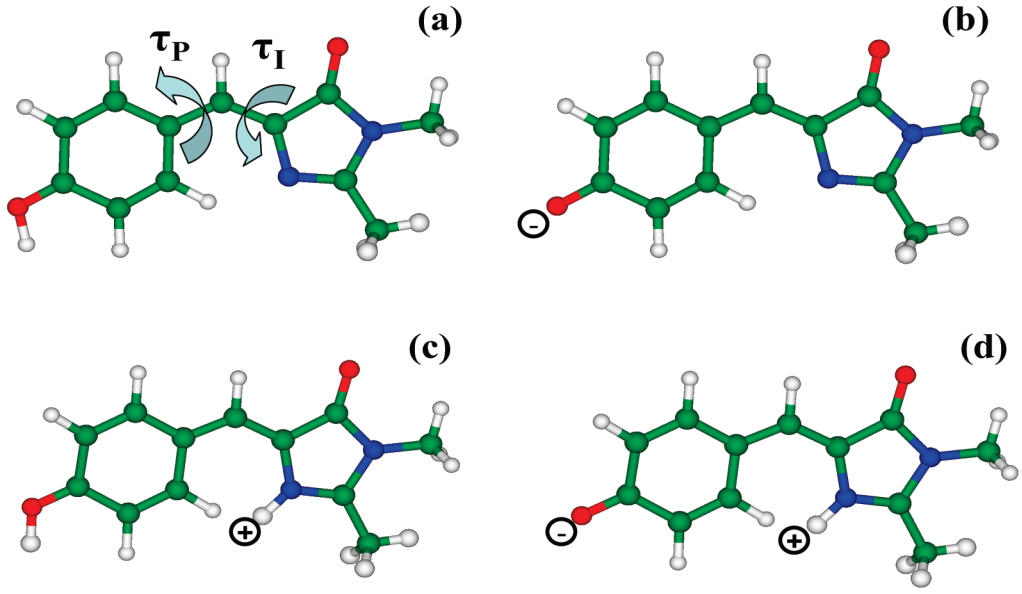


Figure 1. Protonation forms of the GFP chromophore, HBDI, in the cis conformation: (a) neutral, (b) anionic, (c) cationic, and (d) zwitterionic. Here and below, the carbon atoms are colored in green, nitrogen in blue, and oxygen in red. The definition of the twisting angles around the phenolate (P) and imidazolidinone (I) bridge bonds, τ_P and τ_I, is shown in panel a.

cal perspective, the characterization of denatured (isolated) chromophores is the first step toward understanding their photochemical properties in realistic environments. Modeling isolated species involves calculations of the properties of the chromophores in the gas phase and in solution using quantum chemistry methods.

The experimental studies of the GFP-type chromophores in the gas phase[7−11] provide important information on the structure and spectra of these species. The works of Andersen et al.[7−10] characterized the absorption bands corresponding to the S_0-S_1 vertical excitation by the photodestruction spectroscopy of mass-selected ions injected into an electrostatic ion storage ring. The band maxima for the anionic and cationic forms (Figure 1b and c) of the GFP chromophore, 4′-hydroxybenzylidene-2,3-dimethylimidazolinone (HBDI) were reported at 479 and 406 nm (2.59 and 3.05 eV), respectively. To mimic the neutral species (Figure 1a), the so-called, "neutral+" analogs were used. The earlier study reported the value 415 nm (2.99 eV),[9] while in the later paper the authors reported a revised value of 370 nm (3.35 eV).[10] Forbes and Jockusch studied the gaseous HBDI anion in the ion trap measuring laser induced fluorescence and photoactivation action spectra.[11] No fluorescence was detected in these experiments, and electron detachment and fragmentation were found to be the primary modes of ion deactivation.

The absence of fluorescence in the gas phase is consistent with the solution results—the fluorescence quantum yield of HBDI in solution drops by more than 3 orders of magnitude relative to wild-type GFP. This is attributed to ultrafast (∼0.5−2 ps) internal conversion (IC) of the denaturated chromophore to the ground state, in contrast to the nanosecond lifetime of the excited state in the protein environment. The mechanism of IC is still not clear. Solution studies showed that there is no correlation with solvent viscosity (which can be modified by either using different solvents or by varying the temperature), which can only be explained by volume-conserving and strongly exothermic reaction coordinates toward IC. Moreover, the rates of IC were found to be independent of the protonation state of the protein. The theoretical studies[12−17] of the S_1 potential energy surface (PES) and the minimum energy conical intersection (MECI) points did not provide a simple unifying explanation; e.g., volume-conserving (i.e., hula-twist) pathways to MECI were found to have large barriers, whereas lower-energy relaxation coordinates correspond to nonvolume-conserving motions.

Properties of the gas phase chromophores of the GFP type have been considered in several theoretical studies (see the recent review[18]) using methods ranging from semiempirical models[19−24] to time-dependent density functional theory (TD-DFT)[10,21−23,25−31] and different *ab initio* approaches.[10,16,17,32−42] A majority of these papers focused on the absorption spectra of the chromophore (or its oxidized forms[44]) in the ground electronic state, and only a few works described the chromophore properties in the S_1 excited state. Calculations of Martin et al.[16] for the anionic form of HBDI performed using CASSCF(12/11)/6-31G* and CASPT2//CASSCF suggested an extremely flat landscape of the S_1 PES in the region of the potentially fluorescent state (called FS in ref 16). The details of the S_1 PES including the energy minima and the minimum energy conical intersection points were computed with the two-root (S_0, S_1) state-averaged (SA2-CASSCF) procedure with equal weights (w1 = 1, w2 = 1). The calculated S_0-S_1 energy gap at the FS point corresponded to the wavelength of 507 nm (2.45 eV).[16] The paper of Altoe at el.[17] mainly focused on the vibrational properties of the ground state anionic, neutral, and cationic forms of HBDI in solution but also included the results for the gas-phase photoreaction pathway of the HBDI anion. These calculations were carried out using the SA2-CASSCF(12,11)/6-31G* wave functions and single-state perturbation theory (CASPT2) corrections. In both papers,[16,17] a twisted type intersection between the S_1 and S_0 PESs was reported. Olsen and Smith[37] employed SA3-CASSCF(4/3)/DZP to compute stationary points on the S_0 and S_1 surfaces,

as well as to optimize the S_1/S_0 MECI. The energies at these points were recalculated with multireference multistate Rayleigh−Schrödinger perturbation theory (MR-MSRSPT2). The authors of ref 37 compared their conclusions to those reported previously by Martin et al.[16] and Altoe et al.[17] Simulations of photodynamics of the neutral form of HBI and HBDI in vacuo, in solution and in the protein, including calculations of the S_0 and S_1 energies and the MECI points performed by Martinez et al.[15,38] were based either on semiempirical or on the *ab initio* SA2-CASSCF(2,2)/6-31G methods.

A recent study[45] investigated the effect of protonation on the excited-state isomerization of isolated HBI by *ab initio* molecular dynamics with SA-CASSCF(2,2)/6-31G. The authors observed that both neutral (protonated) HBI and the anion forms undergo fast isomerization; however, the former species isomerizes exclusively around τ_I (the imidazolinone CC bond, see Figure 1a), whereas the latter species rotates mostly around the phenolate CC bond (τ_P), although the τ_I channel is also open. The difference was explained in terms of the effect of resonance on the bond alternation pattern in the two forms; i.e., protonation detunes the resonance in HBI. Owing to the well-known tendency of CASSCF to exaggerate bond alternation, the exact branching ratios (and the barriers) may be sensitive to the electronic structure methods employed. This study suggested that excited-state cis−trans isomerization may be strongly coupled with (gated by) protonation.[45]

For the purpose of our study, the results of Olsen and Smith[37] for the PES stationary points in the vicinity of the cis isomer (called Z isomer in ref 37) of the HBDI anion (Figure 1b) at the SA3-CASSCF(4/3)/DZP and the perturbatively corrected levels present the important reference data. We employ a similar strategy, namely, the state-averaged SA-CASSCF/cc-pVDZ and the state-specific CASSCF (w1 = 0, w2 = 1) approaches with a larger active space than that in ref 37 and the versions of the multiconfigurational quasi degenerate perturbation theory of the second order (MCQDPT2)[46,47] on top of SA-CASSCF to characterize all four protonation forms of HBDI in the S_1 state. For selected energy differences, we also report the EOM-CCSD/6-311G* and SOS-CIS(D)/cc-pVTZ values.

It is established that absorption in the protein occurs in the neutral and anionic protonation forms of the chromophore; but, the fluorescence is only due to the anionic form.[2] However, the involvement of the cationic and zwitterionic forms (Figure 1c,d) in the GFP photodynamics cannot be ruled out since the imidazolinone nitrogen may be protonated through a conservative Glu amino acid residue that is hydrogen bonded to the chromophore.[19]

The photocycle of the chromophore is initiated by the $\pi-\pi^*$ transition (i.e., $S_0 \rightarrow S_1$). The $S_1 \rightarrow S_0$ fluorescence may be observed from the presumably planar minimum on the S_1 PES. The radiationless relaxation from S_1 to S_0 is likely to proceed via conformational changes over the twisting angles τ_P and τ_I (Figure 1a). The same twisting motions lead to the sharp drop of the oscillator strength. Thus, even if the chromophore remains in the S_1 state at the twisted configuration, the fluorescence will be lost. The magnitudes of the barriers separating the planar minima in the vicinity of the Franck−Condon region from the twisted conformers and/or MECI points are crucial for understanding the excited-state lifetimes and fluorescent properties of the chromophore.

We report accurate calculations of the absorption and emission band maxima as well as locations of the MECI points and relevant energy barriers for all four protonated forms of HBDI. For the anionic form, our value of the absorption wavelength is close to the experimental gas phase values,[7,8,10] and the location of the MECI point is close to that obtained by Olsen and Smith.[37] This validates our results for the neutral, cationic, and zwitterionic forms.

These data are necessary to rationalize the properties of the GFP-type chromophores inside the protein by allowing one to separate the intrinsic properties of the chromophore from the role of the protein matrix. A direct comparison of geometry configurations and energies of the various critical points in the excited state in the gas phase and in the protein will help to elucidate the photochemical mechanism.

Methods

Geometry optimizations of the ground electronic states were performed using DFT (PBE0/cc-pVDZ) and RI-MP2/cc-pVTZ. The location of the minimum energy points and a partial scan of the S_1 PES were performed either using CASSCF/cc-pVDZ for the second root of the Hamiltonian or using state-averaged (SA) CASSCF. Energy differences between the S_0 and S_1 states at the selected geometries were computed with the perturbatively corrected CASSCF-based methods with the cc-pVDZ basis set as well as by SOS-CIS(D)/cc-pVTZ and EOM-EE-CCSD/6-311G* (the core electrons were frozen in the EOM calculations).

EOM-CC[48] and SOS-CIS(D)[49,50] calculations were performed using Q-Chem.[51] Multireference calculations were carried out with the Firefly computer program.[47] We employed two new procedures that are important for MECI point calculations, which were recently developed and implemented in Firefly: (I) analytic gradients for SA-CASSCF and (II) corrected MCQDPT2. The latter development corrected the bugs in the MCQDPT2 procedure,[46] resulting in a new version of the program called XMC-QDPT2. We refer to the Firefly Web site[47] for the technical details relevant to the calculations of MECI points and XMCQDPT2.

The calculations for each protonation form were initiated from the coordinates of the respective cis isomer. We then considered deformations over the angles τ_I and τ_P relaxing all other coordinates until the conical intersection or twisted minimum on S_1 was reached. We did not attempt to describe the photoinduced cis−trans isomerization of the chromophore.

Molecular Orbital Framework

Figure 2 illustrates relevant molecular orbitals (MOs) of different protonated forms of HBDI.

For all four forms, the excited S_1 state is derived from the HOMO to the valence LUMO transition (when using large bases, the valence LUMO may appear above a manifold of diffuse orbitals that fill the HOMO−LUMO gap). Both

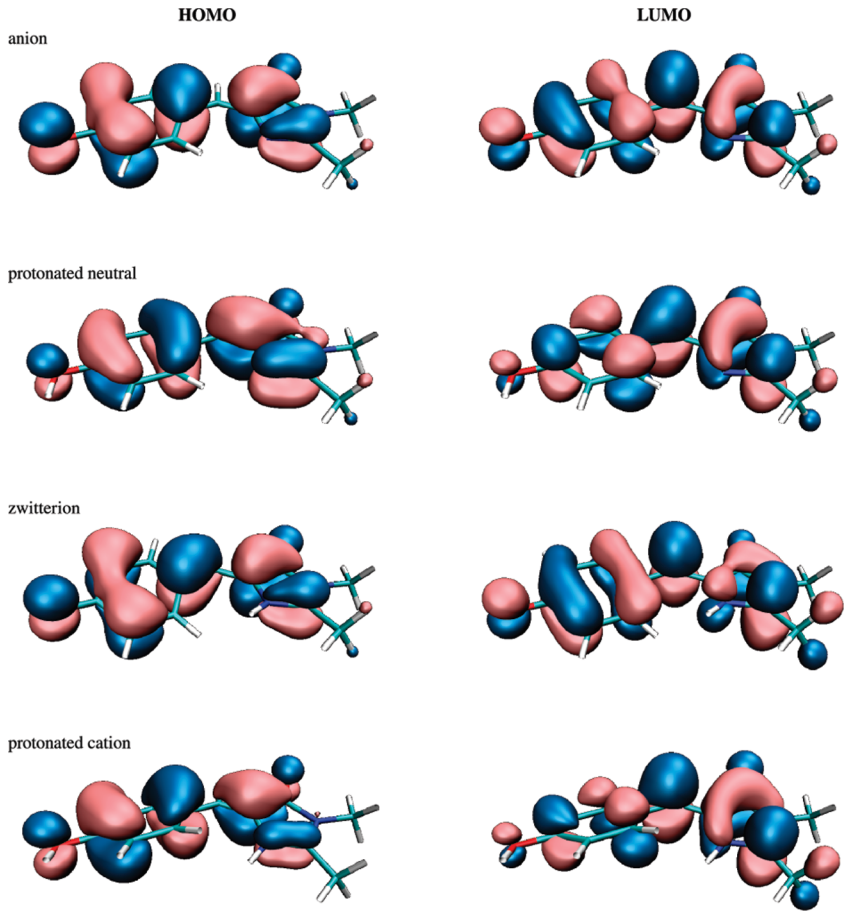


Figure 2. Frontier molecular orbitals of the anionic, neutral, zwitterionic, and cationic forms of HBDI, HOMO (left) and LUMO (right), computed by HF/6-311G*.

orbitals are of the π type and exhibit allylic character in the bridge region.[44] Despite the delocalized nature of the orbitals, the oscillator strength of S_1 and structural changes (e.g., in bond alternation and permanent dipole moment) induced by excitation can be explained by focusing on the bridge moiety (see refs 37, 39, 40, and 44).

The shapes of the MOs of the anionic and neutral HBDI are similar; however, the protonation stabilizes the HOMO considerably. For example, the vertical detachment energy (VDE) of the anionic form is 2.54 eV,[44] whereas the respective ionization energy for the neutral is 7.29 eV. Moreover, the protonation detunes the resonance. Consequently, instead of an almost perfectly allylic pair of orbitals in the anionic forms, the HOMO and LUMO in the neutral form develop π(C−C(I)) and π^*(C−C(I)) characters, respectively [C denotes the bridge carbon atom, and C(I) denotes an imidazolinone cycle carbon atom bonded directly to the bridge carbon]. This difference explains more pronounced bond alternation in the neutral form, in which the C−C(I) bond is much shorter than the C−C(P) one (see Table 1) [C(P) denotes the carbon atom of the phenolic ring bonded directly to the bridge carbon].

This difference in frontier orbitals also leads to larger changes in bond alternation in the S_1 state resulting in the sharper-shaped PES for the neutral form as compared to the extremely flat surface of the anionic form.

Stabilization of the HOMO in the protonated (neutral) form is responsible for a higher $S_0 \rightarrow S_1$ excitation energy (see below), which can be explained within the Hückel model. Alternatively, these changes in the electronic character of the bright state may be explained with the three-state valence-bond-like model.[37]

Table 1. Relevant Geometric Parameters of the Four Forms of HBDI Optimized with RI-MP2/cc-pVTZ[a]

	anion	neutral	zwitterion	cation
C(I)−C	1.378	1.350	1.395	1.350
C(P)−C	1.394	1.435	1.374	1.424
C=O (I)	1.233	1.216	1.222	1.201
C=O (P)	1.249	1.359	1.233	1.340
Twist	0	0	17.2	22.1

[a] Angles in degrees, bond lengths in Å.

The picture of MOs for the zwitterionic form is consistent with an inverted bond alternation pattern (see Table 1); namely, the HOMO and LUMO bear π(C−C(I)) and π^*(C−C(I)) characters, respectively. The structure and MOs of the protonated cationic form are similar to those of the neutral form.

Results for the Anionic Form

The cis-anionic form of the GFP chromophore (Figure 1b) is believed to be responsible for emission inside the protein matrix.[2] No fluorescence of the HBDI anion is detected in the gas phase due to rapid radiationless decay of the excited state population.[11] The position of the optical absorption band

Table 2. Calculated Vertical S_0-S_1 Excitation Energies of the Anionic (Deprotonated) Form of the GFP Chromophore[a]

system, calculation details	λ, nm	ΔE, eV	ref
HBI, CASPT2//CASSCF(12/11)/6-31G*	465	2.67	16
HBDI, CASPT2/cc-pVTZ//DFT(BLYP)[b]	431	2.88	41
HBDI, EOM-CCSD/cc-pVDZ//DFT(BLYP)	408	3.04	41
HBDI, EOM-CCSD/6-311G*//RI-MP2/cc-pVTZ	400	3.10 (1.25)	t.w.
HBDI, SOS-CIS(D)/cc-pVTZ//RI-MP2/cc-pVTZ	473	2.62 (1.54)	39
HBDI, XMCQDPT2/cc-pVDZ//DFT(PBE0)/cc-pVDZ	494	2.51 (1.19)	t.w.

[a] Oscillator strength is given in parentheses. The experimental gas phase value[5,6] is 2.59 eV or 479 nm. [b] CASPT2/cc-pVTZ//DFT(BLYP) calculations from ref 41 employed the default IPEA corrected zero-order Hamiltonian (the MOLCAS 7.2 program).

in the gas phase corresponding to the S_0-S_1 vertical excitation was measured by photodestruction spectroscopy of mass-selected ions injected into an electrostatic ion storage ring.[7,8] The band is centered at 479 nm (or 2.59 eV) extending from 2.4 to 2.8 eV (440−520 nm). Several theoretical papers reported fairly good results for this band either for the HBDI molecule or for a slightly simplified HBI model system by using different ab initio methods; however, the resonance (autoionizing) nature of this state[11,39] complicates the interpretation of the experimental spectrum and makes it difficult to obtain a converged theoretical value of the band maximum. Table 2 summarizes the majority of the previously published excitation energies for the vertical $S_0 \rightarrow S_1$ transition. The detailed discussion of the methods' performance can be found elsewhere;[39,41] we only note that the value computed in this work with XMCQDPT2/SA2-CASSCF(14/12)/cc-pVDZ//DFT(PBE0)/cc-pVDZ (494 nm or 2.51 eV) is within 15 nm (0.08 eV) from the experimental maximum at 479 nm (2.59 eV).[7,8]

It is important to compare our results for the selected critical points on the S_0 and S_1 PES to those obtained by Olsen and Smith.[37] Figure 3 presents the energy diagram for the anionic form of HBDI showing the energies at the ground state minimum (Min-S_0), at the planar minimum on the excited state (Plan-S_1), at the S_0/S_1 MECI point attainable by the τ_I twist (MECI−I-$S_{0/1}$), and at the twisted over angles τ_I and τ_P minimum energy points on the excited state (TwI-S_1 and TwP-S_1). Owing to the very flat shape of the S_1 PES, we were able to obtain only approximate values of the barriers on S_1.

The coordinates and the total energies of the critical points on the PESs calculated in this work using DFT(PBE0)/cc-pVDZ for S_0 and SA-CASSCF(12/11)/cc-pVDZ for S_1 are given in the Supporting Information. First, we note that the geometries computed in this work are consistent with those shown in Figure 8 and Figure 9 of ref 37. Table 3 presents the selected geometric parameters of the Min-S_0, Plan-S_1, TwI-S_1, and MECI−I-$S_{0/1}$ structures of the deprotonated HBDI anion. We observe significant elongation of the C(I)−C and C(P)−C bonds in the Plan-S_1 structure relative

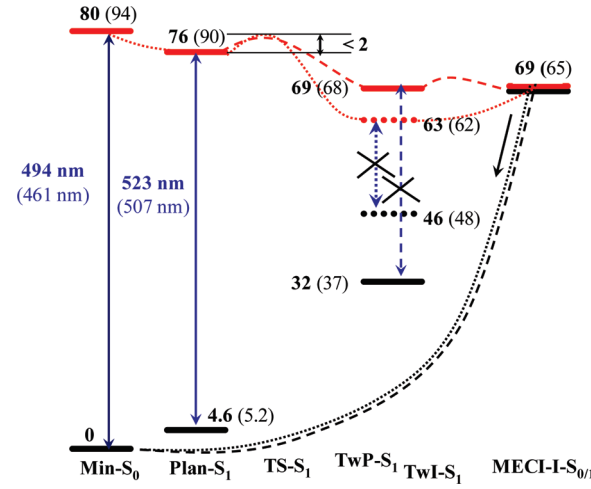


Figure 3. Energy diagram of the anionic (deprotonated) form of HBDI at the cis conformation (Figure 1b). Thick black horizontal lines denote the ground state (S_0) levels; thick red lines - the excited state (S_1) levels. The structures corresponding to the minimum on S_0 (Min-S_0), the planar minimum energy point on S_1 (Plan-S_1), the MECI point (MECI−I-$S_{0/1}$), and the minimum on S_1 twisted over τ_I (TwI-S_1) and τ_P (TwP-S_1) are indicated. The energy values (in kcal/mol) in parentheses refer to the data of Olsen and Smith obtained at the SA3-CASSCF(4,3)/DZP level.[37] The energy values in bold (in kcal/mol) refer to the present calculations with SA-CASSCF(12/11)/cc-pVDZ. The SA means equal weighting average over the first two states (w1 = 1, w2 = 1) for all of the structures when computing the S_0 and S_1 state energies. The TwP-S_1, TwI-S_1, and MECI structures were optimized using the SA procedure; the (0,1) weighting state-specific treatment was employed for locating the Plan-S_1 and TS-S_1 structures' location. We also show the wavelengths for the vertical transition energies corresponding to the S_0-S_1 gap at Min-S_0 and Plan-S_1 computed in this work (in bold) and previously[37] (in parentheses) by using the highest theoretical levels in both works: XMCQDPT2/CASSCF(14/12)/cc-pVDZ//DFT-(PBE0)/cc-pVDZ at Min-S_0; XMCQDPT2/CASSCF(14/12)/cc-pVDZ//SS-CASSCF(12/11)/cc-pVDZ at Plan-S_1; and MS-CASPT2//SA3-CASSCF(4,3)/DZP.[37]

Table 3. Selected Geometric Parameters of the HBDI Anion Optimized with CASSCF(12/11)/cc-pVDZ[a]

	Min-S_0	Plan-S_1	TwI-S_1	MECI−I-$S_{0/1}$
C(I)−C	1.376	1.423	1.461	1.468
C(P)−C	1.414	1.455	1.414	1.457
C=O (I)	1.213	1.222	1.209	1.192
C=O (P)	1.222	1.227	1.236	1.255
Twist-I	0	0	91.612	115.921
Twist-P	0	0	0.88	11.20

[a] Angles in degrees, bond lengths in Å. "Twist-I" is the N−C(I)−C−H dihedral angle value; "Twist-P" is the H−C−C(P)−C dihedral angle value.

to Min-S_0. This change is due to partial charge transfer from the C(I)−C and C(P)−C bonds (HOMO) to the CH bridge (LUMO) (see Figure 2 and the Supporting Information). The TwI-S_1 structure has an even longer C(I)−C distance because the N−C(I)−C−H dihedral angle is close to 90°, which disrupts the conjugation between the rings, thus making C(I)−C a single bond. The MECI−I-$S_{0/1}$ structure develops new features (as compared to TwI-S_1): notable are a Twist-P

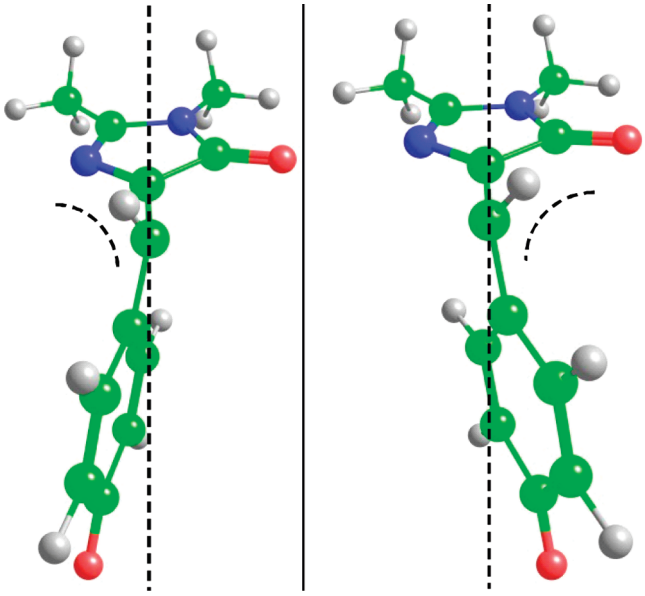


Figure 4. Two I-twisted MECI structures found in this work. The left one was obtained by starting from the trans structure, and the right one - from the cis structure.

angle value of ~10° and visible pyramidalization at the bridge carbon.

Our energy gaps also agree fairly well with those from ref 37 despite the differences in the calculation schemes (the size of the active space, the state-averaging procedure, and the treatment of perturbation corrections). The flat landscape of the S_1 PES along the τ_I angle shown in Figure 10 of ref 37 is consistent with an almost negligible barrier height (less than 2 kcal/mol) required to climb out of the planar S_1 minimum estimated in our work. Similar values of barrier heights on S_1 were reported in refs 16 and 17. For the I-Twist structure, we found two different MECI points depending on the starting point for optimization (either from the trans or cis structures). The difference between these MECI points, which is due to asymmetry of the imidazolinone ring, can easily be seen in Figure 4. The energy difference between these structures is less than 1 kcal/mol (see the Supporting Information). We were not able to locate the MECI-P structure—all of the searches eventually ended up at MECI-I.

The S_0-S_1 energy gap at Plan-S_1 computed in this work is 523 nm (2.37 eV), and this transition has a large oscillator strength. As expected from the MO shapes (see Figure S5 in the Supporting Information), the S_0-S_1 transitions at TwI-S_1 and TwP-S_1 are forbidden, as indicated by the crosses in Figure 3.

Overall, these results are consistent with the experimental findings;[7,8] namely, fluorescence from the anionic GFP chromophore in the gas phase is hardly possible since upon excitation the system can evolve almost freely on the S_1 surface toward optically dark structures and the conical intersection point (MECI−I-$S_{0/1}$). The dramatic change in fluorescence properties and excited state lifetime in the protein suggests that the protein environment strongly affects the shape of the S_1 PES, increasing the barriers along the twisting coordinate.

Results for the Neutral Form

The neutral form of the GFP chromophore (Figure 1a) is responsible for the shorter wavelength absorption at ~400 nm (3.10 eV), but not for steady-state emission inside the protein matrix.[2] Attempts to determine the gas phase absorption using the photodestruction spectroscopy through the positively charged model systems mimicking the true neutral chromophore are presented in refs 9 and 10. The most recent paper from the Andersen group[10] reports a revised value of the band maximum, which is closer to 370 nm (3.35 eV) rather than to 415 nm (2.99 eV) as claimed earlier.[9]

The wavelength of 370 nm (3.35 eV) is also consistent with the results of the most recent quantum chemical calculations.[10,24,41] It should be noted that neutral HBDI (or HBI) presents a difficult case for the CASSCF-based methods. Unlike the anionic form, the desired excited $\pi\pi^*$ state of the neutral does not appear as the second root of the CAS-CI Hamiltonian for various choices of the active space (up to the CASSCF(16/14) partitioning[35]) with different basis sets. Consequently, the state averaging should be performed for more than two states. In particular, Bravaya et al.[36] applied the SA4-CASSCF(16/14) method to optimize the orbitals and the expansion coefficients followed by aug-MCQDPT2[52] to compute the excitation energy at 3.11 eV (or 399 nm). Unfortunately, later, bugs in the original MCQDPT2 code were found,[47] and the predictions of these calculations should be revisited. In this work, we applied SA3-CASSCF(14/12)/cc-pVDZ followed by the XMC-QDPT2 method[47] to compute the S_0-S_1 energy gap at Min-S_0 (which in turn was optimized at the PBE0/cc-pVDZ level). The computed wavelength, 375 nm (3.31 eV), practically coincides with the most recent experimental measurement of 370 nm (3.35 eV).[10]

The EOM-CCSD and SOS-CIS(D) calculations yield considerably higher excitation energies, i.e., 3.83 and 4.12 eV, respectively. While there are discrepancies between the excitation energy values, all methods agree that the absorption of the protonated form is strongly blue-shifted (1.0−1.2 eV) relative to that of the anion. Table 4 summarizes the results of the excitation energy calculations.

The geometry of the planar minimum (Plan-S_1) in the S_1 state was optimized with the CASSCF(12/11)/cc-pVDZ method for the true second root (w1 = 0, w2 = 1). The S_1-S_0 energy gap at this point was calculated by using the XMCQDPT2/SA2(w1=1,w2=1)CASSCF(14/12) approach, that is, at the same level as for the anionic form. Thus, we report here an accurate theoretical estimate for the wavelength, 459 nm (2.70 eV), of the allowed $S_1 \rightarrow S_0$ transition of neutral HBDI in the gas phase. We were able to locate saddle points on the S_1 PES at the SA2-CASSCF(12/11)/cc-pVDZ level for the true second root (w1 = 0, w2 = 1). The first one corresponds to the twisting over angle τ_P with the only imaginary frequency of 64i; the second corresponds to the twisting over angle τ_I with the only imaginary frequency of 209i. Their energies are about 8−9 kcal/mol above those of Plan-S_1;

Table 4. Calculated Vertical S_0-S_1 Excitation Energies of the Neutral Form of the GFP Chromophore[a]

system, calculation details	λ, nm	ΔE, eV	ref
HBDI, CASPT2/cc-pVDZ// DFT(BLYP)[b]	346	3.58	41
HBDI, EOM-CCSD/cc-pVDZ// DFT(BLYP)	310	4.00	41
HBDI, R1-CC2/aug-cc-pVDZ// CC2/aug-cc-pVDZ	349	3.55	10
HBDI, aug-MCQDPT2/aug-cc-pVDZ//MP2/(aug)-cc-pVDZ	399	3.11	10, 36
HBDI, EOM-CCSD/6-311G*// RI-MP2/cc-pVTZ	301	4.12 (0.82)	this work
HBDI, SOS-CIS(D)/cc-pVTZ// RI-MP2/cc-pVTZ	324	3.83 (1.10)	this work
HBDI, XMCQDPT2/cc-pVDZ// DFT(PBE0)/cc-pVDZ	375	3.31 (0.51)	this work

[a] Oscillator strength is given in parentheses. The experimental gas-phase value[10] is 3.35 eV or 370 nm. [b] The CASPT2/cc-pVTZ// DFT(BLYP) calculation from ref 41 employed the default IPEA zero-order Hamiltonian (by using the MOLCAS 7.2 program).

therefore, the planar minimum on S_1 can, in principle, be populated in the gas phase, given that the excited-state proton transfer is less favorable in the absence of a hydrogen-bonding network.

The coordinates of the twisted minimum (TwI-S_1) and of the MECI point (MECI−I-$S_{1/0}$) were located at the SA2-(w1=1,w2=1)CASSCF(12/11) level. The S_1-S_0 transition from TwI-S_1 has negligible oscillator strength. The energy of the conical intersection point (MECI−I-$S_{0/1}$) is approximately equal to the energy of the planar minimum. The diagram in Figure 5 summarizes the results. Selected geometric parameters of the Min-S_0, Plan-S_1, TwI-S_1, and MECI−I-$S_{0/1}$ structures are collected in Table 5.

In the Plan-S_1 structure, the C(I)−C bond stretches by 0.077 Å due to redistribution of the electronic density in the bridge region in the excited state relative to the ground state. For the same reason, the C(P)−C bridge bond shortens by 0.042 Å. The change in other bond lengths is less pronounced, but still notable. The TwI-S_1 structure is similar to that of the anion with the Twist-I angle being close to 90° and Twist-P being close to 0°. MECI−I-$S_{0/1}$ differs from that of the anionic form: the Twist-P angle is closer to zero, and there is no visible pyramidalization at the bridge carbon; instead, this structure features pyramidalization at the C(I) carbon. The MECI-$S_{0/1}$ structure suggests that the decay to the ground state can lead to both isomers, cis and trans, with almost equal probabilities.

Results for the Cationic Form

Table 6 and the diagram in Figure 6 show the results for the cationic form (Figure 1c) using the same methodology as for the anionic and neutral forms. First of all, we note that the S_0-S_1 transition wavelength at the ground state minimum energy point computed with XMCQDPT2/SA2-CASSCF(14/12)/cc-pVDZ//DFT(PBE0)/cc-pVDZ equals 406 nm (3.05 eV), which coincides with the experimental value of 406 nm.[8,9]

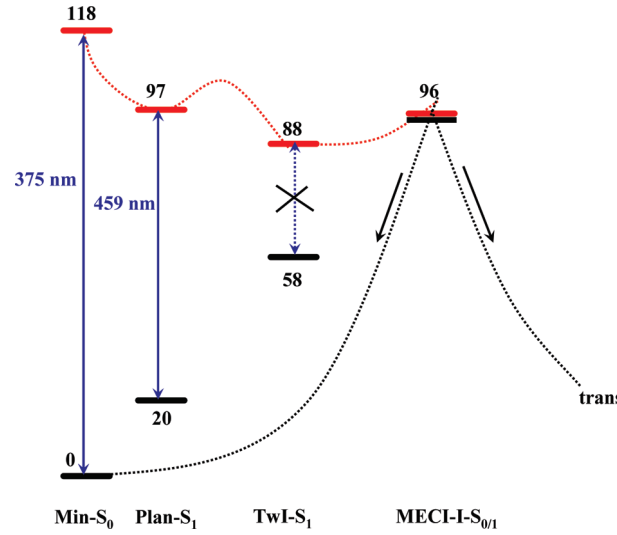


Figure 5. Energy diagram for the neutral form of HBDI in the vicinity of the cis conformation (Figure 1a). Thick black horizontal lines denote the ground state (S_0) levels; thick red lines - the excited state (S_1) levels. Structures corresponding to the minimum energy point on S_0 (Min-S_0), the planar minimum energy point on S_1 (Plan-S_1), the MECI point (MECI−I-$S_{0/1}$), and the minimum energy point on S_1 twisted over τ_I (TwI-S_1) are shown. The energy values (kcal/mol) refer to the present calculations using SA-CASSCF(12/11)/cc-pVDZ. The SA means equal weighting average over the first three states for Min-S_0, two states for TwI-S_1 and MECI, and the (0,1) weighting for Plan-S_1. We also show the wavelengths for the vertical S_0-S_1 transition energies at Min-S_0 and Plan-S_1 computed using the highest theoretical level: XMCQDPT2/SA3-CASSCF(14/12)/cc-pVDZ//DFT(PBE0)/cc-pVDZ at Min-S_0; XMCQDPT2/SA2-CASSCF(14/12)/cc-pVDZ//SA2-CASSCF(12/11)/cc-pVDZ at Plan-S_1.

Table 5. Selected Geometric Parameters of Neutral HBDI Optimized with CASSCF(12/11)/cc-pVDZ[a]

	Min-S_0	Plan-S_1	TwI-S_1	MECI−I-$S_{0/1}$
C(I)−C	1.351	1.428	1.440	1.428
C(P)−C	1.468	1.426	1.398	1.423
C=O (I)	1.197	1.211	1.237	1.222
C=O (P)	1.361	1.366	1.332	1.340
Twist-I	0	0	90.23	68.28
Twist-P	0	0	1.067	0.319

[a] Angles in degrees, bond lengths in Å. "Twist-I" is the N−C(I)−C−H dihedral angle value. "Twist-P" is the H−C−C(P)−C dihedral angle value.

Table 6. Calculated Vertical S_0-S_1 Excitation Energies of the Cationic Form of the GFP Chromophore[a]

system, calculation details	λ, nm	ΔE, eV
HBDI, EOM-CCSD/6-311G*// RI-MP2/cc-pVTZ	345	3.60
HBDI, SOS-CIS(D)/cc-pVTZ// RI-MP2/cc-pVTZ	380	3.27 (1.08)
HBDI, XMCQDPT2/cc-pVDZ// DFT(PBE0)/cc-pVDZ	406	3.06 (0.87)

[a] Oscillator strength is given in parentheses.

The selected geometric parameters of the Min-S_0, Plan-S_1, TwI-S_1, and MECI−I-$S_{0/1}$ structures are collected in Table 7.

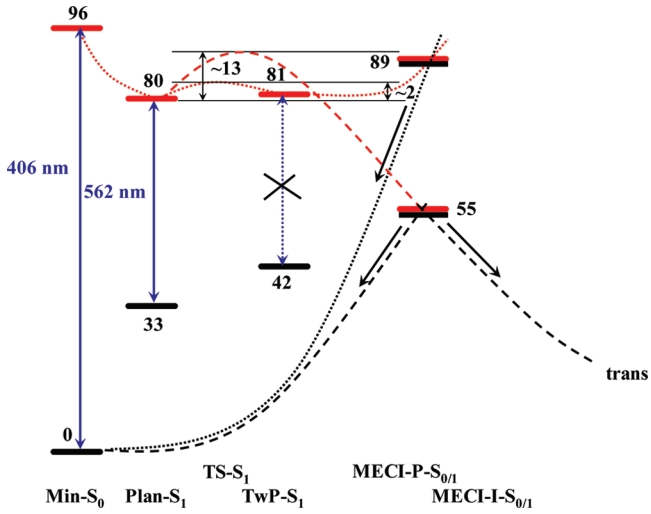


Figure 6. Energy diagram for the cationic form of HBDI at the cis conformation (Figure 1c). Thick black horizontal lines denote the ground state (S_0) levels; thick red lines - the excited state (S_1) levels. The structures corresponding to the minimum energy point on S_0 (Min-S_0), the planar minimum energy point on S_1 (Plan-S_1), the MECI points (MECI−P-$S_{0/1}$ and MECI−I-$S_{0/1}$), and the minimum energy point on S_1 twisted over τ_P (TwP-S_1) are shown. The energy values (kcal/mol) refer to the present calculations with SA2-CASSCF(12/11)/cc-pVDZ. The MECI structures were calculated with equal weighting average over the first two states, whereas structures of Plan-S_1, Tw−P-S_1, TS−P-S_1, and TS−I-S_1 were calculated with the (0,1) weighting. We also show the wavelengths for the vertical S_0−S_1 transition energies at Min-S_0 and Plan-S_1 computed using the highest theoretical level: XMCQDPT2/SA2-CASSCF(14/12)/cc-pVDZ//DFT(PBE0)/cc-pVDZ at Min-S_0; XMCQDPT2/SA2-CASSCF(14/12)/cc-pVDZ//SA2-CASSCF(12/11)/cc-pVDZ at Plan-S_1.

Table 7. Selected Geometric Parameters of the HBDI Cation Optimized with CASSCF(12/11)/cc-pVDZ[a]

	Min-S_0	Plan-S_1	MECI−I-$S_{0/1}$	MECI−P-$S_{0/1}$
C(I)−C	1.346	1.354	1.460	1.344
C(P)−C	1.470	1.453	1.389	1.473
C=O (I)	1.186	1.201	1.213	1.205
C=O (P)	1.341	1.300	1.313	1.264
Twist-I	5.32	2.99	86.90	3.17
Twist-P	32.768	8.378	2.313	82.36

[a] Angles in degrees, bond lengths in Å. "Twist-I" is the N−C(I)−C−H dihedral angle value. "Twist-P" is the H−C−C(P)−C dihedral angle value.

We also note that emission from the planar minimum on S_1 (Plan-S_1) corresponds to the wavelength of 562 nm (2.21 eV), which is the longest wavelength among various forms of HBDI. This planar minimum is separated from other stationary points on S_1 by the low-lying saddle point (less than 2 kcal/mol) with a single imaginary frequency of 92i. Another saddle point is more than 10 kcal/mol higher. There are two MECI points (MECI−P-$S_{0/1}$ and MECI−I-$S_{0/1}$) through which the decay to the ground state may occur. Emission from the twisted minimum on S_1, TwP-S_1, is forbidden.

The MECI−I-$S_{0/1}$ structure is similar to the TwI-S_1 structure of the anionic and neutral forms: Twist-I is close to 90°, whereas the Twist-P value is close to 0°. In MECI−P-$S_{0/1}$, the picture is reversed: Twist-P is close to 90°, and Twist-I is almost zero.

Table 8. Calculated Vertical S_0−S_1 Excitation Energies of the Zwitterionic Form of the GFP Chromophore[a]

system, calculation details	λ, nm	ΔE, eV
HBDI, EOM-CCSD/6-311G*// RI-MP2/cc-pVTZ	397	3.13 (1.09)
HBDI, SOS-CIS(D)/cc-pVTZ// RI-MP2/cc-pVTZ	477	2.60 (1.47)
HBDI, XMCQDPT2/cc-pVDZ// DFT(PBE0)/cc-pVDZ	503	2.46 (0.99)

[a] Oscillator strength is given in parentheses.

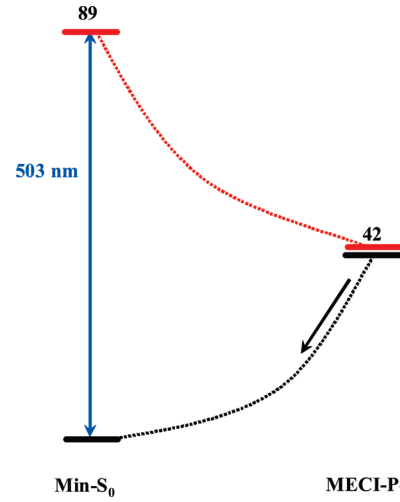


Figure 7. Energy diagram for the zwitterionic form of HBDI at the cis conformation (Figure 1d). Thick black horizontal lines denote the ground state (S_0) levels; thick red lines - the excited state (S_1) levels. The energy values (kcal/mol) refer to the present SA-CASSCF(12/11)/cc-pVDZ calculations. The SA means equal weighting average over the first two states (1,1). We also show the wavelength for the vertical $S_0 \rightarrow S_1$ transition at Min-S_0 computed with XMCQDPT2/CASSCF(14/12)/cc-pVDZ//DFT(PBE0)/cc-pVDZ.

Results for the Zwitterionic Form

The zwitterionic form of the GFP chromophore (Figure 1d) was considered in earlier studies[19] as a possible candidate for the emitting species. Our results are presented in Table 8 and in Figure 7.

Our calculations yield the simplest-shaped S_1 PES for this form, as shown in Figure 7. No minimum energy points on S_1 were located. Thus, the zwitterionic form would undergo barrierless relaxation to MECI, which is 47 kcal/mol below the Franck−Condon region. Zwitterionic MECI−P-$S_{0/1}$ resembles the structure of the anion's TwP-S_1: the Twist-P angle is almost 90°, the Twist-I is 0°, there is no pyramidalization at the bridge atoms, and C(I)−C is close to a double bond (1.36 Å as compared to 1.42 Å in Min-S_0), whereas C(P)−C stretches by 0.12 Å (see Table 9).

Results for the S_1−S_0 and S_0−S_1 Optical Transitions

The S_1−S_0 transition energies and oscillator strengths were calculated for the anionic, neutral, and cationic forms of

Table 9. Selected Geometric Parameters of the HBDI Zwitterion Optimized with CASSCF(12/11)/cc-pVDZ[a]

	Min-S_0	MECI−P-$S_{0/1}$
C(I)−C	1.420	1.363
C(P)−C	1.358	1.476
C═O (I)	1.214	1.204
C═O (P)	1.217	1.235
Twist-I	10.93	0
Twist-P	5.23	89.97

[a] Angles in degrees, bond lengths in Å. "Twist-I" is the N−C(I)−C−H dihedral angle value. "Twist-P" is the H−C−C(P)−C dihedral angle value.

HBDI at the Plan-S_1 geometry configurations (Table 10). For the zwitterionic form, the Plan-S_1 structure was not found, as described above.

The Stokes shift increases in a series: anion < neutral < cation. The shift for the anionic form in the gas phase (29 nm) is very close to the protein value of 32 nm.[1,2] Such a small value is due to strong resonance effects in the ground and excited states resulting in small geometry distortions of the respective equilibrium geometries (Min-S_0 and Plan-S_1). In the neutral form, the resonance is partially broken, and the structural differences are more pronounced. The largest shift (156 nm) is observed for the cationic form and can be attributed to significant differences between the Min-S_0 and Plan-S_1 structures described above. We note that the cationic structure shows not only the largest Stokes shift but also the longest S_1-S_0 transition wavelength, which is red-shifted by 39 nm relative to the anion.

Reference 43 reported the energies of the bright S_0-S_1 transitions obtained with the SA-CASSCF, CASPT2/SA-CASSCF, TDDFT, and EOM-CC methods employing different basis sets. Table 11 compares our XMCQDPT2/SA-CASSCF(14/12)/cc-pVDZ results with the CASPT2/SA-CASSCF(2/2)/6-31G* values from ref 43.

The results for the anionic form agree well, whereas for other forms the differences are significant. The discrepanices can be explained by the features of configurational composition of the S_1 CASSCF bright states of these forms. The anionic wave function is dominated by the HOMO−LUMO transition, whereas for other protonated forms several other important configurations appear with considerable weights, which is not adequately accounted for in small active spaces.

Conclusions

We presented the results of quantum chemical calculations of the transition energies, equilibrium geometries, and conical intersection points for the first two singlet electronic states of the model GFP chromophore, HBDI, in the vicinity of its cis conformation in the gas phase. We described all four possible protonation states of the chromophore at the uniform level of theory. The anionic form has been studied in detail before;[16,17,37] however, the high-level results for other forms are reported for the first time. We note a very good agreement of our results with the experimental data for the absorption band wavelengths of the anionic, cationic, and neutral forms.

Table 11. Comparison of the Present and Previously Reported Computed Excitation Energies for the S_0-S_1 Transition of Different Protonated Forms of HBDI

	this work		ref 43	
system	λ, nm	ΔE, eV	λ, nm	ΔE, eV
HBDI, anion	494	2.51	490	2.53
HBDI, neutral	375	3.31	340	3.65
HBDI, cation	406	3.06	435	2.85
HBDI, zwitterion	503	2.46	440	2.82

The principal difference between the (deprotonated) anionic and neutral forms stems from the changes in the frontier MOs: the protonation detunes the resonance and breaks the almost perfectly allylic character of the anionic MOs. Consequently, the bond alternation in the ground state and its change in S_1 are more pronounced in the neutral.

The flat shape of the S_1 PES of the anionic chromophore found in our and previous calculations is consistent with the experimentally observed lack of fluorescence in the gas phase. Given the shape of the surface, the excited state should rapidly decay to twisted conformations thus preventing fluorescence from the Plan-S_1 structure. Apparently, the shape of the S_1 PES should be considerably different in the protein environment.

In contrast, in the neutral form, the planar minimum on the S_1 surface may, in principle, be populated in the gas phase. Fluorescence decay in this case would involve the I-twist motion with a notable barrier (∼8 kcal/mol), and therefore, a longer lifetime is expected relative to the anion. The zwitterionic form has no planar minimum on the S_1 PES and features a low-lying MECI, thus enabling an efficient S_1 decay pathway through a space-conserving P-twist. The cationic S_1 surface turns out to be very flat along the P-twist coordinate featuring the P-twist minimum with a negligible S_1-S_0 oscillator strength. On the contrary, the barrier for I-twist is more than 13 kcal/mol, which is the largest value in our calculations. However, there is the I-twisted MECI point at lower energy relative to the plan-S_1 structure, so we cannot rule out this radiationless decay pathway. We also

Table 10. Calculated Vertical S_0-S_1 and S_1-S_0 Energies and the Respective Adiabatic Values of the Different Protonated Forms of the GFP Chromophore[a]

	S_0-S_1		S_1-S_0		
system and calculation details	λ, nm	ΔE, eV	λ, nm	ΔE, eV	Stokes shift, nm
HBDI, anion XMCQDPT2/SA2-CASSCF(14/12)/cc-pVDZ//CASSCF(12/11)/cc-pVDZ	494 (1.19)	2.51	523 (1.08)	2.37	29
HBDI, neutral XMCQDPT2/SA2-CASSCF(14/12)/cc-pVDZ//CASSCF(12/11)/cc-pVDZ	375 (0.51)	3.31	459 (0.36)	2.70	84
HBDI, cation XMCQDPT2/SA2-CASSCF(14/12)/cc-pVDZ//CASSCF(12/11)/cc-pVDZ	406 (0.87)	3.06	562 (0.44)	2.21	156

[a] Oscillator strength is given in parentheses.

note an enormous Stokes shift for the cationic form (156 nm) due to significant geometry relaxation in the Plan-S_1 state.

The I-twist energy profiles and the MECI structures suggest different excited-state isomerization pathways depending on the protonation state of the chromophore. The barrier for the I-twist motion decreases in the series: cation (13 kcal/mol) > neutral (8 kcal/mol) > anion (>2 kcal/mol). In the zwitterionic form, there is no Plan-S_1 stationary point, and a P-twist deactivation pathway is barrierless. The above barriers suggest that the motion along the I coordinate leading to twisted structures (minima and MECI) occurs much faster in the anionic form than in the neutral and cationic species. The gap between S_0 and S_1 at the Tw−I-S_1 geometry is 17 and 30 kcal/mol for the anionic and the neutral forms, respectively. With such small values (especially, 17 kcal/mol), one can expect a sufficient overlap between the S_0 and S_1 vibrational levels facilitating radiationless S_1-S_0 transition. Moreover, the SA-CASSCF calculations most likely overestimate the energy gap.[37] Overall, this result is in agreement with the observed ultrafast decay. Another pathway for the radiationless decay of the anionic form is via MECI, which is located almost at the same energy as Tw−I-S_1 but has a slightly different structure. Whereas the geometry of Tw−I-S_1 does not favor cis to trans isomerization over the return to the planar cis configuration on S_0, the MECI-I structure favors relaxation to the cis structure. In the neutral form, the MECI-I point is 8 kcal/mol above the Tw−I-S_1 structure. Both structures do not favor the relaxation to the cis structure over the isomerization. In the cationic form, MECI-I is the only twisted "minimum" on the S_1 surface along the I twist coordinate.

To summarize, the motion along the I coordinate for the anionic form is expected to be fast and may lead to cis to trans isomerization (with small yield) following S_1-S_0 radiationless transition via MECI-I. For the neutral and cationic forms, the relaxation should be slower due to larger energy gaps, and the structures do not favor any specific pathway, e.g., cis to trans isomerization versus the return the initial Plan-S_0 structure. Thus, we expect the isomerization yield to be higher for the neutral and cationic forms than for the anion. The I-Twisted conical intersections of the anionic, neutral, and cationic forms of the chromophore suggest possible pathways for the cis−trans isomerization, which is relevant to photoinduced dynamics in photoswitchable proteins.

Acknowledgment. This work was conducted under the auspices of the iOpenShell Center for Computational Studies of Electronic Structure and Spectroscopy of Open-Shell and Electronically Excited Species (http://iopenshell.usc.edu) supported by the National Science Foundation through the CRIF:CRF CHE-0625419 + 0624602 + 0625237 grant, as well as through the CHE-0951634 grant (A.I.K.). We thank Alex Granovsky and Dr. Ksenia Bravaya for generous help and valuable discussions. This work is also supported by the joint grant from the U.S. Civilian Research and Development Foundation (project RUC1-2914-MO-07) and the Russian Foundation for Basic Research (10-03-00085). I.V.P., B.L.G., and A.V.N. thank the SKIF-GRID program for providing resources at the SKIF-MSU, SKIF-SIBERIA, and UGATU computational facilities.

Supporting Information Available: Figures of the HBDI anion SA2-CASSCF(14/12)/cc-pVDZ, HBDI neutral SA3-CASSCF(14/12)/cc-pVDZ, HBDI zwitterion SA3-CASSCF(14/12)/cc-pVDZ, and HBDI cation SA3-CASSCF(14/12)/cc-pVDZ natural orbitals and their occupancies and the HBDI anion I-twisted minimum HOMO and LUMO orbitals. Tables of Cartesian coordinates. Details for the neutrals form S_0-S_1 transition XMCQDPT/CASSCF calculation. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Tsien, R. Y. *Annu. Rev. Biochem.* **1998**, *67*, 509.

(2) Zimmer, M. *Chem. Rev.* **2002**, *102*, 759.

(3) Nienhaus, G. U. *Angew. Chem., Int. Ed.* **2008**, *47*, 8992.

(4) Tonge, P. J.; Meech, S. R. *J. Photochem. Photobiol. A: Chem.* **2009**, *205*, 1.

(5) Meech, S. R. *Chem. Soc. Rev.* **2009**, *38*, 2922.

(6) van Thor, J. J.; Gensch, T.; Hellingwerf, K. J.; Johnson, L. N. *Nat. Struct. Biol.* **2002**, *9*, 37.

(7) Nielsen, S. B.; Lapierre, A.; Andersen, J. U.; Pedersen, U. V.; Tomita, S.; Andersen, L. H. *Phys. Rev. Lett.* **2001**, *87*, 228102.

(8) Andersen, L. H.; Lappierre, A.; Nielsen, S. B.; Nielsen, I. B.; Pedersen, S. U.; Pedersen, U. V.; Tomita, S. *Eur. Phys. J. D* **2002**, *20*, 597.

(9) Lammich, L.; Petersen, M. A.; Nielsen, M. B.; Andersen, L. H. *Biophys. J.* **2007**, *92*, 201.

(10) Rajput, J.; Rahbek, D. B.; Andersen, L. H.; Rocha-Rinza, T.; Christiansen, O.; Bravaya, K. B.; Erokhin, A. V.; Bochenkova, A. V.; Solntsev, K. M.; Dong, J.; Kowalik, J.; Tolbert, L. M.; Petersen, M. A.; Nielsen, M. B. *Phys. Chem. Chem. Phys.* **2009**, *11*, 9996.

(11) Forbes, M. W.; Jockusch, R. A. *J. Am. Chem. Soc.* **2009**, *131* (47), 17038.

(12) Olsen, S.; Manohar, L.; Martinez, T. J. *Biophys. J.* **2002**, *82*, 359.

(13) Toniolo, A.; Granucci, G.; Martinez, T. J. *J. Phys. Chem. A* **2003**, *107*, 3822.

(14) Megley, C. M.; Dickson, L. A.; Maddalo, S. L.; Chandler, G. J.; Zimmer, M. *J. Phys. Chem. B* **2009**, *113*, 302.

(15) Toniolo, A.; Olsen, S.; Manohar, L.; Martinez, T. J. *Faraday Discuss.* **2004**, *129*, 149.

(16) Martin, M. E.; Negri, F.; Olivucci, M. *J. Am. Chem. Soc.* **2004**, *126*, 5452.

(17) Altoe, P.; Bernardi, F.; Garavelli, M.; Orlandi, G.; Negri, F. *J. Am. Chem. Soc.* **2005**, *127*, 3952.

(18) Nemukhin, A. V.; Grigorenko, B. L.; Savitsky, A. P. *Acta Natur.* **2009**, *2*, 25.

(19) Voityuk, A. A.; Michel-Beyerle, M. E.; Rösch, N. *Chem. Phys. Lett.* **1997**, *272*, 162.

(20) Weber, W.; Helms, V.; McCammon, J. A.; Langhoff, P. W. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6177.

(21) Gross, L. A.; Baird, G. S.; Hoffman, R. C.; Baldridge, K. K.; Tsien, R. Y. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 11990.

(22) Schäfer, L. V.; Groenhof, G.; Klingen, A. R.; Ullmann, G. M.; Boggio-Pasqua, M.; Robb, M. A.; Grubmüller, H. *Angew. Chem., Int. Ed.* **2007**, *46*, 530.

(23) Wan, S.; Liu, S.; Zhao, G.; Chen, M.; Han, K.; Sun, M. *Biophys. Chem.* **2007**, *129*, 218.

(24) Topol, I.; Collins, J.; Polyakov, I.; Grigorenko, B.; Nemukhin, A. *Biophys. Chem.* **2009**, *145*, 1.

(25) Marques, M. A. L.; Lopez, X.; Varsano, D.; Castro, A.; Rubio, A. *Phys. Rev. Lett.* **2003**, *90*, 258101.

(26) Lopez, X.; Marques, M. A. L.; Castro, A.; Rubio, A. *J. Am. Chem. Soc.* **2005**, *127*, 12329.

(27) Xie, D.; Zeng, X. *J. Comput. Chem.* **2005**, *26*, 1487.

(28) Nemukhin, A. V.; Topol, I. A.; Burt, S. K. *J. Chem. Theory Comput.* **2006**, *2*, 292.

(29) Sun, M. *Int. J. Quantum Chem.* **2006**, *106*, 1020.

(30) Amat, P.; Granucci, G.; Buda, F.; Persico, M.; Tozzini, V. *J. Phys. Chem. B* **2006**, *110*, 9348.

(31) Timerghazin, Q. K.; Carlson, H. J.; Liang, C.; Campbell, R. E.; Brown, A. *J. Phys. Chem. B* **2008**, *112*, 2533.

(32) Helms, V.; Winstead, C.; Langhoff, P. W. *THEOCHEM* **2000**, *506*, 179.

(33) Das, A. K.; Hasegawa, J.-Y.; Miyahara, T.; Ehara, M.; Nakatsuji, H. *J. Comput. Chem.* **2003**, *24*, 1421.

(34) Olsen, S.; Smith, S. C. *J. Am. Chem. Soc.* **2007**, *129*, 2054.

(35) Bravaya, K. B.; Bochenkova, A. V.; Granovsky, A. A.; Nemukhin, A. V. *Russ. J. Phys. Chem. B* **2008**, *2*, 671.

(36) Bravaya, K. B.; Bochenkova, A. V.; Granovsky, A. A.; Savitsky, A. P.; Nemukhin, A. V. *J. Phys. Chem. A* **2008**, *112*, 8804.

(37) Olsen, S.; Smith, S. C. *J. Am. Chem. Soc.* **2008**, *130*, 8677.

(38) Virshup, A. M.; Punwong, C.; Pogorelov, T. V.; Lindquist, B. E.; Ko, C.; Martínez, T. D. *J. Phys. Chem. B* **2009**, *113*, 3280.

(39) Epifanovsky, E.; Polyakov, I.; Grigorenko, B.; Nemukhin, A.; Krylov, A. I. *J. Chem. Theory Comput.* **2009**, *5*, 1895.

(40) Polyakov, I.; Epifanovsky, E.; Grigorenko, B.; Krylov, A. I.; Nemukhin, A. *J. Chem. Theory Comput.* **2009**, *5*, 1907.

(41) Filippi, C.; Zaccheddu, M.; Buda, F. *J. Chem. Theory Comput.* **2009**, *5*, 2047.

(42) Ma, Y.; Rohlfing, M.; Molteni, C. *J. Chem. Theory Comput.* **2010**, *6*, 257.

(43) Olsen, S. The electronic excited states of green fluorescent protein chromophore models, 2004. Stanford University Web Site. http://mtzweb.stanford.edu/resources/theses/olsen/olsenthesis.htm (accessed July 5, 2010).

(44) Epifanovsky, E.; Polyakov, I.; Grigorenko, B.; Nemukhin, A.; Krylov, A. *J. Chem. Phys.* **2010**, *132*, 115104.

(45) Olsen, S.; Lamothe, K.; Martínez, T. D. *J. Am. Chem. Soc.* **2010**, *132*, 1192.

(46) Nakano, H. *J. Chem. Phys.* **1993**, *99*, 7983.

(47) Granovsky, A. A. Firefly (former PC GAMESS) Home Page. http://classic.chem.msu.su (accessed July 5, 2010).

(48) Krylov, A. I. *Annu. Rev. Phys. Chem.* **2008**, *59*, 433.

(49) Grimme, S. *J. Chem. Phys.* **2003**, *118*, 9095.

(50) Rhee, Y. M.; Head-Gordon, M. *J. Phys. Chem. A* **2007**, *111*, 5314.

(51) Shao, Y.; Molnar, L. F.; Jung, Y.; Kussmann, J.; Ochsenfeld, C.; Brown, S.; Gilbert, A. T. B.; Slipchenko, L. V.; Levchenko, S. V.; O'Neil, D. P.; Distasio, R. A., Jr.; 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.; Korambath, P. P.; Adamson, R. D.; Austin, B.; Baker, J.; Bird, E. F. C.; Daschel, H.; Doerksen, R. J.; Drew, A.; Dunietz, B. D.; Dutoi, A. D.; Furlani, T. R.; Gwaltney, S. R.; Heyden, A.; Hirata, S.; Hsu, C.-P.; Kedziora, G. S.; Khalliulin, R. Z.; Klunziger, P.; Lee, A. M.; Liang, W. Z.; 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 III, H. L.; Zhang, W.; Bell, A. T.; Chakraborty, A. K.; Chipman, D. M.; Keil, F. J.; Warshel, A.; Herhe, W. J.; Schaefer III, H. F.; Kong, J.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172.

(52) Bravaya, K.; Bochenkova, A.; Granovsky, A.; Nemukhin, A. *J. Am. Chem. Soc.* **2007**, *129*, 13035–13042.

CT100227K

JCTC Journal of Chemical Theory and Computation

Discrete Variable Representation Implementation of the One-Electron Polarization Model

Tae Hoon Choi,[†] Thomas Sommerfeld,[‡] S. Levent Yilmaz,[†] and Kenneth D. Jordan*,[†]

Department of Chemistry, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, Southeastern Louisiana University, Hammond, Louisiana 70402, and Center for Simulation and Modeling, University of Pittsburgh, Pittsburgh, Pennsylvania 15260

Received May 18, 2010

Abstract: A discrete variable representation (DVR) implementation of an one-electron polarization model (OPEM) for characterizing $(H_2O)_n^-$ clusters is described. For the $(H_2O)_{90}^-$ cluster, evaluation of the energy and gradient using a suitable DVR basis sets is about a 2 orders of magnitude faster than corresponding calculations using a Gaussian orbital basis set. The DVR version of the code has been parallelized using OpenMP to enable molecular dynamics (MD) simulations of large $(H_2O)_n^-$ clusters.

1. Introduction

The interaction of excess electrons with water clusters has been the subject of numerous experimental and theoretical studies.[1–40] A key to much of the computational work in this area has been the development of the electron−water potentials for use in one-electron model Hamiltonians.[27–40] Our group has introduced both a quantum Drude oscillator approach[33,34] and a computationally simpler one-electron polarizable model (OEPM)[39] for describing $(H_2O)_n^-$ clusters. The GTO implementation of the OEPM model, including the evaluation of analytical gradients, has been described in earlier publications.[39,41]

When using Gaussian type orbital (GTO) basis sets, the construction of the matrix elements $\mathbf{H}_{kl} = \int \varphi_k^*(x)\mathbf{H}\varphi_l(x)\mathrm{d}x$ is the major time-consuming part of the calculation of the energy and of the analytical gradients. We consider here an alternative approach involving the discrete variable representation (DVR) method.[42–56] Although the use of the DVR basis sets results in much larger Hamiltonian matrices, the matrix elements are rapid to evaluate, and the sparsity of the matrices allows for rapid diagonalization. In the following section, we describe the DVR implementation of the OEPM algorithm and compare its performance with the GTO implementation for $(H_2O)_n^-$ clusters with n as large as 90. We note that Jacobson and Herbert have also used a DVR approach for calculating the electron binding energies of their one-electron polarization model.[40] However, there are several differences between the two approaches, making a detailed examination of a DVR approach to our one-electron model instructive. Parallelization of the OEPM-DVR code using shared memory and OpenMP[57] is also described.

2. Computational Methods

2.1. DVR with the OEPM Hamiltonian. The OEPM method is built on top of the DPP water model developed in our group.[58] In this model, each water monomer carries three point charges, $+Q$ at each H atom and $-2Q$ at a so-called M site located 0.25 Å from the O atom on the rotational axis and displaced toward the H atoms. In the DPP model, each atom is polarizable, with Thole-type[59] damping of the charge-induced dipole and induced dipole−induced dipole interactions. Finally, there are exponential repulsive interactions between atoms of different monomers and attractive van der Waals interactions between the O atoms. In the OEPM method, the model Hamiltonian (in atomic units) for an excess electron interacting with a water cluster is

$$\mathbf{H} = -\frac{1}{2}\nabla^2 - \sum_i \frac{Q_i}{r_i} f_{\mathrm{pc}}(r_i) + \sum_j V_j^{\mathrm{rep}} + \sum_j \frac{\boldsymbol{\mu}_j \cdot \mathbf{r}_j}{r_j^3} f_{\mathrm{ind}}(r_j) - \sum_j \frac{\alpha}{2r_j^4} f_{\mathrm{pol}}(r_j) \quad (1)$$

* To whom correspondence should be addressed: E-mail: jordan@pitt.edu.

[†] University of Pittsburgh.

[‡] Southeastern Louisiana University.

10.1021/ct100263r © 2010 American Chemical Society
Published on Web 07/12/2010

where the first sum is over the charge sites in the DPP model, and the third term represents a short-ranged repulsive potential (V_{rep}), expanded in terms of six *s* GTOs on each monomer. The sums in the fourth and fifth terms, which correspond to the interaction of the excess electron with the induced dipoles (from the water−water interactions) and the induced dipoles from the electron−water interaction, are over the *M* sites. (Although each water molecule employs three polarizable sites for the water−water interaction, only a single polarizable site is used for the electron−water interaction. This reduces the computational effort while having little impact on the accuracy of the model.) It should be noted that for computational efficiency we have collapsed the induced dipoles on the atoms of each monomer to the associated *M* site. The f_{pc}, f_{ind}, and f_{pol} damping functions are represented as $f_{pc}(r_i) = 1 - \exp(-b_1 r_i^2)$, $f_{ind}(r_j) = (1 - \exp(-b_2 r_j^2))^2$, and $f_{pol}(r_j) = (1 - \exp(-b_3 r_j^2))^2$, respectively. The f_{pc} damping term is employed in the DVR but not in the GTO version of the code.

The DVR method is well documented in the literature,[42–50] and only a brief summary is given here. We adopt the sine DVR approach, which uses particle-in-a-box eigenfunctions as a basis. For simplicity we describe the DVR method for one-dimensional case; extension to three dimensions is straightforward. The underlying basis functions for the sine DVR method are

$$\varphi_j(x) = \begin{cases} \sqrt{\frac{2}{L}}\sin\left(\frac{j\pi(x - x_0)}{L}\right) & \text{for } x_0 \le x \le x_{n+1} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

where x_0 and x_{n+1} refer to the box edges and $L = x_{n+1} - x_0$. The DVR kinetic energy matrix elements can be evaluated analytically by transforming to eigenfunctions of the position operator expressed in terms of the φ_j. In this new basis

$$\mathbf{T}_{kl}^{DVR} = -\left(\frac{\pi(n+1)}{L}\right) \times \begin{cases} \frac{1}{3} + \frac{1}{6(n+1)^2} - \frac{1}{2(n+1)^2\sin^2\left(\frac{k\pi}{n+1}\right)} & \text{if } k = l \\ \frac{2(-1)^{k-1}}{(n+1)^2} \frac{\sin\left(\frac{k\pi}{n+1}\right)\sin\left(\frac{l\pi}{n+1}\right)}{\left(\cos\left(\frac{k\pi}{n+1}\right) - \cos\left(\frac{l\pi}{n+1}\right)\right)^2} & \text{if } k \ne l \end{cases} \quad (3)$$

With a δ function basis set, the potential energy matrix would be diagonal. In the finite basis sets used in actual applications, the eigenfunctions of the position operator approximate delta functions. Thus, for sufficiently large DVR basis sets, to a good approximation one can treat *V* as diagonal. For sine DVR basis sets, the diagonal elements of *V* are defined at the *n* uniformly spaced grid points $\{x_i\}$, where *n* is the number of sine functions in the basis set. The total Hamiltonian matrix for the 3-dimensional case has the form

$$\tilde{\mathbf{H}} = \begin{pmatrix} 3N^{4/3} \\ \text{nonzero elements} \end{pmatrix} + \begin{pmatrix} V(q_1) & 0 & 0 & 0 & \cdots & 0 \\ 0 & V(q_2) & 0 & 0 & \cdots & 0 \\ 0 & 0 & V(q_3) & 0 & \cdots & 0 \\ 0 & 0 & 0 & V(q_4) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & V(q_N) \end{pmatrix} \quad (4)$$

where *N* is the total number of DVR grid points ($N = n \times n \times n$), and q_i denotes a point on the three-dimensional grid. The kinetic energy matrix has $3N^{4/3}$ nonzero elements. The resulting $N \times N$ Hamiltonian matrix is very sparse, and the eigenvalue equation, $\mathbf{HC} = \varepsilon C$, is solved using iterative Lanczos diagonalization.[60]

2.2. Analytical Gradients in the DVR Method. In the DVR approach the gradient of the energy assumes a particularly simple form since there are no Pulay terms,[61,62] and the kinetic energy terms do not contribute. The gradient is thus given by

$$\frac{\partial E}{\partial R} = (C_1 \; C_2 \; C_3 \; \cdots \; C_N) \times \begin{pmatrix} \frac{\partial V(q_1)}{\partial R} & 0 & 0 & \cdots & 0 \\ 0 & \frac{\partial V(q_2)}{\partial R} & 0 & \cdots & 0 \\ 0 & 0 & \frac{\partial V(q_3)}{\partial R} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{\partial V(q_N)}{\partial R} \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \\ C_3 \\ \vdots \\ C_N \end{pmatrix} = \sum_{i=1}^{N} \frac{\partial V(q_i)}{\partial R} C_i^2 \quad (5)$$

where the C_i are the coefficients of the converged wave function.

2.3. Box Size and Grid Points. In the applications of the OEPM-DVR method to $(H_2O)_n^-$ clusters, it is important to choose appropriate box sizes and grid point spacing. This can be especially challenging for excess electron systems as the spatial extent of the excess electron can vary appreciable, depending on the arrangement of the water monomers. To illustrate the issues at play we consider the two $(H_2O)_6^-$ clusters and the two $(H_2O)_{45}^-$ clusters shown in Figure 1. One $(H_2O)_6^-$ structure (A) binds the excess electron by about 0.6 eV and the other (B) binds it by only ~0.1 eV. Both $(H_2O)_{45}^-$ structures strongly bind the excess electron, but one has it bound on the surface and the other in the interior. The geometries of the clusters were taken from ref 39.

3. Results

3.1. Convergence with Box Size and Density of Grid Points. The first set of exploratory calculations employed a cubic box of 30 bohr size and 1 bohr spacing of the grid points. In these and all other calculations carried out in this study, the

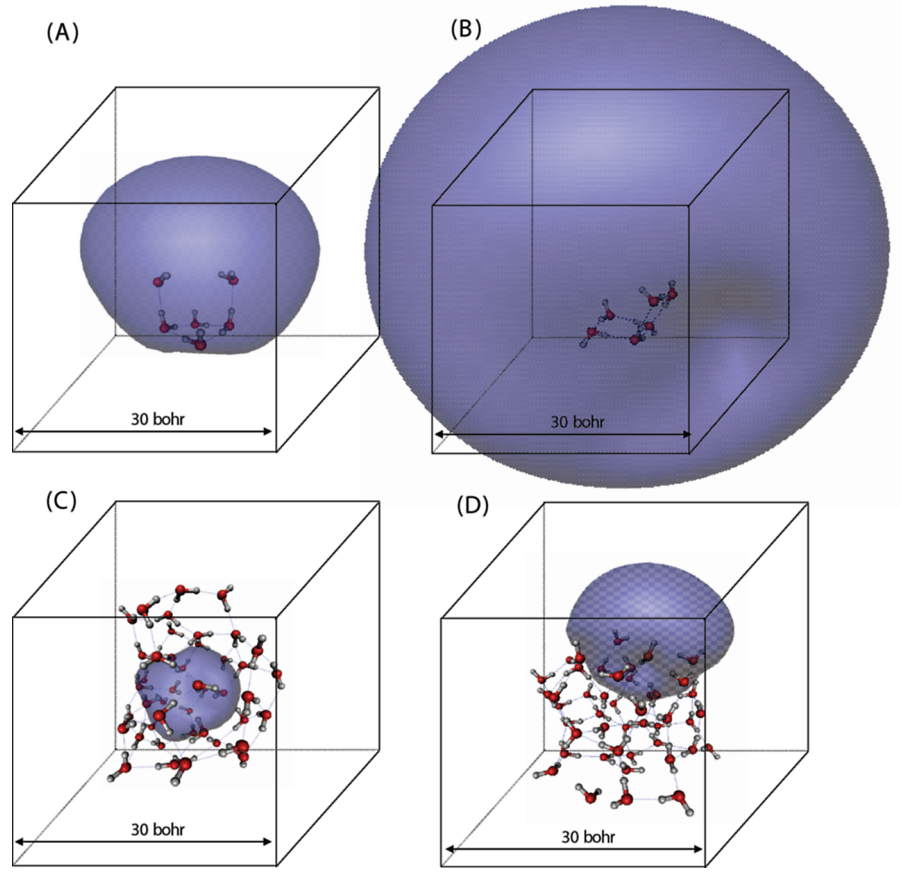


Figure 1. Isosurface containing 90% of the excess electron density calculated by the OEPM-GTO method are plotted for two $(H_2O)_6^-$ and two $(H_2O)_{45}^-$ clusters.

center of the cubic grid of DVR points was located at the center of mass of the cluster. For each of the four test systems the surface enclosing 90% of the density of the excess electron was calculated. The results are shown in Figure 1 from which it is immediately clear that the 30 bohr box is too small for the weak electron binding $(H_2O)_6^-$ species (B) but that it should be appropriate for the two $(H_2O)_{45}^-$ for which the excess electron is much more localized. One cannot tell from this figure alone as to whether this box size is suitable for the $(H_2O)_6^-$ (A) species.

In Figure 2 the electron binding energies of the four test systems are plotted as a function of the number of grid point (n) along one dimension of the cubic box for box lengths of 30, 60, and 90 bohr. For the 30 bohr box, the binding energies are relatively constant as n varies from 30 to 110, but for the larger boxes, the binding energies show appreciable oscillations for grids with spacings between the points greater than about 1.5 bohr. Clearly for convergence of the energies one needs to employ grids with spacings less than this. We also checked the convergence of the forces and find that in most cases, there are well converged for DVR basis sets that correspond to a grid spacings of ≤ 1 bohr.

From Figure 2, it is also seen that the excess electron is not even bound for $(H_2O)_6^-$ (B) when the 30 bohr box is used and that the use of this small box size leads to a sizable error in the EBE of the more strongly bound $(H_2O)_6^-$ species as well. The 30 bohr box, however, is suitable for the two $(H_2O)_{45}^-$ isomers, and the 60 bohr box is suitable for obtaining well converged EBE's for all four of the test systems.

4. Performance of the DVR Method

In evaluating the performance of the DVR method for describing $(H_2O)_n^-$ ions, comparison is made with the timings for calculations using a Gaussian-type basis set which has been optimized for use with the OEPM method.[39] This basis set employs three s and one p function on each water monomer and a 5s 4p set of GTOs at the center of mass of the cluster. Figure 3 reports the CPU time for the evaluation of the energy and gradient of clusters containing two to 90 water molecules with the DVR and the GTO basis sets. DVR results are reported for $30 \times 30 \times 30$, $60 \times 60 \times 60$, and $90 \times 90 \times 90$ grids. The energy and gradient evaluations with the GTO basis set both display $O(M^{2.9})$ scaling, where M is the number of monomers, whereas the corresponding DVR calculations essentially display $O(M^{1.0})$ and $O(M^{1.4})$ scaling, respectively. The weaker scaling in the DVR case is a consequence of the fact that the number of DVR grid points was not increased with increasing the number of molecules. Of course for very large clusters, if one did not know in advance where the electron is localized, it would be necessary to adopt larger DVR grids. For the $60 \times 60 \times 60$ grid, the energy plus gradient calculations with the DVR basis set are faster than calculations with the GTO basis set for $M \geq 17$.

With the DVR basis set, the most time-consuming part of the energy calculation is the diagonalization of the Hamiltonian matrix. Thus, the CPU time for the energy calculation in the DVR basis set depends mainly on the number of grid points, and only weakly (linearly) on the number of water

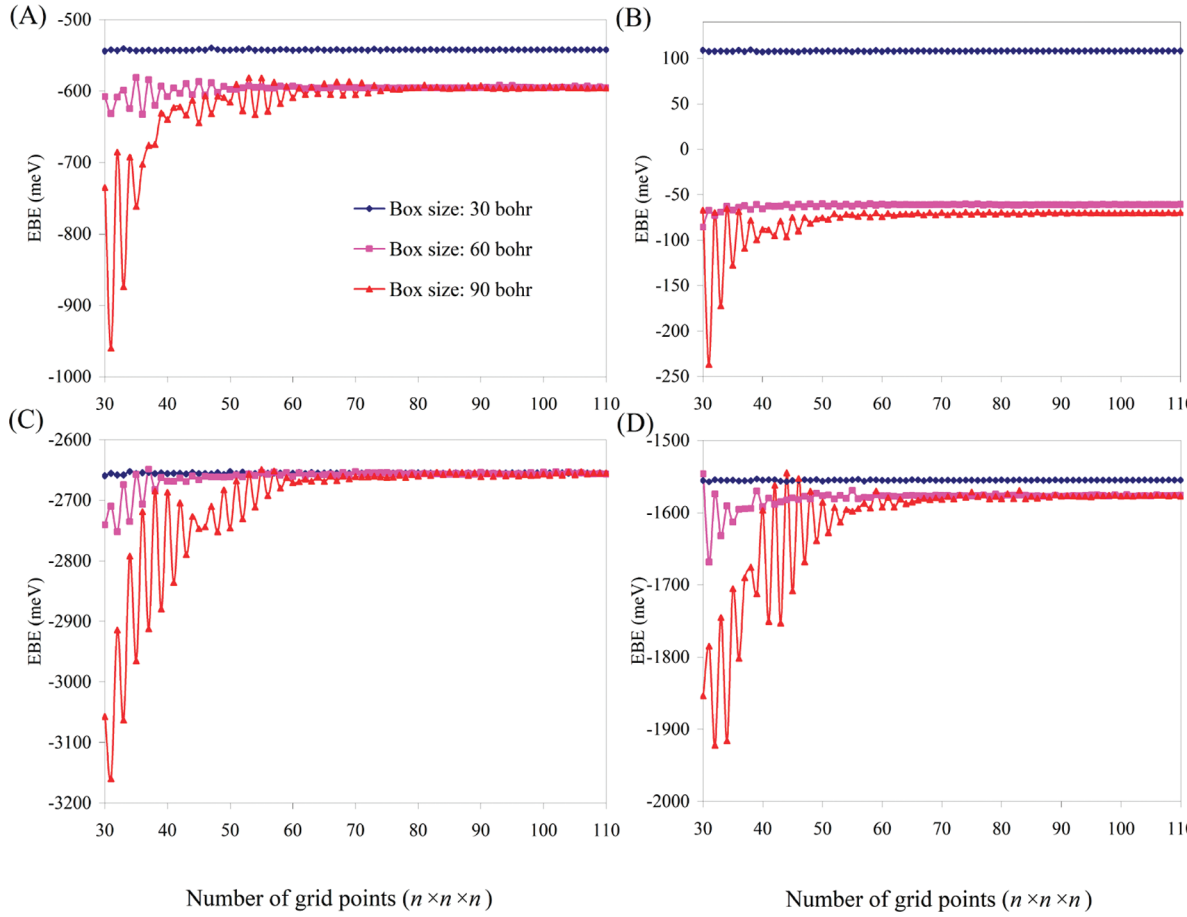


Figure 2. Electron binding energies of the four $(H_2O)_n^-$ clusters shown in Figure 1. Results are reported for 30 × 30 × 30, 60 × 60 × 60, and 90 × 90 × 90 bohr box sizes. The number of grid points in each dimension is varied from 30 to 110.

molecules (see Figures 3 and 4). For the gradient evaluations with the DVR method, there is a slightly stronger ~O($M^{1.4}$) scaling with the number of monomers. Figure 5 reports the computational time associated with the various parts of the gradient evaluation as a function of the number of water monomers. For the $n < 6$ clusters, the most time-consuming part of the gradient evaluation involves the derivatives of the repulsive potential, but for larger clusters, the derivatives of the induced dipole moment term dominate the gradient evaluation. The latter derivatives can be expressed as

$$\frac{\partial V_{\text{ind}}}{\partial R_k} = \frac{\partial}{\partial R_k}\sum_j \frac{\boldsymbol{\mu}_j \cdot \mathbf{r}_j}{r_j^3} = \sum_j \boldsymbol{\mu}_j \cdot \frac{\partial}{\partial R_k}\left(\frac{\mathbf{r}_j}{r_j^3}\right) + \sum_j \frac{\mathbf{r}_j}{r_j^3} \cdot \frac{\partial \boldsymbol{\mu}_j}{\partial R_k} \tag{6}$$

where the subscript k denotes the atomic coordinates and $\boldsymbol{\mu}_j$ is the induced dipole moment located at the M site of the jth molecule. The first term on the right-hand side of equation. Equation 6 can be expanded as

$$\begin{aligned}\sum_j \boldsymbol{\mu}_j \cdot \frac{\partial}{\partial R_k}\left(\frac{\mathbf{r}_j}{r_j^3}\right) &= \sum_j \mu_{xj}\frac{\partial}{\partial R_k}\left(\frac{r_{xj}}{r_j^3}\right) + \mu_{yj}\frac{\partial}{\partial R_k}\left(\frac{r_{yj}}{r_j^3}\right) + \mu_{zj}\frac{\partial}{\partial R_k}\left(\frac{r_{zj}}{r_j^3}\right) \\ &= \mu_{xk}\frac{\partial}{\partial R_k}\left(\frac{r_{xk}}{r_k^3}\right) + \mu_{yk}\frac{\partial}{\partial R_k}\left(\frac{r_{yk}}{r_k^3}\right) + \mu_{zk}\frac{\partial}{\partial R_k}\left(\frac{r_{zk}}{r_k^3}\right)\end{aligned} \tag{7}$$

Only the term with $k = j$ is nonzero, making the evaluation of this contribution to the gradient computationally fast. However, the second term on the right-hand side of equation. Equation 6, which can be expanded as

$$\sum_j \frac{\mathbf{r}_j}{r_j^3} \cdot \frac{\partial \boldsymbol{\mu}_j}{\partial R_k} = \sum_j \left[\left(\frac{r_{xj}}{r_j^3}\right)\frac{\partial \mu_{xj}}{\partial R_k} + \left(\frac{r_{yj}}{r_j^3}\right)\frac{\partial \mu_{yj}}{\partial R_k} + \left(\frac{r_{zj}}{r_j^3}\right)\frac{\partial \mu_{zj}}{\partial R_k}\right] \tag{8}$$

does not undergo such a simplification. In the gradient evaluation, only the induced dipole moment term involves a sum over molecules. Of the clusters considered here the computational effort associated with the terms in equations. Equations 7 and 8 are O($M^{1.0}$) and O($M^{1.8}$), respectively. Hence for sufficiently large clusters, the gradient evaluation in the DVR approach will display a computational scaling approach O($M^{1.8}$). (Again this assumes the number of grid points does not grow with system size, which in most applications can be assured by use of a prescreening step.) In general, in the absence of prescreening, the number of grid points would grow linearly water system size giving in overall scaling of O($M^{2.8}$), the same as with the GTO basis set.

The OEPM-DVR code has been parallelized using structured loop-level parallelism implemented via OpenMP.[57] The energy gradient and potential energy computations contain computationally intensive data-parallel loops, parallelization of which is straightforward. The iterative Lanczos algorithm used to calculate the energy, on the other hand, is implemented via a sequential solver

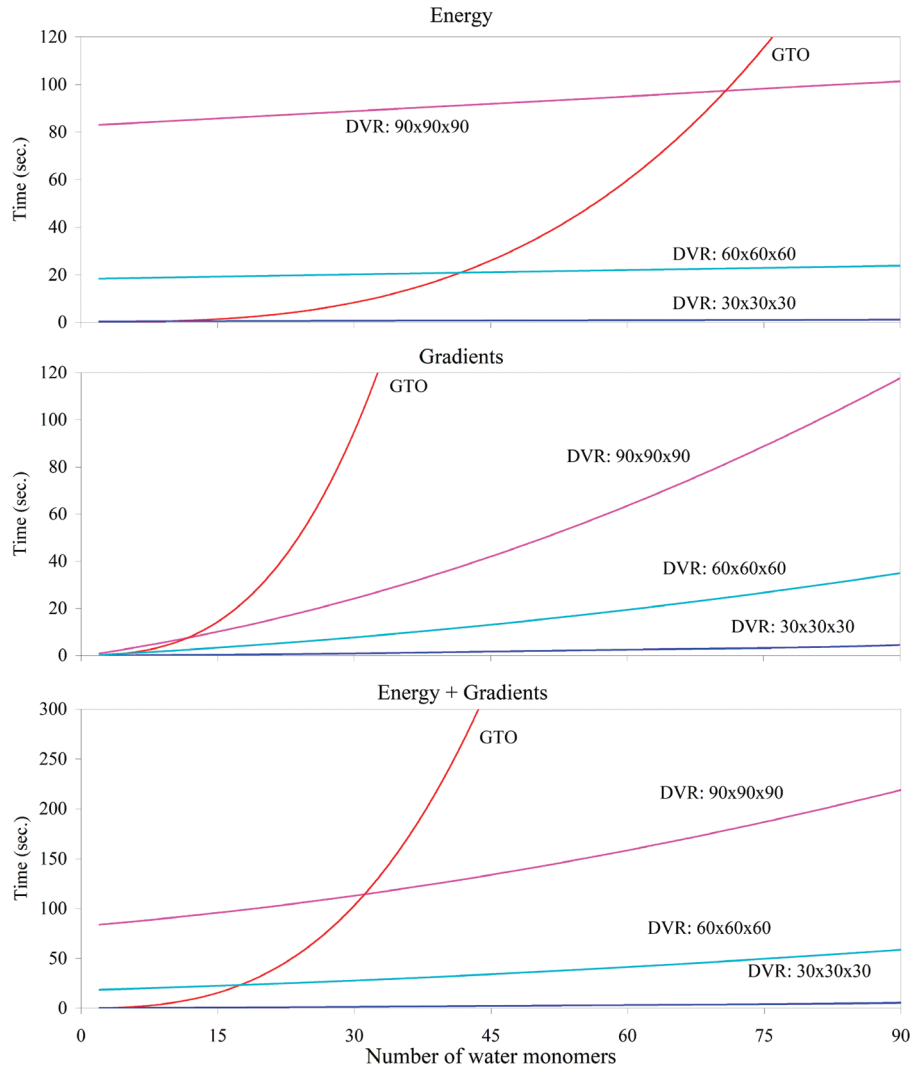


Figure 3. CPU time required to calculate the energy and gradient using GTO basis sets and three different DVR grids, 30 × 30 × 30, 60 × 60 × 60, and 90 × 90 × 90. The last graph shows the time of the energy and gradient together.

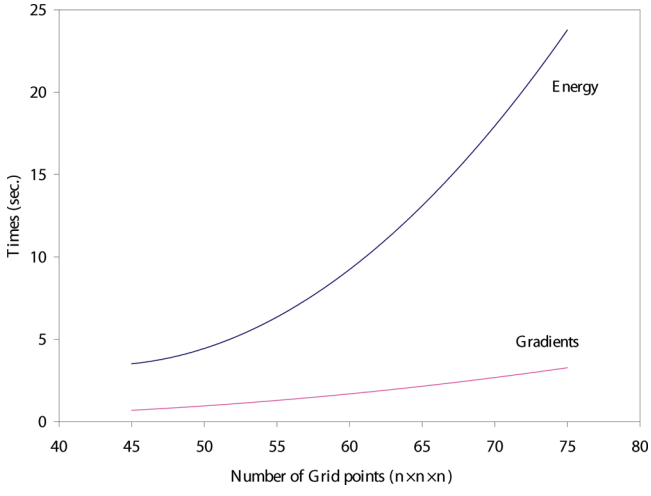


Figure 4. CPU time required to calculate the energy and gradients of a $(H_2O)_6^-$ cluster using 60 × 60 × 60 bohr box size. The number of grid points in the each direction is varied from 45 to 75.

library,[63] and the OpenMP parallelization was limited to only the eigenvector update stages during reverse communication with the solver.

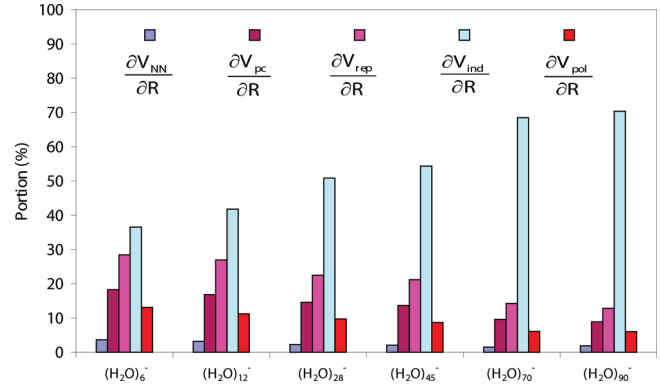


Figure 5. Relative CPU times for the various contributions to the gradient calculation.

Figure 6 reports the results of the parallel speed-up analysis. The runs are made on a SGI Altix 4700 shared-memory system. For the speedup analysis the energy and gradient for the $(H_2O)_{45}^-$ cluster are calculated using a 60 × 60 × 60 DVR grid. Figure 6 shows the normalized wall-time for various sections of the code as well as the overall run. Near-perfect speed-up is obtained for the potential energy and the gradient computations, while the diagonal-

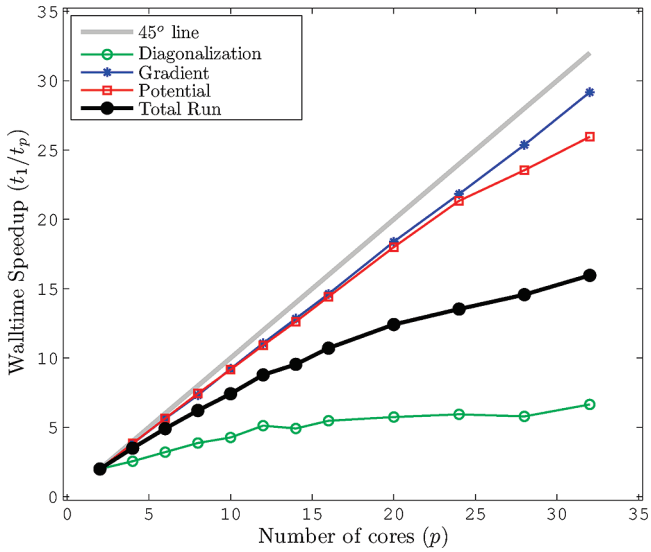


Figure 6. Parallel speed-up analysis of $(H_2O)_{45}^-$ cluster using a 60 × 60 × 60 DVR grid.

ization step does not perform well beyond 12 cores. For $(H_2O)_{45}^-$, the computer time to evaluate the enery exceeds that for the gradient, causing the efficiency of the parallelization of the energy plus gradient to perform poorly on over 12 CPUs as well. However, the computational effort to evaluate the gradient grows more rapidly with increasing cluster size than the time to evaluate the energy (see Figure 3), and for $(H_2O)_{90}^-$ the scaling of the parallelization is quite good up to about 32 cores. As a result nanosecond MD simulations are now possible on large $(H_2O)_n^-$ clusters using the OEPM approach even with the inclusions of interaction with the induced dipoles.

5. Conclusions

In this work, a sine-type DVR basis implementation of the one-electron polarization model for $(H_2O)_n^-$ clusters developed in our group has been described. Even with a grid as large as 60 × 60 × 60 [equivalent to 216 000 sine basis functions], the gradient calculations with the DVR approach are computationally faster than those using the GTO basis sets for clusters, containing more than 12 monomers.

The sine-type particle-in-the-box basis set, generates a uniform set of grid points. Significant computational savings could be achieved by using nonuniform grids, in particular using a sparser density of points in regions of space, where the electron density is relatively low. Exploratory calculations with harmonic oscillator DVR grids, in which the distance between grid points grows with increasing distance from the center of the box, shows that with such grids it is possible to achieve about an order of magnitude reduction in the CPU time required to evaluate the energy and gradients in the DVR approach with no loss of accuracy.

Acknowledgment. The authors acknowledge helpful discussions with Jing Ding. This research was carried out with support from the National Science Foundation under grant number CHE-0809457. T.S. acknowledges the donors of The American Chemical Society Petroleum Research Fund for support of this research through grant PRF no. 47344-GB6. The calculations were carried out on computers in the University of Pittsburgh's Center for Molecular and Materials Simulations and the Pittsburgh Supercomputing Center.

References

(1) Kim, K. S.; Park, I.; Lee, S.; Cho, K.; Lee, J. Y.; Kim, J.; Joannopoulos, J. D. *Phys. Rev. Lett.* **1995**, *76*, 956.

(2) Paik, D. H.; Lee, I. R.; Yang, D. S.; Baskin, J. S.; Zewail, A. H. *Science* **2004**, *306*, 672.

(3) Hammer, N. I.; Shin, J. W.; Headrick, J. M.; Kiden, E. G.; Roscioli, J. R.; Weddle, G. H.; Johnson, M. A. *Science* **2004**, *306*, 675.

(4) Jordan, K. D. *Science* **2004**, *306*, 618.

(5) Borgis, D.; Rossky, P. J.; Turi, L. *J. Chem. Phys.* **2006**, *125*, 064501.

(6) Coe, J. V.; Lee, G. H.; Eaton, J. G.; Arnold, S. T.; Sarkas, H. W.; Bowen, K. H.; Ludewigt, C.; Haberland, H.; Worsnop, D. R. *J. Chem. Phys.* **1990**, *92*, 3980.

(7) Verlet, J. R. R.; Bragg, A. E.; Kammrath, A.; Cheshnovsky, O.; Neumark, D. M. *Science* **2005**, *307*, 93.

(8) Bragg, A. E.; Verlet, J. R. R.; Kammrath, A.; Cheshnovsky, O.; Neumark, D. M. *Science* **2004**, *306*, 669.

(9) Hammer, N. I.; Roscioli, J. R.; Johnson, M. A.; Myshakin, E. M.; Jordan, K. D. *J. Phys. Chem. A* **2005**, *109*, 11526.

(10) Hammer, N. I.; Roscioli, J. R.; Johnson, M. A. *J. Phys. Chem. A* **2005**, *109*, 7896.

(11) Diken, E. G.; Robertson, W. H.; Johnson, M. A. *J. Phys. Chem. A* **2004**, *108*, 64.

(12) Hammer, N. I.; Roscioli, J. R.; Bopp, J. C.; Headrick, J. M.; Johnson, M. A. *J. Chem. Phys.* **2005**, *123*, 244311.

(13) Roscioli, J. R.; Hammer, N. I.; Johnson, M. A.; Diri, K.; Jordan, K. D. *J. Chem. Phys.* **2008**, *128*, 104314.

(14) Roscioli, J. R.; Johnson, M. A. *J. Chem. Phys.* **2007**, *126*, 024307.

(15) Haberland, H.; Schindler, H. G.; Worsnop, D. R. *Ber. Bunsen-Ges.* **1984**, *88*, 270.

(16) Haberland, H.; Ludewigt, C.; Schindler, H. G.; Worsnop, D. R. *J. Chem. Phys.* **1984**, *81*, 3742.

(17) Lee, H. M.; Suh, S. B.; Tarakeshwar, P.; Kim, K. S. *J. Chem. Phys.* **2005**, *122*, 044309.

(18) Suh, S. B.; Lee, H. M.; Kim, J.; Lee, J. Y.; Kim, K. S. *J. Chem. Phys.* **2000**, *113*, 5273.

(19) Lee, H. M.; Suh, S. B.; Kim, K. S. *J. Chem. Phys.* **2003**, *119*, 7685.

(20) Lee, H. M.; Lee, S.; Kim, K. S. *J. Chem. Phys.* **2003**, *119*, 187.

(21) Khan, A. *J. Mol. Struct.: THEOCHEM* **2008**, *850*, 144.

(22) Gutowski, M.; Skurski, P. *Recent Res. Dev. Phys. Chem.* **1999**, *3*, 245.

(23) Herbert, M. J.; Head-Gordon, M. *J. Phys. Chem. A* **2005**, *109*, 5217.

(24) Herbert, M. J.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2006**, *8*, 68.

(25) Barnett, R. N.; Landman, U.; Cleveland, C. L.; Jortner, J. *J. Chem. Phys.* **1988**, *88*, 4429.

(26) Frigato, T.; VandeVondele, J.; Schmidt, B.; Schutte, C.; Jungwirth, P. *J. Phys. Chem. A* **2008**, *112*, 6125.

(27) Wallqvist, A.; Thirumalai, D.; Berne, B. J. *J. Chem. Phys.* **1986**, *85*, 1583.

(28) Wallqvist, A.; Thirumalai, D.; Berne, B. J. *J. Chem. Phys.* **1987**, *86*, 6404.

(29) Schnitker, J.; Rossky, P. J. *J. Chem. Phys.* **1987**, *86*, 3462.

(30) Turi, L.; Borgis, D. *J. Chem. Phys.* **2002**, *117*, 6186.

(31) Turi, L.; Sheu, W. S.; Rossky, P. J. *Science* **2005**, *309*, 914.

(32) Turi, L.; Madarasz, A.; Rossky, P. J. *J. Chem. Phys.* **2006**, *125*, 014308.

(33) Wang, F.; Jordan, K. D. *J. Chem. Phys.* **2002**, *116*, 6973.

(34) Wang, F.; Jordan, K. D. *J. Chem. Phys.* **2003**, *119*, 11645.

(35) Sommerfeld, T.; Jordan, K. D. *J. Phys. Chem. A* **2005**, *109*, 11531.

(36) Sommerfeld, T.; Jordan, K. D. *J. Am. Chem. Soc.* **2006**, *128*, 5828.

(37) Sommerfeld, T.; Gardner, S. D.; DeFusco, A.; Jordan, K. D. *J. Chem. Phys.* **2006**, *125*, 174301.

(38) DeFusco, A.; Sommerfeld, T.; Jordan, K. D. *Chem. Phys. Lett.* **2008**, *455*, 135.

(39) Sommerfeld, T.; DeFusco, A.; Jordan, K. D. *J. Phys. Chem. A* **2008**, *112*, 11021.

(40) Jacobson, L. D.; Williams, C. F.; Herbert, J. M. *J. Chem. Phys.* **2009**, *130*, 124115.

(41) Choi, T. H.; Jordan, K. D. *Chem. Phys. Lett.* **2008**, *464*, 139.

(42) Beck, M.; Jackle, A.; Worth, G.; Meyer, H.-D. *Phys. Rep.* **2000**, *324*, 1.

(43) Lill, J. V.; Parker, G. A.; Light, J. C. *Chem. Phys. Lett.* **1982**, *89*, 483.

(44) Light, J. C.; Hamilton, I. P.; Lill, J. V. *J. Chem. Phys.* **1985**, *82*, 1400.

(45) Baye, D.; Heenen, P.-H. *J. Phys. A* **1986**, *19*, 2041.

(46) Vincke, M.; Malegat, L.; Baye, D. *J. Phys. B* **1993**, *26*, 811.

(47) Szalay, V. *J. Chem. Phys.* **1993**, *99*, 1978.

(48) Szalay, V.; Nemes, L. *Chem. Phys. Lett.* **1994**, *231*, 225.

(49) Baye, D.; Hesse, M.; Vincke, M. *Phys. Rev. E* **2002**, *65*, 026701.

(50) Littlejohn, R. G.; Cargo, M.; Jr., T. C.; Mithell, K. A.; Poirier, B. *J. Chem. Phys.* **2002**, *116*, 8691.

(51) Littlejohn, R. G.; Cargo, M. *J. Chem. Phys.* **2002**, *116*, 7350.

(52) Schneider, B. I.; Nygaard, N. *Phys. Rev. E* **2004**, *70*, 056706.

(53) Buisseret, F.; Semay, C. *Phys. Rev. E* **2005**, *71*, 026705.

(54) Xiao, Y.; Poirier, B. *J. Theoret. Comput. Chem.* **2007**, *6*, 309.

(55) Rayson, M. J. *Phys. Rev. E* **2007**, *76*, 026704.

(56) Yan, D.; Peng, L.-Y.; Gong, Q. *Phys. Rev. E* **2009**, *79*, 036710.

(57) http://www.openmp.org. Accessed March, 2010.

(58) DeFusco, A.; Schofield, D.; Jordan, K. D. *Mol. Phys.* **2007**, *105*, 2681.

(59) Thole, B. T. *Chem. Phys.* **1981**, *59*, 341.

(60) Lanczos, C. J. *Res. Natl. Bur. Stand.* **1950**, *45*, 255.

(61) Pulay, P. *Mol. Phys.* **1969**, *17*, 197.

(62) Meyer, W.; Pulay, P. *J. Chem. Phys.* **1972**, *56*, 2109.

(63) Lehoucq, R. B.; Sorensen, D. C.; Yang, C. *ARPACK Users' Guide: Solution of Large-Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods*; SIAM: Philadelphia, PA, 1998.

CT100263R

JCTC Journal of Chemical Theory and Computation

Study of Proton Coupled Electron Transfer in a Biomimetic Dimanganese Water Oxidation Catalyst with Terminal Water Ligands

Ting Wang, Gary W. Brudvig, and Victor S. Batista*

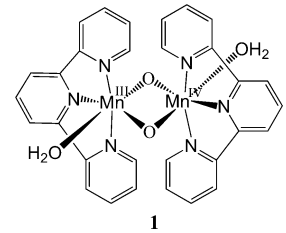
Department of Chemistry, Yale University, P.O. Box 208107, New Haven, Connecticut 06520-8107

Received May 19, 2010

Abstract: The oxomanganese complex $[H_2O(terpy)Mn^{III}(\mu\text{-}O)_2Mn^{IV}(terpy)H_2O]^{3+}$ (**1**, terpy = 2,2′:6-2″-terpyridine) is a biomimetic model of the oxygen-evolving complex of photosystem II with terminal water ligands. When bound to TiO_2 surfaces, **1** is activated by primary oxidants (e.g., Ce^{4+}(aq) or oxone in acetate buffers) to catalyze the oxidation of water yielding O_2 evolution [G. Li et al. *Energy Environ. Sci.* **2009**, *2*, 230−238]. The activation is thought to involve oxidation of the inorganic core $[Mn^{III}(\mu\text{-}O)_2Mn^{IV}]^{3+}$ to generate the $[Mn^{IV}(\mu\text{-}O)_2Mn^{IV}]^{4+}$ state $\mathbf{1}_{ox}$ first and then the highly reactive Mn oxyl species $Mn^{IV}O^{\bullet}$ through proton coupled electron transfer (PCET). Here, we investigate the step $\mathbf{1} \rightarrow \mathbf{1}_{ox}$ as compared to the analogous conversion in an oxomanganese complex without terminal water ligands, the $[(bpy)_2\ Mn^{III}(\mu\text{-}O)_2\ Mn^{IV}\ (bpy)_2]^{3+}$ complex (**2**, bpy = 2,2′-bipyridyl). We characterize the oxidation in terms of free energy calculations of redox potentials and pK_a's as directly compared to cyclic voltammogram measurements. We find that the pK_a's of terminal water ligands depend strongly on the oxidation states of the Mn centers, changing by ∼13 pH units (i.e., from 14 to 1) during the III,IV → IV,IV transition. Furthermore, we find that the oxidation potential of **1** is strongly dependent on pH (in contrast to the pH-independent redox potential of **2**) as well as by coordination of Lewis base moieties (e.g., carboxylate groups) that competitively bind to Mn by exchange with terminal water ligands. The reported analysis of ligand binding free energies, pK_a's, and redox potentials indicates that the III,IV → IV,IV oxidation of **1** in the presence of acetate (AcO^-) involves the following PCET: $[H_2O(terpy)Mn^{III}(\mu\text{-}O)_2Mn^{IV}(terpy)AcO]^{2+} \rightarrow [HO(terpy)Mn^{IV}(\mu\text{-}O)_2Mn^{IV}(terpy)AcO]^{2+} + H^+ + e^-$.

I. Introduction

Studies of redox transitions in synthetic oxomanganese compounds can provide fundamental insights on analogous redox processes in the oxygen-evolving complex (OEC) of photosystem II (PSII),[1−7] and in artificial photosynthetic systems based on biomimetic complexes.[8−10] The main challenge is to gain understanding of 'redox-leveling' processes where proton coupled electron transfer (PCET) leads to the accumulation of multiple oxidizing equivalents over a low range of potential. This paper explores the PCET mechanism associated with the oxidation of the oxomanganese complex $[H_2O(terpy)Mn^{III}(\mu\text{-}O)_2Mn^{IV}(terpy)H_2O]^{3+}$ (**1**, terpy = 2,2′:6-2″-terpyridine), shown in Figure 1, forming

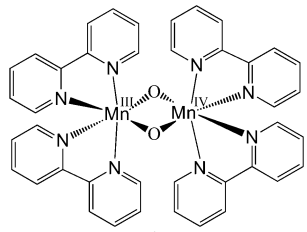


1

Figure 1. Schematic of complex **1** $[(H_2O)(terpy)Mn^{III}(\mu\text{-}O)_2Mn^{IV}(terpy)(H_2O)]^{3+}$ (terpy = 2,2′:6-2″-terpyridine).

* Corresponding author. Fax: +1 203 432 6144, E-mail: victor.batista@yale.edu.

10.1021/ct1002658 © 2010 American Chemical Society
Published on Web 07/15/2010



2

Figure 2. Schematic of complex **2** $[(bpy)_2Mn^{III}(\mu\text{-}O)_2Mn^{IV}(bpy)_2]^{3+}$ (bpy = 2,2′ bipyridyl).

the $\mathbf{1}_{ox}$ state through the oxidation $[Mn^{III}(\mu\text{-}O)_2Mn^{IV}]^{3+} \rightarrow [Mn^{IV}(\mu\text{-}O)_2Mn^{IV}]^{4+}$ of its inorganic core.[8,11−18]

Several Mn complexes have been proposed as structural models of the OEC of PSII where Mn ions are connected by μ-oxo linkages (i.e., deprotonated water molecules).[17,19−34] In addition to complex **1**, other Mn dimers were studied,[25,27,28] including the mixed valence Mn dimer **2** $[(bpy)_2\ Mn^{III}\ (\mu\text{-}O)_2\ Mn^{IV}\ (bpy)_2]^{3+}$ (bpy = 2,2′- bipyridyl), shown in Figure 2 (originally synthesized by Nyholm and Turco[29] and characterized by X-ray diffraction spectroscopy by Plaksin et al)[30] as well as Mn trimers[26] and tetramers.[22,31−34] Oxygen evolution catalyzed by complex **1** has been the subject of several studies.[8,11−16] However, the underlying catalytic mechanism and the nature of the reaction intermediates remain only partially understood.[12,14] It is, therefore, essential to advance our understanding of the reaction pathways, including changes in the protonation states of the ligands and the influence of competitive ligands that affect the redox potentials when binding to Mn.

The computations reported in this paper aim to provide a thorough characterization of the $\mathbf{1} \rightarrow \mathbf{1}_{ox}$ redox state transition, complementing earlier studies on the capabilities of density functional theory (DFT) methods for predictions of redox potentials of transition-metal complexes.[17,35−52] Most of these previous studies investigated functionals based on the generalized gradient approximation (GGA), such as BLYP,[53] BP86,[54] and Perdew−Burke−Ernzerhof (PBE)[55] as well as hybrid functionals (e.g., B3LYP),[56,57] including studies of oxomanganese complexes.[17,58] However, the analysis of PCET mechanisms in biomimetic oxomanganese complexes with terminal water ligands has yet to be reported and is addressed in this paper, including the regulatory effect of carboxylate binding on the redox properties of the complex.

In a recent study, we have shown that **1** binds to TiO_2 nanoparticles (NPs) by direct deposition, exchanging a water ligand by the TiO_2 NP.[8] Attachment to NPs with different degrees of crystallinity, including Degussa P25 (∼85% anatase and ∼15% rutile) and synthetic TiO_2 (sintered at 450 °C, D450, or without thermal treatment D70) was characterized by electron paramagnetic resonance (EPR), UV−vis spectroscopy, and electrochemistry. In addition, O_2 evolution was observed upon activation of **1** deposited on P25 using Ce^{4+} (or oxone in acetate buffer) as a primary oxidant.[8] However, understanding the underlying oxidation process still requires clarification of the nature of the ligation scheme, including the exchange of water by acetate ligands or TiO_2 NP's as well as changes in the protonation states of terminal waters and oxo/hydroxo bridges.

Cyclic voltammograms of **1** deposited on TiO_2 surfaces indicated that the redox activity of **1** was similar to its behavior in solution,[8,12,18,59] including a quasi-reversible redox couple at ∼1 V (P_a^1 and P_c^1)[8] assigned to the $\mathbf{1} \leftrightarrow \mathbf{1}_{ox}$ conversion and a broad cathodic wave at ca. 850 mV (P_c^2) due to reduction of the Mn^{IV} tetramer formed by spontaneous dimerization of $\mathbf{1}_{ox}$. In this study, we focus on the pH dependence of the redox potential associated with the $\mathbf{1} \leftrightarrow \mathbf{1}_{ox}$ conversion. In addition, we analyze the effect of oxidation on the pK_a's of terminal waters and acetate binding to Mn. The results are compared to the analogous (III,IV) → (IV,IV) transition in complex **2** where there are no water ligands involved. The reported analysis builds upon our recent work on the characterization of the oxidation state transition (III,III) → (III,IV) in **2**, as analyzed according to DFT calculations of redox potentials and pK_a's obtained by using the standard thermodynamic cycle formalism applied in conjunction with continuum solvation models.[58] The computational results provided fundamental insight on the mechanism responsible for the pH dependence of redox potentials, as described by the Pourbaix diagrams reported by voltammogram measurements,[25] where the linear dependence of $E_{1/2}$ with pH in the range pH = 3−9 (with a ∼59 mV/pH slope) was consistent with the one-electron one-proton couple $[(bpy)_2Mn^{III}(\mu\text{-}O)_2Mn^{IV}(bpy)_2]^{3+} + H^+ + e^- \rightarrow [(bpy)_2Mn^{III}(\mu\text{-}O)(\mu{-}OH)\text{-}Mn^{III}(bpy)_2]^{3+}$.[58] The agreement between theory and experiments showed that the DFT B3LYP level provided accurate descriptions of the regulatory effect of oxidations on the pK_a's of μ-hydroxo ligands as well as the effect of deprotonation of the bridges on the redox potentials of the metal centers. Here, we extend these earlier studies to explore the role of terminal water ligands in the regulation of redox potentials of complex **1**, including the analysis of deprotonation and carboxylate binding on the $\mathbf{1} \leftrightarrow \mathbf{1}_{ox}$ oxidation state transition. The reported analysis of PCET, coupling oxidation state transitions of Mn and deprotonation of water ligands attached to those metal centers, should provide insights on redox-leveling processes in the OEC of PSII[1−7] and artificial photosynthetic systems.[11,14,60]

The paper is organized as follows. Section II outlines the computational methods applied for calculations of pK_a's and redox potentials. Section III presents our computational results and direct comparisons with experimental measurements. Concluding remarks and future research directions are outlined in Section IV.

II. Computational Methods

The methods implemented in this study have already been described in our previous work on oxomanganese complexes.[17,58] Here, we outline the computational methodology only briefly.

Electronic Structure Calculations. All electronic structures were obtained at the Becke-3−Lee−Yang−Parr (B3LYP) DFT level,[56,57] using the Jaguar suite of electronic structure programs. Minimum energy configurations of complexes **1** and **2** in the reduced (III,IV) and oxidized (IV,IV) states were obtained in broken symmetry (BS) spin

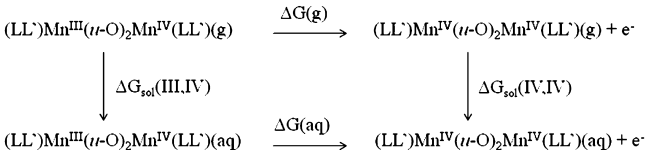


Figure 3. Born−Haber thermodynamic cycle used for calculations of free energy changes $\Delta G(\text{aq}) = -nFE^0$ yielding the standard redox potentials E^0, as described in the text, with L = terpy and L′ = H_2O for complex **1** and L = L′ = bpy for complex **2**.

states with α and β electronic densities localized on different metal centers.[61,62] All optimizations involved unrestricted Kohn−Sham wave functions (UB3LYP) and yielded antiferromagnetically coupled high-spin manganese centers. We applied a mixed basis set, including the LACVP basis that accounts for a nonrelativistic description of effective core potentials (ECPs) for the Mn^{4+} and Mn^{3+} centers, the 6-31G (2df) basis sets including polarization functions for μ-oxo O^{2-} species, and the 6-31G basis sets for the rest of the atoms. All optimizations were followed by UB3LYP single point energy calculations based on Dunning's correlation-consistent triple-ζ basis set[63−65] cc-pVTZ(-f), including a double set of polarization functions. Very similar results were also obtained for the cc-pVTZ(-f)++ basis set, as previously reported for other systems.[43]

Calculations of Standard Potentials. The half-cell standard potentials E^0 for the redox state transitions (III,IV) ↔ (IV,IV) of complexes **1** and **2** were obtained from DFT calculations of free energy changes $\Delta G(\text{aq}) = -nFE^0$, where $n = 1$ is the number of electrons involved in the redox couple and $F = 23.06$ kcal mol^{-1} V^{-1} is the Faraday constant. Redox potentials were reported relative to the normal hydrogen electrode (NHE), experimentally determined to be 4.43 eV.[66] Therefore, we have subtracted 4.43 V from the absolute potentials to make direct comparisons to cyclic voltammogram data referred to the NHE.

Free Energy Calculations of Standard Potentials. Using the Born−Haber cycle, depicted in Figure 3, we computed $\Delta G(\text{aq}) = \Delta G(\text{g}) + \Delta G_{\text{sol}}(\text{III,III}) - \Delta G_{\text{sol}}(\text{III,IV})$ from the free energy change $\Delta G(\text{g}) = \Delta H(\text{g}) - T\Delta S(\text{g})$ due to the redox state transition in the gas phase, where $\Delta H(\text{g}) = \Delta H_{\text{EA}}(\text{DFT}) + \Delta H_{\text{ZPE}} + \Delta H_{\text{T}}$ was obtained from the DFT electron attachment enthalpies $\Delta H_{\text{EA}}(\text{DFT})$ in the gas phase, while the changes in the zero point energy ΔH_{ZPE} and the corrections for molecular entropy changes $\Delta S(\text{g})$ were based on vibrational frequency calculations.

The solvation free energies $\Delta G_{\text{sol}}(\text{III,IV})$ and $\Delta G_{\text{sol}}(\text{IV,IV})$ were computed by using the self-consistent reaction field (SCRF) approach of standard continuum solvation models.[67,68] All calculations were based on gas-phase geometries employing the dielectric constant of water ($\varepsilon = 80.37$) for the continuum medium and on a solvent radius of 1.40 Å. Corrections due to hydrogen bonding with solvent molecules and changes in the thermal enthalpy ΔH_{T} were neglected.[43]

Free Energy Calculations of pK_a's. The pK_a's of water terminal ligands were obtained by computing the DFT free energy changes $\Delta G_a(\text{aq})$ due to deprotonation of the complexes, as follows: $\text{p}K_a = \beta\Delta G_a(\text{aq})$, where $\beta = (k_B T)^{-1}$ with

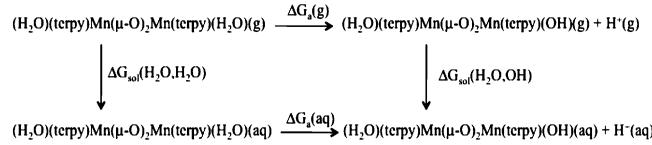


Figure 4. Born−Haber thermodynamic cycle used for calculations of free energy changes $\Delta G_a(\text{aq}) = k_B T\,\text{p}K_a$ yielding the pK_a of terminal water ligands in complex **1**.

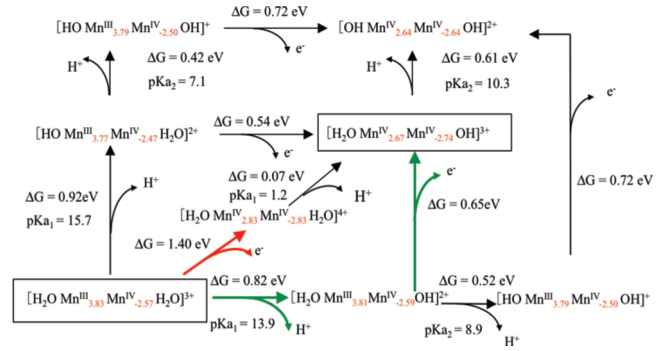


Figure 5. Thermodynamic diagram of PCET for complex **1** in aqueous solutions at pH = 0, obtained from free energy calculations of redox potentials and pK_a's at the DFT B3LYP/cc-pVTZ(-f) level, based on the Haber-Born cycle method applied in conjunction with a continuum solvation model. Formal oxidation numbers are indicated as superscripts in roman numbers, and the spin populations obtained according to the Mulliken population analysis are indicated as subscripts in red. For simplicity, the terpyridine ligands attached to Mn centers have been omitted.

$T = 298.15$ K, and k_B is the Boltzmann constant. The values of $\Delta G_a(\text{aq})$, due to deprotonation of the terminal water ligands for complex **1** in aqueous solutions, were obtained by using the Born−Haber cycle depicted in Figure 4, as follows:

$$\Delta G_a(\text{aq}) = \Delta G_a(\text{g}) + \Delta G_{\text{sol}}(\text{H}_2\text{O,OH}) + \Delta G_{\text{sol}}(\text{H}^+) - \Delta G_{\text{sol}}(\text{H}_2\text{O,H}_2\text{O}) \quad (5)$$

where $\Delta G_a(\text{g}) = \Delta H_a(\text{g}) - T\Delta S_a(\text{g})$ is the free energy change due to deprotonation of **1** in the gas phase, and $\Delta H_a(\text{g}) = \Delta H_a(\text{DFT}) + \Delta H_{\text{ZPE}} + \Delta H_{\text{T}}$ is the total enthalpy change obtained from the DFT energy change $\Delta H_a(\text{DFT})$. The solvation free energies associated with the protonated and deprotonated forms of the complex, $\Delta G_{\text{sol}}(\text{H}_2\text{O,H}_2\text{O})$ and $\Delta G_{\text{sol}}(\text{H}_2\text{O,OH})$, respectively, were estimated by using continuum solvation models, with the solvation free energy of H^+ in water $\Delta G_{\text{sol}}(\text{H}^+)$ taken as −260 kcal mol^{-1}.[66,69,70]

III. Results

Figure 5 shows the free energy diagram for PCET in complex **1** obtained, as described in Section II, from calculations of redox potentials and pK_a's at the DFT B3LYP/cc-pVTZ(-f) level of theory. The diagram shows that the (III,IV) state of complex **1** has terminal water ligands at pH < 13.9, while the oxidized (IV,IV) state spontaneously deprotonates one of its water ligands at pH > 1.2. In addition, the results summarized in Figure 5 show that the (III,IV) → (IV,IV) oxidation of **1** requires only 0.65

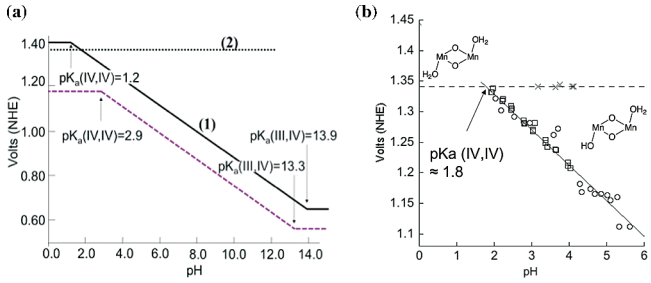


Figure 6. (a) Pourbaix diagrams for complexes **1** and **2**, obtained from free energy calculations of redox potentials and pK_a's at the DFT B3LYP/cc-pVTZ(-f) level of theory, in aqueous solutions (black solid line) and with acetate binding (purple dashed line). (b) Experimental CV data (open circles, ○, and squares, □, for **1** and crosses, ×, for **2**), as reported in ref 18.

$$[H_2O\ Mn^{III}Mn^{IV}\ H_2O]^{3+} + AcO^- \xrightarrow{\Delta G = -0.22eV} [AcO\ Mn^{III}Mn^{IV}\ H_2O]^{2+} + H_2O$$

$$[H_2O\ Mn^{III}Mn^{IV}\ H_2O]^{3+} + AcO^- \xrightarrow{\Delta G = -1.01eV} [H_2O\ Mn^{III}Mn^{IV}\ AcO]^{2+} + H_2O$$

$$[H_2O\ Mn^{III}Mn^{IV}\ AcO]^{2+} + AcO^- \xrightarrow{\Delta G = 0.11eV} [AcO\ Mn^{III}Mn^{IV}\ AcO]^{+} + H_2O$$

Figure 7. Free energy changes upon acetate (AcO^-) binding to complex **1**, obtained at the DFT B3LYP/cc-pVTZ(-f) level of theory. For simplicity, oxo bridges and terpy ligands attached to Mn centers have been omitted.

V at pH > 13.9 after the water ligand bound to Mn^{IV} is deprotonated forming $[H_2O(terpy)Mn^{III}(\mu\text{-}O)_2Mn^{IV}terpy)\text{-}(OH]^{2+}$, while oxidation before deprotonation (at pH < 1.2) is thermodynamically much more demanding ($E^0 = 1.40$ V).

For intermediate values of pH = 1.2−13.9, the oxidation is coupled to deprotonation through PCET. The total energy requirement for the (III,IV) → (IV,IV) oxidation of **1** includes (0.82−0.059 pH) eV to deprotonate the water ligand bound to Mn^{IV}, plus 0.65 eV necessary to oxidize Mn^{III} in the deprotonated complex $[H_2O(terpy)Mn^{III}(\mu\text{-}O)_2Mn^{IV}\text{-}(terpy)OH]^{2+}$. Therefore, by coupling the deprotonation and the oxidation by PCET, the complex can be oxidized at a potential (0.65 + 0.82 − 0.059 pH) V lower than 1.4 V and deprotonated at pH < 13.9. In contrast, in the absence of coupling, the water ligand bound to Mn^{IV} deprotonates only at pH ≥ 13.9, and the complex is oxidized only at voltage ≥1.4 V. In a cyclic voltammogram (CV) measurement, the deprotonation during PCET is driven by the externally applied electric field. Therefore, at pH = 1.2−13.9, the position of the CV peak accounts for both free energy changes, including the deprotonation of the water ligand and the oxidation of the Mn center, yielding a redox potential that changes linearly by 59 mV per pH unit.

Figure 6, left panel, shows the Pourbaix diagram based on ab initio calculations of pK_a's and redox potentials obtained at the DFT B3LYP/cc-pVTZ(-f) level of theory. These calculations predict the redox potential of **1** for the entire pH range, without relying on any kind of experimental data. The experimental CV measurements (Figure 6, right panel) were carried out as described in ref 18.

The calculations for **1** in aqueous solutions predict a linear dependence of $E_{1/2}$ with pH, changing at a rate of 59 mV/pH within the 1.2 < pH < 13.9 range (Figure 6, left panel) in agreement with experimental data (Figure 6, right panel).[18] Beyond this pH range (i.e., at either pH < 1.2 or pH > 13.9), the oxidation of **1** is pH-independent (non-PCET) with the crossover points of pH-independent and -dependent processes defined by the pK_a's of the water ligands in the (III,IV) and (IV,IV) oxidation states. In contrast to the pH-dependent results obtained for complex **1**, the calculated redox potential of **2** is constant ($E^0 = 1.36$ eV) and pH independent since the oxo bridges in **2** are deprotonated at pH > 2.[58] These results are consistent with the CV data for **2** (Figure 6, right panel) which show a reversible one-electron anodic couple at $E_{1/2} = 1.34$ V (vs NHE), assigned to the oxidation of the III,IV complex to the IV,IV state[19,71] that is independent of pH. The favorable comparison between theory and experiments, shown in Figure 6, indicates that the DFT B3LYP/cc-pVTZ(-f) level of theory provides redox potentials for biomimetic oxomanganese complexes with terminal water molecules that agree with the experimental values within an estimated error of ±60 mV, yielding a semiquantitative description of PCET.

In order to investigate the regulatory effect of carboxylate moieties upon competitive binding to Mn by exchange with terminal ligands, we have analyzed the free energy changes due to exchange of water ligands of **1** by acetate. Figure 7 shows that either of the two water ligands (but not both) can spontaneously exchange with acetate, with binding to Mn^{IV} yielding the most stable complex in aqueous solution.

The calculations of binding affinities, shown in Figure 7, indicate that the predominant species in acetate buffers is $[H_2O(terpy)Mn^{III}(\mu\text{-}O)_2Mn^{IV}(terpy)AcO]^{2+}$. The Pourbaix diagram of such a complex, shown in Figure 6 (left panel, purple dashed line), indicates that the acetate binding stabilizes the oxidized form of the complex throughout the whole pH range, decreasing the redox potential of the III→IV transition by as much as 90−220 mV.

Figure 8 shows the detailed analysis of the effect of acetate on the underlying PCET mechanism in complex **1**, as characterized by calculations of redox potentials and pK_a's at the DFT B3LYP/cc-pVTZ(-f) level. It is shown that acetate binding changes the first pK_a of the water ligand from 13.9 to 13.3, in the Mn(III,IV) state and from 1.2 to 2.9 in the Mn(IV,IV) state. Consequently, the pH-range for PCET is shifted from pH = 1.2−13.9 in the absence of acetate to pH = 2.9−13.3 with acetate binding. In addition, Figure 8 shows that acetate binding to Mn^{IV} reduces the pH-independent oxidation potentials from 1.40 to 1.18 V at low pH when H_2O is bound to Mn^{III} and from 0.65 to 0.56 V at high pH when HO^- is ligated to Mn^{III}. Therefore, the redox potential of **1** is 1.18 V at pH < 2.9, since the oxidation takes the red pathway in Figure 8. At pH = 2.9−13.3, the potential decreases linearly at a rate of 59 mV/pH due to PCET (green path with nonspontaneous deprotonation), while at pH > 13.3 it remains constant at 0.56 V, since there is spontaneous deprotonation and the green path in Figure 8 prevails. The first crossover is defined by the pK_a of Mn(IV,IV), the pH at which the red and green pathways are energetically identical. Analogously, the pH of the subsequent crossover point is defined by the pK_a value of Mn(III,IV). The complete

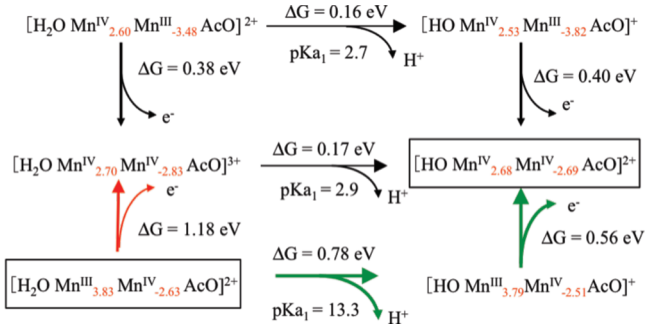


Figure 8. Thermodynamic diagram of PCET for complex **1** with acetate (AcO^-) binding in aqueous solutions at pH = 0, obtained from free energy calculations of redox potentials and pK_a's at the DFT B3LYP/cc-pVTZ(-f) level, based on the Born−Haber cycle method applied in conjunction with a continuum solvation model. Formal oxidation numbers are indicated as superscripts in roman numbers, and the spin populations obtained according to the Mulliken population analysis are indicated as subscripts in red. For simplicity, oxo bridges and terpy ligands attached to Mn centers have been omitted.

description provided by the reported binding energies, pK_a's, and redox potentials indicates that one of the two water ligands in **1** must exchange with acetate, yielding the III,IV → IV,IV oxidation state transition: $[H_2O(terpy)Mn^{III}(\mu\text{-}O)_2Mn^{IV}(terpy)AcO]^{2+} \rightarrow [HO(terpy)Mn^{IV}(\mu\text{-}O)_2Mn^{IV}(terpy)AcO]^{2+} + H^+ + e^-$, involving PCET in the pH = 2.9−13.3 range.

Our analysis of the oxidation mechanism of complex **1**, as influenced by the protonation states of water terminal ligands and the competitive binding of carboxylate moieties, suggests the following trends that might be common to other oxomanganese complexes: (1) at least one carboxylate moiety per high-valent di-μ-oxo-bridged Mn dimer can exchange with terminal water ligands and bind to Mn; (2) the redox potential of the di-μ-oxo-bridged Mn dimer is typically reduced by 100−200 mV upon carboxylate binding and by 620−750 mV upon deprotonation of a terminal water ligand; and (3) carboxylate binding reduces the pH range for PCET by about 2 pH units. Similar effects due to coordination of carboxylate moieties and deprotonation of water ligands might be expected in the OEC of PSII, where the regulatory mechanism based on PCET is thought to prevent the buildup of charge in the cluster through deprotonation of water/hydroxo ligands coupled to oxidation state transitions.

The molecular structure of the OEC of PSII remains only partially understood,[7] and several structural models have been proposed, including the '3 + 1 Mn tetramer'.[2,3] Such a model consists of a cuboidal structure of oxo-bridged high-valent Mn ions chelated by terminal water ligands as well as carboxylate and imidazole moieties from the surrounding protein amino acids. The model is partially consistent with mechanistic studies of water oxidation and high-resolution spectroscopy, although the structural and functional roles played by the ligands, including their potential "redox leveling" role during PCET, are yet to be established. The results reported for complex **1** suggest that ligation of carboxylate moieties and deprotonation of water ligands might be responsible for a similar redox leveling processes, leading to redox steps over a narrow range of potential during the accumulation of four oxidizing equivalents in the OEC.

IV. Conclusions

We have shown that the density functional theory (DFT) B3LYP/cc-pVTZ(-f) analysis of redox potentials and pK_a's of oxomanganese complexes with terminal water ligands provides a useful description of proton coupled electron transfer (PCET) at the detailed molecular level as well as a fundamental understanding of the regulatory effect of carboxylate moieties that bind competitively to Mn by exchange with terminal water ligands. The analysis provides understanding on the regulatory effect of oxidation state transitions on the pK_a's of terminal water ligands as well as the effect of deprotonation of water ligands, or exchange by carboxylate moieties, on the redox potentials of the Mn centers.

We conclude that PCET reduces the redox potential of **1** by coupling the deprotonation of the water ligand bound to Mn^{IV} to the oxidation of the Mn^{III} center, yielding oxidation at a potential (0.65 + 0.82 − 0.059 pH) V lower than the voltage 1.4 V required in the absence of deprotonation at pH < pK_a = 13.9. In the absence of such coupling, both of these processes are not favorable in the corresponding range of voltage and pH, since the water ligand deprotonates only at pH ≥ 13.9 and the complex is oxidized only at ≥1.4 V. We conclude that the exchange of water by acetate reduces the redox potential of the III,IV → IV,IV transition by 100−200 mV and the pH-range for PCET by about 2 pH units. Upon acetate binding to Mn^{IV}, PCET involves deprotonation of water ligated to Mn^{III} along the following oxidation state transition: $[H_2O(terpy)Mn^{III}(\mu\text{-}O)_2Mn^{IV}(terpy)AcO]^{2+} \rightarrow [HO(terpy)Mn^{IV}(\mu\text{-}O)_2Mn^{IV}(terpy)AcO]^{2+} + H^+ + e^-$. The resulting PCET mechanism thus prevents the accumulation of charge in the complex during the oxidation of the complex even at pH below the pK_a of the water ligand.

The observed regulatory effects, due to coordination of carboxylate moieties or deprotonation of water ligands, are expected to be common to PCET in other oxomanganese complexes, including the oxygen-evolving complex (OEC) of photosystem II (PSII) where deprotonation of water/hydroxo ligands is thought to be crucial to prevent charge buildup during the accumulation of multiple oxidation equivalents in the catalytic site. However, a direct simulation of PCET processes in the OEC of PSII is expected to be more challenging due to the presence of an inhomogeneous protein environment surrounding the metal complex, a problem that might require a combined quantum mechanics/molecular mechanics (QM/MM) method, as implemented in our previous studies of PSII.[1−7]

The predicted exchange of water by acetate in complex **1** suggests that other Lewis bases could readily exchange with terminal water ligands in oxomanganese complexes and reduce the redox potential of the complex, including buffer moieties commonly used in electrochemical measurements (e.g., phosphate), oxide groups on semiconductor nanoparticles surfaces (e.g., yielding direct deposition of the complex on the surface), or oxide ligands from other complexes that might yield oligomerization.

Acknowledgment. V. S. Batista acknowledges supercomputer time from NERSC and financial support from the grant NIH 1R01-GM-084267-01. G. W. Brudvig acknowledges support from the grant NIH GM32715.

Supporting Information Available: Description of the computational methods, nuclear coordinates, the spin population analysis, and the effect of oxidation state transition coupled to protonation/deprotonation events on the electrostatic potential atomic charges. Thermodynamic data, including solvation enthalpies and entropies are also provided. This information is available free of charge via the Internet at http://pubs.acs.org/.

References

(1) McEvoy, J. P.; Brudvig, G. W. *Chem. Rev.* **2006**, *106* (11), 4455–4483.

(2) Sproviero, E. M.; Gascon, J. A.; McEvoy, J. P.; Brudvig, G. W.; Batista, V. S. *J. Am. Chem. Soc.* **2008**, *130* (21), 6728–6730.

(3) Sproviero, E. M.; Gascon, J. A.; McEvoy, J. P.; Brudvig, G. W.; Batista, V. S. *J. Am. Chem. Soc.* **2008**, *130* (11), 3428–3442.

(4) Sproviero, E. M.; Gascon, J. A.; McEvoy, J. P.; Brudvig, G. W.; Batista, V. S. *J. Chem. Theory Comput.* **2006**, *2* (4), 1119–1134.

(5) Sproviero, E. M.; Gascon, J. A.; McEvoy, J. P.; Brudvig, G. W.; Batista, V. S. *Curr. Opin. Struct. Biol.* **2007**, *17* (2), 173–180.

(6) Ferreira, K. N.; Iverson, T. M.; Maghlaoui, K.; Barber, J.; Iwata, S. *Science* **2004**, *303* (5665), 1831–1838.

(7) Sproviero, E. M.; Gascon, J. A.; McEvoy, J. P.; Brudvig, G. W.; Batista, V. S. *Coord. Chem. Rev.* **2008**, *252* (3−4), 395–415.

(8) Li, G. H.; Sproviero, E. M.; Snoeberger, R. C.; Iguchi, N.; Blakemore, J. D.; Crabtree, R. H.; Brudvig, G. W.; Batista, V. S. *Energy Environ. Sci.* **2009**, *2* (2), 230–238.

(9) McNamara, W. R.; Snoeberger, R. C.; Li, G.; Schleicher, J. M.; Cady, C. W.; Poyatos, M.; Schmuttenmaer, C. A.; Crabtree, R. H.; Brudvig, G. W.; Batista, V. S. *J. Am. Chem. Soc.* **2008**, *130* (43), 14329–14338.

(10) Abuabara, S. G.; Cady, C. W.; Baxter, J. B.; Schmuttenmaer, C. A.; Crabtree, R. H.; Brudvig, G. W.; Batista, V. S. *J. Phys. Chem. C* **2007**, *111* (32), 11982–11990.

(11) Limburg, J.; Vrettos, J. S.; Liable-Sands, L. M.; Rheingold, A. L.; Crabtree, R. H.; Brudvig, G. W. *Science* **1999**, *283* (5407), 1524–1527.

(12) Baffert, C.; Romain, S.; Richardot, A.; Lepretre, J. C.; Lefebvre, B.; Deronzier, A.; Collomb, M. N. *J. Am. Chem. Soc.* **2005**, *127* (39), 13694–13704.

(13) Collomb, M. N.; Deronzier, A.; Richardot, A.; Pecaut, J. *New J. Chem.* **1999**, *23* (4), 351–353.

(14) Yagi, M.; Narita, K. *J. Am. Chem. Soc.* **2004**, *126* (26), 8084–8085.

(15) Chen, H. Y.; Tagore, R.; Olack, G.; Vrettos, J. S.; Weng, T. C.; Penner-Hahn, J.; Crabtree, R. H.; Brudvig, G. W. *Inorg. Chem.* **2007**, *46* (1), 34–43.

(16) Limburg, J.; Vrettos, J. S.; Chen, H. Y.; de Paula, J. C.; Crabtree, R. H.; Brudvig, G. W. *J. Am. Chem. Soc.* **2001**, *123* (3), 423–430.

(17) Sproviero, E. M.; Gascon, J. A.; McEvoy, J. P.; Brudvig, G. W.; Batista, V. S. *J. Inorg. Biochem.* **2006**, *100* (4), 786–800.

(18) Cady, C. W.; Shinopoulos, K. E.; Crabtree, R. H.; Brudvig, G. W. *Dalton Trans.* **2010**, *39* (16), 3985–3989.

(19) Cooper, S. R.; Calvin, M. *J. Am. Chem. Soc.* **1977**, *99* (20), 6623–6630.

(20) Kirby, J. A.; Robertson, A. S.; Smith, J. P.; Thompson, A. C.; Cooper, S. R.; Klein, M. P. *J. Am. Chem. Soc.* **1981**, *103* (18), 5529–5537.

(21) Mukhopadhyay, S.; Mandal, S. K.; Bhaduri, S.; Armstrong, W. H. *Chem. Rev.* **2004**, *104* (9), 3981–4026.

(22) Ruettinger, W. F.; Campana, C.; Dismukes, G. C. *J. Am. Chem. Soc.* **1997**, *119* (28), 6670–6671.

(23) Cady, C. W.; Crabtree, R. H.; Brudvig, G. W. *Coord. Chem. Rev.* **2008**, *252* (3-4), 444–455.

(24) Hagen, K. S.; Westmoreland, T. D.; Scott, M. J.; Armstrong, W. H. *J. Am. Chem. Soc.* **1989**, *111* (5), 1907–1909.

(25) Thorp, H. H.; Sarneski, J. E.; Brudvig, G. W.; Crabtree, R. H. *J. Am. Chem. Soc.* **1989**, *111* (26), 9249–9250.

(26) Sarneski, J. E.; Thorp, H. H.; Brudvig, G. W.; Crabtree, R. H.; Schulte, G. K. *J. Am. Chem. Soc.* **1990**, *112* (20), 7255–7260.

(27) Meyer, T. J. *Acc. Chem. Res.* **1989**, *22* (5), 163–170.

(28) Sarneski, J. E.; Didiuk, M.; Thorp, H. H.; Crabtree, R. H.; Brudvig, G. W.; Faller, J. W.; Schulte, G. K. *Inorg. Chem.* **1991**, *30* (14), 2833–2835.

(29) Nyholm, R. S.; Turco, A. *Chem. Ind. (London)* **1960**, (3), 74–75.

(30) Plaksin, P. M.; Palenik, G. J.; Stoufer, R. C.; Mathew, M. *J. Am. Chem. Soc.* **1972**, *94* (6), 2121.

(31) Thorp, H. H.; Sarneski, J. E.; Kulawiec, R. J.; Brudvig, G. W.; Crabtree, R. H.; Papaefthymiou, G. C. *Inorg. Chem.* **1991**, *30* (5), 1153–1155.

(32) Ruettinger, W. F.; Ho, D. M.; Dismukes, G. C. *Inorg. Chem.* **1999**, *38* (6), 1036+.

(33) Chen, H. Y.; Faller, J. W.; Crabtree, R. H.; Brudvig, G. W. *J. Am. Chem. Soc.* **2004**, *126* (23), 7345–7349.

(34) Brimblecombe, R.; Bond, A. M.; Dismukes, G. C.; Swiegers, G. F.; Spiccia, L. *Phys. Chem. Chem. Phys.* **2009**, *11* (30), 6441–6449.

(35) Tsai, M. K.; Rochford, J.; Polyansky, D. E.; Wada, T.; Tanaka, K.; Fujita, E.; Muckerman, J. T. *Inorg. Chem.* **2009**, *48* (10), 4372–4383.

(36) Cheng, T. Y.; Szalda, D. J.; Hanson, J. C.; Muckerman, J. T.; Bullock, R. M. *Organometallics* **2008**, *27* (15), 3785–3795.

(37) Muckerman, J. T.; Fujita, E.; Hoff, C. D.; Kubas, G. J. *J. Phys. Chem. B* **2007**, *111* (24), 6815–6821.

(38) Fujita, E.; Brunschwig, B. S.; Creutz, C.; Muckerman, J. T.; Sutin, N.; Szalda, D.; van Eldik, R. *Inorg. Chem.* **2006**, *45* (4), 1595–1603.

(39) Hou, H.; Muckerman, J. T.; Liu, P.; Rodriguez, J. A. *J. Phys. Chem. A* **2003**, *107* (44), 9344–9356.

(40) Roy, L. E.; Batista, E. R.; Hay, P. J. *Inorg. Chem.* **2008**, *47* (20), 9228–9237.

(41) Roy, L. E.; Jakubikova, E.; Guthrie, M. G.; Batista, E. R. *J. Phys. Chem. A* **2009**, *113* (24), 6745–6750.

(42) Li, J.; Fisher, C. L.; Chen, J. L.; Bashford, D.; Noodleman, L. *Inorg. Chem.* **1996**, *35* (16), 4694–4702.

(43) Baik, M. H.; Friesner, R. A. *J. Phys. Chem. A* **2002**, *106* (32), 7407–7412.

(44) Uudsemaa, M.; Tamm, T. *J. Phys. Chem. A* **2003**, *107* (46), 9997–10003.

(45) Moens, J.; Geerlings, P.; Roos, G. *Chem.—Eur. J.* **2007**, *13* (29), 8174–8184.

(46) Moens, J.; Jaque, P.; De Proft, F.; Geerlings, P. *J. Phys. Chem. A* **2008**, *112* (26), 6023–6031.

(47) Moens, J.; Roos, G.; Jaque, P.; Proft, F.; Geerlings, P. *Chem.—Eur. J.* **2007**, *13* (33), 9331–9343.

(48) Yang, X.; Baik, M. H. *J. Am. Chem. Soc.* **2006**, *128* (23), 7476–7485.

(49) Ayala, R.; Sprik, M. *J. Chem. Theory Comput.* **2006**, *2* (5), 1403–1415.

(50) Galstyan, A.; Knapp, E. W. *J. Comput. Chem.* **2009**, *30* (2), 203–211.

(51) De Groot, M. T.; Koper, M. T. M. *Phys. Chem. Chem. Phys.* **2008**, *10* (7), 1023–1031.

(52) Wang, T.; Friesner, R. A. *J. Phys. Chem. C* **2009**, *113* (6), 2553–2561.

(53) Becke, A. D. *Phys. Rev. A: At., Mol., Opt. Phys.* **1988**, *38* (6), 3098–3100.

(54) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37* (2), 785–789.

(55) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77* (18), 3865–3868.

(56) Becke, A. D. *J. Chem. Phys.* **1993**, *98* (7), 5648–5652.

(57) Becke, A. D. *J. Chem. Phys.* **1993**, *98* (2), 1372–1377.

(58) Wang, T.; Brudvig, G.; Batista, V. S. *J. Chem. Theory Comput.* **2010**, *6* (3), 755–760.

(59) Tagore, R.; Crabtree, R. H.; Brudvig, G. W. *Inorg. Chem.* **2008**, *47* (6), 1815–1823.

(60) Dismukes, G. C.; Brimblecombe, R.; Felton, G. A. N.; Pryadun, R. S.; Sheats, J. E.; Spiccia, L.; Swiegers, G. F. *Acc. Chem. Res.* **2009**, *42* (12), 1935–1943.

(61) Noodleman, L. *J. Chem. Phys.* **1981**, *74* (10), 5737–5743.

(62) Benard, M. *J. Chem. Phys.* **1979**, *71* (6), 2546–2556.

(63) Dunning, T. H. *J. Chem. Phys.* **1989**, *90* (2), 1007–1023.

(64) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96* (9), 6796–6806.

(65) Woon, D. E.; Dunning, T. H. *J. Chem. Phys.* **1993**, *98* (2), 1358–1371.

(66) Reiss, H.; Heller, A. *J. Phys. Chem.* **1985**, *89* (20), 4207–4213.

(67) Rashin, A. A.; Honig, B. *J. Phys. Chem.* **1985**, *89* (26), 5588–5593.

(68) Marten, B.; Kim, K.; Cortis, C.; Friesner, R. A.; Murphy, R. B.; Ringnalda, M. N.; Sitkoff, D.; Honig, B. *J. Phys. Chem.* **1996**, *100* (28), 11775–11788.

(69) Jang, Y. H.; Sowers, L. C.; Cagin, T.; Goddard, W. A. *J. Phys. Chem. A* **2001**, *105* (1), 274–280.

(70) Lim, C.; Bashford, D.; Karplus, M. *J. Phys. Chem.* **1991**, *95* (14), 5610–5620.

(71) Morrison, M. M.; Sawyer, D. T. *J. Am. Chem. Soc.* **1977**, *99* (1), 257–258.

CT1002658

JCTC Journal of Chemical Theory and Computation

Generalized X-Pol Theory and Charge Delocalization States

Jiali Gao,*,† Alessandro Cembran,† and Yirong Mo‡

Department of Chemistry, Digital Technology Center, and Supercomputing Institute, University of Minnesota, Minneapolis, Minnesota 55455, and Department of Chemistry, Western Michigan University, Kalamazoo, Michigan 49008

Received June 2, 2010

Abstract: The mixed molecular orbital and valence bond (MOVB) method has been used to generalize the explicit polarization (X-Pol) potential to incorporate charge delocalization resonance effects in the framework of valence bond theory. In the original X-Pol method, a macromolecular system is partitioned into individual fragments or blocks, and the molecular orbitals of the system are strictly localized within each block. Consequently, these block-localized molecular orbitals (BLMOs) are nonorthogonal across different blocks. In the generalized X-Pol (GX-Pol) theory, we construct charge delocalization VB states by expanding the localization space from monomer blocks into pairwise delocalized blocks. Thus, the expansion of the basis space leads to charge delocalization between monomer pairs, and a series of pairwise delocalization states can be constructed. In general, *L*-body delocalized states can be analogously defined by grouping *L* monomer blocks into one. The Hartree product wave function for each state can be fully antisymmetrized, which introduces explicitly exchange repulsion among all blocks. The GX-Pol wave function is a linear combination of all *L*-body charge transfer (valence bond) states, which incorporates charge delocalization and their resonance as well as static correlation effects. The GX-Pol method provides a general and rigorous theory to incorporate charge delocalization explicitly into these fragment-based electronic structural methods for macromolecular systems.

1. Introduction

The explicit polarization (X-Pol) method is the first practical fragment-based molecular orbital[1–5] approach for macromolecular simulations[2,6–8] in which a Hartree product wave function is used on the basis of antisymmetric wave functions of individual subsystems. By construction, molecular orbitals in the X-Pol wave function are strictly localized within the subspace defined by each individual block; that is, these are block-localized molecular orbitals (BLMOs). However, the use of a Hartree product wave function neglects exchange repulsion and charge transfer effects between different blocks. Previously, an empirical Lennard-Jones potential was used to account for the repulsive interactions,[1–5,8] and we have presented an approach to incorporate the exchange repulsion explicitly into the X-Pol method by antisymmetrizing its wave function.[9–13] It was found that the total exchange repulsion is short-ranged, as is well-known, and is essentially pairwise additive for two water trimer complexes examined.[9] The latter is somewhat surprising in view of the need for orthogonalization of the BLMOs. Nevertheless, these findings suggest that the use of pairwise energy terms could be a very good approximation for the treatment of exchange repulsion interactions.

The strict block localization of molecular orbitals within each subsystem in the X-Pol wave function also excludes charge transfer (CT) contributions between different blocks.[11,14–20] Although the amount of charge transferred is relatively small, the energy component due to CT effects can be significant and it is important in hydrogen-bonded interactions,[15,20] critical for biomolecular modeling. The inclusion of charge transfer effects in the X-Pol method can be easily accomplished, but it necessarily requires the

* Corresponding author: e-mail gao@jialigao.org.
† University of Minnesota.
‡ Western Michigan University.

10.1021/ct100292g © 2010 American Chemical Society
Published on Web 07/15/2010

expansion of the orbital space of the localization blocks, and a number of approaches can be envisioned. Here, we describe one approach to systematically include CT configurations into a generalized X-Pol (GX-Pol) wave function.

Methods for treating fragmental electronic structures may be traced to the work of McWeeny[21] or even earlier, who discussed the density matrices of orthogonal group functions; however, Stoll and Preuss[22] were the first to describe a procedure based on many-body interaction energy correction to improve the energy of such a fragment molecular orbital approach as an approximation to the Hartree−Fock (HF) or density functional energy:

$$E_{\mathrm{HF}} \approx E_0 + \sum_{ab}^{\mathrm{dimers}} (E_{ab} - E_0) + \sum_{abc}^{\mathrm{trimers}} \Delta E_{abc} + \ldots \quad (1)$$

where E_0 is the total monomer energy and the subsequent summations correspond to the dimer, trimer, etc., corrections. The many-body interaction energy correction approach of Stoll and Preuss treats the monomer, dimer, trimer, etc., terms in the electrostatic field of the rest of the system.[22,23] This is different from the many-body decomposition scheme of Stillinger and co-workers,[24] in which the many-body terms do not include polarization by the rest of the system. The energy correction at each order of expansion (dimer, trimer, etc.) is obtained by subtracting the interaction energies of the preceding order, making the computational procedure a fast-converging, build-up approach that is exceedingly simple and straightforward.[25,26] However, the molecular wave function is not available from such a fragment molecular orbital approach, and the somewhat ad hoc energy addition and subtraction scheme makes it difficult to obtain analytical gradients of the total energy (of course, it is always possible to obtain the gradients by a variety of procedures, at the expense of higher computational costs).[4] It is important to note that the configuration weights of different many-body terms are not identical, although it is implicitly assumed to be the same in this approach.

In this paper, we use the mixed molecular orbital and valence bond (MOVB) approach to generalize the X-Pol wave function into a multistate X-Pol wave function in the framework of valence bond theory.[27−33] In the generalized explicit polarization (GX-Pol) wave function, charge transfer as well as exchange repulsion effects can be systematically determined by the self-consistent field (SCF) method for the entire system. Rather than an ad hoc energy correction expansion, the approach is similar to the traditional multiconfiguration self-consistent field (MCSCF), and complete-active-space self-consistent field (CASSCF) or equivalently ab initio valence bond self-consistent field (VBSCF) methods,[34] applied to treating the resonance of charge transfer VB states. Consequently, static correlation effects are, at least partially, included into the GX-Pol method. In the following, we first present the theoretical background in section 2. Then we summarize the computational details to illustrate the generalized X-Pol method. Results and discussion are given in section 4, followed by a summary of the main conclusions from this work.

2. Method

For completeness, we first briefly outline the "monomeric" X-Pol wave function.[1−5] Here, we use the phrase "monomer X-Pol" to emphasize that the total molecular wave function of a condensed-phase system including proteins is constructed from monomer wave functions to distinguish it from the generalization to many-body delocalized states to be described below. Then, we introduce the dimeric charge delocalization state to construct a two-body X-Pol wave function and its generalization to many-body X-Pol wave functions. Finally, we propose to employ these charge delocalization states as effective valence bond (VB) configurations in VBSCF optimization.[27,28,30−32] We emphasize here that although the theory and illustrative examples are given in terms of wave function theory, *the method and algorithm are identically applicable to density functional theory (DFT), by use of block-localized density functional theory (BLDFT) to define VB states, as described in ref 27*. We also discuss the distinction between consistent diabatic configuration (CDC) and variational diabatic configuration (VDC) methods.[32]

2.1. Block Localization and the Monomeric X-Pol Wave Function. In the X-Pol method,[1−5] a macromolecular system is partitioned into M blocks. The ath block contains k_a basis functions and n_a electrons, and there are a total of K primitive basis functions and N electrons in the system:

$$K = \sum_{a=1}^{M} k_a \quad \text{and} \quad N = \sum_{a=1}^{M} n_a \quad (2)$$

Molecular orbitals in a given block are written as linear combinations of the primitive basis functions located on atoms in that specific subspace $\{\chi_\mu^a;\ \mu = 1, \ldots, k_a\}$:

$$\phi_j^a = \sum_{\mu=1}^{k_a} c_{j\mu}^a \chi_\mu^a \quad (3)$$

The X-Pol wave function is constructed from the monomer blocks as a Hartree product of the determinant wave functions of individual blocks:[1]

$$\Psi_x = R_x \hat{A}(\Phi_1)\hat{A}(\Phi_2)\ldots\hat{A}(\Phi_M) \quad (4)$$

where $\hat{A}$ is an antisymmetrizing operator, R_x is the normalization constant, and Φ_a is a successive product of the occupied spin orbitals in the ath subsystem (eq 3):

$$\Phi_a = \phi_1^a \phi_2^a \ldots \phi_{n_a}^a \quad (5)$$

For convenience in the following discussion, eq 4 is called the "monomer" X-Pol wave function, which itself can be antisymmetrized as a block-localized wave function (BLW):[10−13,30,31]

$$\Psi_x^A = R_x^A \hat{A}\left(\prod_a^M \Phi_a\right) \quad (6)$$

The antisymmetrized X-Pol wave function, also known as BLW,[10] includes explicitly the interfragment eXchange interactions (X-Pol-X)[9] that are ignored in the Hartree

product X-Pol wave function (eq 4) but are approximated empirically by Lennard-Jones terms.[1−3]

2.2. Charge Transfer (Delocalization) States and Many-Body X-Pol Wave Functions. Since the molecular orbitals (MOs) are strictly block-localized within each subsystem by construction, there is no possibility of charge transfer between different monomer blocks.[10,11,27] The charge delocalization energy due to intermolecular (or interfragment) charge transfer can be very important for hydrogen-bonding complexes and biomolecular interactions.[11,15,20,35–39] For example, the energy component due to charge transfer in the dimer complex of water amounts to −0.4 kcal/mol[11] of the total hydrogen bonding energy (−5 kcal/mol).[40] Note that the term charge transfer used here is more rigorously related to electronic delocalization to distinguish it from the diabatic states in electron transfer reaction; however, in this paper, we use both charge transfer and charge delocalization without specific distinction since the method is equally applicable to electron transfer processes. In the monomer X-Pol method (eq 4), charge transfer effects are modeled *effectively* as electrostatic interactions as in molecular mechanics, which is adequate in the spirit of force field development to keep the formalism simple and the computational procedure efficient.[1–4,8] On the other hand, in situations where the explicit treatment of charge transfer effects is of interest,[20,27] it is desirable to define a general approach to treat this effect in the X-Pol theory.

To this end, we generalize the monomer X-Pol wave function to dimer, and generally, many-body, X-Pol wave functions. First, we define a charge transfer state between monomers *a* and *b*, whose wave function is written as a Hartree product of the antisymmetric (determinant) wave function of the dimer (*ab*) and the antisymmetric wave functions of all other monomers:

$$\Psi_{(ab)} = R_{(ab)}\hat{A}(\Phi_1)...\hat{A}(\Phi_{ab})...\hat{A}(\Phi_M) \tag{7}$$

where $R_{(ab)}$ is the normalization constant for the wave function defined above. Clearly, it is straightforward to define analogously a fully antisymmetrized dimer X-Pol (X-Pol-X) wave function:

$$\Psi^A_{(ab)} = R^A_{(ab)}\hat{A}(\Phi_1...\Phi_{ab}...\Phi_M) \tag{8}$$

An important distinction between the wave functions defined by eqs 4 and 7 (and by eqs 6 and 8) is that the molecular orbitals in the dimer, indexed by (*ab*), Φ_{ab}, are expanded over the basis functions of the entire dimer space, in contrast to that in eq 4 with strict block localization within each monomer space. Thus, there are $M - 1$ blocks remaining in the state defined by eq 7 (and eq 8) since two blocks are combined to form a single dimeric subsystem. Importantly, since the wave function specified in eq 7 or 8 represents the expansion of basis space, it introduces charge delocalization effects between monomers *a* and *b* in comparison with that of eq 4. The energy difference between these two states (eqs 6 and 8) corresponds to the X-Pol charge transfer energy between the two subsystems according to our interaction energy decomposition analysis (EDA) based on block-localized wave function (BLW):[11,41]

$$\Delta E^{\mathrm{CT}}_{(ab)} = \langle \Psi^A_{(ab)}|H|\Psi^A_{(ab)}\rangle - \langle \Psi^A_x|H|\Psi^A_x\rangle \tag{9}$$

Consequently, the individual dimer X-Pol wave function in eq 7 (or eq 8) represents a CT state between two monomer blocks, in the presence of the electrostatic field of the rest of the system. We note that, unlike other EDA approaches,[41,42] the intermediate wave functions at all stages of the decomposition analysis are fully defined and *variationally optimized*,[11] providing the most relevant reference states in polarization and charge transfer analysis.[12,20,43]

We define the total dimeric generalized X-Pol (GX-Pol) wave function as a linear combination of all dimeric charge transfer states:

$$\Theta_{\mathrm{X2}} = \sum_{a=1}^{M}\sum_{b=a+1}^{M} c_{(ab)}\Psi^A_{(ab)} \tag{10}$$

where the subscript X2 specifies a GX-Pol wave function at the dimer CT level and $c_{(ab)}$ is a configuration coefficient to be optimized along with all the BLMO coefficients in exactly the same fashion as in standard MCSCF or VBSCF methods.[34,44] In fact, the interpretation of the configuration specified by eq 7, and the fully antisymmetrized counterpart (eq 8), is an effective valence bond state, and eq 10 is a VB wave function defined by these CT states (the valence bond resonance integrals between determinant states are given below). Thus, the generalization of the X-Pol method to the treatment of multiconfigurational states is equivalent to the previously described MOVB theory[30,31] which has been used in the study of chemical reactions in solution,[20,28–31] cluster analysis,[29] and a range of charge transfer analysis applications.[12,15–17,20,43,45] MOVB has been presented with the use of ab initio molecular orbital thoery,[28,30–32] semiempirical methods,[33,46] and density functional theory.[27]

The ground-state energy of the entire system, including all or some dimeric CT contributions, is minimized by a valence bond self-consistent field (VBSCF) method, which can be expressed as follows:

$$E_{\mathrm{X2}} = \langle \Theta_{\mathrm{X2}}|H|\Theta_{\mathrm{X2}}\rangle \tag{11}$$

A general, many-body X-Pol wave function can be defined similarly. Thus, for a full system separated into *M* blocks (monomers), the *L*-body GX-Pol wave function is

$$\Theta_{XL} = \underbrace{\sum_{a=1}^{M}\sum_{b=a+1}^{M}\cdots\sum_{f=a+L-1}^{M}}_{L\text{ summation terms}} c_{ab\cdots f}\Psi^A_{(ab\cdots f)} \tag{12}$$

where the subscript X*L* denotes the *L*-body GX-Pol wave function, and the individual effective VB configuration is defined analogously as in eq 7 by grouping *L* blocks into one delocalized subsystem. Equation 12 defines a full range of many-body GX-Pol wave functions, from the simplest case in which each monomer block is fully separated and the molecular orbitals are strictly block-localized (the monomer X-Pol wave function) to the full delocalization of the entire system, which is the Hartree−Fock result. Of course, it is clear from eqs 10 and 12 that, in contrast to the method of

Stoll and Preuss,[22] there is no redundancy to have to determine lower-body states to construct an *L*-body GX-Pol wave function since the GX-Pol method is not a build-up model and the charge delocalization of the lower-order contributions is fully encompassed in the GX-Pol states. For example, it is not necessary to construct monomer and dimer states to form a "trimer correction" term.

It is of interest to point out that the intermediate VB wave functions, generally defined by eq 12, are multiconfigurational wave functions, and they include partially static correlation effects and dispersion contributions.[27] An exploration of this property is of considerable interest and remains for future investigations.

2.3. Consistent and Variational Diabatic Configurations. Previously, we have classified the effective VB states optimized by the VBSCF method as consistent diabatic configurations (CDC) in that the individual VB states, called diabatic configurations, are consistently optimized to yield the minimum ground-state energy.[32] An alternative method is to use a configuration interaction (CI) approach to optimize the configuration coefficients in eq 10 with fixed BLMOs in each effective VB state that has been variationally optimized individually.[28,30–33] We call these VB states the variational diabatic configurations (VDC).[32] Obviously, the CDC and VDC states have different energies for the same state defined by eq 7 (and eq 8) since they are obtained by different energy minimization targets.[32] In the latter case, one solves the generalized secular determinant equation to yield the configuration coefficients and the ground-state energy:[27,28,30–32]

$$\begin{vmatrix} H_{11} - ES_{11} & \cdots & H_{1,M2} - ES_{1,M2} \\ \cdots & \cdots & \cdots \\ H_{M2,1} - ES_{M2,1} & \cdots & H_{M2,M2} - ES_{M2,M2} \end{vmatrix} = 0 \quad (13)$$

where $M2 = M(M - 1)/2$ is the total number of dimer configurations, $S_{(ab),(st)}$ is the overlap matrix between dimer configurations (ab) and (st), and the Hamiltonian matrix element and VB resonance integral is defined by[27,30–32,47]

$$H_{(ab),(st)} = \langle \Psi^A_{(ab)} | H | \Psi^A_{(st)} \rangle \quad (14)$$

We note that the explicit expression of eq 14 has been given previously,[30,31] and again, the approach described here is applicable to both molecular orbital theory and density functional theory as described by Mo et al.[14] and by Cembran et al.[27]

3. Computational Details

All computations have been performed by use of a locally modified GAMESS program[48] and the Xiamen University Valence Bond (XMVB) program.[34] The valence double-ζ 6-31+G(d) basis set with polarization and diffuse functions was used to optimize the trimer geometries, and the aug-cc-pCVDZ[49] basis set was used in energy calculations. The goal here is to illustrate the significance of resonance effects by coupling dimeric charge delocalization VB states in the GX-Pol theory. A more thorough examination with a larger set of systems will be published later. The minimum energy configuration of a cyclic water trimer structure, **c-W$_3$**, along with another trimer configuration, **s-W$_3$**, which was constructed by minimizing the dimer water complex first followed by placing a third water molecule at the C_2 image about the bisection of the acceptor water molecule, are adopted in this study. Both structures are depicted in Figure 1.

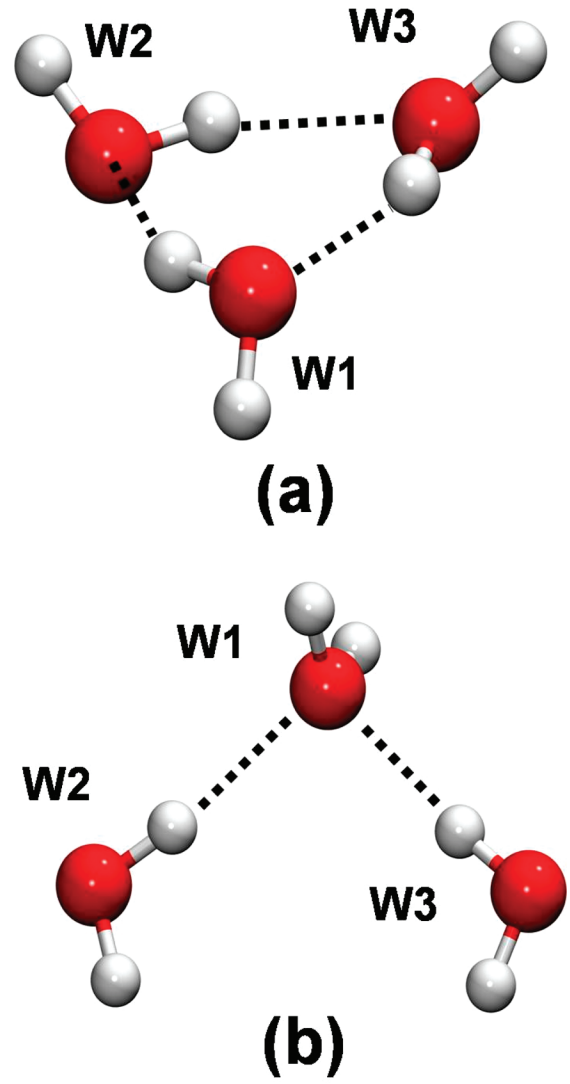


Figure 1. Schematic representation of (a) the minimum water trimer structure (**c-W$_3$**) and (b) a symmetric configuration (**s-W$_3$**).

Throughout the following discussion, each water monomer in the trimer complexes is partitioned as a monomer block, and their geometries are kept as those in the optimized configuration at the HF/6-31+G(d) level. The notation $\Psi^A_{(ab)} = \hat{A}(\Phi_{ab}\Phi_c\}$, where a, b, $c = 1$, 2, or 3, is used to represent a charge delocalization state between water molecules a and b in the electrostatic field of the third water monomer c. We use the fully antisymmetrized wave function that also includes explicitly exchange repulsion between different blocks. To estimate the dimeric charge delocalization energy, the block-localized wave function for the triblock system $\Psi_{\text{X-Pol-X}} = \hat{A}\{\Phi_1\Phi_2\Phi_3\}$ is also determined.[9] The structural weight is determined as described by Chirgwin and Coulson.[50]

Table 1. Computed Total Energies for the Cyclic Water Trimer Minimum Structure and for a Symmetric Trimer Geometry[a]

method	total energy (hartrees)	
	c-W$_3$	**s-W$_3$**
$\Phi_1^g + \Phi_2^g + \Phi_3^g$	−228.128 79	−228.129 19
Ψ_{123}^{HF}	−228.147 55	−228.139 35
$\Psi_{X\text{-Pol}}^A = \hat{A}(\Phi_1\Phi_2\Phi_3)$	−228.142 58	−228.136 26
$\Psi_{23}^A = \hat{A}(\Phi_{23}\Phi_1)$	−228.144 37	−228.136 34
$\Psi_{13}^A = \hat{A}(\Phi_{13}\Phi_2)$	−228.144 13	−228.137 75
$\Psi_{12}^A = \hat{A}(\Phi_{12}\Phi_3)$	−228.144 31	−228.137 75
VDC-MOVB(3)	−228.145 37	−228.139 09
CDC-MOVB(3)	−228.151 16	−228.141 02

[a] All calculations are performed with the aug-cc-pCVDZ basis set at the HF/6-31+G(d) geometry.

Table 2. Computed Relative Energies for the Cyclic Water Trimer Minimum Structure and for a Symmetric Trimer Geometry[a]

method	relative energy (kcal/mol)	
	c-W$_3$	**s-W$_3$**
$\Phi_1^g + \Phi_2^g + \Phi_3^g$	0.00	0.00
$\Psi_{X\text{-Pol}}^A = \hat{A}(\Phi_1\Phi_2\Phi_3)$	−8.65 (0.00)	−4.43 (0.00)
$\Psi_{23}^A = \hat{A}(\Phi_{23}\Phi_1)$	−9.78 (−1.12)	−4.48 (−0.05)
$\Psi_{13}^A = \hat{A}(\Phi_{13}\Phi_2)$	−9.62 (−0.97)	−5.37 (−0.93)
$\Psi_{12}^A = \hat{A}(\Phi_{12}\Phi_3)$	−9.74 (−1.09)	−5.37 (−0.93)
$\Psi_{X\text{-Pol}}^A$ + CT	−11.83 (−3.18)	−6.35 (−1.92)
Ψ_{123}^{HF}	−11.77 (−3.12)	−6.37 (−1.94)
VDC-MOVB(3)	−10.40 (−1.75)	−6.21 (−1.78)
CDC-MOVB(3)	−14.03 (−5.38)	−7.42 (−2.99)

[a] All calculations are performed with the aug-cc-pCVDZ basis set at the HF/6-31+G(d) geometry. Values in parentheses are interaction energies due to charge transfer [$\Delta E_{(ab)}^{CT}$] without basis-set superposition error correction.

4. Results and Discussion

Listed in Tables 1 and 2 are the total and relative energies for the isolated monomers at their complex configurations, the antisymmetrized X-Pol (i.e., X-Pol-X) trimer ($\Psi_{X\text{-Pol-X}}$), and the dimeric charge delocalization states [$\Psi_{ab}^A = \hat{A}(\Phi_{ab}\Phi_c)$]. The upper limit of charge transfer interaction energy between a pair of water molecules, in the presence of the third water in the trimer complex, is the energy difference between the two variational diabatic configurations Ψ_{ab}^A and $\Psi_{X\text{-Pol}}^A$. Using these VDC states as the effective VB configurations, one can carry out a configuration interaction calculation by optimizing only the configurational coefficients of the MOVB wave function.[28,30,31] In this case, the individual charge delocalization states remain unchanged, and thereby the resonance among these states, or the diabatic coupling, is not necessarily optimal. The energy from this procedure (eq 13) is called the VDC-MOVB(3) method, where the number in parentheses indicates the number of configurations used in the VB optimization. Alternatively, MOVB wave function can be fully optimized following standard procedures such as MCSCF and VBSCF methods in which both the orbital and configurational coefficients are simultaneously varied.[32] Since the individual CT diabatic states are obtained consistently with the ground-state energy minimization, the result is denoted by CDC-MOVB(3). Here, resonance effects make important contributions.

Table 2 shows that the X-Pol binding energies for the two water trimer complexes, **c-W$_3$** and **s-W$_3$**, are −8.65 and −4.43 kcal/mol, respectively, significantly smaller than the fully delocalized HF results (−11.77 and −6.37 kcal/mol). The difference represents the total charge delocalization effects (also called charge transfer in energy decomposition analysis) due to block localization imposed by the X-Pol wave function, which are −3.12 and −1.94 kcal/mol, respectively. The energies due to charge transfer between a pair of water molecules are shown in parentheses, which are in the range of −0.97 to −1.12 kcal/mol for the **c-W$_3$** complex. Apparently, the charge transfer effect between two water molecules is significantly enhanced in the presence of the polarization by the third water in comparison with the water dimer alone, which has a CT energy of −0.2 and −0.4 kcal/mol from the 6-31++G(d,p) and aug-cc-pVTZ basis sets. The pairwise charge delocalization energies are very similar for the three pairs in **c-W$_3$** since each water accepts and donates a hydrogen bond from the other two water molecules. On the other hand, monomers W2 and W3 (Figure 1) are placed in a repulsive orientation in the trimer complex **s-W$_3$**, which exhibits little charge transfer effects (−0.05 kcal/mol), which are depicted in Figure 2. These two water monomers (W2 and W3) have a C_2 symmetry about the molecular axis of the central water (W1), which accepts a hydrogen bond from each of the W2 and W3 monomers. The charge transfer energies are −0.93 kcal/mol for each of the two hydrogen bonds.

The resonance effects due to charge delocalization from VDC-MOVB(3) calculations are modest, lowering the X-Pol energy by −1.75 and −1.78 kcal/mol in the two complexes. Specifically, in the cyclic complex, **c-W$_3$**, which involves a sequence of donor and acceptor hydrogen bonds, 56% of the total charge delocalization effect is obtained in the configuration interaction approach, suggesting that there is a strong cooperative effect in the overall charge delocalization interaction that is not fully included in the VDC approach. In the case of **s-W$_3$**, in which there is little cooperative interaction, 92% of charge delocalization contributions are determined. Full relaxation of the GX-Pol wave function, by simultaneous optimization of both the orbital and configurational coefficients in the CDC-MOVB method, achieves the greatest extent of valence bond resonance effect. The computed total stabilization energies, relative to the monomer X-Pol result, in the two trimer complexes are −5.38 and −2.99 kcal/mol for **c-W$_3$** and **s-W$_3$**, respectively, far greater than the total charge delocalization energy at the Hartree−Fock level of theory. The MOVB method is a multiconfigurational approach, which also includes partial static correlation effects; this is reflected by the total electronic energy, lower than the corresponding HF value (Table 1). Thus, of the total stabilization energies in the trimer complexes, −2.2 and −1.1 kcal/mol, the amounts exceeding the corresponding HF binding energies, may be attributed to dispersion correlation effects. For comparison, Schutz et al.[51] estimated that the MP2 contribution to the binding energy of the **c-W$_3$** complex is −3.60 kcal/mol with the HF optimized geometry and the 6-311++G(d,p) basis

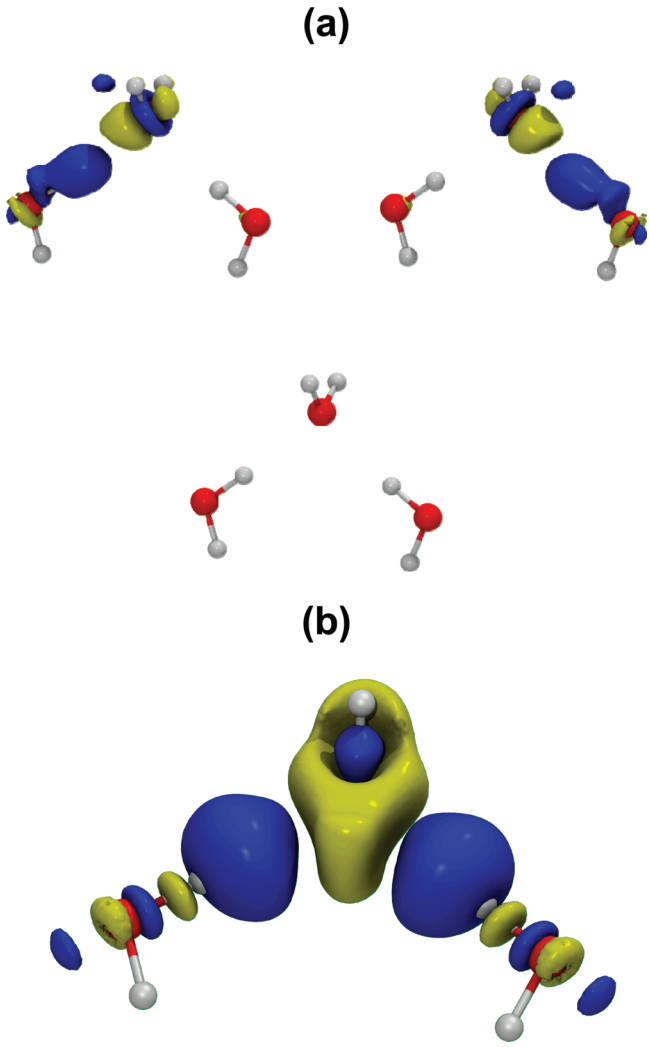


Figure 2. Electron density difference contour between the charge delocalized and the strictly localized systems, $\rho(\hat{A}\psi_1\psi_2\psi_3)$, in the **s-W$_3$** complex, (a) for the three pairs of dimer charge transfer state, $\rho(\hat{A}\psi_{ab}\psi_c)$, and (b) for the resonance state of the fully delocalized system from the two-body generalized explicit polarization (GX-Pol) wave function, $\rho(\Theta_{X2})$. The structure orientation is shown in Figure 1b and the contour levels are at 0.0002 au with blue contours representing gain in electron density and yellow contours representing charge depletion.

set. The binding energy for the **c-W$_3$** complex has been estimated to be −15.8 kcal/mol by use of CCSD(T)/CBS.[40]

Table 2 shows that the charge delocalization energies are roughly additive in both cases. In particular, the sum of the X-Pol energy and the total VDC charge delocalization energy ($\Psi^A_{X\text{-Pol}}$+ CT) is very close to the corresponding HF interaction energy. Here, the total charge delocalization energy is the sum of the three pairwise charge transfer contributions (e.g., −3.18 kcal/mol in the **c-W$_3$** structure). The fast converging property, that is, a monomeric X-Pol calculation followed by a variational dimeric charge delocalization energy correction, in a many-body interaction decomposition scheme[22] can be attributed to the fact that the dominant, nonadditive polarization effects[11] have already been included in the X-Pol wave function.[1] However, one should be cautious about the seemingly good agreement in this approach because the additive CT energy (−3.18 kcal/mol) is greater than the total charge delocalization effect (−3.12 kcal/mol). Note that, inasmuch as the difference is small, the sum exceeds the total charge delocalization energy, without inclusion of correlation contributions, is significant to indicate the nonvariational discrepancy resulting from a lack of consideration of cooperative effects of the entire system. Furthermore, the difference between the VDC and CDC results presented above indicates that good agreement between the additive results and the total HF interaction energies is fortuitous since orbital relaxation is essential in full VBSCF (MCSCF) charge resonance. It appears to be important to further analyze the additive properties on a wider range of structures and functionalities by use of the methods described here.

Table 3. Configuration Weights from the CDC and VDC-MOVB(3) Wave Functions

	c-W3		s-W$_3$	
configuration	CDC	VDC	CDC	VDC
$\Psi^A_{23} = \hat{A}(\Phi_{23}\Phi_1)$	0.384	0.375	0.000	−0.875
$\Psi^A_{13}\hat{A}(\Phi_{13}\Phi_2)$	0.305	0.272	0.500	0.937
$\Psi^A_{12}\hat{A}(\Phi_{12}\Phi_3)$	0.311	0.353	0.500	0.937

Nevertheless, in practice, it appears to be a reasonable strategy to use the X-Pol potential with an empirical estimate for the exchange repulsion to carry out Monte Carlo and molecular dynamics simulations of a condensed-phase system,[1−3,5,8] followed by the pairwise CT and exchange repulsion energy corrections to obtain more accurate results. The most systematic approach is to employ the many-body GX-Pol wave function in CDC-MOVB calculations to determine the ground-state potential energy surface in dynamics simulations.

The Chirgwin−Coulson structural weights[50] from MOVB calculations are given in Table 3 for both trimer complexes. The structural weight can be formulated in a number of ways, and the results in Table 3 are determined as follows:[34,50]

$$w_{(ab)} = c_{(ab)}{}^2 + \sum_{(st)\neq(ab)}^{3} c_{(ab)}c_{(st)}\langle\Psi^A_{(ab)}|\Psi^A_{(st)}\rangle \qquad (15)$$

Obviously, eq 15 should not be confused with the familiar Mulliken population analysis.[52] The structural weights provide key insight into the nature of chemical bonding in valence bond theory, and are an indication of the resonance contributions of charge delocalization states in the present analysis. In the case of the cyclic configuration, **c-W$_3$**, the hydrogen-bond strengths are very similar; thereby, both the CDC and VDC methods yield similar structural contributions in the total wave function. However, for the symmetric trimer complex, **s-W$_3$**, the charge transfer state $\Psi^A_{(23)}$ has negligible contribution to the charge delocalization of the entire system in the CDC-MOVB wave function. The overlap integrals between $\Psi^A_{(23)}$ and the other two CT states are essentially zero, and the structural weight for $\Psi^A_{(23)}$ is zero. However, the variationally optimized diabatic states are strongly overlapping, with calculated overlap integrals of 0.999 between $\Psi^A_{(23)}$ and two other states. As a result, there is strong out-of-phase mixing in these VDC states. The difference

between the CDC and VDC is a further indication of the lack of cooperative effects in the nonvariational optimization of the total VDC wave function.

Figure 2 illustrates the individual pairwise charge delocalization effects in the electrostatic field of the other monomer and the resonance delocalization results in the GX-Pol wave function for the **s-W$_3$** trimer complex. The electron density difference (EDD) isosurface[11,53] is obtained by subtracting the strictly localized, but fully polarized, X-Pol-X electron density from the corresponding delocalized wave functions, for the dimer delocalized pairs (Figure 2a) and for the valence bond resonance state (Figure 2b). In the dimer delocalization states between W1 and W2 and between W1 and W3 (see also Figure 1), electron densities are depleted (yellow contours) from the hydrogen-bond acceptor water (W1), predominantly from oxygen, whereas charge densities are accumulated (blue contours) along the hydrogen-bond donor H−O bond vector, principally located on the hydrogen atom. There is no noticeable charge density variation (the contour level was set to be 0.0002 au) in the W2−W3 delocalization state, consistent with the negligibly small (−0.05 kcal/mol) CT energy. The resonance of these three states, with nearly 50% contributions from the W1−W2 and W1−W3 delocalization states and essentially zero structural weight from the W2−W3 complex, shows the cooperative effect of these states. The charge density loss from the hydrogen-bond acceptor W1 water is spread out over the entire molecule and symmetrized, along with some compensating polarization gains in the inner part of the density distribution. It is interesting to notice the alternating pattern along the donor O−H bonds due to gain in charge density from CT and polarization delocalization along the bond vector. The overall molecular charge delocalization from individual charge transfer states is well represented in the total molecular electron density from the CDC optimization of the GX-Pol wave function.

Figure 3 shows the local two-body charge transfer (CT) in the presence of polarization of the third water, relative to the fully localized monomer state in the cyclic minimum energy complex, **c-W$_3$**. The resonance delocalization of the three states in Figure 3 determined by the GX-Pol model is depicted in Figure 4a, which is compared with the charge transfer (CT) effects in the fully delocalized HF determinant wave function in Figure 4b. It is aesthetically pleasing to visualize that the traditional Heitler−London−Slater−Pauling valence bond resonance theory of localized configurations as modeled by the GX-Pol method can provide an excellent description of the charge delocalization (i.e., charge transfer) as illustrated by the fully delocalized Hartree−Fock wave function. Importantly, such analyses coupled with quantitative structural weight can provide a deeper understanding of intermolecular interactions, including charge transfer effects in condensed phases.[20]

5. Conclusions

The explicit polarization (X-Pol) method has been generalized to incorporate charge delocalization resonance effects in the framework of valence bond theory. In the original X-Pol method, a macromolecular system is partitioned into

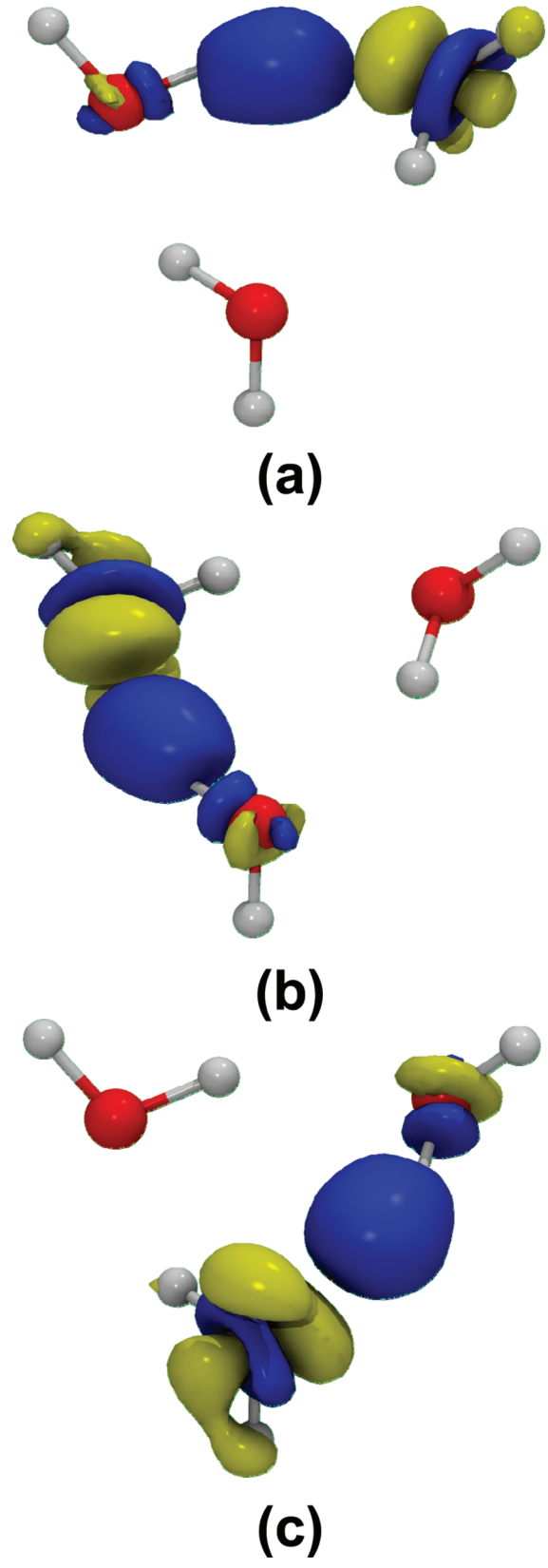


Figure 3. Electron density difference isosurface between the pairwise charge delocalized and the strictly localized system, $\rho(\hat{A}\psi_1\psi_2\psi_3)$, in the **c-W$_3$** complex: (a) W2−W3 dimer pair, $\rho(\hat{A}\psi_{23}\psi_1)$; (b) W1−W2 dimer pair, $\rho(\hat{A}\psi_{12}\psi_3)$; and (c) W1−W3 dimer pair, $\rho(\hat{A}\psi_{13}\psi_2)$. The structure orientation is shown in Figure 1a and the contour levels are at 0.0002 au, with blue contours representing gain in electron density and yellow contours representing charge depletion.

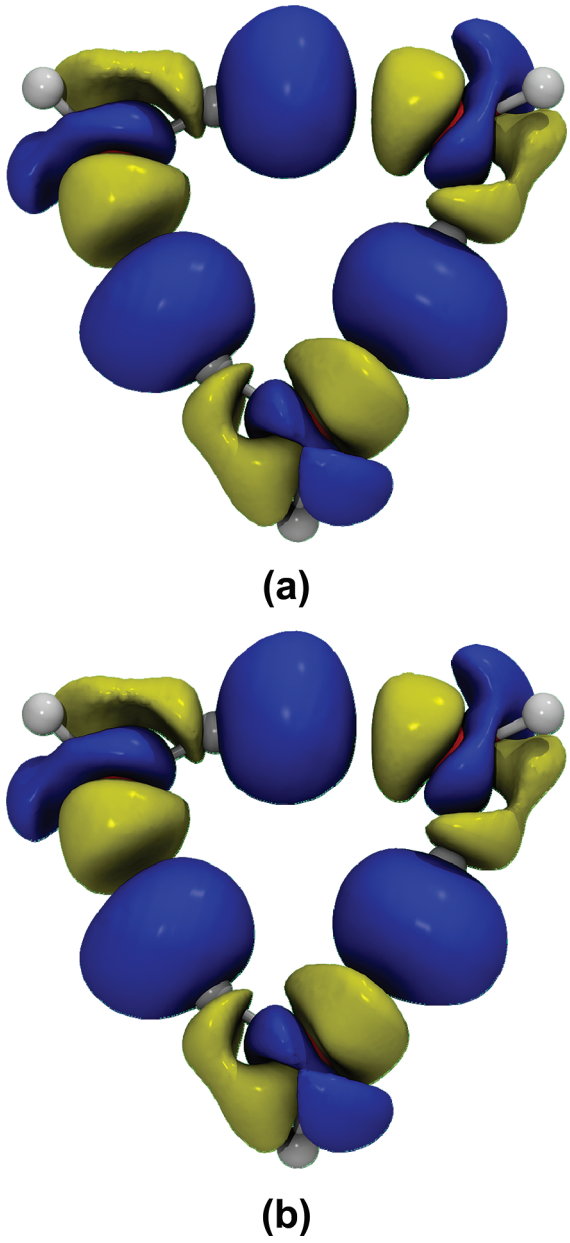


Figure 4. Comparison of (a) resonance charge delocalization modeled by the two-body GX-Pol wave function, Θ_{X2}, for **c-W$_3$**, which is a combination of the three states illustrated in Figure 3, and (b) charge transfer effects from Hartree−Fock theory. Electron density difference from the three-block X-Pol-X state, $\rho(\hat{A}\psi_1\psi_2\psi_3)$, is depicted at a contour level of 0.0002 au, with blue contours representing gain in electron density and yellow contours representing charge depletion. Note that the two figures look identical, which is the main point illustrated here.

individual blocks, which may be a single water molecule, a residue, a group of residues and molecules, or a subset of atomic orbitals on the same or different atoms. Molecular orbitals are strictly localized within each block, by virtue of expansion of MOs over basis functions within each subsystem only. Consequently, these block-localized molecular orbitals (BLMOs) are nonorthogonal across different blocks, but they can be orthogonal within each block. The X-Pol wave function is constructed as a Hartree product of the individually determinant wave functions of all blocks, which means that the exchange repulsion, dispersion correlation, and charge transfer between different blocks are neglected but are approximated empirically in X-Pol.

In the generalized X-Pol (GX-Pol) theory, we construct charge delocalization VB states by expanding the block localization space from individual blocks into pairwise delocalized blocks. Thus, the expansion of the basis space leads to charge delocalization between monomer pairs, and a series of $M \times (M - 1)/2$ pairwise charge delocalization states (which can also be called charge transfer states) can be constructed, where M is the total number of blocks (subsystems). The wave function for each of these CT states is a Hartree product of $M - 1$ blocks since two blocks have been grouped into a single CT unit. In general, L-body delocalized states can be analogously defined by grouping L monomer blocks into one. The Hartree product wave function for each state can be fully antisymmetrized, which introduces explicitly exchange repulsion interactions among all blocks. The GX-Pol wave function is a linear combination of all L-body delocalization VB states, which incorporates charge delocalization and their resonance as well as static correlation effects. The GX-Pol theory is illustrated by considering two water trimer complexes, one with a cooperative hydrogen-bonding network and another consisting of repulsive pair interactions. The illustrative examples show that the GX-Pol method can effectively incorporate charge delocalization and exchange repulsion explicitly in these fragment-based electronic structural methods for macromolecular systems.

Acknowledgment. We are grateful to Dr. Lingchun Song for computational assistance and valuable discussions. This work is supported by the National Science Foundation (Grant CHE09-57162).

References

(1) Gao, J. *J. Phys. Chem. B* **1997**, *101*, 657.

(2) Gao, J. *J. Chem. Phys.* **1998**, *109*, 2346.

(3) Xie, W.; Gao, J. *J. Chem. Theory Comput.* **2007**, *3*, 1890.

(4) Xie, W.; Song, L.; Truhlar, D. G.; Gao, J. *J. Chem. Phys.* **2008**, *128*, 234108.

(5) Song, L.; Han, J.; Lin, Y. L.; Xie, W.; Gao, J. *J. Phys. Chem. A* **2009**, *113*, 11656.

(6) Wierzchowski, S. J.; Kofke, D. A.; Gao, J. *J. Chem. Phys.* **2003**, *119*, 7365.

(7) Xie, W.; Song, L.; Truhlar, D. G.; Gao, J. *J. Phys. Chem. B* **2008**, *112*, 14124.

(8) Xie, W.; Orozco, M.; Truhlar, D. G.; Gao, J. *J. Chem. Theory Comput.* **2009**, *5*, 459.

(9) Cembran, A.; Bao, P.; Wang, Y.; Song, L.; Truhlar, D. G.; Gao, J. *J. Chem. Theory Comput.*, accepted; DOI: 10.1021/ct100268p.

(10) Mo, Y.; Peyerimhoff, S. D. *J. Chem. Phys.* **1998**, *109*, 1687.

(11) Mo, Y.; Gao, J.; Peyerimhoff, S. D. *J. Chem. Phys.* **2000**, *112*, 5530.

(12) Mo, Y.; Zhang, Y.; Gao, J. *J. Am. Chem. Soc.* **1999**, *121*, 5737.

(13) Stoll, H.; Wagenblast, G.; Preuss, H. *Theor. Chim. Acta* **1980**, *57*, 169.

(14) Mo, Y.; Song, L.; Lin, Y. *J. Phys. Chem. A* **2007**, *111*, 8291.

(15) Mo, Y.; Gao, J. *J. Phys. Chem. A* **2001**, *105*, 6530.

(16) Mo, Y.; Subramanian, G.; Gao, J.; Ferguson, D. M. *J. Am. Chem. Soc.* **2002**, *124*, 4832.

(17) Mo, Y.; Schleyer, P. v. R.; Wu, W.; Lin, M.; Zhang, Q.; Gao, J. *J. Phys. Chem. A* **2003**, *107*, 10011.

(18) Cubero, E.; Luque, F. J.; Orozco, M.; Gao, J. *J. Phys. Chem. B* **2003**, *107*, 1664.

(19) Brauer, C. S.; Craddock, M. B.; Kilian, J.; Grumstrup, E. M.; Orilall, M. C.; Mo, Y.; Gao, J.; Leopold, K. R. *J. Phys. Chem. A* **2006**, *110*, 10025.

(20) Mo, Y.; Gao, J. *J. Phys. Chem. B* **2006**, *110*, 2976.

(21) McWeeny, R. *Proc. R. Soc. London, A* **1959**, *253*, 242.

(22) Stoll, H.; Preuss, H. *Theor. Chem. Acc.* **1977**, *46*, 12.

(23) Shukla, A.; Dolg, M.; Stoll, H.; Fulde, P. *Chem. Phys. Lett.* **1996**, *262*, 213.

(24) Hankins, D.; Moskowitz, J. W.; Stillinger, F. H. *J. Chem. Phys.* **1970**, *53*, 4544.

(25) Zhang, D. W.; Xiang, Y.; Zhang, J. Z. H. *J. Phys. Chem. B* **2003**, *107*, 12039.

(26) Ding, Y.; Mei, Y.; Zhang, J. Z. H. *J. Phys. Chem. B* **2008**, *112*, 11396.

(27) Cembran, A.; Song, L.; Mo, Y.; Gao, J. *J. Chem. Theory Comput.* **2009**, *5*, 2702.

(28) Gao, J.; Garcia-Viloca, M.; Poulsen, T. D.; Mo, Y. *Adv. Phys. Org. Chem.* **2003**, *38*, 161.

(29) Gao, J.; Mo, Y. *Prog. Theor. Chem. Phys.* **2000**, *5*, 247.

(30) Mo, Y.; Gao, J. *J. Comput. Chem.* **2000**, *21*, 1458.

(31) Mo, Y.; Gao, J. *J. Phys. Chem. A* **2000**, *104*, 3012.

(32) Song, L.; Gao, J. *J. Phys. Chem. A* **2008**, *112*, 12925.

(33) Song, L.; Mo, Y.; Gao, J. *J. Chem. Theory Comput.* **2009**, *5*, 174.

(34) Song, L.; Mo, Y.; Zhang, Q.; Wu, W. *J. Comput. Chem.* **2005**, *26*, 514.

(35) Nadig, G.; Van Zant, L. C.; Dixon, S. L.; Merz, K. M., Jr. *J. Am. Chem. Soc.* **1998**, *120*, 5593.

(36) Van der Vaart, A.; Merz, K. M., Jr. *J. Phys. Chem. A* **1999**, *103*, 3321.

(37) Van der Vaart, A.; Merz, K. M., Jr. *J. Am. Chem. Soc.* **1999**, *121*, 9182.

(38) Van Der Vaart, A.; Merz, K. M., Jr. *Int. J. Quantum Chem.* **2000**, *77*, 27.

(39) van der Vaart, A.; Merz, K. M., Jr. *J. Chem. Phys.* **2002**, *116*, 7380.

(40) Bryantsev, V. S.; Diallo, M. S.; van Duin, A. C. T.; Goddard, W. A. I. *J. Chem. Theory Comput.* **2009**, *5*, 1016.

(41) Chen, W.; Gordon, M. S. *J. Phys. Chem.* **1996**, *100*, 14316.

(42) Kitaura, K.; Morokuma, K. *Int. J. Quantum Chem.* **1976**, *10*, 325.

(43) Mo, Y.; Wu, W.; Song, L.; Lin, M.; Zhang, Q.; Gao, J. *Angew. Chem., Int. Ed.* **2004**, *43*, 1986.

(44) Song, L.; Song, J.; Mo, Y.; Wu, W. *J. Comput. Chem.* **2009**, *30*, 399.

(45) Mo, Y.; Gao, J. *Acc. Chem. Res.* **2007**, *40*, 113.

(46) Cembran, A.; Payaka, A.; Lin, Y. L.; Xie, W.; Song, L.; Mo, Y.; Gao, J. *J. Chem. Theory Comput.* **2010**, *6*, 2242.

(47) King, H. F.; Staton, R. E.; Kim, H.; Wyatt, R. E.; Parr, R. G. *J. Chem. Phys.* **1967**, *47*, 1936.

(48) 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. S. *J. Comput. Chem.* **1993**, *14*, 1347.

(49) Woon, D. E.; Dunning, T. H., Jr. *J. Chem. Phys.* **1995**, *103*, 4572.

(50) Chirgwin, H. B.; Coulson, C. A. *Proc. R. Soc. London, Ser. A* **1950**, *2*, 196.

(51) Schutz, M.; Burgi, T.; Leutwyler, S.; Burgi, H. B. *J. Chem. Phys.* **1993**, *99*, 5228.

(52) Mulliken, R. S. *J. Chem. Phys.* **1964**, *61*, 20.

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

CT100292G

JCTC Journal of Chemical Theory and Computation

Maximum Flux Transition Paths of Conformational Change

Ruijun Zhao,*,† Juanfang Shen,‡ and Robert D. Skeel*,†,‡

Department of Computer Science, Purdue University, West Lafayette, Indiana 47907-2107, and Department of Mathematics, Purdue University, West Lafayette, Indiana 47907-2067

Received December 22, 2009

Abstract: Given two metastable states A and B of a biomolecular system, the problem is to calculate the likely paths of the transition from A to B. Such a calculation is more informative and more manageable if done for a reduced set of collective variables chosen so that paths cluster in collective variable space. The computational task becomes that of computing the "center" of such a cluster. A good way to define the center employs the concept of a committor, whose value at a point in collective variable space is the probability that a trajectory at that point will reach B before A. The committor "foliates" the transition region into a set of isocommittors. The maximum flux transition path is defined as a path that crosses each isocommittor at a point which (locally) has the highest crossing rate of distinct reactive trajectories. This path is based on the same principle used in the minimum resistance path of Berkowitz et al., but it has two advantages: (i) the path is invariant with respect to a change of coordinates in collective variable space, and (ii) the differential equations that define the path are simpler. It is argued that such a path is nearer to an ideal path than others that have been proposed, with the possible exception of the finite-temperature string method path. To make the calculation tractable, three approximations are introduced, yielding a path that is the solution of a nonsingular two-point boundary-value problem. For such a problem, one can construct a simple and robust algorithm. One such algorithm and its performance are discussed.

1. Summary

1.1. Introduction. Considered here is the problem of computing transition paths of conformational change, given two different metastable states of a biomolecule. One motivation for this is to facilitate the accurate calculation of free energy differences. Another motivation is to determine the existence and structure of transition states and intermediate metastable states. The latter are possible targets for inhibitors of enhanced specificity in cases where a family of proteins have active sites with very similar structure. A good example of this situation is the Src tyrosine kinase family,[1] which has long been implicated in the development of cancer. For this system, there are already computational results,[2−4] supported by experimentation,[5] for the transition path from an active catalytic domain to an inactive catalytic domain.

Some approaches to this problem generate ensembles of trajectories based on the equations of motion. Notable examples are transition path sampling[6] and Markov state models.[7] Applying such methods to large proteins (without compromise) would appear to require exceptional computing capabilities, so here we pursue a more theoretical approach that avoids "direct numerical simulation." Such an approach seeks to characterize one (or several isolated) "representative" reaction path connecting two given metastable states, each path representing a bundle or cluster of trajectories. Here, we adopt a well developed and tested theory, namely, transition path theory (TPT).[8−11] Additional references on computing transition paths are found in ref 12. In general, it

* To whom correspondence should be addressed. E-mail: rzhao@purdue.edu (R.Z.); skeel@cs.purdue.edu (R.D.S.).

† Department of Computer Science.

‡ Department of Mathematics.

10.1021/ct900689m © 2010 American Chemical Society
Published on Web 07/23/2010

may also be of interest to calculate (i) the reaction rate for each bundle, or, at least, the relative rate for different bundles, and (ii) the potential of mean force. Here, we consider only the calculation of the path itself.

It happens that the path proposed here is based on the same principle as the minimum resistance path[13] but differs in an important respect. Here, we make the path independent of the choice of coordinates in collective variable space by introducing a metric tensor to measure distance. This enables one to choose these coordinates on the basis of convenience, and, in particular, makes it easier to adapt to limitations imposed by a simulation program. Additionally, an appropriate choice of metric yields differential equations for defining the path that are significantly simpler than those for the minimum resistance path.

The minimum resistance path is offered as an alternative to the minimum energy path (MEP; also known as the steepest descent path), to capture finite temperature effects. The MaxFlux method[14] computes the minimum resistance path, as does the nudged elastic band implementation of ref 15. Other temperature-dependent paths have been proposed, including one defined as a most probable stochastic trajectory in the sense of a path integral,[16] one based on mean first-passage times,[17] and one whose tangent is the most probable direction determined from a swarm of trajectories.[12] A compelling case for including temperature effects is given in ref 18, which shows for alanine dipeptide the considerable extent to which the MaxFlux path from α_R to C_{7ax1} differs from the minimum free energy path (MFEP).[9] This same article also shows that this same MaxFlux path well represents an ensemble of trajectories generated with transition path sampling. However, the ability to locate the center of such a bundle depends on a selection of collective variables that avoid fine-grained roughness in the free energy landscape. Otherwise, actual averaging is needed, as in the finite-temperature string (FTS) method (see section IV.B of ref 19 and ref 20).

In a nutshell, this article embraces a certain aspect of TPT and carries it to a logical conclusion, obtaining a formula, an implementation, and a proof of concept. The aim is to compute a path that is closer to the ideal than the MFEP and that, in a couple of respects, is better than the path of the FTS method. Additionally, the formula for the path is computationally more attractive than the formula that underlies either the path of the FTS method *or the MFEP*.

1.2. Outline and Discussion. There are two distinct steps in getting a solution: The first is to define the problem without concern for the methods to be employed (other than taking into account the intrinsic difficulty of the problem). Defining a problem apart from a method gives a more concise definition. Also, by not guessing about what is feasible computationally, one may avoid unnecessary compromises. The second step is to construct a method and algorithm.

Given two metastable states A and B of a biomolecular system, the aim is to calculate the likely paths of the transition from A to B. Such a calculation is more informative and more manageable if done for a reduced set of *collective variables*, functions of the system configuration x,

$$\zeta_1 = \xi_1(x),\ \zeta_2 = \xi_2(x), ..., \zeta_\nu = \xi_\nu(x), \quad \text{abbreviated as } \zeta = \xi(x)$$

chosen so that paths cluster in collective variable space. The computational task becomes that of computing the "center" of such a cluster. A good way to define the center employs the concept of a committor, whose value at a point in collective variable space is the probability that a trajectory at that point will reach B before A. The committor "foliates" the transition region into a set of committor isosurfaces known as isocommittors. The maximum flux transition path (MFTP) is defined as a path that intersects each isocommittor at a point which (locally) has the highest crossing rate of distinct reactive trajectories. A more detailed account of the problem definition is given in section 2.

The minimum free energy path has been used for some time to represent reactive trajectories in collective variable space. Only fairly recently has its relationship to reactive trajectories been explained. The article[9] applies large deviation theory to show that the MFEP is the most probable path in the zero temperature limit of dynamics on a free energy surface defined at finite temperature. Hence, the MFEP (though not the MEP) is an inherently inconsistent construct, and it is *useful only to the extent that it represents fully finite-temperature trajectories*. In fact, it does this fairly well on the simple tests reported here.

To make the calculation tractable, three approximations are introduced. To make the committor a more accessible quantity, the set of paths is approximated by a Brownian dynamics model, resulting in a boundary value problem in ν-dimensional space. Then, the number of space dimensions is reduced to one by assuming that most of the transition paths are contained in a tube, resulting in a two-point boundary-value problem with 2ν unknowns. A third approximation reduces this to ν unknowns, whose solution is a *maximum flux transition path*. The resulting equations involve a free energy gradient term and an explicitly temperature-dependent curvature term. Specifically, the maximum flux transition path $\zeta = Z(s)$, $0 \le s \le 1$, is defined by the condition that

$$-\beta \nabla F^+(Z) - \frac{\nabla_\zeta c(Z, Z_s)}{c(Z, Z_s)} + \frac{(D(Z)^{-1} Z_s)_s}{c(Z, Z_s)^2} \,||\, D(Z)^{-1} Z_s$$

holds for $\zeta = Z(s)$, where

$$F^+(\zeta) = F(\zeta) - \frac{1}{2\beta} \log(\det D(\zeta)), \qquad c(\zeta, \omega) = (\omega^T D(\zeta)^{-1} \omega)^{1/2}$$

β is the inverse temperature, $F(\zeta)$ is the free energy profile, $D(\zeta)$ is a proto-diffusion tensor depending on masses and ξ, and the subscript s denotes differentiation (d/ds). In the high temperature limit, the path becomes a straight line. In the low temperature limit, the path becomes an MFEP. At zero temperature, the path will have cusps at some intermediate local minima, which presents difficulties if free energy profiles or relative reaction rates are to be determined. This formula is a key result of this article. Details are given in section 3. By contrast, the minimum resistance path satisfies

a more complicated condition, $-\beta\nabla F - 2c^{-1}\nabla_\xi c + c^{-2}(WD^{-1}Z_s)_s \parallel WD^{-1}Z_s$, where $W = 2I - Z_s(Z_s^TZ_s)^{-1}Z_s^T$, which differs from the MFTP except in very special cases such as $D(\zeta) = I$.

The temperature-dependent curvature term not only provides a finite temperature correction to the MFEP but yields a nonsingular second-order ordinary differential equation, amenable to standard techniques—except for the need to do computationally intensive sampling to evaluate terms in the differential equation. An existing set of algorithms for the MFEP[9,21] applies equally well to the MFTP. In section 4, the equation of a simplified MFTP, eq 12, is discretized using upwinded differencing and solved using the semi-implicit simplified string method.[22] (A notable alternative is the nudged elastic band method, introduced in ref 23.) Algorithmic details are provided in section 4.

Section 5 compares the MFTP to the MFEP on numerical examples. First, an artificial problem in full configuration space is solved to demonstrate the effect of the curvature term of the MFTP. (A problem in full configuration space is equivalent to a problem in collective variable space with perfect sampling.) In particular, the necessity of using an adaptive mesh for the MFEP is demonstrated. Then, alanine dipeptide in a vacuum is solved using the ϕ,ψ dihedrals as collective variables. For the transition path from C_{7ax} to C_{7eq} as in ref 9, the computational cost for calculating the MFTP and the MFEP is almost the same. However, for a transition path from C_{7eq} to C'_{7eq} through C_{7ax} shown in ref 19, the MFEP has a cusp at C_{7ax}, and the computational cost for finding such a cusp is expensive. On the other hand, the MFTP smooths out the cusp, and the computational cost is reduced. Finally, the MFTP is calculated for alanine dipeptide in explicit water and for alanine decapeptide.

An open source implementation of the MFTP method is available[24] as a relatively simple set of Python modules with examples using pure Python, CHARMM,[25] and NAMD.[26]

1.3. Conclusions. For alanine dipeptide, the MFEP, MFTP, and FTS method paths are quite similar. On a contrived problem with a rough energy landscape, e.g., Figure 2 in ref 20, the FTS method path gives a much better result. On a different contrived problem given in section 5.1, the MFTP gives a much better result. Contrived examples are relevant because computational techniques are sometimes applied in extreme situations for which they may not have been designed. In terms of quality, the MFTP ranks higher than the MFEP but lower than the FTS method path (because the latter addresses the more serious difficulty of multiple local minima).

The minimum free energy path (and that of the FTS method) can have cusps at local minima of free energy, which makes it unsuitable for defining an isocommittor at these points and harder to compute. Computational difficulties include the need for an adaptive mesh and a greater number of iterations until convergence.

2. What Is the Problem?

We begin by defining an ensemble of transition paths from A to B: For simplicity, assume the molecular system obeys Newtonian dynamics with potential energy function $U(x)$ and a diagonal matrix M of atomic masses. Positions x and momenta p satisfy $x = X(t)$ and $p = P(t)$, where $(\mathrm{d}/\mathrm{d}t)X(t) = M^{-1}P(t)$ and $(\mathrm{d}/\mathrm{d}t)P(t) = -\nabla U(X(t))$. Initial values are drawn from a Boltzmann–Gibbs distribution $\rho(x, p)$: positions x from probability density const$\cdot \mathrm{e}^{-\beta U(x)}$ and momenta p from a Maxwell distribution. Imagine an extremely long trajectory. The trajectory enters and leaves A and B many times, yielding a huge set of reactive paths from A to B. (A reactive path is a piece of the trajectory outside of A and B that comes from A and goes to B.)

Generating an ensemble of trajectories is extremely demanding computationally. And, even if this were possible, what would the user do with all the data? By answering such a question, we might well avoid the task of computing trajectories. It is likely that one would cluster the trajectories to produce a concise description. Therefore, one might instead directly determine such a concise description. Specifically, if the paths cluster into one or several distinct isolated bundles/tubes/channels/pathways, one might compute a "representative path" for each cluster. This idea is developed in the paragraphs that follow.

However, transition paths might not cluster adequately—in full configuration space. Assume, though, that there is a smaller set of collective variables, $\zeta = \xi(x)$, such that in ζ space, paths cluster into one or several distinct isolated channels connecting two separated subsets A_ξ and B_ξ of collective variable space. Otherwise, there is little of interest to compute. A typical example of collective variables is ϕ/ψ angles along a peptide backbone. Once the collective variables are specified, the problem is to calculate a path in collective variable space, $\zeta = Z(s)$, $0 \le s \le 1$, connecting A_ξ to B_ξ where the transition paths are concentrated. Along with a parametrization of the path in collective variable coordinates would be a realization of it in Cartesian coordinates, so once the path is generated, structures can be studied as well. A drawback of this approach is the need to identify an appropriate set of collective variables. Indeed, defining suitable collective variables is an important research problem.[27]

We want a minimal set of collective variables subject to two conditions: First, the coordinates ζ must suffice to describe states A_ξ and B_ξ in ζ space corresponding to A and B. Second, coordinates ζ must also be rich enough to "express the mechanism of conformational change" along the transition path. To make the second condition more precise, we introduce the notion of "quasi-committor."

To measure the progress of a transition, there is a natural reaction coordinate, known as the *committor*. This concept of a commitment probability was introduced by Onsager,[28] and the abbreviated term "committor" was introduced in ref 29, which they defined as follows: For each point x in configuration space, consider a trajectory starting with $X(0) = x$ and velocities drawn at random from a Maxwell distribution and define the committor $q(x)$ to be the probability of reaching B before A. Since it is the coordinates of the collective variables that are of interest, it is natural also to define a *quasi-committor*: For each point ζ, consider a trajectory starting with random initial values *conditioned on*

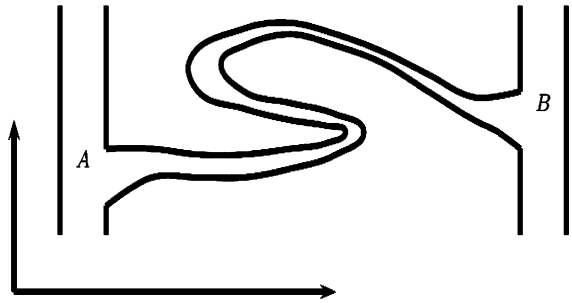


Figure 1. Schematic illustration of a poor choice of collective variables. The horizontal axis is collective variables, and the vertical axis is unrepresented degrees of freedom. The collective variables fail to indicate the progress of the reaction.

$\xi(x) = \zeta$ and define the quasi-committor $\hat{q}(\zeta)$ to be the probability of reaching B_ξ before A_ξ:

$$\hat{q}(\zeta) = \Pr(\xi(X(t)) \text{ reaches } B_\xi \text{ before } A_\xi | \xi(X(0)) = \zeta)$$

We could say that the variables $\zeta = \xi(x)$ are rich enough to express the mechanism of conformational change if the quasi-committor $\hat{q}(\zeta)$ has no local minima or maxima outside of A_ξ and B_ξ (except for regions of negligible probability). Otherwise, there is some unexpressed degree of freedom important to the transition. As an example, suppose that virtually all trajectories stay within a narrow tube having a geometry in full configuration space illustrated by Figure 1. Suppose that the free energy profile *as a function of arc length* along the transition tube is much higher in the backward section than it is in the two forward sections. Then, most of the increase in the quasi-committor as a function of arc length occurs in the middle section of the tube. Consequently, the variation in the quasi-committor, as a function of the ill-chosen collective variable ζ corresponding to the horizontal axis, will be dominated by this middle section of the tube. This results in a graph of $\hat{q}(\zeta)$ that increases at the beginning and end of its range but decreases in the middle part. In addition to $\hat{q}(\zeta)$ having no local extrema, it is desirable that $\hat{q}(\xi(x)) \approx q(x)$. The quality of the collective variables can be checked in principle by calculating quasi-committor values $\hat{q}(\zeta)$ at points along the path from dynamics trajectories.

Two approaches have been proposed for defining the center of a cluster of paths in ζ space:

(i) A most probable path, e.g., a swarm-of-trajectories string method[12] path

(ii) A path that intersects each isosurface of the quasi-committor at a center of the collection of points where reactive trajectories cross that isosurface, e.g., a finite temperature string method[19] path and a *maximum flux transition path*.

An MFEP is a limiting case of both approaches. (The MFEP is obtained from these various approaches by letting $\beta \to \infty$ in the path formula but not in the definition of the free energy profile.) Defining a representative path is a compromise between (i) best capturing the object of interest and (ii) simplicity.

One problem with seeking the most probable path is that it is unclear how to assign relative probabilities to paths. More importantly, the most probable path tends to be a path of minimum energy, and it is not clear—a priori—that this is a "representative" path. For Hamiltonian dynamics, it would seem that the probability that we attach to a path would be proportional to $\exp(-\beta E)$, where E is the energy. Hence, the most probable path is the one with just enough energy to surmount the potential energy barriers. For stochastic dynamics, the explanation of how to assign probability to paths is quite complicated—if paths of different durations are being compared. An explanation for Brownian dynamics is possible using Freidlin−Wentzell theory and the assumption of vanishingly small noise (see Appendix A of ref 9). It is reassuring though that the results of Freidlin−Wentzell theory agree with those of TPT in the zero-temperature limit (for $F(\zeta)$ held fixed).

For defining a path in terms of an intersecting point on each isosurface of a quasi-committor, one needs

(i) A definition for the distribution of crossing points of reactive trajectories through a quasi-commitor isosurface

(ii) A definition of centrality, e.g., mode, median, or mean

We consider each of these in turn.

The finite-temperature string method defines the distribution of crossing points of reactive trajectories in a way that includes recrossings. A subsequent article[10] illustrates the dramatic distortions that arise by including recrossings, and it emphasizes crossings of a surface by distinct reactive trajectories instead of all crossings by reactive trajectories. They define such a distribution in terms of the net crossings of reactive trajectories across each infinitesimal piece of a surface. It is not obvious, however, that this necessarily gives non-negative values everywhere on the isosurface of a quasi-committor, so, instead, we use the density of last crossings by reactive trajectories, called *last hitting points* in ref 8. For the Brownian dynamics approximation developed in the next section, these two measures are identical.

Consider now the question of defining the center. Let $j(\zeta)$ denote the density associated with a definition for the distribution of crossing points of reactive trajectories through a quasi-committor isosurface. One choice for the center is the point of highest probability, In other words, seek the path $\zeta = Z(s)$, $0 \le s \le 1$, each of whose points $Z(s)$ is a local maximum of the density $j(\zeta)$ on the quasi-committor isosurface Σ passing through $Z(s)$. This is what we use for the MFTP. Another choice, associated with the *finite-temperature string method*, is to construct the path from the mean value ζ' on each quasi-committor isosurface Σ: the point ζ' that minimizes $\int_\Sigma |\zeta' - \zeta|^2 j(\zeta)\, d\zeta$. Although this notion is a superior measure of centrality, it is more complicated to explain. In practice, methods for finding a maximum are designed only to find a local maximum, which is what we do for the MFTP. This is satisfactory if there is a choice of collective variables that produces a free energy landscape free of roughness at the scale of the thermal energy.[20] In any case, the equations defining a center-of-density path are intrinsically more expensive computationally to solve than those for the MFTP, because they require averaging on quasi-committor isosurfaces $\hat{q}(\zeta) =$ constant (in addition to conditional averages on collective variable isosurfaces $\xi(x)$

$= \zeta$ in full configuration space) rather than merely determining a (local) maximum.

These definitions for the center are flawed, however. "Collective variable space" is a set of points, each point ζ representing a manifold $\xi(x) = \zeta$ in configuration space of dimension $3N - \nu$. A change of variables *within* collective variable space, $\xi(x) = \chi(\xi'(x))$, should not change the path, only its representation. This can be achieved by defining a metric tensor $M(\zeta)$ for measuring distances—the distance from ζ to $\zeta + \mathrm{d}\zeta$ being given by $(\mathrm{d}\zeta^T M(\zeta)\mathrm{d}\zeta)^{1/2}$—which somehow measures the distance between two infinitesimally close manifolds in configuration space. (Being a tensor means that $M_{\xi'}(\zeta') = (\partial\chi/\partial\zeta')^T M_{\xi}(\chi(\zeta'))\partial\chi/\partial\zeta')$. Appendix A derives a correction factor for hypersurface area, and accordingly, we define a corrected last hitting point density:

$$j_c(\zeta) = \frac{|\nabla\hat{q}(\zeta)|}{|M(\zeta)^{-1/2}\nabla\hat{q}(\zeta)|}(\det M(\zeta))^{-1/2} j(\zeta)$$

A natural choice for $M(\zeta)$ arises during the course of deriving the hitting point density.

3. A Method

As stated previously, computing $\hat{q}(\zeta)$ is not feasible. Consequently, we derive a method, which employs three uncontrolled approximations—a controlled approximation being one that can be made arbitrarily accurate with sufficient computational effort. Section 3.1 approximates paths in collective variable space by those of Brownian dynamics. Section 3.4 assumes most paths lie in a tube where isocommittors are planar, and section 3.5 assumes that on average the trajectories are parallel to the path. The basic ingredients of much of this development are present in the literature but scattered among several articles. Here, they are combined to produce equations from which we derive the MFTP.

3.1. Brownian Dynamics Approximation of Collective Variable Paths. The probability density function (pdf) for $\xi(x)$ is

$$\rho_\xi(\zeta) = \langle\delta(\xi(x) - \zeta)\rangle = \int\int \delta(\xi(x) - \zeta)\rho(x,p)\,\mathrm{d}x\,\mathrm{d}p$$

where $\delta(\zeta) = \delta(\zeta_1)\delta(\zeta_2)...\delta(\zeta_\nu)$. Let $\langle...\rangle_\zeta$ be the expectation for the conditional density $\rho(x, p|\xi(x) = \zeta)$:

$$\langle O(x)\rangle_\zeta = \frac{\langle\delta(\xi(x) - \zeta)O(x)\rangle}{\langle\delta(\xi(x) - \zeta)\rangle}$$

In Appendix B is an adaptation of an argument from ref 9 (Section III, A and B) suggesting that as an approximation to $\hat{q}(\zeta)$, we should seek a function $q(\zeta)$ that minimizes a certain functional $I(q)$ that can be expressed in terms of collective variables ζ. Define the free energy $F(\zeta)$ for coordinates $\zeta = \xi(x)$ by

$$\mathrm{const}_\xi \mathrm{e}^{-\beta F(\zeta)} = \rho_\xi(\zeta) = \langle\delta(\xi(x) - \zeta)\rangle \qquad (1)$$

Also, define a proto-diffusion tensor D by

$$D(\zeta) = \frac{1}{2}\beta^{-1}\langle\xi_x(x)M^{-1}\xi_x(x)^T\rangle_\zeta$$

(There is freedom in the scaling of D. We use this freedom to make eq 4 below agree with an alternative derivation of the Brownian dynamics, in which one assumes instantaneous relaxation of the degrees of freedom not represented by the collective variables. The tensor $D(\zeta)$ fails to be a diffusion tensor because it is missing a time scale factor.) The functional is then

$$I(q) = \mathrm{const}_\xi \int \mathrm{e}^{-\beta F(\zeta)}\nabla q(\zeta)^T D(\zeta)\nabla q(\zeta)\,\mathrm{d}\zeta \qquad (2)$$

where the integral is over the transition region outside of A_ξ and B_ξ subject to $q(\zeta) = 0$ on the boundary of A_ξ and $q(\zeta) = 1$ on the boundary of B_ξ.

The corresponding Euler–Lagrange equation for $q(\zeta)$ is the Smoluchowski (backward Kolmogorov) equation:

$$-\nabla\cdot \mathrm{e}^{-\beta F(\zeta)} D(\zeta)\nabla q(\zeta) = 0 \qquad (3)$$

subject to $q(\zeta) = 0$ on the boundary of A_ξ and $q(\zeta) = 1$ on the boundary of B_ξ.

The function q that satisfies the Smoluchowski equation subject to the given boundary conditions can be shown to be the exact committor function for paths $\zeta = \zeta(\tau)$ in collective variable space generated by the Brownian dynamics

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\zeta = -\beta D(\zeta)\nabla F(\zeta) + (\nabla\cdot D(\zeta))^T + \sqrt{2}D_{1/2}(\zeta)\eta(\tau) \qquad (4)$$

where $D_{1/2}D_{1/2}^T = D$ and $\eta(\tau)$ is a collection of standard white noise processes. The fact that τ is an artificial time does not affect the committor. In principle, the assumption $q(\zeta) \approx \hat{q}(\zeta)$ can be checked a posteriori by comparing committor values of the Brownian dynamics to the quasi-committor values of actual dynamics.

Reference 9 (Section III.C) appears to suggest that the Smoluchowski equation uniquely specifies dynamics except for scaling of time: If the Smoluchowski eq 3 is satisfied by committors $q(\zeta)$ for arbitrary sets A'_ξ and B'_ξ in collective variable space, then trajectories whose committor functions satisfy eq 3 must have paths that are those of the Brownian dynamics. Hence, paths in collective variable space can be generated with the proper probabilities from the system of stochastic differential equations.

3.2. Last Hitting-Point Distribution. Appendix C considers the rate at which reactive trajectories cross an arbitrary surface Σ that separates collective variable space into two parts, one containing A_ξ and the other containing B_ξ. The result given there is that the rate of the *last* crossing of Σ by reactive trajectories is given by the integral

$$\int_\Sigma J(\zeta)\cdot\hat{n}(\zeta)\,\mathrm{d}S_\zeta$$

where $\hat{n}(\zeta)$ points to the side containing B_ξ and

$$J(\zeta) = \rho_\xi(\zeta)D(\zeta)\nabla q(\zeta)$$

is the last hitting-point flux. The choice of the last hitting point to represent the point where a reactive trajectory crosses an isocommittor is somewhat arbitrary. Therefore, it is gratifying to know that the expression for $J(\zeta)$ also gives the net flux and the first hitting-point flux of reactive trajectories.

The normal to an isocommittor is given by $\hat{n}(\zeta) = \nabla q(\zeta)/|\nabla q(\zeta)|$, so the distribution of last hitting points on an isocommittor is proportional to

$$j(\zeta) = \rho_{\xi}(\zeta)\nabla q(\zeta)^T D(\zeta)\nabla q(\zeta)/|\nabla q(\zeta)|$$

In particular, of interest is the flow rate through a tiny disk at ζ. Thus, we consider here the corrected last hitting point density $j_c(\zeta)$. A natural choice for $M(\zeta)$ is $D(\zeta)^{-1}$ because it simplifies the expression for $j_c(\zeta)$ and it can be shown to be a metric tensor. ($D(\zeta)$ is the "harmonic" Boltzmann-weighted average of $(\xi_x M^{-1}\xi_x^T)^{-1}$, and $(\Delta\zeta^T(\xi_x M^{-1}\xi_x^T)^{-1}\Delta\zeta)^{1/2}$ is the distance from the hyperplane $\xi(x) = \zeta$ to the hyperplane $\xi(x) = \zeta + \Delta\zeta$ at the point x.) Hence, the corrected last hitting point density becomes

$$j_c(\zeta) = |D(\zeta)^{1/2}\nabla q(\zeta)|(\det D(\zeta))^{1/2}\rho_{\xi}(\zeta)$$

3.3. Defining the Path. For computation, it is convenient to label the isocommittors with the path parameter s. In particular, denote by $\Sigma(s)$, $0 \le s \le 1$, the isocommittor passing through $\zeta = Z(s)$. Write $\bar{q}(s) = q(Z(s))$ and define $\sigma(\zeta)$ implicitly by

$$q(\zeta) = \bar{q}(\sigma(\zeta)) \tag{5}$$

In this way, the committor $q(\zeta)$ is decomposed into two independent parts: one part $\sigma(\zeta)$ specifies the isocommittor label and the other part $\bar{q}(s)$ calibrates the isocommittors. Thus, $\nabla q(\zeta) = \bar{q}_s(\sigma(\zeta))\nabla\sigma(\zeta)$, and the corrected normal flux is

$$j_c(\zeta) = \bar{q}_s(\sigma(\zeta))(\nabla\sigma(\zeta)^T D(\zeta)\nabla\sigma(\zeta)\det D(\zeta))^{1/2}\rho_{\xi}(\zeta) \tag{6}$$

(recalling that the subscript s denotes differentiation d/ds). Note that $\bar{q}_s(\sigma(\zeta))$ is constant on an isocommittor $\Sigma(s)$, so it can be neglected when determining the center of intensity of $j_c(\zeta)$.

Each point $Z(s)$ on the desired path maximizes the last hitting-point flux $j_c(\zeta)$ on the isocommittor $q(\zeta) = q(Z(s))$. Hence, $\nabla j_c(Z(s)) \parallel \nabla q(Z(s))$. To keep the derivation independent of the calibration $\bar{q}(s)$, introduce a vector $n(s)$, not necessarily normalized, such that $n(s) \parallel \nabla q(Z(s))$. Hence,

$$\nabla j_c(Z(s)) \parallel n(s) \tag{7}$$

3.4. The Localized Tube Assumption. Assume there exists a tube connecting A_{ξ} to B_{ξ} such that (i) on each isocommittor, regions of high $j_c(\zeta)$ are concentrated in the tube and (ii) each isocommittor is nearly planar in the tube. This scenario is illustrated in Figure 2.

Exploit the localized tube assumption by approximating the isocommittor through $Z(s)$ as a plane $\Pi(s)$ with normal $n(s)$. Hence, the isocommittor surface $\Sigma(s)$: $\sigma(\zeta) = s$ has the simple description of a hyperplane

$$\Pi(s): n(s)\cdot(\zeta - Z(s)) = 0 \tag{8}$$

These approximations (see ref 11 (section 6.6.1)) are sufficient to define a practical method (see ref 30 (section 12)). The unknown direction vector $n(s)$ is to be chosen to minimize the integral $I(q)$ of eq 2 restricted to some tube. For simplicity, the boundary points $Z(0)$ and $Z(1)$ can be moved to points in A_{ξ} and B_{ξ} that locally minimize $F(\zeta)$. In this way, the problem of solving for a committor of many variables is reduced to that of a one-dimensional calculation along the length of the tube.

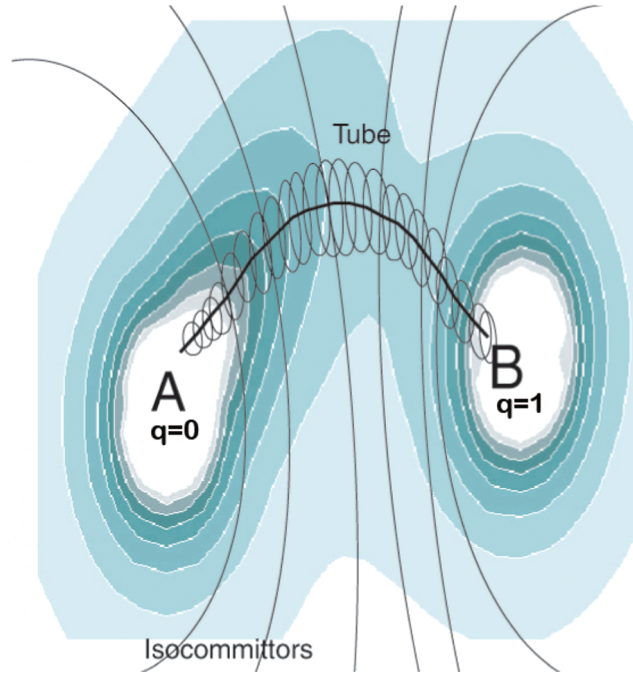


Figure 2. Localized tube assumption. Shading indicates contours of free energy. Thin curves denote isocommittors. Ellipses enclose concentrations of crossing points from reactive trajectories, and the thick curve is the center.

It remains to derive the condition that determines $Z(s)$. This is done in Appendix D, where it is shown that the condition is

$$-\beta\nabla F^{+}(Z) + \left.\frac{\nabla_{\zeta}c(\zeta, D(\zeta)n)}{c(\zeta, D(\zeta)n)}\right|_{\zeta=Z} + \frac{n_s}{n^T Z_s} \parallel n$$

A further simplification is to assume that (iii) $D(\zeta)$ is nearly constant on each isocommittor within the tube. More specifically, approximate $D(\zeta)$ by $D(Z(\sigma(\zeta)))$. Then, as shown in Appendix D, the above condition simplifies to

$$-\beta\nabla F(Z) + \frac{n_s}{n^T Z_s} \parallel n \tag{9}$$

3.5. The Maximum Flux Transition Path. Although the localized tube assumption is sufficient for defining a practical method, the method would not be simple, so we make an additional simplifying assumption: Assume the flux $J(\zeta)$ points in the direction of the path so that $J(Z(s)) \parallel Z_s(s)$ or $D(Z(s))\nabla q(Z(s)) \parallel Z_s(s)$, whence

$$n(s) \parallel D(Z(s))^{-1}Z_s(s)$$

As it can be shown, the result is a *maximum flux transition path*

$$-\beta\nabla F^{+}(Z) - \frac{\nabla_{\zeta}c(Z,Z_s)}{c(Z,Z_s)} + \frac{(D(Z)^{-1}Z_s)_s}{c(Z,Z_s)^2} \,||\, D(Z)^{-1}Z_s \tag{10}$$

(The simplifying assumption is justified, for example, if the probability is strongly peaked around the path, resulting in most of the probability contained in a narrow tube with a flux $J(\zeta)$ pointing in the direction of the tube and the path.) This assumption is also made for the FTS method, see eq 14 of ref 19 and section II.A. of ref 20. Geometrically, this condition means that instead of having the free energy gradient vanish orthogonal to the path, it is balanced by a "centripetal" force, which reduces curvature and avoids cusps.

To express eq 10 as an equation, write it as $-\beta c(Z,Z_s)^2\nabla F^{+} - c(Z,Z_s)\nabla_{\zeta}c(Z,Z_s) + (D^{-1}Z_s)_s = \lambda D(Z)^{-1}Z_s$ where λ is a scalar and premultiply by Z_s^T to obtain an expression for λ. After eliminating λ, the equation becomes

$$(I - \Pi)(c(Z,Z_s)^2\beta\nabla F^{+}(Z) + c(Z,Z_s)\nabla_{\zeta}c(Z,Z_s) - (D(Z)^{-1}Z_s)_s) = 0 \tag{11}$$

where $\Pi = c^{-2}D^{-1}Z_sZ_s^T$.

Note that, if D is constant, the limit $\beta \to 0$ for eq 11 gives a geodesic $Z_{ss} = 0$, which is the desired result.

In the two-dimensional case with $D = I$, the Euclidean length of $(I - \Pi)D(D^{-1}Z_s)_s/(Z_s^TD^{-1}Z_s)$ is exactly equal to the curvature, which is defined to be the reciprocal of the radius of curvature. To see this, note that this is true if we parametrize with the (actual) arc length, and note also that the curvature term is independent of parametrization (which can be checked analytically).

If we normalize the parametrization using $(\mathrm{d}/\mathrm{d}s)c(Z,Z_s) = 0$, this implies

$$\Pi(\nabla_{\zeta}c(Z,Z_s) - c(Z,Z_s)^{-1}(D(Z)^{-1}Z_s)_s) = 0$$

Combining with eq 11, we have

$$(I - \Pi)c(Z,Z_s)^2\beta\nabla F^{+}(Z) + c(Z,Z_s)\nabla_{\zeta}c(Z,Z_s) - (D(Z)^{-1}Z_s)_s = 0$$

(If the normalization $(Z_s^TZ_s)_s$ is used instead, the equation has the same form but with the projector $\Pi = (Z_s^TZ_s)^{-1}D^{-1}Z_sZ_s^TD$.)

The presence of partial derivatives of $D(\zeta)$ can be awkward. From the simplified condition eq 9, we get the simplified maximum flux transition path (SMFTP):

$$-\beta\nabla F(Z) + \frac{(D(Z)^{-1}Z_s)_s}{c(Z,Z_s)^2}||D(Z)^{-1}Z_s \tag{12}$$

We will implement the SMFTP in the following section; a similar procedure can be devised for the MFTP.

Values obtained from constructing the path can be used to calculate the free energy $F(Z(s))$ along the path:

$$F(Z(s)) - F(Z(0)) = \int_0^s \nabla F(Z(s'))^TZ_s(s')\,\mathrm{d}s' \tag{13}$$

However, $F(Z(s))$ is not a potential of mean force for the transition.

3.6. The Minimum Free Energy Path. The simplifying assumption of the preceding subsection, which is used to derive the MFTP, is valid in the limit $\beta \to \infty$ in the Brownian dynamics approximation, see refs 11 (section 6.6) and ref 9 (Appendix A). A more systematic derivation might therefore neglect the curvature term. The result would be a *minimum free energy path*:

$$Z_s||-\beta D(Z)\nabla F(Z)$$

Each point $\zeta = Z(s)$ on the MFEP is a local minimum of $F(\zeta)$ in the hyper-plane orthogonal to $D(Z(s))^{-1}Z_s(s)$.

One difference from an MFTP is that an MFEP can have a cusp at an intermediate local minimum. If the path passes sufficiently close to a local minimum $\zeta = \zeta_0$ of $F(\zeta)$, then for a short section of the path, $\zeta = Z(s)$, $a \le s \le b$, a quadratic approximation to $F(\zeta)$ is accurate. Assume $D =$ constant and $F(\zeta) = 1/2(\zeta - \zeta_0)^TA(\zeta - \zeta_0) +$ constant, where A is symmetric positive definite. The MFEP is then defined by $Z_s||-\beta DA(Z - \zeta_0)$. Perform a change of variables, $Y = \beta^{-1/2}Q^TD_{1/2}^{-1}(Z - \zeta_0)$ where $Q\Lambda Q^T$ is a diagonalization of $D_{1/2}^TAD_{1/2}$. The MFEP for $Y(s)$ is hence given by $Y_s||-\Lambda Y$. For simplicity, suppose that $Y = [x, y]^T$, that $x(a) < 0 < x(b)$, and that $\Lambda = \mathrm{diag}(\lambda, \mu)$ with $\lambda > \mu$. The path is hence defined by $y_s/(\mu y) = x_s/(\lambda x)$, which can be integrated to yield the path

$$y = \begin{cases} (x/x(a))^{\mu/\lambda}y(a) & x(a) \le x \le 0 \\ (x/x(b))^{\mu/\lambda}y(b) & 0 \le x \le x(b) \end{cases}$$

which has a cusp at $x = 0$.

The FTS method path is also likely to suffer from the presence of cusps, because for a harmonic potential, the average position is the same as the most probable position.

The presence of cusps undermines the localized tube assumption. In particular, the assumption of isocommittors being approximately planar breaks down at a cusp. This poses a difficulty when computing quantities that are averages on isocommittors. Additionally, cusps complicate the numerical approximation of paths.

4. An Algorithm

An algorithm for calculating a transition path employs a progression of four controlled approximations: discretization of the path $\zeta = Z(s)$ and the equations that define it, a finite number of iterations for the solution of nonlinear discrete equations, the use of restraints for constrained sampling, and finite sampling.

4.1. Discretization. The path $Z(s)$, $0 \le s \le 1$, is approximated as a piecewise polynomial with break points $0 = s_0 < s_1 < \ldots < s_J = 1$. Here, we choose a uniform mesh $s = 0, \Delta s, \ldots, 1$ and obtain the path by piecewise linear interpolation. Thus, the problem is reduced to determining *unknown* nodal values $Z_j \approx Z(s_j)$, $j = 0, 1, \ldots, J$, each representing a replica of the system in a different configuration.

It is convenient for computation to use for the path parameter s the arc length along the path divided by the total length of the path. In such a case, $|Z_s(s)|$ is constant. The arc length normalization becomes

$$|Z_{j+1} - Z_j|/\Delta s = |Z_j - Z_{j-1}|/\Delta s, \quad j = 1, 2, ..., J-1$$

Equation 12 is written as

$$-\beta D \nabla F - \frac{1}{c^2} D_s D^{-1} Z_s + \frac{1}{c^2} Z_{ss} \parallel Z_s$$

This is discretized by the finite difference scheme

$$(Z_s)_j \parallel g_j, \text{ where } g_j \stackrel{\text{def}}{=} -\beta D_j (\nabla F)_j - \frac{1}{c_j^2}(D_s D^{-1} Z_s)_j + \frac{1}{c_j^2}\frac{Z_{j+1} - 2Z_j + Z_{j-1}}{\Delta s^2}$$

and where

$$c_j^2 = \frac{1}{2}\Delta s^{-2}(\Delta_- Z_j^T D_j^{-1} \Delta_- Z_j + \Delta_+ Z_j^T D_j^{-1} \Delta_+ Z_j) \quad (14)$$

$$(D_s D^{-1} Z_s)_j = \frac{1}{2}\Delta s^{-2}(\Delta_- D_j D_j^{-1} \Delta_- Z_j + \Delta_+ D_j D_j^{-1} \Delta_+ Z_j) \quad (15)$$

with

$$\Delta_\pm D_j = \mp(D_j - D_{j\pm1}), \text{ and } \Delta_\pm Z_j = \mp(Z_j - Z_{j\pm1})$$

We choose upwinded differencing for $(Z_s)_j$ based on the direction of the modified mean force g_j:

$$(Z_s)_j = \begin{cases} (Z_j - Z_{j-1})/\Delta s & \text{if } g_j^T(Z_j - Z_{j-1}) > 0 \\ (Z_{j+1} - Z_j)/\Delta s & \text{if } g_j^T(Z_j - Z_{j+1}) > 0 \end{cases} \quad (16)$$

In the unlikely event that both conditions are satisfied, the choice is dictated by the arc length normalization step of the simplified string method to be discussed next.

Analogous to the development in section 3.5, it is possible to use the parametrization normalization to get a nonsingular second order difference equation for the values Z_j, but this is omitted because the solution method uses the formulation given here.

For the MFEP, cusps can occur at some intermediate local minima, requiring an adaptive mesh to resolve.

4.2. Solution of Nonlinear Discrete Equations. A second component of the algorithm is an iterative method for achieving rapid local convergence given a plausible initial guess.

Because of its simplicity and demonstrated effectiveness, we adopt the semi-implicit simplified string method used in ref 22 (eq 11). To determine a path, begin with an initial guess and generate successive improvements by alternating between moving the points of the curve Z_j in the direction g_j and reparameterizing.

The first step of each iteration is to solve the following equations for Z_j^*:

$$\frac{Z_j^* - Z_j}{\tau^2} = \frac{1}{c_j^2}\frac{Z_{j+1}^* - 2Z_j^* + Z_{j-1}^*}{\Delta s^2} - \frac{1}{c_j^2}(D_s D^{-1} Z_s)_j - \beta D_j (\nabla F)_j, \quad j = 1, 2, ..., J-1$$

$$\frac{Z_j^* - Z_j}{\tau^2} = -\beta D_j (\nabla F)_j, \quad j = 0, J$$

where c_j^2 and $(D_s D^{-1} Z_s)_j$ are given in eqs 14 and 15. (The extra factor τ provides the time scale factor missing from D.)

Then, the normalization adjustment is to choose the $\{Z_j\}$ to be equidistant along the resulting curve:

$s_0^* = 0, s_j^* = s_{j-1}^* + |Z_j^* - Z_{j-1}^*|$
$Z^*(s) =$ piecewise linear interpolation of $\{(s_j^*/s_J^*, Z_j^*)\}$, $0 \le s \le 1$
$Z_j^{\text{new}} = Z^*(j/J)$

It can be shown that if the semi-implicit simplified string method converges, the resulting points Z_j satisfy a nonstandard discretization of the differential equation containing τ as a parameter. In the limit $\tau \to 0$, the discretization becomes upwinded differencing.

For large systems, targeted molecular dynamics[31] has been used to get an initial path.[2,32] Another potentially promising but quite different approach is rigidity analysis.[33]

4.3. Conditional Averages. Evaluation of ∇F and D at break points involves sampling on hyper-surfaces $\{x:\xi(x) = Z_j\}$ of configuration space.

For calculating such conditional expectations, the Dirac delta function $\delta(s)$ can be approximated by the pdf of a Gaussian $\delta_\varepsilon(s) = (2\pi\varepsilon^2)^{-1/2} \exp(-s^2/(2\varepsilon^2))$. Note that

$$\delta_\varepsilon(\xi(x) - \zeta) e^{-\beta U(x)} = (2\pi\varepsilon^2)^{-\nu/2} e^{-\beta U(x;\zeta)}$$

where

$$U(x;\zeta) = U(x) + \sum_{i=1}^{\nu} u_i(x, \zeta_i), \text{ and } u_i(x, \zeta_i) = \frac{1}{2\beta\varepsilon^2}(\xi_i(x) - \zeta_i)^2 \quad (17)$$

Then, $\langle O(x)\rangle_\zeta = \langle O(x)\delta_\varepsilon(\xi(x) - \zeta)\rangle / \langle \delta_\varepsilon(\xi(x) - \zeta)\rangle$ is nothing but an average using $U(x;\zeta)$. The effect is that of using restraining potentials instead of constraints. These restraints should be as strong as possible without restricting the step size used in the sampling. From $\text{const}_\xi \exp(-\beta F(\zeta)) = \langle \delta_\varepsilon(\xi(x) - \zeta)\rangle$, we have

$$\nabla F(\zeta) = -\frac{1}{\beta\varepsilon^2}\langle \xi(x) - \zeta \rangle_\zeta$$

4.4. Sampling. We would like to estimate the statistical error of Z_j^*. Ideally, we want the standard deviation of the estimate smaller than some given tolerance. The major contribution to the sampling error of Z_j^* comes from that of $(\nabla F)_j$, because of the cancelation and subsequent multiplication by ε^{-2}. Thus, we neglect the statistical error of D_j in estimating the error of g_j. So then, the statistical error of Z_j^* comes from the sample average of $\Delta_j = \beta D_j (\nabla F)_j^n$, $n = 1, 2, ..., N$, where N is the sample size. The statistical error is defined by $(\max_{0\le j\le J} \text{error bar of } \Delta_j)\tau^2$, where an error bar is an estimate of 1 standard deviation. Such an estimate can be obtained using block averaging as in ref 34 (Appendix D.3). In general, 32 blocks is a reasonable choice.

At each iteration, the configuration x from the previous iteration could be used to start the equilibration of the molecular dynamics. Thus, it is necessary that values of x be stored such that $\xi(x) = Z_j$, $j = 0, 1, ..., J$. It is reasonable

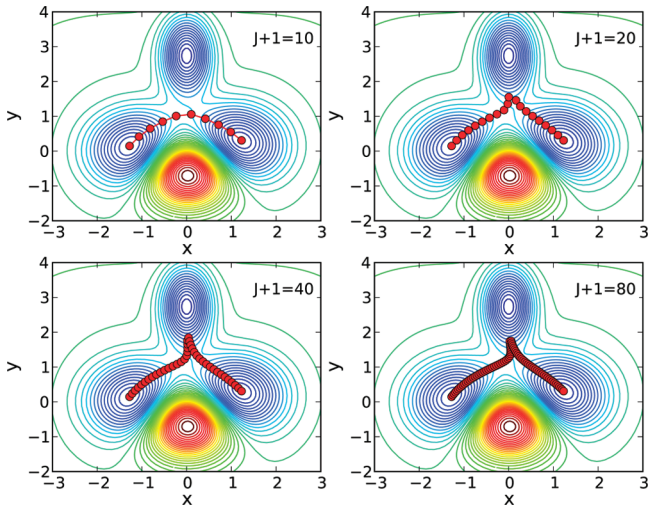


Figure 3. Minimum energy path obtained using the simplified string method. The initial path is the straight line between (−1, 0) and (1, 0). The path is discretized into $J + 1$ images. Four figures are generated using $J + 1 = 10$, 20, 40, 80 images, respectively.

to expect that less equilibration time is needed in later iterations as the path converges.

5. Numerical Tests

5.1. An Artificial Problem. As an example to illustrate our method, consider a problem finding the MFTP and MFEP for the potential energy function

$$U(x, y) = -4\exp(-4x^2 - (y - 2.75)^2) - 5\exp(-(x - 1)^2 - (y - 0.15)^2) - 5\exp(-(x + 1)^2 - y^2) + 8\exp(-x^2 - (y + 0.5)^2) + 0.001(x^4 + y^4)(x^4 + y^4)$$

where the energy unit is kcal/mol and the mass matrix $\mathbf{M}$ has identical diagonal entries. Unless specifically mentioned, the inverse temperature $\beta^{-1} = 0.59595$ kcal/mol, corresponding to 300 K. In particular, we take collective variables $\zeta = \xi(x, y) = (x, y)$. In this case, the MFEP becomes a minimum energy path (MEP). Alternatively, an MEP can be considered as an MFEP, for which we have an accurate estimate of $F(\zeta)$.

In Figure 3, we show an MEP connecting two local minima through the third local minimum. The MEP has a cusp at the intermediate minimum. The MEPs are computed using the simplified string method with piecewise linear interpolation and equal arc length normalization. The time step $\tau^2 = 0.01$. The iteration is stopped if $d < 0.00005$, where $d = \max_{0\le j\le J}|Z_j^{\text{new}} - Z_j|$. From the figure, we can see that the cusp is missing if the number of images ($J + 1 = 10$) is too small. Also, the MEP does not go through the intermediate local minimum as it should, even with many images ($J + 1 = 80$).

A calculation (not shown here) similar to that for Figure 3 was done for the MFTP. The MFTP is calculated using the semi-implicit simplified string method described in section 4. The MFTP can be resolved using a relatively small set of images; for example, the MFTP calculated by only 10 images ($J = 9$) is almost indistinguishable from the one

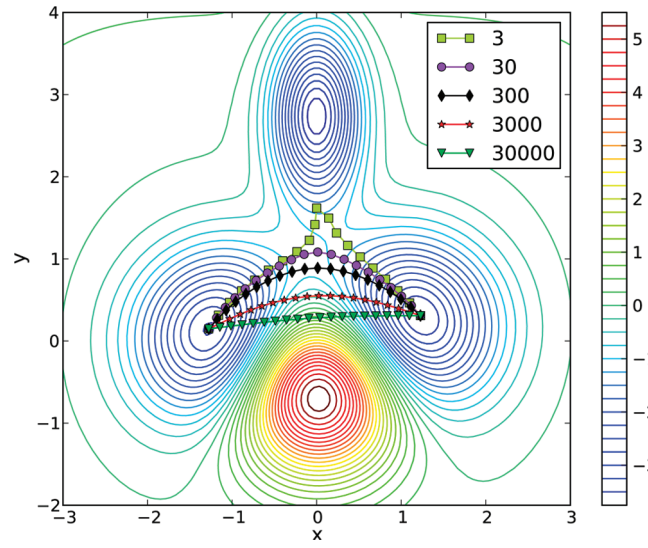


Figure 4. Maximum flux transition path obtained using the semi-implicit simplified string method. Here, we used the same initial path and the same stopping criterion for convergence as for 3. The MFTPs are generated using 20 images at 3, 30, 300, 3000, and 30 000 K (which roughly correspond to $\beta^{-1} = 0.006$, 0.06, 0.6, 6, and 60 kcal/mol.) The contour lines are separated by 0.25 kcal/mol.

calculated using 80 images ($J = 79$). The MFTP avoids the cusp problem.

The MFTP generates different paths at different temperatures. Figure 4 shows MFTPs at 3, 30, 300, 3000, and 30 000 K. It is clear that the MFTP is close to the MEP at a low temperature (3 K) and is close to a straight line at a high temperature (30 000 K), which is what we expect.

An FTS method path is expected to be similar to an MFEP for this example.

5.2. Φ, Ψ for Alanine Dipeptide in a Vacuum. For comparison with the MFEP, we study alanine dipeptide at 300 K in a vacuum.[9] We compare the MFEP and the MFTP with two dihedral angles ϕ and ψ as collective variables. All simulations were performed using the CHARMM simulation program[25,35] and the full-atom representation of the molecule in the CHARMM force field.[36,37] Langevin dynamics with a friction coefficient of 10.0 ps^{-1} and a time step of 1.0 fs was used. For the calculation of ∇F and D, harmonic potentials as in eq 17 were added involving the dihedral angles ϕ and ψ with force constant $k = 2000$ kcal/(mol rad^2) ($k = 1/\beta\varepsilon^2$ so $\varepsilon = 1°$).

The initial path in collective variable space is a straight line between two points in (ϕ, ψ) space. The path is discretized into $J + 1$ images. The configuration of alanine dipeptide at each image along the initial path is built using the IC module in CHARMM with dihedral angles fixed at the interpolated values. Then 1000 steps of minimization and 50 000 steps of heating are followed before the iteration starts. Each iteration of the path involves 50 000 steps of equilibration and 500 000 steps of sampling (per image). The configuration at the final step of sampling in the previous iteration is used as the initial configuration for the equilibration in the next iteration.

We begin by comparing the MFTP and MFEP from $C_{7\text{eq}}$ to $C_{7\text{ax}}$. The MFEP is calculated using the simplified string

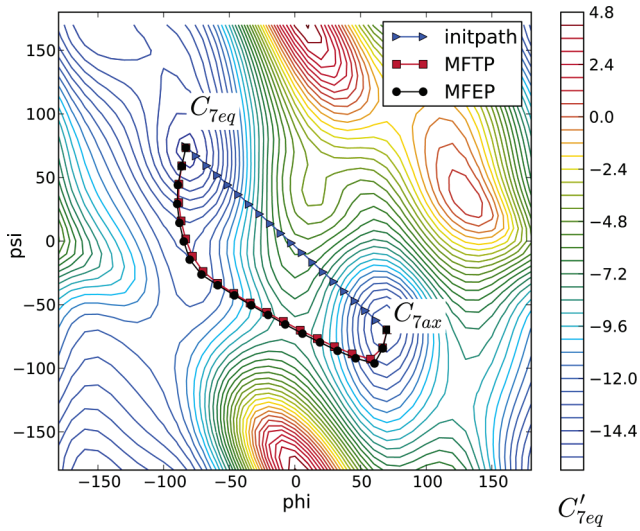


Figure 5. Maximum flux transition path and minimum free energy path from C_{7eq} to C_{7ax} for alanine dipeptide in a vacuum at 300 K. Triangles are images of the initial path, rectangles are the images of the maximum flux transition path, and circles are the images of the minimum free energy path. The contours are those for the zero-temperature free energy (adiabatic energy). The contour lines are separated by 0.6 kcal/mol.

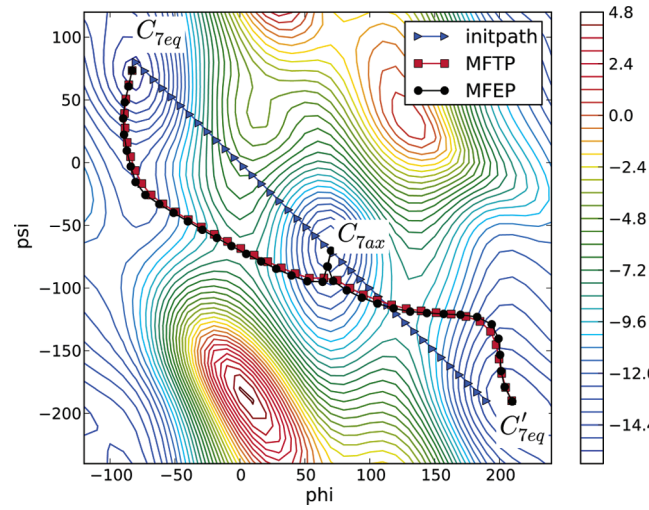


Figure 6. Maximum flux transition path and minimum free energy path for alanine dipeptide from C_{7eq} to C'_{7eq} passing through C_{7ax} in a vacuum at 300 K. The figure is generated using 40 images. Triangles are the images for the initial path, rectangles are the images of the maximum flux transition path, and circles are the images of the minimum free energy path. The contours are those for the zero-temperature free energy. The contour lines are separated by 0.6 kcal/mol.

method with linear interpolation between images and equal arc length normalization. The MFTP is calculated using the semi-implicit simplified string method. In Figure 5, the initial path is the straight line between (−83.2°, 74.5°) and (70°, −70°), which were determined as C_{7eq} and C_{7ax} in ref 9. The path is discretized into 20 images. The time step $\tau^2 = 0.16$ in CHARMM time units squared, or $\tau^2 = (19.56\ \text{fs})^2$. The statistical error estimated by block averaging using 32 blocks is ±0.00577°. The iteration is stopped if $d < 0.18°$. (The tolerance value should be chosen properly since the statistical error will eventually dominate the other errors so that d fluctuates about a positive number.) It takes 34 and 31 iterations to converge for the MFTP and MFEP, respectively. The computational cost for two methods is comparable. The path calculated for this problem by the FTS method using the CHARMM force field is given in Figure 5 of ref 19.

Next, we compare the MFTP and MFEP from C_{7eq} to C'_{7eq}. In particular, we calculate the transition path $C_{7eq}-C_{7ax}-C'_{7eq}$, in which C_{7ax} serves as an intermediate metastable state. The initial path is taken to be the straight line between (−80°, 80°) and (190°, −190°). Figure 6 shows the MFTP and MFEP generated using 40 images. The time step $\tau^2 = 0.16$ in CHARMM time units squared. The iteration is stopped if $d < 0.18°$. It takes 35 and 44 iterations for the MFTP and MFEP to converge, respectively. It is evident that the MFTP is more efficient than the MFEP in this case.

5.3. Φ, Ψ for Alanine Dipeptide in Solution. We also test our method for alanine dipeptide solvated in explicit water. Again, the backbone dihedrals ϕ and ψ are used as collective variables to describe the transition. The initial paths are straight lines connecting two points among (−77°, 138°), (55°, 48°), (60°, −72°), and (−77°, −39°) in (ϕ, ψ) space.

For preparing the simulation, each starting structure for alanine dipeptide with constrained ϕ and ψ angles is solvated in a (20 × 18 × 15) Å^3 box with 191 TIP3[38] water molecules

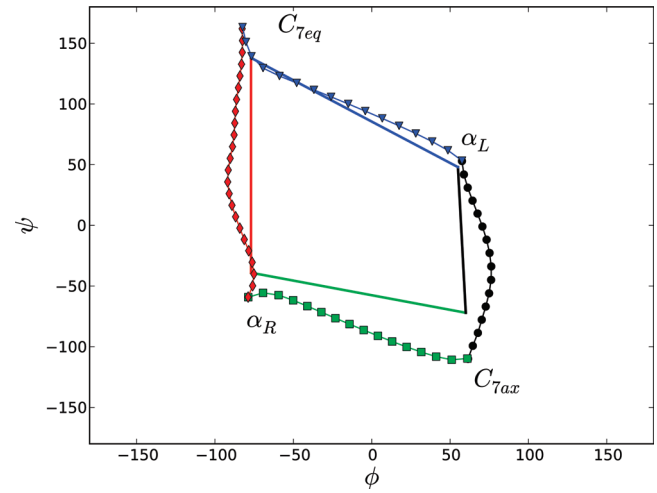


Figure 7. Maximum flux transition paths for alanine dipeptide in solution. The transition paths are calculated by the semi-implicit simplified string method with the nearby straight lines as initial paths.

and equilibrated for 50 000 ps. The molecular dynamics are carried out with the CHARMM program under the CHARMM22 force field. Periodic boundary conditions are used, and the electrostatic interactions are treated with the particle-mesh Ewald method.[39] The system is simulated at a constant pressure of 1.0 atm and a constant temperature 300 K with the algorithm based on Hoover's methods. We use a 1 fs time step with the SHAKE[40] algorithm to keep all bonds involving hydrogen atoms at fixed lengths.

In Figure 7, four MFTPs are calculated using the semi-implicit simplified string method. Each iteration involves 50 000 steps of equilibration and 500 000 steps of sampling. The transition paths are the result of 50 iterations. The path $C_{7eq}-\alpha_R$ calculated for this problem by the FTP method using

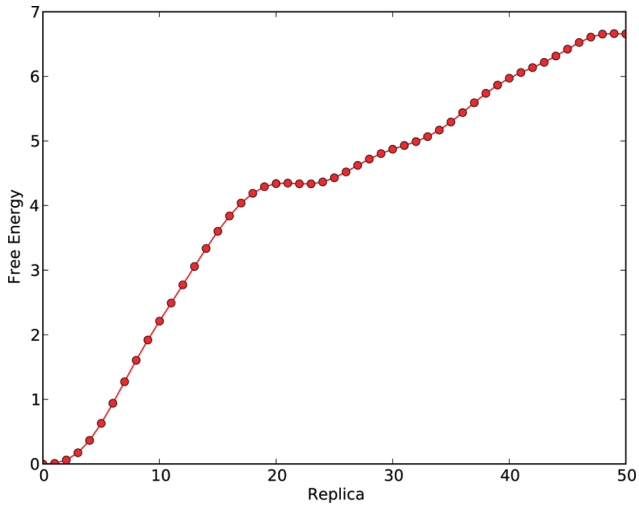


Figure 8. Free energy profile along the transition path.

the CHARMM force field is given in Figure 12 of ref 9. The MFTP is similar to the FTS method path.

5.4. α Carbon Coordinates for Alanine Decapeptide in a Vacuum. In this example, conformational change of alanine decapeptide from an α-helical structure to a π helix is studied. Alanine decapeptide is a small molecule with nine residues and 102 atoms, terminated with methyl groups.[41] The transition from one state to another involves breaking and forming hydrogen bonds. In particular, the α helix forms hydrogen bonds between residues i and $i + 4$, and the π helix forms hydrogen bonds between residues i and $i + 5$.

In defining the reaction, the backbone atoms are more important than side chain atoms.[41] In this study, we use the Cartesian coordinates of the nine α carbon atoms as collective variables. The MFTP is calculated using the semi-implicit simplified string method, where the time step $\tau^2 = 0.2$ in CHARMM time units squared and the iteration is stopped if $d < 0.004$ Å. The initial path is a straight line connecting coordinates of the α helix and π helix in collective variable space, and it is discretized into 51 replicas. The molecular simulations are done using CHARMM program with CHARMM22 all-atom force field. The harmonic force constant is 200 kcal/mol/Å^2 ($\varepsilon = 0.0547$ Å). At each iteration, samples are taken from a 500 ps simulation following 50 ps equilibration. It takes 234 iterations before the path converges. The more stringent stopping criterion requires more computational time, but it gives a better quality path.

Figure 8 shows the free energy profile along the MFTP. The free energy is calculated using eq 13. Specifically, $F_0 = 0$, $F_j = F_{j-1} + {}^1/_2(\nabla F_{j-1} + \nabla F_j)(Z_j - Z_{j-1})$, $j = 1, 2, ..., J$. There might be transition states at replicas 20 and 49 and a marginal intermediate metastable state at replica 23. Computation using backbone ϕ and ψ angles as collective variables yields a similar free energy plot. The free energy plot along the transition path is similar to one of the three MEPs in ref 41. For the other two paths, the transition takes place by passing through a higher energy barrier than the energy difference between the α helix and π helix. It appears that the transition from the α helix to π helix is not a barrier-crossing event.

Acknowledgment. This material is based upon work supported by grant R01GM083605 from the National Institute of General Medical Sciences, award A5286056128 from the University of Minnesota, and by a 2007 Purdue Research Foundation Special Incentive Research Grant. We would like to thank Carol Post for the collaboration that nurtured this work. Also, thanks to He Huang for an initial implementation of the string method and an early demonstration of cusps for alanine dipeptide, and to Voichita Dadarlat for Figure 2. Additionally, thanks to Eric Vanden-Eijnden for helpful information about transition path methods and theory, and for suggestions that improved the original manuscript; thanks to Jhih-Wei Chu for supplying coordinates of α and π helices; and thanks to reviewers for suggestions that spurred extensive improvements. Finally, thanks to the Center for Biological Physics at Arizona State University and the Institute for Mathematics and Its Applications at the University of Minnesota for providing environments that facilitated this work.

Appendix

A. Correction Factor for Hypersurface Area. The flow rate through a tiny disk at ζ is $j(\zeta)$ times the area of the disk. Let $D(M; \varepsilon)$ be a disk of radius ε in the $M(\cdot)$ metric on a quasi-committor isosurface at some point ζ. Then, $D(M; \varepsilon)$ is the intersection of the ball $\{\zeta' \mid (\zeta' - \zeta)^T M(\zeta)(\zeta' - \zeta) \le \varepsilon^2\}$ with the surface $\{\zeta' | \hat{n} \cdot (\zeta' - \zeta) \approx 0\}$, where $\hat{n} = \nabla\hat{q}/|\nabla\hat{q}|$. Let $1_\varepsilon(u) = 1$, $0 \le u \le \varepsilon$, and $1_\varepsilon(u) = 0$, $u > \varepsilon$. Then,

$$\text{area} = \int_{D(M;\varepsilon)} dS_\zeta = \int 1_\varepsilon((\zeta' - \zeta)^T M(\zeta)(\zeta' - \zeta)) \times \delta(\hat{n} \cdot (\zeta' - \zeta))\, d\zeta'$$

After a change of variables $\zeta' = \zeta + M(\zeta)^{-1/2}\zeta''$, we have

$$\begin{aligned}\text{area} &= \int 1_\varepsilon((\zeta'')^T\zeta'')\delta(\hat{n} \cdot M^{-1/2}\zeta'')\, \det(M^{-1/2})\, d\zeta'' \\ &= |M^{-1/2}\hat{n}|^{-1}(\det M)^{-1/2} \int 1_\varepsilon((\zeta'')^T\zeta'')\delta(\zeta'' \cdot \hat{t})\, d\zeta'' \\ &= |M^{-1/2}\hat{n}|^{-1}(\det M)^{-1/2} \int_{D(I;\varepsilon)} d\xi\end{aligned}$$

where $\hat{t} = M^{-1/2}\hat{n}/|M^{-1/2}\hat{n}|$. Accordingly, we obtain a correction factor

$$|M^{-1/2}\hat{n}|^{-1}(\det M)^{-1/2} = |\nabla\hat{q}|(\det M)^{-1/2}/|M^{-1/2}\nabla\hat{q}|$$

B. Derivation of Brownian Dynamics Approximation. The quasi-committor is related to a full phase-space committor q^* defined in ref 11 (section 6.2) as follows:

$$q^*(x, p) = \Pr(X(t) \text{ reaches B before A} | X(0) = x, P(0) = p)$$

Note that $q^*(x, p) = 0$ or 1, because the dynamical equation is deterministic. By definition, the quasi-committor $\hat{q}(\zeta) = \langle q^*(x, p)\rangle_\zeta$.

It is not difficult to show that $\hat{q}(\xi(x))$ approximates $q^*(x, p)$ in the sense that it minimizes $\langle |q(\xi(x)) - q^*(x, p)|^2\rangle$ over all $q(\zeta)$. However, this is not useful for determining $\hat{q}(\zeta)$ because $q^*(x, p)$ is too costly to compute. On the other hand, it is

possible to find a best approximation to $q^*(x, p)$ in another sense. Because q^* is constant on a trajectory, we have

$$0 = \frac{\mathrm{d}}{\mathrm{d}t} q^*(X(t), P(t)) = (Lq^*)(X(t), P(t)) \text{ where}$$

$$L = (M^{-1}p) \cdot \nabla_x - U_x \cdot \nabla_p$$

Consequently, q^* satisfies the stationary Liouville equation

$$Lq^* = 0, \quad q^* = 0 \text{ on A}, \quad q^* = 1 \text{ on B}$$

Since we do know that $Lq^* = 0$, we seek instead an approximation q that minimizes $I(q) = \langle |L(q(\xi(x)) - q^*(x, p))|^2 \rangle$, a standard tactic in numerical analysis. As shown in section III.B of ref 9, this simplifies to

$$I(q) = \frac{1}{\beta} \langle |M^{-1/2} \nabla_x q(\xi(x))|^2 \rangle$$

which is to be as small as possible. A low value for $I(q)$ is attained by having $q(\zeta)$ increase monotonically from the value 0 on A_ξ to the value 1 on B_ξ, which is consistent with the prescription given earlier that $\xi(x)$ be chosen so that the quasi-committor has no local minima or maxima outside of A_ξ and B_ξ.

The functional $I(q)$ can be expressed in terms of collective variables ζ as given by eq 2 and shown in eq 15 of ref 9.

C. Derivation of Lasting Hitting-Point Distribution. The proof of Proposition 5 in ref 8 (p 158) analyzes the flux of reactive trajectories. The flux $J(\zeta)$ gives the rate at which such trajectories cross an arbitrary surface Σ that divides collective variable space into two parts, one containing A_ξ and the other containing B_ξ, via the integral $\int_\Sigma J(\zeta) \cdot \hat{n}(\zeta)\, \mathrm{d}S_\zeta$ where $\hat{n}(\zeta)$ points to the side containing B_ξ. The proof actually examines not all crossings but only those occurring within a vanishingly small time interval before the last crossing—see eq 50 of ref 8. Therefore, it considers the net flux only in this limiting sense. As the length of the time interval $\tau \to 0$, the positions of these crossings all converge to the position of the last crossing. So, indeed, one gets the flux of the last hitting point from Proposition 5 of ref 8. The result given in ref 8 (eq 39), as well as in refs 10 (eq 6, A12) and 11 (eq 62), is that the last hitting-point flux for reactive trajectories is $J(\zeta) = \rho_\xi(\zeta) D(\zeta) \nabla q(\zeta)$. (Proposition 4 of ref 8 does not apply to the infinitely damped case of Langevin dynamics.) The expression for $J(\zeta)$ also gives the net flux of reactive trajectories, see eq 32 of ref 11. Also, the formula for $j(\zeta)$ in section 3.2 agrees in the special case $D = I$ with that for the *first* hitting point distribution given in ref 42 (Appendix B). Last and first are the same for reversible dynamics like Brownian dynamics. Finally, there is an example in Metzner et al., section III.C, where it is suggested to use $\hat{n} \cdot J$.

D. Derivation of the Maximum Flux Condition. We have from eqs 1 and 6 that the normal flux is

$$j_c(\zeta) = \mathrm{const}_\xi \exp(-\beta F(\zeta))\, \bar{q}_s(\sigma(\zeta)) (\nabla\sigma(\zeta)^T \times D(\zeta) \nabla\sigma(\zeta) \det D(\zeta))^{1/2} \quad (18)$$

where $\sigma(\zeta)$ is defined implicitly by $q(\zeta) = \bar{q}(\sigma(\zeta))$. And for each point $Z(s)$ on the desired path, the condition to be satisfied, eq 7, is $\nabla j_c(Z(s)) \| n(s)$. Furthermore, the assumption, eq 8, that isocommittors are planar implies

$$n(\sigma(\zeta)) \cdot (\zeta - Z(\sigma(\zeta))) = 0 \quad (19)$$

Differentiating eq 19 wrt ζ, we get

$$(n_s(\sigma) \cdot (\zeta - Z(\sigma)) - n(\sigma) \cdot Z_s(\sigma)) \nabla\sigma + n(\sigma) = 0$$

where the argument ζ of σ has been omitted, whence

$$\nabla\sigma = (n(\sigma) \cdot Z_s(\sigma) - n_s(\sigma) \cdot (\zeta - Z(\sigma)))^{-1} n(\sigma) \quad (20)$$

Substituting eq 20 into eq 18, the normal flux becomes

$$j_c(\zeta) = \varphi(\sigma(\zeta), \zeta)$$

where

$$\varphi(s, \zeta) = \mathrm{const}_\xi \exp(-\beta F^+(\zeta))\, \bar{q}_s(s) (n(s) \cdot Z_s(s) - n_s(s) \cdot (\zeta - Z(s))^{-1} c(\zeta, D(\zeta) n(s))$$

Note that

$$\frac{\nabla_\zeta \varphi}{\varphi} = -\beta \nabla F^+ + \frac{\nabla_\zeta c(\zeta, D(\zeta) n)}{c(\zeta, D(\zeta) n)} + \frac{n_s}{n \cdot Z_s - n_s \cdot (\zeta - Z)}$$

and

$$\left.\frac{\nabla_\zeta \varphi}{\varphi}\right|_{\zeta=Z} = -\beta \nabla F^+(Z) + \left.\frac{\nabla_\zeta c(\zeta, D(\zeta) n)}{c(\zeta, D(\zeta) n)}\right|_{\zeta=Z} + \frac{n_s}{n^T Z_s}$$

Thus, we have

$$\frac{\nabla j}{j} = \frac{(\nabla_\zeta \varphi)(\sigma(\zeta), \zeta)}{\varphi(\sigma(\zeta), \zeta)} + \frac{\varphi_s(\sigma(\zeta), \zeta)}{\varphi(\sigma(\zeta), \zeta)} \nabla\sigma(\zeta)$$

and

$$\left.\frac{\nabla j}{j}\right|_{\zeta=Z} = -\beta \nabla F^+(Z) + \left.\frac{\nabla_\zeta c(\zeta, D(\zeta) n)}{c(\zeta, D(\zeta) n)}\right|_{\zeta=Z} + \frac{n_s}{n^T Z_s} + \frac{\varphi_s(s, Z) n}{\varphi(s, Z) n^T Z_s}$$

Hence, the maximum flux condition is that

$$-\beta \nabla F^+(Z) + \left.\frac{\nabla_\zeta c(\zeta, D(\zeta) n)}{c(\zeta, D(\zeta) n)}\right|_{\zeta=Z} + \frac{n_s}{n^T Z_s} \| n$$

If we approximate $D(\zeta)$ by $\bar{D}(\sigma(\zeta))$, where

$$\bar{D}(s) \stackrel{\mathrm{def}}{=} D(Z(s))$$

then

$$\varphi(s, \zeta) = \mathrm{const}_\xi (\det \bar{D}(s)\, n(s)^T \bar{D}(s)\, n(s))^{1/2} \times \exp(-\beta F(\zeta))\, \bar{q}_s(s) (n(s) \cdot Z_s(s) - n_s(s) \cdot (\zeta - Z(s))^{-1}$$

and

$$\left.\frac{\nabla_\zeta \varphi}{\varphi}\right|_{\zeta=z} = -\beta \nabla F(Z) + \frac{n_s}{n^T Z_s}$$

Thus, the maximum flux condition simplifies to

$$-\beta \nabla F(Z) + \frac{n_s}{n^T Z_s} \mid\mid n$$

References

(1) Zhang, J.; Yang, P. L.; Gray, N. S. *Nat. Rev. Cancer* **2009**, *9*, 28–39.

(2) Gan, W.; Yang, S.; Roux, B. *Biophys. J.* **2009**, *97*, L8–L10.

(3) Ozkirimli, E.; Post, C. B. *Protein Sci.* **2006**, *15*, 1051–1062.

(4) Yang, S.; Banavali, N. K.; Roux, B. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 3776–3781.

(5) Ozkirimli, E.; Yadav, S. S.; Miller, T. W.; Post, C. B. *Protein Sci.* **2008**, *17*, 1871–1880.

(6) Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L. *Annu. Rev. Phys. Chem.* **2002**, *53*, 291–318.

(7) Singhal, N.; Pande, V. S. *J. Chem. Phys.* **2005**, *123*, 204909.

(8) E, W.; Vanden-Eijnden, E. *J. Stat. Phys.* **2006**, *123*, 503–523.

(9) Maragliano, L.; Fischer, A.; Vanden-Eijnden, E.; Ciccotti, G. *J. Chem. Phys.* **2006**, *125*, 024106.

(10) Metzner, P.; Schtte, C.; Vanden-Eijnden, E. *J. Chem. Phys.* **2006**, *125*, 084110.

(11) Vanden-Eijnden, E. Transition Path Theory. In *Computer Simulations in Condensed Matter: From Materials to Chemical Biology.* Vol. 2; Ferrario, M.; Ciccotti, G.; Binder, K., Eds.; Springer: New York, 2006, Vol. 704, pp 453−493.

(12) Pan, A. C.; Sezer, D.; Roux, B. *J. Phys. Chem. B* **2008**, *112*, 3432–3440.

(13) Berkowitz, M.; Morgan, J. D.; McCammon, J. A.; Northrup, S. H. *J. Chem. Phys.* **1983**, *79*, 5563–5565.

(14) Huo, S.; Straub, J. E. *J. Chem. Phys.* **1997**, *107*, 5000–5006.

(15) Crehuet, R.; Field, M. J. *J. Chem. Phys.* **2003**, *118*, 9563–9571.

(16) Elber, R.; Shalloway, D. *J. Chem. Phys.* **2000**, *112*, 5539–5545.

(17) Park, S.; Sener, M. K.; Lu, D.; Schulten, K. *J. Chem. Phys.* **2003**, *119*, 1313–1319.

(18) Jiménez, A.; Crehuet, R. *Theor. Chem. Acc.* **2007**, *118*, 769–775.

(19) Ren, W.; Vanden-Eijnden, E.; Maragakis, P.; E, W. *J. Chem. Phys.* **2005**, *123*, 134109.

(20) Vanden-Eijnden, E.; Venturoli, M. *J. Chem. Phys.* **2009**, *130*, 194103.

(21) E, W.; Ren, W.; Vanden-Eijnden, E. *J. Chem. Phys.* **2007**, *126*, 164103.

(22) Vanden-Eijnden, E.; Heymann, M. *J. Chem. Phys.* **2008**, *128*, 061103.

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

(24) Zhao, R. "MFTP code", http://bionum.cs.purdue.edu/mftp (accessed Dec 22, 2009).

(25) Brooks, B. R. *et al. J. Comput. Phys.* **2009**, *30*, 1545–1614.

(26) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. *J. Comput. Phys.* **2005**, *26*, 1781–1802.

(27) Ma, A.; Dinner, A. R. *J. Phys. Chem. B* **2005**, *109*, 6769–6779.

(28) Onsager, L. *Phys. Rev.* **1938**, *54*, 554–557.

(29) Bolhuis, P. G.; Dellago, C.; Chandler, D. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5877–5882.

(30) E, W.; Vanden-Eijnden, E. Metastability, conformation dynamics, and transition pathways in complex systems. In *Multiscale Modelling And Simulation,*; Attinger, S.; Koumoutsakos, P. , Eds.; Springer-Verlag: New York, 2004; Vol. 39, pp 35−68.

(31) Schlitter, J.; Engels, M.; Krger, P. *J. Mol. Graphics* **1994**, *12*, 84–89.

(32) Huang, H.; Ozkirimli, E.; Post, C. B. *J. Chem. Theory Comput.* **2009**, *5*, 1304–1314.

(33) Lei, M.; Zavodszky, M. I.; Kuhn, L. A.; Thorpe, M. F. *J. Comput. Phys.* **2004**, *25*, 1133–1148.

(34) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*; Academic Press: CA, 2002.

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

(36) MacKerell, A. D., Jr.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.

(37) Mackerell, Jr, A. D.; Feig, M.; Brooks, III, C. L. *J. Comput. Phys.* **2004**, *25*, 1400–1415.

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

(39) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577–8593.

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

(41) Chu, J.; Trout, B. L.; Brooks, B. R. *J. Chem. Phys.* **2003**, *119*, 12708–12717.

(42) Vanden-Eijnden, E.; Venturoli, M.; Ciccotti, G.; Elber, R. *J. Chem. Phys.* **2008**, *129*, 174102.

CT900689M

JCTC Journal of Chemical Theory and Computation

Atomistic details of the Catalytic Mechanism of Fe(III)−Zn(II) Purple Acid Phosphatase

Marta E. Alberto,[†] Tiziana Marino,[†] Maria J. Ramos,[‡] and Nino Russo*,[†]

Dipartimento di Chimica, Università della Calabria, Via P. Bucci, cubo 14c, 87036 Arcavacata di Rende (CS), Centro di Calcolo ad Alte Prestazioni per Elaborazioni Parallele e Distribuite−Centro d'Eccellenza MIUR, Italy and REQUIMTE, Departamento de Química, Faculdade de Ciências, Universidade do Porto, Rua do Campo Alegre, 687, 4169-007 Porto, Portugal

Received April 08, 2010

Abstract: In the present work, we performed a theoretical investigation of the reaction mechanism of the Fe(III)−Zn(II) purple acid phosphatase from red kidney beans (rkbPAP), using the hybrid density functional theory and employing different exchange−correlation potentials. Characterization of the transition states and intermediates involved and the potential energy profiles for the reaction in different environments (gas phase, protein environment, and water) are reported. Our results show that the Fe(III)−Zn(II)PAP catalyzes the hydrolysis of methylphosphate via direct attack by a bridging metals-coordinated hydroxide leading to the cleavage of the ester bond. From our study emerges that the rate-limiting step of the reaction is the nucleophilic attack followed by the less energetically demanding release of the leaving group. Furthermore, we provide insights into some important points of contention concerning the precatalytic complex and the substrate coordination mode into the active site prior to hydrolysis. In particular: (i) Two models of enzyme−substrate with different orientations of the substrate into the active site were tested to evaluate the possible roles played by the conserved histidine residues (His 202 and His 296); (ii) Different protonation states of the substrate were taken into account in order to reproduce different pH values and to verify its influence on the catalytic efficiency and on the substrate binding mode; (iii) The metals role in each step of the catalytic mechanism was elucidated. We were also able to ascertain that the activation of the leaving group by the protonated His 296 is decisive to reach an optimal catalytic efficiency, while the bond scission without activation requires higher energy to occur.

1. Introduction

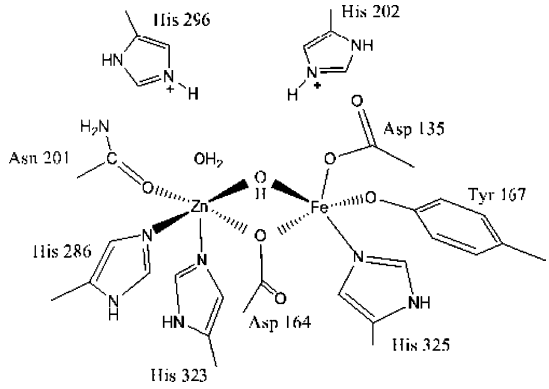
Purple acid phosphatases (PAPs)[1−3] belong to binuclear metallohydrolases, an interesting family of enzymes that have received in the last years considerable attention.[4−11] Binuclear sites in enzymes appear to have several potentially useful properties, not found in mononuclear centers, which play essential roles in catalysis. Members of this class have been recognized as potential targets for the development of chemotherapeutics and for drug design against a wide variety of human disorders[12−20] and represent also promising candidates in bioremediation.[21]

PAPs are the only binuclear metallohydrolases that use a heterovalent active site Fe(III)−M(II) to catalyze the hydrolysis of monophosphates at acidic to neutral pH.[1−3,22−26] The identity of the divalent metal ion varies with the source of the enzyme. Mammalian PAPs contain an antiferromagnetically coupled binuclear iron center Fe(III)−Fe(II) in the active site,[27−33] while plant PAPs most typically have Fe(III)−Zn(II) centers,[26,34−36] but an interesting example of an enzymatically active binuclear Fe(III)−Mn(II) center was found in the case of PAP from the sweet potato.[24,25,37] In spite of the scarce similarity in their primary sequences, the

* Corresponding author. E-mail: nrusso@unical.it.
[†] Università della Calabria.
[‡] Universidade do Porto.

10.1021/ct100187c © 2010 American Chemical Society
Published on Web 07/01/2010

Scheme 1. Red Kidney Bean PAP Active Site



active site structure and the residues coordinating the metal ions in the active site are identical in mammalian and plant PAPs, displaying similar enzymatic and spectroscopic properties.[38] For example, the intense purple color that distinguishes PAPs from the other phosphatases is due to a charge-transfer transition (λ_{max} = 510−560 nm) in the active site from a conserved tyrosinate to the Fe(III) ion.[39−41] Nevertheless, while plant PAPs have a chromophoric center similar to that of their animal counterparts (λ_{max} = 550 nm), their activity cannot be regulated by reversible oxidation/reduction, as it indeed occurs for mammalian PAPs which are reversibly inactivated by the oxidation of the divalent iron ion, suggesting that PAP may be regulated by changes in redox potential within animal cells.[42,43] The identity of their active sites has been demonstrated in several studies by using metal ion replacement[44−52] and by spectroscopic and magnetochemical measurements,[53,54] giving evidence that they employ a similar working mechanism. Moreover other studies suggest that rkbPAP is a good model for the structure and mechanism of other acid phosphatases, such as the human one.[57] PAP from red kidney beans, the object of our investigation, was the first member of this family for which a crystal structure became available[55,56] and is by far the most explored so far. The metal atoms first-shell ligands include seven invariant aminoacid residues: three histidines, two aspartates, a tyrosine, and an asparagine (Scheme 1).[56] The presence of the water molecule terminally coordinated to the divalent metal ion was supported by ENDOR measurements[58] as well as electron density maps,[59,60] and EXAFS[61] studies sustained the presence of the bridging hydroxide, the most likely candidate as the nucleophile.[58] Although in many earlier studies the presence of a terminal Fe(III)-bound water molecule as a nucleophile candidate has been proposed,[62−64] ENDOR measurements of pig PAP and rkbPAP−SO_4 crystal structures indicate that there is no water ligand at the iron site, suggesting also that the trivalent metal ion is five coordinated.[58] Such evidence lead to the exclusion of a mechanism in which a terminal Fe(III)-bound hydroxide acts as nucleophile. Moreover, experimentally it was observed that the reduced nucleophilicity of the bridging hydroxide is compensated by an increased electrophilicity of the substrate when comparing the possibilities of: (i) a terminally coordinated nucleophile and terminally coordinated substrate, with (ii) a bridging nucleophile and bridging substrate.[58]In addition, the crystal structures of PAPs, determined in the presence of a phosphate group,[56] show a network of hydrogen bonds with uncoordinated and conserved histidine residues of the second shell of the ligands (His 202 and His 296 for rkbPAP), which, therefore, could play essential roles in catalysis. Although the importance of these conserved histidines has been demonstrated by site-directed mutagenesis studies,[14,65,66] it is still not clear the effective role of each one in the stabilization of the substrate in the catalytic cavity.

Despite the presence in the literature of structural spectroscopic and kinetic data, the individual steps of the PAPs-catalyzed hydrolysis reaction remain a matter of controversy. The general mechanism proposed in the literature for PAP-catalyzed reactions involves a nucleophilic attack performed by the bridging hydroxide to the phosphorus center leading to hydrolytic cleavage of the phosphate−ester bond.[4,67] Nevertheless, several crucial aspects are still uncertain, such as the binding mode of the substrate in the active site prior to the hydrolysis, the effective role of each metal atom, and also the function of the two histidine residues located in the second-shell ligands (His 202 and His 296).[4,5,7,37,46,48,58,64,67]

In order to better characterize the several steps of the catalytic mechanism of the Fe(III)−Zn(II)−PAP (from rkbPAP), for the first time a detailed theoretical exploration of the hydrolysis process is reported in this work, providing characterization of the transition states and intermediates involved and presenting the potential energy profiles for the reaction in different environments (gas phase, protein environment, and water). Moreover our computations can give insights into some interesting aspects. In particular: (i) Two models of enzyme−substrate (ES) with different orientations of the substrate into the active site were tested to evaluate the possible roles played by the conserved histidine residues (His 202 and His 296); (ii) Different protonation states of the substrate were taken into account in order to reproduce different pH values and to verify its influence on the catalytic efficiency and on the substrate binding mode; (iii) The metals role in each step of the catalytic mechanism was elucidated.

2. Computational Details and Models

The theoretical investigation of the catalytic mechanism of the Fe(III)−Zn(II)−PAP was performed at density functional theory (DFT) level by means of GAUSSIAN 03[68] suite of programs. Geometry optimizations in the gas phase were carried out using the hybrid B3LYP functional, composed by Becke's[69] three-parameter hybrid exchange functional (B3), and the correlation functional of Lee, Yang and Parr (LYP)[70] using 6-31G(d,p) basis sets for all atoms except the iron and Zn ions, which were described by the quasi-relativistic Stuttgart−Dresden pseudopotentials.[71] In order to confirm proper convergence to equilibrium and transition-state geometries, vibrational frequency analysis was done based on analytical second derivatives of the Hamiltonian at the same level of theory. Solvent effects were introduced in the framework of the self-consistent reaction field conductor-like polarizable continuum model (SCRF-CPCM),[72,73] using two dielectric constants to simulate water and reproduce the protein environment. For the latter, an empirical

value of $\varepsilon = 4$ accounts for the average effect of both the protein and the water medium surrounding the protein. For the water, $\varepsilon = 80$ was used. In the CPCM method, the continuum is modeled as a conductor, instead of a dielectric. This simplifies the electrostatic computations, and corrections are made a posteriori for dielectric behavior. The UA0 radii were used to build the cavity.[74] In order to obtain more accurate energies in the gas phase, in the protein environment, and in water, single points calculations were performed on the optimized geometries using the larger basis set 6-311+G(2d,2p) and testing several exchange−correlation functionals (B3LYP,[69,70] PBE0PBE,[75] BB1K,[76] MPWB1K,[77] and MW1B95).[77] Our benchmark reported in Section 3.3, shows that the meta-hybrid functional (BB1K, MPWB1K, MW1B95) reproduces with good accuracy experimental values found on similar reactions catalyzed by PAPs enzymes.[24,78,79] In addition, many studies have demonstrated that the hybrid-meta generalized gradient approximation (GGA) methods give more reliable results in the reproduction of the reaction kinetics showing great ability in the calculation of barrier heights with high accuracy, also compared to higher-level post-Hartree−Fock methods, supporting our evidence.[80,81] The energy values discussed in the text are those obtained with the hybrid-meta MPWB1K functional, previously employed successfully also in other enzymatic studies.[80−82] The crystal structure of red kidney bean PAP, determined in the presence of phosphate acting as an inhibitor (PDB code 4kbp, chain A),[56] was used to devise a model of the active site of the enzyme. The latter contains the metal ions Fe(III) and Zn(II) and the seven amino acids of their first-shell ligands. In particular, the Fe(III) ion is coordinated to Asp 135, Tyr 167, His 325, and Asp 164 that bridges both of the metals, while the Asn 201, His 286, and His 323 represent the divalent metal ion Zn(II) ligands. In order to study their roles in the catalytic cycle, two histidine residues (His 296 and His 202) of the second shell of the ligands were also included in the model. No water molecule was introduced to complete the coordination environment of the Zn ion because the formation of the precatalytic complex has been suggested to cause the expulsion of this labile terminal water ligand.[67] Only the side chains of the residues were kept in our model. Moreover the carbon atoms where truncation was done were kept fixed to their X-ray crystal positions during the optimizations to avoid an unrealistic expansion of the cluster during the calculations. This procedure gives rise to smaller imaginary frequencies, but the latter does not significantly contribute to the zero point energies (ZPE) and can thus be tolerated. A bridging hydroxide as a nucleophile was included in the cluster model, in agreement with previous studies that support the presence of this group.[58,61] In the absence of crystallographic data for the ES complex, the latter was modeled in the active site as methyl-phosphate since this kind of enzyme catalyzes the hydrolysis of monophosphates. The resulting cluster contains 114 atoms. In order to establish the most stable spin multiplicity of the system, preliminary calculations on the ES complex with different values of spin multiplicity ($2S + 1 = 2$, 4, 6, and 8) were performed. The lowest energy was obtained with a value of 6, which arises from a high-spin configuration of the Fe(III) center, in agreement with electron paramagnetic resonance (EPR) measurements that support our result.[25] The doublet, quartet, and octet electronic states lie at 36.3, 23.8, and 98.3 kcal/mol with respect to the same structure obtained with sextet multiplicity, respectively. For the quartet spin state, the entire catalytic cycle was reproduced in order to consider eventual involvement of different spin states along the potential energy surface (PES). The results clearly show that no crossing occurs between the two energetic profiles. (See Figure S1 in the Supporting Information). Spin density distribution and molecular orbital pictures for all the stationary points along the PES are reported in Figures S2 and S3 of the Supporting Information, respectively. As expected, from these data, it is clear that the higher spin densities are localized on the iron center.

3. Results and Discussion

The steps involved in the whole process and determined through our calculations are sketched in Scheme 2. According to our results, a mechanism in which the metal-bridging hydroxide performs a nucleophilic attack on the phosphorus atom leading to the cleavage of the ester bond, activated by a proton shift from the histidine residue 296, is proposed. The potential energy profiles for the considered reaction cycle are depicted in Figure 5 of Section 3.3.

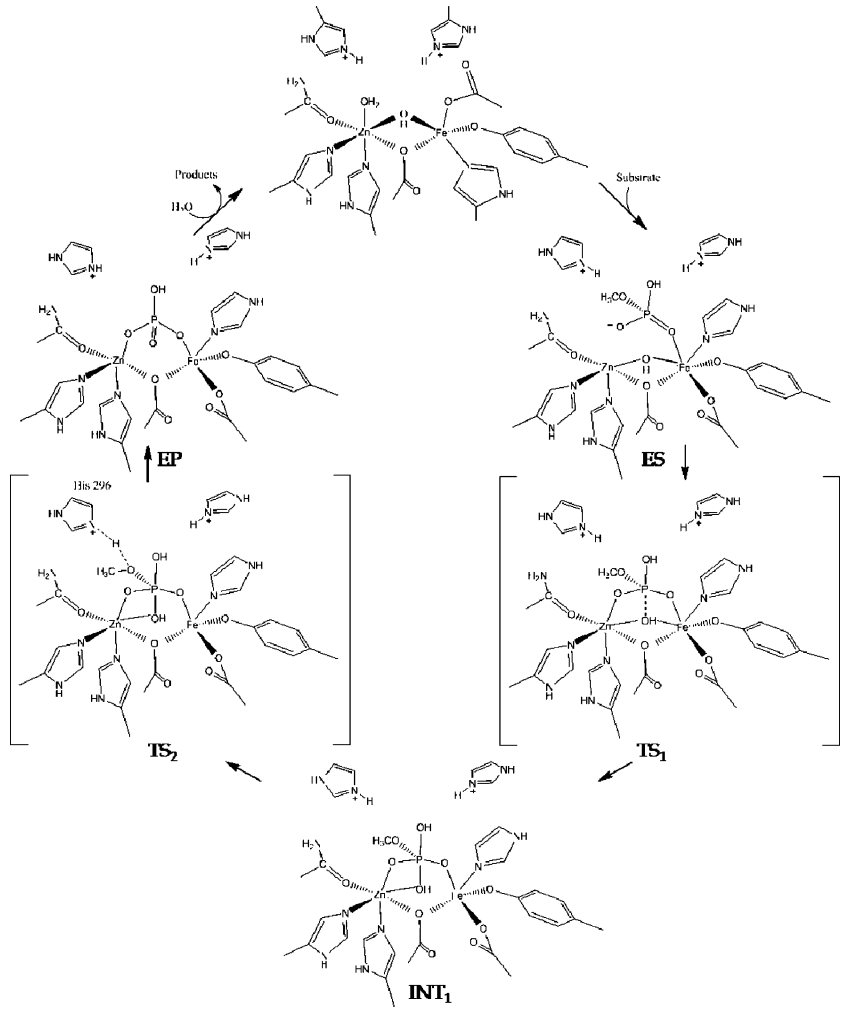
The possibility of a concerted SN_2 mechanism has also been considered in our investigation, but we found that two transition states are required to complete the catalytic process.

3.1. ES Complex. Since the X-ray structure of rkbPAP used to devise the cluster model of the active site was determined in the presence of an inhibitor, to obtain a valid initial structure for our exploration of the catalytic mechanism, we evaluated several crucial aspects in the construction of our model. First of all, the substrate's orientation into the active site, and then its protonation state and the coordination binding mode to the dimetallic center. In order to consider the right orientation of the substrate into the catalytic cavity, we took into account some experimental indications suggesting that the His 296 may be responsible for the proton shift to the leaving group during the catalytic cycle, while a second conserved residue His 202 should be involved only in the transition-state stabilization.[14,65,66] According to this evidence, the substrate was modeled in the catalytic cavity with the methoxy group oriented toward the His 296 (Scheme 3a). In addition, we tried to study a mechanism in which the His 202 worked as proton carrier to the leaving group, considering the other stereoisomer of the methyl phosphate (Scheme 3b), but all the efforts to locate the saddle points for the reaction with that model failed.

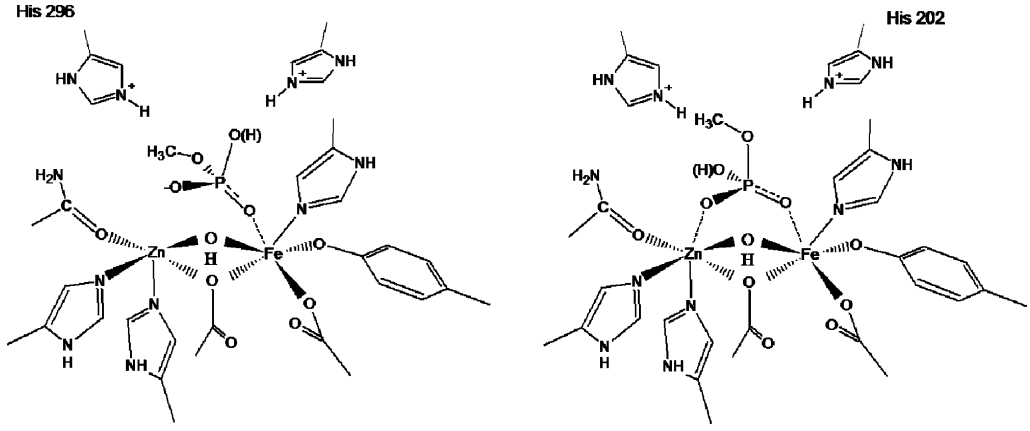
From geometric point of view, the results obtained using models (a) and (b) show an important difference. The substrate binds in a different manner the dimetallic center in the optimized structures inducing different electronic distributions, probably responsible of the different activity of the two stereoisomers. The optimized ES complex for model (b) is reported in Figure S4 of the Supporting Information.

Another point of conjecture is the mode of substrate binding. Although several studies of this kind of enzyme

Scheme 2. Proposed Mechanism for the Hydrolysis of Methyl-Phosphate Promoted by Red Kidney Bean PAP



Scheme 3. Considered Substrate Orientations into the Catalytic Cavity



exists, the exact coordination of the substrate to the dimetallic center it is still uncertain. It has been previously proposed that the precise substrate coordination may be pH dependent.[4] In order to better understand this aspect, we took into consideration that the Fe−Zn−PAP from red kidney beans display a maximum catalytic rate at pH $\approx$ 6.[78] At that value of pH (4.5 < pH < 6.5), our substrate is likely in its monoanionic form $(CH_3)HPO_4^-$. At pH slightly higher, the substrate could be present in its dianionic form. In order to provide insight into this important point of controversy, we tested two different protonation states of the substrate into the active site. We were able to verify that the activation barrier associated with the model containing $CH_3OPO_3^{2-}$ as the substrate is higher by more than 10 kcal/mol with respect to the other one relative to the monoanionic species, confirming that the reaction proceeds faster at a lower pH. Moreover, we prove that the pH influences the coordination mode. The $(CH_3)HPO_4^-$ substrate binds in a monodentate-like manner the bimetallic center, while the dianionic form of the substrate $CH_3OPO_3^{2-}$results in a bicoordinated ES

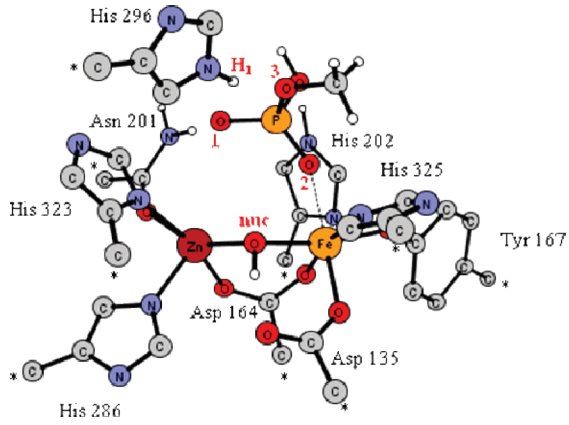


Figure 1. Optimized structure of the ES complex. Only the most significant hydrogen atoms are reported. Stars indicate the atoms kept fixed to their crystallographic positions during optimizations.

Table 1. Main Bond Lengths (Å) of the Optimized Stationary Points along the Reaction Path

	ES	TS_1	INT_1	TS_2	EP
Zn−O_1	3.47	2.10	2.03	2.03	1.97
Fe−O_2	2.30	2.11	2.00	2.01	2.07
Zn−Onuc	1.98	2.28	2.17	2.17	–
Fe−Onuc	2.01	2.24	3.36	3.43	–
P−Onuc	3.64	2.00	1.85	1.80	1.55
Fe−Zn	3.54	4.00	4.58	4.83	4.36
O_3−H_1	1.74	1.66	1.76	1.20	0.98

complex (see Figure S5 of the Supporting Information). As a result of this preliminary investigation, we report in Figure 1 the optimized structure of the ES complex. Main geometrical parameters of ES, together with those for the other stationary points encountered along the reaction path, are collected in Table 1.

As shown in Figure 1, the overall geometry of the metal centers predicted by our calculation is of a distorted square pyramidal geometry for the pentacoordinated Zn(II) cation, while the iron center adopts an octahedral geometry. The internuclear distance between the two metal ions is 3.54 Å. In the optimized complex, the substrate is stabilized in the active site by a network of hydrogen-bonding interactions with the Asn 201 and with conserved histidine residues located in the second coordination sphere (His 202 and His 296). The distance between the bridging hydroxide and the phosphorus atom of the incoming substrate is 3.64 Å. Thanks to this disposition in the active site, the substrate is in an ideal near-attack configuration.[83] During the optimization, the substrate orients the methoxy group toward the protonated His 296, resulting in a monodentate coordination to the Fe(III) ion. In such a complex, the leaving group is optimally oriented along the axis of the access channel to easily leave the active site after the hydrolysis process. In Figure 2 is shown the orientation of the methoxy group with respect to the whole chain A of the enzyme.

3.2. Reaction Mechanism. The nucleophilic attack of the bridging hydroxide (O_{nuc}) to the substrate phosphorus atom (TS_1) induces a cascade of geometric changes leading to the first enzyme−intermediate complex (INT_1). The subsequent proton shift from the His 296 to the leaving group (TS_2)

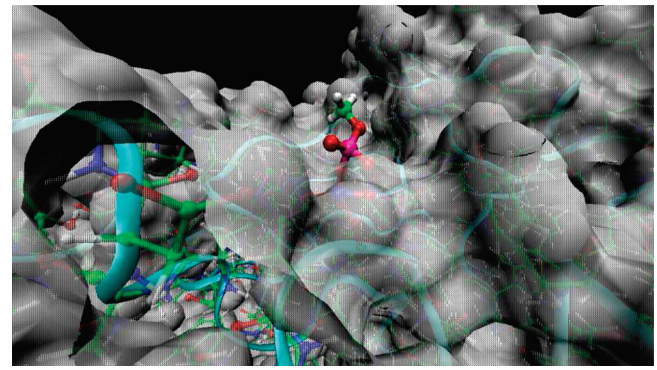


Figure 2. Orientation of the methyl phosphate group into the enzymatic channel. The B3LYP-optimized coordinates of the active site were introduced in the crystallographic data of the rkbPAP (pdb code 4kbp) by superimposition of the carbon atoms kept fixed to their X-ray crystal positions during the optimizations.

culminates in the enzyme−product complex (EP). The geometries of the optimized stationary points (see Table 1) are displayed in Figure 3.

The first transition state (TS_1) involves not only the shortening of the P−O_{nuc} distance, which in the ES varies from 3.64 to 2.00 Å but implies also important changes in the substrate coordination mode to the active site. The methyl-phosphate group binds now in a bidentate manner the dimetallic center, displaying distances of 2.10 and 2.11 Å from Zn and Fe, respectively. The bridging coordination geometry adopted led to an elongation of the Zn−Fe distance from 3.54 Å in the ES complex to 4.00 Å in the TS_1, consenting to increase the nucleophilicity of the metal-bound hydroxide by lengthening the distances Zn−O_{nuc} and Fe−O_{nuc} to 2.28 and 2.24 Å, respectively. In addition, modification in the substrate geometry was observed. The three oxygen atoms of the substrate (O_1, O_2, and O_3) occupy equatorial positions with respect to P, adopting a pentacoordinate geometry. The analysis of the vibrational mode clearly indicates the stretching of the incoming P−O_{nuc} bond, as a result of the nucleophilic attack of the bridging hydroxide to the phosphorus atom, and was confirmed to be a first-order saddle point with only one imaginary frequency (88 cm^{-1}). The hydrogen-bonding network found in the enzyme−substrate complex is still retained in the transition state. In particular the distance between the leaving group's oxygen atom and the hydrogen of the conserved His 296 (O_3−H_1) decreases in the transition state, reaching a distance of 1.66 Å. The latter not only contributes to stabilize the transition state but allows an optimal orientation of the leaving group toward the subsequent protonation step. TS_1 lies at 15.6 kcal/mol above ES, as can be observed from Figure 5. Upon addition of solvent effects, the barrier becomes slightly lower (13.1 and 13.3 kcal/mol for protein environment and water, respectively).

A trigonal bipyramidal geometry was found for the phosphorus atom in the intermediate complex (INT_1) in which the P−O_{nuc} bond is completely formed being 1.85 Å. This mode of coordination has been observed in the crystal

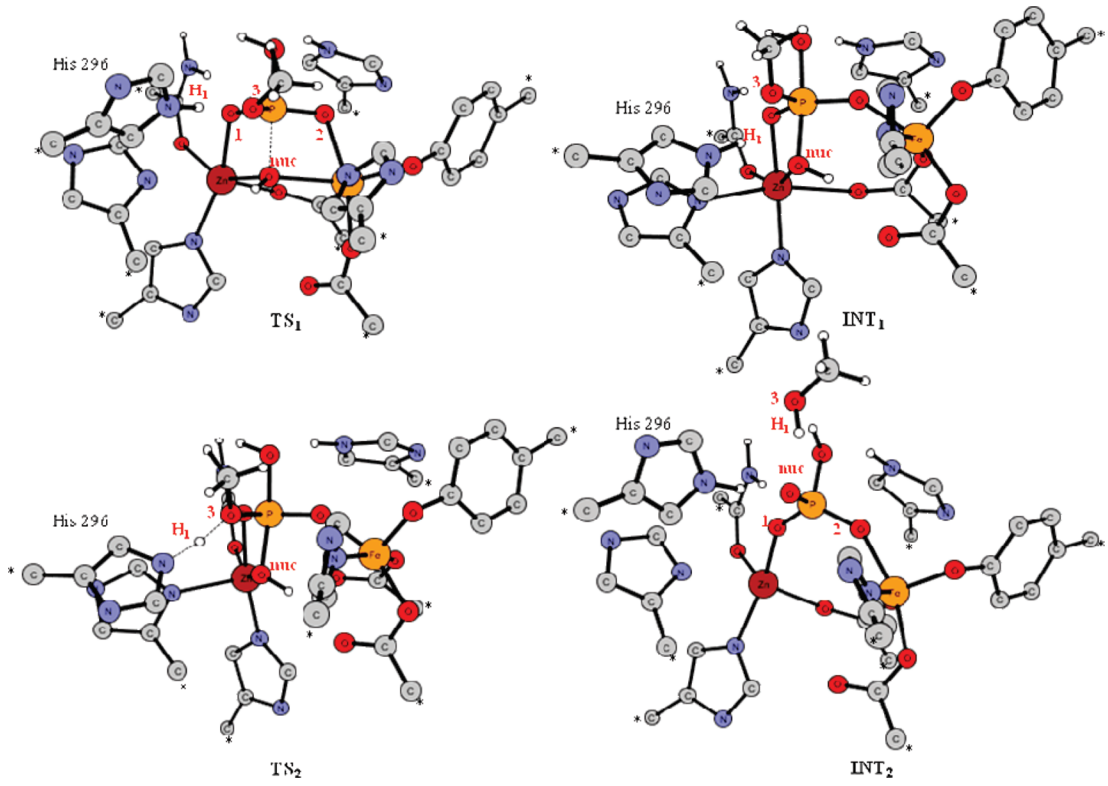


Figure 3. Optimized structures of the stationary points TS_1, INT_1, TS_2, and EP for the hydrolysis of the methyl phosphate promoted by the rkbPAP enzyme. Stars indicate the atoms kept fixed to their crystallographic positions during optimizations. Only the most significant hydrogen atoms are reported in the figure.

structures of sweet potato PAP[78] and di-Ni(II) urease,[84] both with bound phosphate, and of di-Mn(II) λPP,[85] with bound sulfate.

Although the substrate remains still bicoordinated to the dimetallic center, the distance between the Zn and Fe ions increases significantly becoming 4.58 Å. The lengthening of the metal−metal distance was already suggested in a previous work, in which the binding of the substrate has been shown to decrease the coupling interaction between the two metal centers, as indicated by a decrease in J from 20 to 6 cm^{-1}.[58] The bridging hydroxide is completely shifted away from the trivalent metal ion toward the divalent one, showing distances of 3.36 and 2.17 Å from them, respectively. From an energetic point of view, in both the gas phase and the solvent, the intermediate INT_1 shows only a small difference in the energy values with respect to TS_1.

To get the final product, a proton shift from the His 296 to the substrate's methylated oxygen atom was observed in the second transition state TS_2. An imaginary frequency of 376i cm^{-1} confirms its nature as first-order saddle point. The vibrational mode indicates the motion of the N−H_1 and H_1−O_3 couple of bonds with critical lengths of 1.27 and 1.20 Å, respectively. In such a transition state, the proton shift happens before the just formed methanol group leaves the substrate. This stationary point lies at 19.3 kcal/mol above ES in the gas phase, while solvation effects increase this difference becoming 21.0 and 22.7 kcal/mol in protein environment and water, respectively. The requirement of an activated leaving group to reach an optimal catalytic efficiency has been the subject of previous studies.[78,24] For red kidney bean and pig PAP, it was demonstrated that

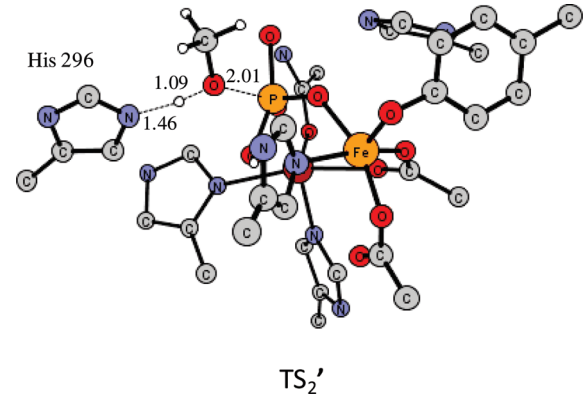


Figure 4. Concerted transition state in which the release of the methoxy group occurs at the same time of the proton shift from the His 296

reactivity declines as the leaving group's pK_a increases, while no dependence of catalytic parameters from the leaving group's pK_a was found from the sweet potato PAP. This evidence suggests that sweet potato PAP can catalyze efficiently a range of activated and inactivated phosphate esters in contrast to red kidney bean and pig PAPs. The possibility that the bond between the phosphorus atom and the methoxy group breaks without activation was also considered in our investigation. We were able to characterize a further transition state (TS_2') in which no activation takes place before the leaving group leaves and in which a shift from the His 296 occurs in a concerted manner with the scission (Figure 4). The analysis of the obtained vibrational frequency (138i cm^{-1}) clearly indicates the rupture of the

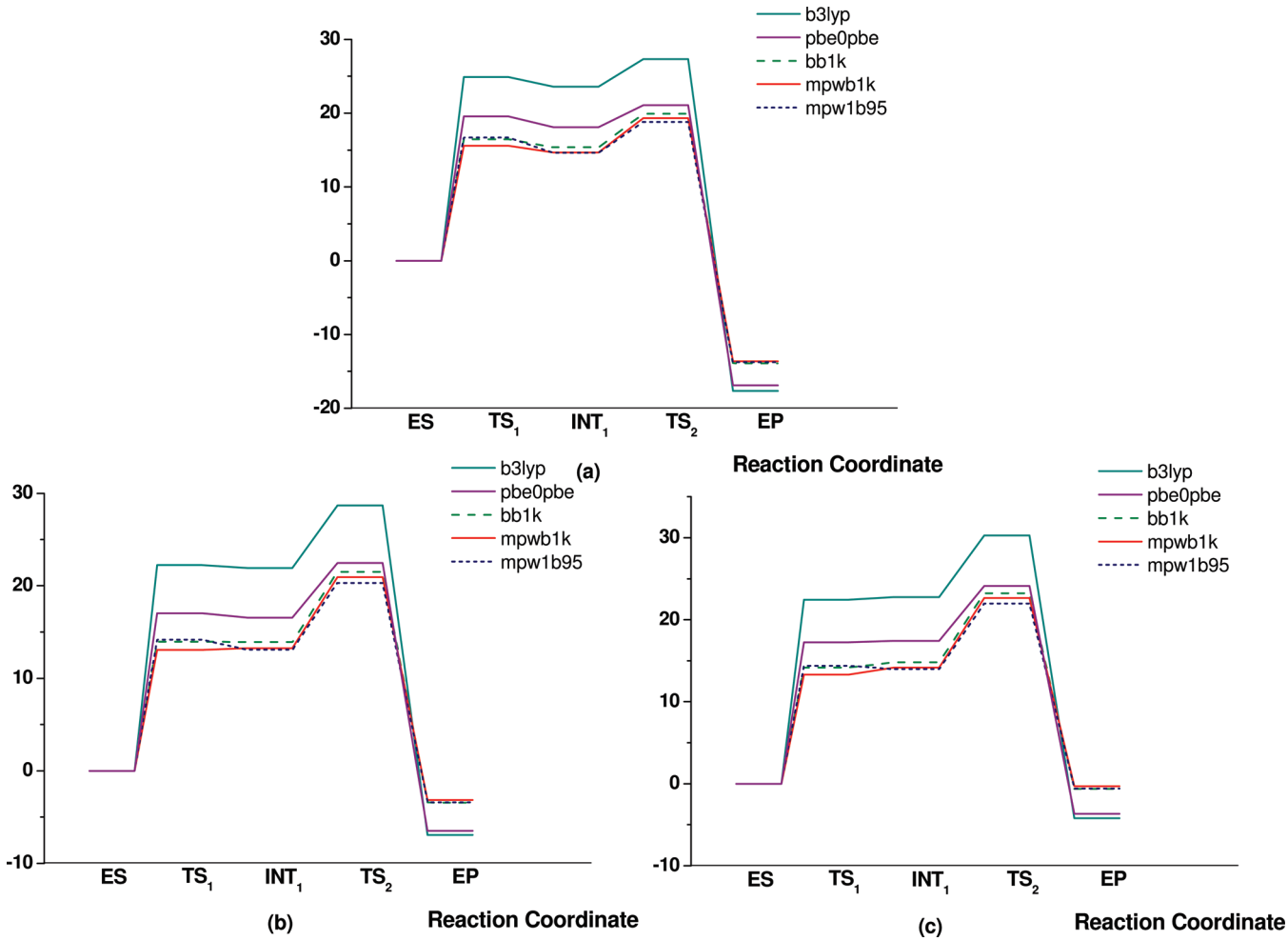


Figure 5. Potential energy profiles for the reaction in: (a) gas phase, (b) protein environment ($\varepsilon = 4$), and (c) water.

P−O_3 bond and the formation of the O_3−H_1 one. Nevertheless, this hypothesis of a concerted transition state is energetically unfavorable requiring a higher amount of energy with respect to the other transition state TS_2. This stationary point lies at 32.0 kcal/mol with respect to the ES in the gas phase, while with the introduction of solvent effects, the differences become 36.8 and 39.0 kcal/mol in protein environment and water. According to our calculation for red kidney bean PAP, we can observe that the activation of the leaving group is crucial to get the final enzyme−product complex.

The last point encountered along the reaction path corresponds to the enzyme−product complex (EP). The release of the leaving group induces significant structural rearrangements. The phosphorus atom is likely to regenerate its optimal tetrahedral geometry requiring a reorganization of the active site. During the optimization, a rotation of the bond between the phosphorus atom and the native nucleophile (P−O_{nuc}) occurs to reach a reasonable tetra-coordination. The phosphate remains tightly bound to the dimetallic center by the oxygen atoms O_1 and O_2 and with distances of 1.97 and 2.07 Å from Zn and Fe, respectively. The OH_{nuc} group oriented toward the His 296 spontaneously transfers the proton to the His 296, regenerating the active site. The distance between the metal ions decreases in the EP complex reaching 4.36 Å, from 4.83 Å in the previous TS_2 geometry. Such a complex is stabilized by a network of hydrogen bonding involving the His 296 and His 202 of the second coordination sphere, the His 325 Fe(III)-bound, and Asn 201 Zn(II)-bound. The EP complex is obtained exothermically with a great energy gain (13.6 kcal/mol in gas phase). Interestingly, the solvent effect seems to destabilize the product of reaction that is found at −3.1 and −0.3 kcal/mol lower in energy than the reactants.

3.3. Energetic of the Reaction. Different exchange−correlation functionals were used to estimate the energies of the stationary points located on the paths in the gas phase and in different environments (water and protein). This benchmark could be interesting since a large number of density functionals at different levels of sophistication have become available, and a well-established computation protocol for the enzymatic reaction is still lacking.

A detailed analysis of the performance of the popular B3LYP functional in the reproduction of a large variety of chemical properties and system types was performed in a recent study.[80] The latter shows that the B3LYP performance in the prediction of geometrical structures including closed- and open-shell structures is satisfactory and is able to compete in accuracy with the other proposed XC functional as well as with other ab initio methods.

On the contrary, for some properties several new density functionals significantly outperform this popular hybrid

Table 2. Energies (kcal/mol) of the Stationary Points along the Reaction Paths[a]

		ES	TS_1	INT_1	TS_2	EP
Gas phase	B3LYP	0	24.93	23.58	27.33	−17.65
	PBE0PBE	0	19.58	18.11	21.07	−16.87
	BB1K	0	16.47	15.36	19.92	−13.91
	MPWB1K	0	15.59	14.66	19.34	−13.64
	MPW1B95	0	16.68	14.61	18.83	−13.76
$\varepsilon = 4$	B3LYP	0	22.26	21.94	28.67	−6.94
	PBE0PBE	0	17.03	16.55	22.48	−6.48
	BB1K	0	13.94	13.92	21.52	−3.45
	MPWB1K	0	13.06	13.24	20.96	−3.14
	MPW1B95	0	14.16	13.09	20.32	−3.39
$\varepsilon = 80$	B3LYP	0	22.44	22.77	30.27	−4.18
	PBE0PBE	0	17.24	17.42	24.11	−3.67
	BB1K	0	14.17	14.82	23.22	−0.62
	MPWB1K	0	13.3	14.15	22.67	−0.3
	MPW1B95	0	14.39	13.97	21.98	−0.57

[a] In the gas phase, in the protein environment ($\varepsilon = 4$), and in water ($\varepsilon = 80$).

functional. Many studies have demonstrated that the hybrid-meta GGA methods give more reliable results in the reproduction of the reaction kinetics showing great ability in the calculation of barrier heights with high accuracy.[80] Taking these aspects under consideration, in our investigation we chose the B3LYP functional for the optimizations of all the stationary points along the reaction path and, to obtain energetic values, we performed single point calculation with a larger basis set testing the hybrids B3LYP and PBE0PBE and the meta-hybrids BB1K, MPWB1K, and MW1B95. The results are shown in Figure 5.

From Table 2 and Figure 5 it is possible to observe that the hybrid-meta GGA functionals (BB1K, MPWB1K, and MPW1B95) give very similar results concerning the energy of the stationary points along the paths in the gas phase, in the protein environment, and in water. In addition, the calculated values well agree with the energy barriers extracted from the k_{cat} experimentally obtained for reactions catalyzed by PAPs enzymes (13−15 kcal/mol depending on the substrate).[24,78,79]

On the contrary, the profiles obtained at B3LYP and PBE0PBE levels of theory show higher energy results. Furthermore, we can observe that the protein environment has a slightly higher stabilization effect on the structures than the more polar water environment, and moreover only the first part of the mechanism results stabilized. The TS_2 and the product of the reaction seem to be destabilized with the introduction of the environment.

Conclusion

The catalytic mechanism of the Fe(III)−Zn(II) purple acid phosphatase (PAP) enzyme was investigated at the density functional (DF) level of theory. The model cluster used to simulate the active site of the enzyme, made up of 114 atoms, was large enough to reliably reproduce the hydrolysis of the methyl-phosphate monoanionic substrate ($CH_3OHPO_3^-$).

The main results can be summarized as follows:

- The PAP hydrolysis reaction involves the formation of the trigonal bypiramidal intermediate, arising from nucleophilic addition of the bridging hydroxide molecule to the substrate phosphorus atom, followed by the protonation of the substrate methoxy group leading to the final P−O bond cleavage.
- The role of the two conserved histidine residues was elucidated. The protonated His 296 has been confirmed to play a fundamental role acting as a proton carrier and activating the methoxy leaving group, while the His 202 takes part in the optimal orientation of the substrate during the catalysis by establishing a hydrogen bond with it.
- From our study emerges that the rate-limiting step of the reaction is the nucleophilic attack followed by the less energetically demanding release of the leaving group.
- We were able to establish that the activation of the leaving group by the His 296 is actually crucial in order to reach an optimal catalytic efficiency. A transition state (TS_2') in which no activation takes place before the leaving group leaves and in which a shift from the His 296 occurs in a concerted manner with the scission was further characterized, but it requires a higher amount of energy to be overcome than TS_2.
- The dianionic form of the substrate was also considered in our investigation. We found that the activation barrier for the nucleophilic attack associated with that model is considerably higher than the other one relative to the monoanionic species, confirming that the reaction proceeds faster at lower pH. This result is in agreement with the experimental indications that support a monoanionic substrate being pH = 6, the optimum value for the achievement of the maximum catalytic rate for rkbPAP.[78]
- We found that the monoanionic form of methylphosphate binds in a monodentate-like manner, the bimetallic center in the precatalytic complex, while the dianionic form $CH_3OPO_3^{2-}$ results in a bicoordinated ES complex. Our evidence supports the hypothesis that the mode of substrate binding is pH dependent, as it has been previously suggested in other works.[4]
- Different XC functionals were used to calculate the barrier heights in the gas phase, in the protein environment,and in water. The meta-hybrid functionals give results that better agree with the energy barriers calculated from the k_{cat} obtained for similar reactions catalyzed by PAPs enzymes, confirming their ability in the reproduction of the reaction kinetics.

Acknowledgment. The University of Calabria, the Regione Calabria (FSE-POR 2000/2006, misura 3.7), the MIUR PRIN 2008, and the University of Porto are gratefully acknowledged.

Supporting Information Available: Further information regarding the chosen spin state along the PES, spin density distribution, and molecular orbital picture for the all stationary points, optimized structure, and selected geometrical parameters of the ES complex using Model (b), and the ES and TS_1 containing $CH_3OPO_3^{2-}$ substrates are provided. This information are available free of charge via the Internet at http://pubs.acs.org/.

References

(1) Twitchett, M. B.; Sykes, A. G. *Eur. J. Inorg. Chem.* **1999**, 2105.

(2) Klabunde, T.; Krebs, B. *Struct. Bonding (Berlin, Ger.)* **1997**, *89*, 177.

(3) Oddie, G. W.; Schenk, G.; Angel, N. Z.; Walsh, N.; Guddat, L. W.; de Jersey, J.; Cassady, A. I.; Hamilton, S. E.; Hume, D. A. *Bone* **2000**, *27*, 575.

(4) Mitić, N.; Smith, S. J.; Neves, A.; Guddat, L. W.; Gahan, L. R.; Schenk, G. *Chem. Rev.* **2006**, *106*, 3338.

(5) Wilcox, D. E. *Chem. Rev.* **1996**, *96*, 2435.

(6) Sträter, N.; Lipscomb, W. N.; Klabunde, T.; Krebs, B. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 2024.

(7) Dismukes, G. C. *Chem. Rev.* **1996**, *96*, 2909.

(8) Barford, D.; Das, A. K.; Egloff, M. P. *Annu. Rev. Biophys. Biomol. Struct.* **1998**, *27*, 133.

(9) Rusnak, F.; Mertz, P. *Physiol. Rev.* **2000**, *80*, 1483.

(10) Jackson, M. D.; Denu, J. M. *Chem. Rev.* **2001**, *101*, 2313.

(11) Lowther, W. T.; Matthews, B. W. *Biochim. Biophys. Acta* **2000**, *1477*, 157.

(12) Nuttleman, P. R.; Roberts, R. M. *J. Biol. Chem.* **1990**, *265*, 12192.

(13) Sibille, J. C.; Doi, K.; Aisen, P. *J. Biol. Chem.* **1987**, *262*, 59.

(14) Kaija, H.; Alatalo, S. L.; Halleen, J. M.; Lindqvist, Y.; Schneider, G.; Väänänen, H. K.; Vihko, P. *Biochem. Biophys. Res. Commun.* **2002**, *292*, 128.

(15) Räisänen, S. R.; Alatalo, S. L.; Ylipahkala, H.; Halleen, J. M.; Cassady, A. I.; Hume, D. A.; Väänänen, H. K. *Biochem. Biophys. Res. Commun.* **2005**, *331*, 120.

(16) Moss, D. W.; Raymond, F. D.; Wile, D. B. *Crit. Rev. Clin. Lab. Sci.* **1995**, *32*, 431.

(17) Angel, N. Z.; Walsh, N.; Forwood, M. R.; Ostrowski, M. C.; Cassady, A. I.; Hume, D. A. *J. Bone Miner. Res.* **2000**, *15*, 103.

(18) Hayman, A. R.; Jones, S. J.; Boyde, A.; Foster, D.; Colledge, W. H.; Carlton, M. B.; Evans, M. J.; Cox, T. M. *Development* **1996**, *122*, 3151.

(19) Ek-Rylander, B.; Flores, M.; Wendel, M.; Heinegard, D.; Andersson, G. *J. Biol. Chem.* **1994**, *269*, 14853.

(20) Valizadeh, M.; Schenk, G.; Nash, K.; Oddie, G. W.; Guddat, L. W.; Hume, D. A.; de Jersey, J.; Burke, T. R., Jr.; Hamilton, S. *Arch. Biochem. Biophys.* **2004**, *424*, 154.

(21) Cashikar, A. G.; Kumaresan, R.; Rao, N. M. *Plant Physiol.* **1997**, *114*, 907.

(22) Antanaitis, B. C.; Aisen, P. *J. Biol. Chem.* **1982**, *257*, 5330.

(23) Doi, K.; Bradley, C.; Aisen, P. *Struct. Bonding (Berlin, Ger.)* **1988**, *70*, 1.

(24) Schenk, G.; Ge, Y.; Carrington, L. E.; Wynne, C. J.; Searle, I. R.; Carroll, B. J.; Hamilton, S.; de Jersey, J. *Arch. Biochem. Biophys.* **1999**, *370*, 183.

(25) Durmus, A.; Eicken, C.; Sift, B. H.; Kratel, A.; Kappi, R.; Hütterman, J.; Krebs, B. *Eur. J. Biochem.* **1999**, *260*, 709.

(26) Merkx, M.; Averill, B. A. *Biochemistry* **1998**, *37*, 8490.

(27) Chen, T. T.; Bazer, F. W.; Cetorelli, J. J.; Pollard, W. E.; Roberts, R. M. *J. Biol. Chem.* **1973**, *248*, 8560.

(28) Campbell, H. D.; Zerner, B. *Biochem. Biophys. Res. Commun.* **1973**, *54*, 1493.

(29) Campbell, H. D.; Dionysius, D. A.; Keough, D. T.; Wilson, B. E.; de Jersey, J.; Zerner, B. *Biochem. Biophys. Res. Commun.* **1978**, *82*, 615.

(30) Robinson, D. B.; Glew, R. H. *J. Biol. Chem.* **1980**, *255*, 5864.

(31) Hayman, A. R.; Warburton, M. J.; Pringle, J. A.; Coles, B.; Chambers, T. J. *Biochem. J.* **1989**, *261*, 601.

(32) Allen, S. H.; Nuttleman, P. R.; Ketcham, C. M.; Roberts, R. M. *J. Bone Miner. Res.* **1989**, *4*, 47.

(33) Janckila, A. J.; Woodford, T. A.; Lam, K. W.; Li, C. Y.; Yam, L. T. *Leukemia* **1992**, *6*, 199.

(34) Beck, J. L.; McConachie, L. A.; Summors, A. C.; Arnold, W. N.; de Jersey, J.; Zerner, B. *Biochim. Biophys. Acta* **1986**, *869*, 61.

(35) Bozzo, G. G.; Raghothama, K. G.; Plaxton, W. C. *Biochem. J.* **2004**, *377*, 419.

(36) Bozzo, G. G.; Raghothama, K. G.; Plaxton, W. C. *Eur. J. Biochem.* **2002**, *269*, 6278.

(37) Schenk, G.; Boutchard, C. L.; Carrington, L. E.; Noble, C. J.; Moubaraki, B.; Murray, K. S.; de Jersey, J.; Hanson, G. R.; Hamilton, S. *J. Biol. Chem.* **2001**, *276*, 19084.

(38) Schenk, G.; Guddat, L. W.; Ge, Y.; Carrington, L. E.; Hume, D. A.; Hamilton, S.; de Jersey, J. *Gene* **2000**, *250*, 117.

(39) Antanaitis, B. C.; Aisen, P.; Lilienthal, H. R. *J. Biol. Chem.* **1983**, *258*, 3166.

(40) Averill, B. A.; Davis, J. C.; Burman, S.; Zirino, T.; Sanders-Loehr, J.; Loehr, T. M.; Sage, J. T.; Debrunner, P. G. *J. Am. Chem. Soc.* **1987**, *109*, 3760.

(41) Yang, Y.-S.; McCormick, J. M.; Solomon, E. I. *J. Am. Chem. Soc.* **1997**, *119*, 11832.

(42) Wang, D. L.; Holz, R. C.; David, S. S.; Que, L.; Stankovich, M. T. *Biochemistry* **1991**, *30*, 8187.

(43) Bernhardt, P. V.; Schenk, G.; Wilson, G. J. *Biochemistry* **2004**, *43*, 10387.

(44) Beck, J. L.; Keough, D. T.; de Jersey, J.; Zerner, B. *Biochim. Biophys. Acta* **1984**, *791*, 357.

(45) Funhoff, E. G.; Ljusberg, J.; Wang, Y.; Andersson, G.; Averill, B. A. *Biochemistry* **2001**, *40*, 11614.

(46) Funhoff, E. G.; Klaassen, C. H. W.; Samyn, B.; Van Beeumen, J.; Averill, B. A. *ChemBioChem* **2001**, *2*, 355.

(47) Merkx, M.; Averill, B. A. *Biochemistry* **1998**, *37*, 11223.

(48) Merkx, M.; Pinkse, M. W. H.; Averill, B. A. *Biochemistry* **1999**, *38*, 9914.

(49) Pinkse, M. W. H.; Merkx, M.; Averill, B. A. *Biochemistry* **1999**, *38*, 9926.

(50) Funhoff, E. G.; Bollen, M.; Averill, B. A. *J. Inorg. Biochem.* **2005**, *99*, 521.

(51) Beck, J. L.; McArthur, M. J.; de Jersey, J.; Zerner, B. *Inorg. Chim. Acta* **1988**, *153*, 39.

(52) Beck, J. L.; de Jersey, J.; Zerner, B. *J. Am. Chem. Soc.* **1988**, *110*, 3317.

(53) Gehring, S.; Fleischhauer, P.; Behlendorf, M.; Huber, M.; Lorösch, J.; Haase, W.; Dietrich, M.; Witzel, H.; Locke, R.; Krebs, B. *Inorg. Chim. Acta* **1996**, *252*, 13.

(54) Battistuzzi, G.; Dietrich, M.; Locke, R.; Witzel, H. *Biochem. J.* **1997**, *323*, 593.

(55) Sträter, N.; Klabunde, T.; Tucker, P.; Witzel, H.; Krebs, B. *Science* **1995**, *268*, 1489.

(56) Klabunde, T.; Sträter, N.; Fröhlich, R.; Witzel, H.; Krebs, B. *J. Mol. Biol.* **1996**, *259*, 737.

(57) Klabunde, T.; Sträter, N.; Krebs, B.; Witzel, H. *FEBS Lett.* **1995**, *367*, 56.

(58) Smoukov, S. K.; Quaroni, L.; Wang, X.; Doan, P. E.; Hoffman, B. M.; Que, L., Jr. *J. Am. Chem. Soc.* **2002**, *124*, 2595.

(59) Guddat, L. W.; McAlpine, A. S.; Hume, D.; Hamilton, S.; de Jersey, J.; Martin, J. L. *Structure* **1999**, *7*, 757.

(60) Lindqvist, Y.; Johansson, E.; Kaija, H.; Vihko, P.; Schneider, G. *J. Mol. Biol.* **1999**, *291*, 135.

(61) Wang, X.; Que, L. *Biochemistry* **1998**, *37*, 7813.

(62) Merkx, M.; Averill, B. A. *J. Am. Chem. Soc.* **1999**, *121*, 6683.

(63) Merkx, M.; Averill, B. A. *Biochemistry* **1998**, *37*, 8490.

(64) Twitchett, M. B.; Schenk, G.; Aquino, M. A. S.; Yiu, D. T. Y.; Lau, T. C.; Sykes, A. G. *Inorg. Chem.* **2002**, *41*, 5787.

(65) Funhoff, E. G.; Wang, Y.; Andersson, G.; Averill, B. A. *FEBS J.* **2005**, *272*, 2968.

(66) Truong, N. T.; Naseri, J. I.; Vogel, A.; Rompel, A.; Krebs, B. *Arch. Biochem. Biophys.* **2005**, *440*, 38.

(67) Schenk, G.; Elliott, T. W.; Leung, L.; Carrington, L. E.; Mitić, N.; Gahan, L. R.; Guddat, L. W. *BMC Struct. Biol.* **2008**, *8*, 6.

(68) 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 A.1; Gaussian, Inc.: Pittsburgh, PA, 2003.

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

(70) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785.

(71) Andrae, D.; Haussermann, U.; Dolg, M.; Stoll, H.; Preuss, H. *Theor. Chim. Acta* **1990**, *77*, 123.

(72) Barone, V.; Cossi, M. *J. Phys. Chem. A* **1998**, *102*, 1995.

(73) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. Comput. Chem.* **2003**, *24*, 669.

(74) Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A., III; Skiff, W. M. *J. Am. Chem. Soc.* **1992**, *114*, 10024.

(75) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158.

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

(77) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2004**, *108*, 6908.

(78) Schenk, G.; Gahan, L. R.; Carrington, L. E.; Mitić, N.; Valizadeh, M.; Hamilton, S. E.; de Jersey, J.; Guddat, L. W. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 273.

(79) Cox, R. S.; Schenk, G.; Mitic, N.; Gahan, L. R.; Hengge, A. C. *J. Am. Chem. Soc.* **2007**, *129*, 9550.

(80) Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. *J. Phys. Chem. A* **2007**, *111*, 10439.

(81) Bras, N. F.; Moura-Tamames, S. A.; Fernandes, P. A.; Ramos, M. J. *J. Comput. Chem.* **2008**, *29*, 2565.

(82) Leopoldini, M.; Russo, N.; Toscano, M. *Chem.—Eur. J.* **2009**, *15*, 8026.

(83) Hur, S.; Brulce, T. C. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12015.

(84) Voegtli, W. C.; White, D. J; Reiter, N. J.; Rusnak, F.; Rosenzweig, A. C. *Biochemistry* **2000**, *39*, 15365.

(85) Benini, S.; Rypniewski, W. R.; Wilson, K. S.; Ciurli, S.; Mangani, S. *J. Biol. Inorg. Chem.* **2001**, *6*, 778.

CT100187C

JCTC Journal of Chemical Theory and Computation

Transferability of Nonbonded Interaction Potentials for Coarse-Grained Simulations: Benzene in Water

Alessandra Villa,[†] Christine Peter,[‡] and Nico F. A. van der Vegt*[,§]

Karolinska Institutet, SE-14183 Huddingen, Sweden, Max-Planck-Institute for Polymer Research, D-55128 Mainz, Germany, and Center of Smart Interfaces, Technische Universität Darmstadt, D-64287 Darmstadt, Germany

Received May 3, 2010

Abstract: Methods to parametrize coarse-grained simulation models for molecular fluids frequently either attempt to match the fluid structure (e.g., pair correlation functions) previously obtained with detailed atomistic models or aim at reproducing macroscopically observable thermodynamic properties. In either case, the coarse-grained models are state-point-dependent, and it is unclear to what extent the models obtained at a given state point are transferable, for example, to different compositions in the case of solution mixtures. Usually, it remains unclear as well whether structure-based potentials reproduce macroscopic thermodynamic properties and, vice versa, if thermodynamics-based potentials reproduce microscopic structural properties. In this paper, we use the Kirkwood−Buff theory of solutions in order to link local structural information and thermodynamic properties sampled with structure-based potentials. We investigate benzene/water mixtures at varying concentrations as a model hydrophobic/hydrophilic system and study the transferability of a coarse-grained model that describes the water and benzene molecules as single interaction sites. The coarse-grained model, parametrized at a high aqueous dilution of benzene, reproduces the Kirkwood−Buff integrals of mixtures obtained with the detailed-atomistic model, and it reproduces the change in the benzene chemical potential with composition up to the concentration of thermodynamic instability. The observed transferability of the potential supports the idea that hydrophobic interactions between small molecules are pairwise additive.

1. Introduction

Simulation of complex molecular systems frequently faces the challenge that time and length scales required to compare to experimentally measurable properties are hardly accessible. This is especially problematic in simulation methods that operate at a high level of resolution, for example, including atomistic or even electronic degrees of freedom. In order to bridge this time- and length-scale gap, so-called coarse-grained (CG) simulation models are being developed, where groups of atoms are treated together as larger units, thus reducing the number of degrees of freedom and speeding up the simulation. Interaction potentials for CG models are usually parametrized in such a way that properties known from experiments or from simulations at higher levels of resolution, e.g., atomistic simulations, are reproduced. In recent years, a large amount of effort has been made to investigate how far CG models are—in spite of omitting degrees of freedom—capable of reproducing both thermodynamic and structural properties of the high resolution (real) system.[1–9]

A variety of approaches has been used to develop CG nonbonded interaction potentials.[10–15] CG nonbonded potentials can be derived such that they reproduce a target structure described by a set of radial distribution functions obtained from experimental or all-atom simulation data,[16–19] or can be derived on the basis of calculation of reversible

* To whom correspondence should be addressed. E-mail: vandervegt@csi.tu-darmstadt.de.

[†] Karolinska Institutet.

[‡] Max-Planck-Institute for Polymer Research.

[§] Technische Universität Darmstadt.

10.1021/ct100228t © 2010 American Chemical Society
Published on Web 07/21/2010

work.[20–23] Alternatively, thermodynamic data can be used to parametrize the interaction between "super-atoms",[24–30] a procedure which is often based on hydrophilicity/hydrophobicity arguments.[31,32] A third approach relies on the sampling of a high-resolution system and projecting the forces onto a smaller set of coarse-grained forces.[8,33,34] It is a priori not clear how well structure-based CG potentials are suited to reproducing macroscopic thermodynamic properties and, on the other hand, how well thermodynamics-based CG models perform in terms of reproducing liquid structure. In all cases, the resulting CG potentials are state-point-dependent and their transferability to other thermodynamic states at different temperatures, compositions, etc. is limited and needs to be evaluated. This transferability, in particular to different concentrations of liquid mixtures or solutions, is of vital importance for the simulation of processes such as (bio)molecular aggregation which are characterized by spatially varying structure and fluctuating concentrations.

Recently, binary mixtures have been used as a model system to explore various aspects of the transferability of CG models obtained with different coarse-graining approaches using atomistic simulations as a reference.[1,6,8,22,35–37] Among others, Silbermann et al.[1] addressed the problem of the transferability using implicit solvent CG potentials at different ethanol/water concentrations, while Fischer et al.[6] developed implicit solvent CG models for polyoxyethylene solutions and explored the potential transferability to different chain lengths and mixture concentrations. Hess et al.[22] reported a transferable, effective two-body potential for ionic interactions in aqueous electrolytes, which could be obtained by accounting for multibody effects of the ions exerted on the dielectric behavior of the implicit solvent. In all of these models, the solvent effects are implicitly included in the solute−solute interaction potential. That makes the potentials not suitable to describing the pure liquid systems or the solute partitioning between different solvents. On the other hand, Mullinax and Noid[8] determined CG transferable potentials that are targeted at reproducing structural properties of alkane/alcohol mixtures at different concentrations. In this paper, the authors aimed at getting a single force field capable of representing solution properties at different mixtures.

In the present manuscript, we investigate to what extent a CG model can reproduce properties of both pure liquid and liquid mixtures, more precisely the liquid structure and solution thermodynamics, and to what extent a CG model can be transferred to mixtures of different concentrations of the compounds. We address these questions and use the Kirkwood−Buff (KB) theory of solution to link quantities related to the local solution structure to changes of the solute chemical potential with solution composition.[38,39] To do that, we investigate a hydrophobic/hydrophilic mixture of benzene and water. We develop two CG models based on the same atomistic description, and we apply them to different mixture conditions. In the CG models, both benzene and water molecules are treated as explicit interaction sites.

In the following sections, we will first describe how structure-based CG interaction potentials can be determined for pure fluids, mixtures, and dilute solutions. Then, we will briefly review Kirkwood−Buff theory and the relation between solution structure, KB integrals, and solution thermodynamics before we turn to the results of CG simulations of several benzene/water mixtures.

2. Nonbonded Interactions in CG Models

Structure-based coarse graining of molecular liquids relies on determining CG potentials in such a way that a predefined target function (or multiple target functions) that structurally characterizes the molecular liquid is reproduced in the CG simulation. In the present case, this means that effective two-body interaction potentials are determined to reproduce the (target/reference) radial distribution function $g_{\mathrm{ref}}(r)$ (RDF). The Boltzmann-inverted radial distribution function is the potential of mean force (PMF) between the particles:

$$V_{\mathrm{PMF}}(r) = -k_{\mathrm{B}}T \ln g_{\mathrm{ref}}(r) \equiv V_0^{\mathrm{CG}}(r) \tag{1}$$

where k_{B} denotes the Boltzmann constant and T the temperature. The PMF cannot be used directly as a two-body interaction potential in a CG model, since it incorporates multibody contributions of all the other particles in the system in a statistically averaged way. But the PMF can serve as a first guess ($V_0{}^{\mathrm{CG}}(r)$) in an iterative procedure, where these multibody contributions are eliminated and an effective two-body interaction potential is determined that reproduces the target structure. One possibility for such an iterative procedure is the so-called iterative Boltzmann inversion (IBI) method, where the initial guess potential $V_0{}^{\mathrm{CG}}$ is self-consistently refined:

$$V_{i+1}^{\mathrm{CG}}(r) = V_i^{\mathrm{CG}}(r) + k_{\mathrm{B}}T \ln\left[\frac{g_i(r)}{g_{\mathrm{ref}}(r)}\right] \tag{2}$$

Here, $g_i(r)$ denotes the RDF sampled with potential V_i^{CG} in the ith iteration. For pure fluids (where each molecule corresponds to a single CG site), the determination of the two-body potential through IBI is straightforward and converges to a unique solution (reproducing the reference radial distribution function between the molecular centers of mass) according to the Henderson theorem.[40] For liquid mixtures or solutions, or for pure fluids composed of molecules that are described by several different CG beads, the situation is more complex.[19,41,42] One reason is that in principle several RDFs need to be simultaneously reproduced, which means that several potentials, that mutually affect each other, need to be refined.[19] In the case of solutions or mixtures, one may more easily encounter convergence issues of the iteration process if the PMFs of the mixture (inverted radial distribution functions, eq 1) are a bad initial guess.

In addition, for dilute solutions, IBI becomes inefficient, because the solute−solute RDFs (solute is here referring to the low concentration component) $g_{\mathrm{ref}}(r)$ and $g_i(r)$ converge very slowly in atomistic and CG simulations. In this case, a direct determination of the PMF between two solute molecules in atomistic simulations can alternatively be obtained using free-energy calculation methods such as umbrella sampling or constraint dynamics. Because these methods are

computationally expensive, it will however be prohibitive to use them in an iterative optimization approach. Therefore, an alternative, approximate method that had been previously developed[43] is chosen here to estimate the CG solute−solute interaction potential (V_{BB}^{CG}, since in our case the solute is benzene (B) and the solvent is water (W)).

In this approach, the PMF between two solute molecules in a solvent box, V_{pmf}^{AA}, is calculated at first by all-atom (AA) simulations using constraint dynamics as described in section 4.3. The so obtained PMF incorporates thermally averaged contributions from solute degrees of freedom and from the solvent degrees of freedom; the only remaining degree of freedom is the benzene−benzene center of mass distance. Thus, this PMF cannot be directly used as a CG interaction potential if the CG model has an explicit (CG) solvent representation. To remove the solvent contributions from the PMF, a "correcting potential" is needed. To obtain this "correcting potential" (denoted as $V_{pmf,excl}^{CG}(r)$), we rerun the PMF calculation with CG potentials but exclude the direct solute−solute interaction. That means CG simulations are performed using the solvent−solvent and solute−solvent interaction potentials which had been previously obtained (for example, using the IBI method). Subtracting $V_{pmf,excl}^{CG}(r)$ from the all-atom PMF, V_{pmf}^{AA}, yields an effective two-body interaction potential that can be used as a solute−solute (i.e., benzene−benzene) potential, V_{BB}^{CG} in the CG simulations:

$$V_{BB}^{CG}(r) = V_{pmf}^{AA}(r) - V_{pmf,excl}^{CG}(r) \quad (3)$$

Note that the CG potential $V_{BB}^{CG}(r)$ does not exactly reproduce V_{pmf}^{AA}, since iteration steps in analogy to IBI—which could in principle be made—are left out since the PMF calculations are computationally expensive. Only for a neat liquid (pure B or pure W) would the above procedure yield an exact solution, as explained in more detail in ref 43. There, it is also shown for several polar and apolar solutes in aqueous solution that the so obtained approximate CG potentials reproduce the atomistic target potentials of mean force very well.

In the present study, we will investigate dilute benzene−water mixtures of various concentrations with two types of CG benzene−benzene interaction potentials. The first is derived for pure benzene using IBI, and the second is derived using the above subtraction procedure, eq 3, i.e., with potentials ideally suited for very low benzene concentrations. In both cases, the same CG water−water and water−benzene potentials are used, which are determined by IBI on the basis of atomistic simulations of pure water and a single benzene molecule in water, respectively.

3. Kirkwood−Buff Theory of Solutions

We want to address the question of transferability of CG models to various concentrations, and we want to investigate how structure-based CG models perform at reproducing certain aspects of the thermodynamic behavior of solutions. This question is of essential importance if one thinks of using such CG models for the investigation of structure formation or aggregation processes in biological systems, materials science, etc., since in these processes both structure and association thermodynamics have to be correctly represented.

To this end, we will employ the Kirkwood−Buff (KB) theory of solutions, which provides a link between the microscopic structure of the solution and its thermodynamics.[38,39] Kirkwood−Buff theory has been worked out in detail by Ben-Naim[44] and has previously been used by Smith and Weerasinghe,[45,46] as well as by others,[47−49] in parametrizations of atomistic force fields for aqueous solutions.

Usually, thermodynamic mixture properties are studied at the macroscopic level with no direct reference to molecular-scale properties. In KB theory,[44] thermodynamic properties of a solution or mixture of molecules i and j can be expressed in terms of the KB integrals, G_{ij}:

$$G_{ij} = \int_0^R [g_{ij}(r) - 1]4\pi r^2 \, dr \quad (4)$$

where $g_{ij}(r)$ is the radial distribution function between the center of mass of molecules i and j and R is a correlation distance beyond which the integrand vanishes. The quantity $\rho_j G_{ij}$, with ρ_j being the number density of molecules j, equals the change in the average number of molecules j in a spherical region of radius R caused by placing a molecule i at the center of the region. The quantity $\rho_j G_{ij}$ will in the following be referred to as the excess coordination number ΔN_{ij}:

$$\Delta N_{ij} = \rho_j \int_0^R [g_{ij}(r) - 1]4\pi r^2 \, dr \quad (5)$$

where $\rho_j = N_j/V$, N_j is the number of molecules j in the system, and V is the average volume in the (N_j, P, T) system. Since the correlation radius R typically corresponds to a few molecular diameters, the quantities G_{ij} and ΔN_{ij} are local quantities determined by the local fluid structure.

In our analysis of benzene (B)/water (W) mixtures, we use the preferential solvation parameter defined as

$$\Delta_{BW} = G_{BB} + G_{WW} - 2G_{BW} \quad (6)$$

This quantity depends on all three KB integrals in the system and provides a link with the solution thermodynamics through the KB-derived expression for the chemical potential derivative:

$$\left(\frac{\partial \mu_B}{\partial x_B}\right)_{P,T} = \frac{k_B T}{x_B(1 + \rho_B x_W \Delta_{BW})} \quad (7)$$

where μ_B is the chemical potential of compound B in the mixture at pressure P, temperature T, and composition (i.e., mole fraction) x_B. Integration of eq 7 yields

$$\mu_B(P, T, x_B) = \mu_B^p(P, T) + k_B T \ln x_B + k_B T \int_0^{x_w} \frac{\rho x'_W \Delta_{BW}}{(1 + \rho x'_B x'_W \Delta_{BW})} \, dx'_W \quad (8)$$

where $\mu_B^p(P, T)$ is the chemical potential of pure B at the same pressure P and temperature T as those of the mixture, $\rho = \rho_B + \rho_W$, and x_B and x_W are the mole fractions of B and W, respectively. When $\Delta_{BW} = 0$ at all B mole fractions between zero and x_B, we obtain $\mu_B = \mu_B^p + k_B T \ln x_B$, which is the chemical potential of B in a symmetrical ideal solution.[44] The last term on the right side of eq 8 describes

the deviation from thermodynamic ideality and is determined by the preferential solvation parameter Δ_{BW} that in turn describes the deviation of the local solution composition from the global solution composition.

Another thermodynamic quantity that will be studied is the activity coefficient of compound B, γ_B, which can be obtained from liquid/vapor equilibria according to $\gamma_B = p_B/p_B^0 x_B$. Here, ideal gas behavior of the vapor has been assumed, p_B is the equilibrium vapor pressure of component B, p_B^0 the saturation vapor pressure of pure B, and x_B is the mole fraction of B in the liquid phase. The derivative with respect to the composition x_B is given by

$$\left(\frac{\partial \ln \gamma_B}{\partial \ln x_B}\right)_{P,T} = -\frac{\rho_W x_B \Delta_{BW}}{1 + \rho_W x_B \Delta_{BW}} \quad (9)$$

Clearly, γ_B only varies with x_B, if $\Delta_{BW} \neq 0$; i.e,. the thermodynamic affinities between the solution components (B−B, W−W, and B−W) differ. Equation 9 provides a way to evaluate whether the CG potentials correctly describe the thermodynamic changes that occur in response to variations in the local solvent environment of a given molecule.

4. Computational Details

4.1. Simulations. In the atomistic simulations, the GROMOS 53a6 force field[50] was used for the benzene molecule and the SPC/E[51] model for water. All simulations were performed using the GROMACS suite of programs (version 4.0).[52,53] All of the simulations were performed in a periodic cubic box with dimensions longer than twice the cutoff distance. A cutoff distance of 1.0 nm was used for the Lennard-Jones interactions, and a long-range dispersion correction was applied for energy and pressure. The particle mesh Ewald method[54] was employed to treat Coulomb interactions, using a switching distance of 1.0 nm and a grid spacing of 0.12 nm and a Gaussian width of 0.32 nm. Constant pressure P and temperature T were maintained by coupling the system to an external bath at 1 bar and 300 K, using the Parrinello−Rahman barostat[55] and the Nose−Hoover thermostat,[56,57] respectively. The temperature coupling time was 0.5 ps. The pressure coupling time was 3 ps and the isothermal compressibility 4.5×10^{-5} bar^{-1}. In the atomistic simulations, a leapfrog integrator with a integration time step of 2 fs was used. The bond distances and the bond angle of the solvent water were constrained using the SETTLE algorithm.[58] All other bond distances were constrained using the LINCS algorithm.[59,60] In the coarse-grained simulations, a leapfrog stochastic dynamics integrator[61] was used to integrate the equations of motion. The inverse friction constant was set to 1.0 ps. In the parametrization procedure, all CG simulations were performed at constant NVT conditions at the average volume of the corresponding atomistic simulation. For the data collections, CG simulations were performed at constant pressure and temperature. Atomistic simulations were performed for 60−100 ns and coarse-grained simulation for 10−20 ns. Table 1 reports a summary of the simulated systems discussed in section 5.

4.2. Iterative Boltzmann Inversion. The IBI method was used to generate numerically tabulated potentials that reproduce the RDFs of pure liquid systems (pure benzene and pure water) or a dilute solution (one benzene molecule in water solution) at 300 K. The RDFs were calculated on the basis of the center of mass position of the each molecule. Simulations of 40, 55, and 140 ns were performed for liquid water (895 molecules), liquid benzene (2744 molecules), and a benzene/water solution (one benzene in 888 water molecules) to generate the reference radial distribution functions, $g_{ref}(r)$. In the IBI procedure, a 10 ns coarse-grained simulation was performed at each iteration. A value of $\sum(g(r) - g_{ref}(r))^2 < 0.01$ was chosen as a convergency criterion (where the summation runs over all distance bins). The pressure (with a target value of 1 atm) was included into the iteration process by applying a linear correction term to the tabulated potentials.[10]

Table 1. Summary of the Simulated Benzene (B)/Water (W) Mixture[a]

ρ_B	x_B	volume
	benzene−water mixture	
0.1	0.0019	402.36
0.2	0.0038	401.416
0.3	0.0057	400.276
0.4	0.0076	399.095
0.5	0.0095	407.766
0.6	0.0116	402.771
9.5	0.8390	423.345
	liquid benzene	
11.5	1	397.654

[a] The benzene molar concentration is denoted with ρ_B (M) and the benzene mole fraction with x_B. The atomistic average volumes (nm^3) are also reported.

4.3. Potential of Mean Force. Potentials of mean force, V_{pmf}, were calculated from n distance constraint simulations using

$$V_{pmf}(r) = \int_{r_m}^{r} \left[\langle f_c \rangle_s + \frac{2k_B T}{s}\right] ds \quad (10)$$

f_c is the constraint force, $\langle -f_c \rangle_s$ the average (mean) force exerted on the two particles separated by a fixed distance s, and r_m is the maximum distance between the center of mass of the two benzene molecules evaluated in the constraint simulations. When the distance between the two mass centers is constrained, free rotation of the solute−solute connecting vector remains possible, and larger volume elements are sampled at larger distances. This leads to an entropic contribution to the average constraint force that must be subtracted out. The second term of eq 10 takes care of this contribution. To keep the distance between the two molecules fixed, a linear constraint solver (LINCS)[59,60] was used.

Initial configurations were generated by moving the center of mass of benzene molecules apart with a speed of 0.001 nm/ps in vacuo starting from an equilibrated conformation at a distance of 0.28 nm. Configurations of 0.02 nm each were solvated and simulated. A total of 60 constraint simulations were performed. Each simulation was performed for 10 ns at the atomistic and coarse-grained levels.

5. Results

5.1. Coarse-Grained Models. In our CG models, each molecule is described through a single interaction site located

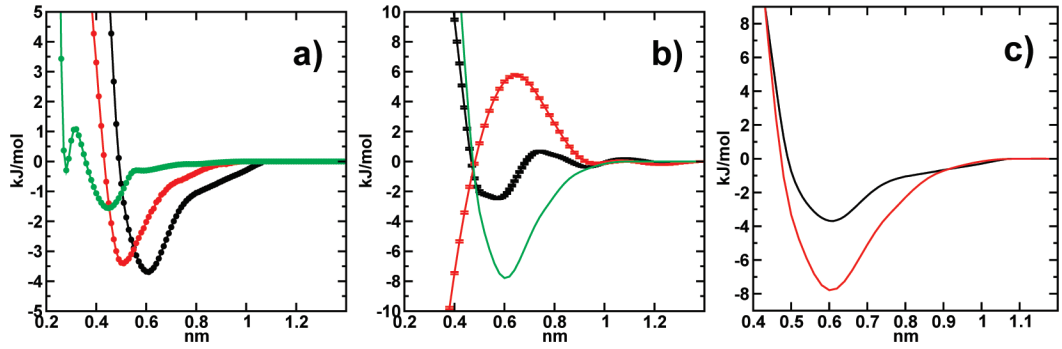


Figure 1. (a) Tabulated potentials obtained from IBI: benzene−benzene (from simulations of pure liquid benzene; black line), benzene−water (from simulations of one benzene molecule in water; red line), and water−water (from simulations of pure liquid water; green line). This model is referred to as CG(IBI). (b) Potential of mean force between two benzene molecules in aqueous solution: V_{pmf}^{AA} in black, $V_{pmf,\ excl}^{CG}$ in red, and V_{BB}^{CG} in green; this model is referred to as CG(PMF). (c) Comparison between tabulated potentials for benzene−benzene: CG(IBI) in black and CG(PMF) in red.

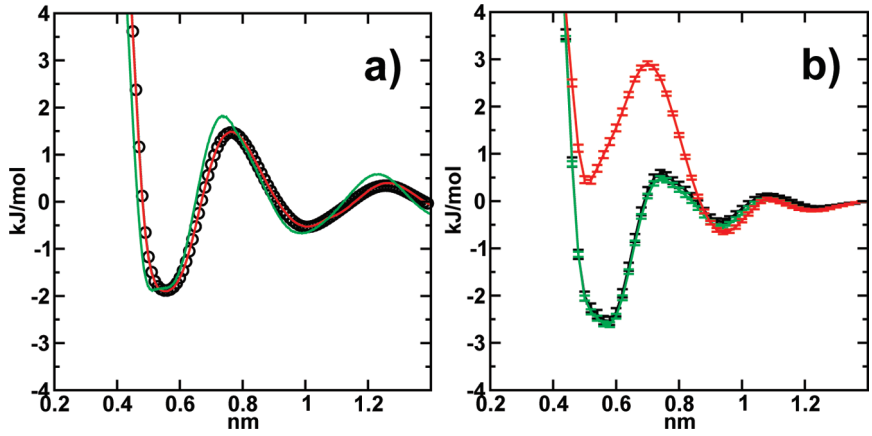


Figure 2. Potential of mean force between two benzene molecules as a function of the benzene's center of mass distance in (a) liquid benzene and (b) a water solution. Results from atomistic simulations, in black, and from CG simulations, CG(IBI) in red and CG(PMF) in green.

on the center of mass of the molecule. The potentials for solvent−solvent and solute−solvent interactions (water−water and benzene−water) were generated by IBI to reproduce the RDFs of pure liquid water and a very dilute solution (one benzene molecule in aqueous solution). The so-obtained tabulated potentials for water−water, V_{WW}^{CG}, and benzene−water, V_{BW}^{CG}, interactions are reported in Figure 1a.

Two different methods were used to determine V_{BB}^{CG} as discussed above. The first method is based on applying IBI to a pure liquid benzene system. The second method relies on benzene in water at infinite dilution where the CG nonbonded interaction potential is computed through direct determination of the potential of mean force between two benzene molecules (eq 3) with constraint dynamics. The potential obtained with the first method is shown in Figure 1a. Results from simulations performed with this potential will be labeled as CG(IBI). Figure 1b shows the PMFs together with the corresponding CG nonbonded potential obtained with the second method. Results from simulations performed with this potential will be labeled as CG(PMF). Since the subtraction method (eq 3) is only approximate, the ability of the resulting nonbonded potential to reproduce the atomistic benzene−benzene PMF in water was checked by recalculating PMFs between pairs of the solutes at the CG level. The differences between the atomistic and CG PMFs fall within the statistical errors as shown in Figure 2b.

The two CG sets of nonbonded potentials, CG(IBI) and CG(PMF), differ only in the benzene−benzene potentials. Both the benzene−benzene potentials have one minimum at 0.6 nm and a long tail that goes to zero between 1.0 and 1.1 nm, but the CG(PMF) potential shows a deeper minimum than CG(IBI) one as shown in Figure 1c.

By construction, the "pressure-corrected" CG(IBI) potentials reproduce the density of the atomistic reference of the respective pure liquids at NPT conditions.[62] In contrast to CG(IBI), the B−B interaction in the CG(PMF) model is not "pressure-corrected". At 1 atm, liquid benzene, modeled with the CG(PMF) model, has a density of 1.052 kg/L, which compares to the 0.895 kg/L obtained with the atomistic model. In dilute aqueous solutions, the effect of this is probably negligible.

5.2. Potential of Mean Force between Benzene Molecules. The two benzene−benzene CG potentials have been derived in two different environments, one in liquid benzene—CG(IBI)—and the other in a dilute benzene solution—CG(PMF). We want to see to what extent the potentials are transferable when the environment around the

benzene molecules changes (i.e., from a liquid benzene environment to a liquid water environment for the CG(IBI) model and from a liquid water environment to a liquid benzene environment for the CG(PMF) model). Figure 2 shows the PMF between two benzenes in liquid benzene and in water solution. In this figure, the atomistic results (black line) are compared with the CG(IBI) (red) and CG(PMF) (green) simulations. Figure 2a clearly shows that the CG(PMF) model reproduces the benzene−benzene PMF in liquid benzene with reasonable accuracy. We observed a tiny shift of the minima toward a shorter distance and slightly higher barriers. Figure 2b shows the PMF between two benzene molecules in water. The figure clearly shows that with the CG(IBI) model the shape of the benzene−benzene PMF deviates strongly from the atomistic reference PMF at distances smaller than 0.85 nm. The AA model predicts a global minimum at a benzene−benzene distance of around 0.5 nm, and a second minimum at 0.9 nm (corresponding to the solvent separated minimum), while the CG(IBI) PMF shows the global minimum at 0.9 nm in agreement with the second minimum of the AA potential of mean force. From the curve shown in Figure 2b, we deduce that the aggregation between benzene molecules in water solution is not favorable when the CG(IBI) model is used. Note again that with the CG(PMF) model, the calculated benzene−benzene PMF closely resembles the atomistic reference PMF.

5.3. KB Analysis of Benzene/Water Mixtures. To better understand to what extent the developed CG potentials are transferable, we simulate benzene/water mixtures at different concentrations. Atomistic and coarse-grained simulations have been performed for benzene−water mixtures at low (from 0.1 to 0.6 M) and high (9.5 M) benzene concentrations. Note that a benzene concentration of 0.5 M (40 g/L) is higher than the experimental benzene solubility in water (1.78 g/L at 298 K[63]). In the simulations presented here, benzene and water remain miscible above the experimental solubility limit. This observation reflects the fact that the GROMOS 53A6 force field underpredicts the benzene hydration free energy.[64]

Usually, thermodynamic mixture properties are studied at the macroscopic level with no direct reference to molecular-scale properties. Here, we want to compare the properties of the mixture simulated with the atomistic and CG models without losing the link between liquid structure and thermodynamics. This can be achieved by means of the Kirkwood−Buff theory of solution, as described in section 3. In Table 2, the values of the benzene−benzene (BB), benzene−water (BW), and water−water (WW) KB integrals (eq 4) for benzene/water mixtures at low benzene concentrations (from 0.1 to 0.5 M) are reported. Interestingly, G_{BB}, calculated from simulations with the atomistic and CG(PMF) models, shows a positive value, except at very low benzene concentrations in the case of the atomistic model, while the values obtained with the CG(IBI) model are all negative. A negative value of G_{BB} corresponds to a negative excess coordination number, $\rho_B G_{BB}$. This is an indication that the aggregation between benzene molecules in water solutions is not favorable, when the CG(IBI) model is used. This result is in line with the potential of mean force shown in Figure 2b.

Table 2. Kirkwood−Buff integrals, $G_{ij}(R)$, and Preferential Solvation Parameters (in cm^3/mol) for Different Benzene−Water Mixtures[a]

	x_B	G_{BB}	G_{BW}	G_{WW}	Δ_{BW}	δ_B	δ_W
AA	0.0019	−343	−77	−15	−204	−266	62
	0.0038	137	−83	−16	288	221	67
	0.0057	369	−91	−13	538	460	78
	0.0076	292	−93	−14	464	385	79
	0.0095	529	−107	−13	730	636	94
CG(IBI)	0.0019	−1014	−94	−10	−836	−920	84
	0.0038	−720	−92	−9	−545	−628	83
	0.0057	−636	−86	−8	−472	−550	78
	0.0076	−592	−80	−8	−440	−512	72
	0.0095	−585	−77	−7	−438	−508	70
CG(PMF)	0.0019	280	−107	−10	484	387	97
	0.0038	179	−107	−9	384	286	98
	0.0057	354	−117	−7	581	471	110
	0.0076	400	−121	−6	636	521	115
	0.0095	491	−129	−3	746	620	126

[a] Results from the atomistic and coarse-grained simulations. KB integrals have been obtained by averaging the $G_{ij}(R)$ over the R interval 1.4−1.8 nm.

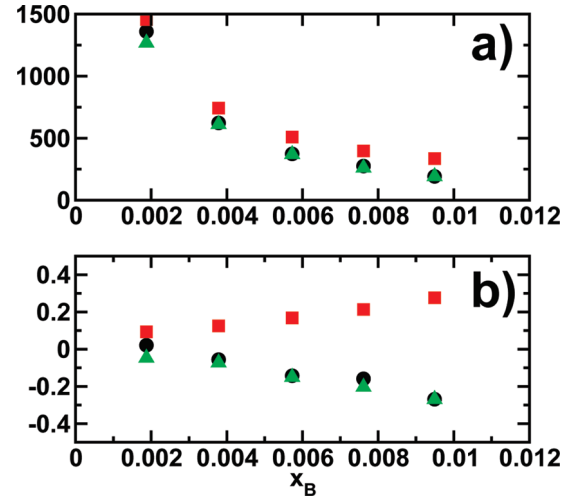


Figure 3. (a) $(\partial\mu_B/\partial x_B)_{P,T}$ (in kJ/mol) and (b) $(\partial \ln \gamma_B/\partial \ln x_B)_{P,T}$ as a function of the benzene mole fraction. Results from the simulations with the atomistic model (AA) in black, with the CG(IBI) model in red, and with the CG(PMF) model in green.

Figure 3 shows the chemical potential derivatives (eq 7) and the derivatives of the activity coefficient (eq 9) for all mixtures with a benzene concentration ranging from 0.1 to 0.5 M. The derivatives $(\partial\mu_B/\partial x_B)_{P,T}$ are all positive, indicating that the solution is thermodynamically stable at benzene concentrations ranging between 0.1 and 0.5 M. With an increasing benzene mole fraction, the derivative of the chemical potential in Figure 3a asymptotically approaches zero for all of the models. Since $(\partial\mu_B/\partial x_B)_{P,T}$ is dominated by an x_B^{-1} dependency (eq 7), we also investigated $(\partial \ln \gamma_B/\partial \ln x_B)_{P,T}$, a quantity which is more sensitive to differences in the intermolecular potentials (Figure 3b). While again the agreement between the AA and CG(PMF) models is very good at all compositions investigated, the data obtained with the CG(IBI) model increase with x_B as opposed to the data obtained with the other two models. The agreement between the CG(PMF) (corresponding to parametrization at infinite dilution) and the atomistic model indicates that the CG

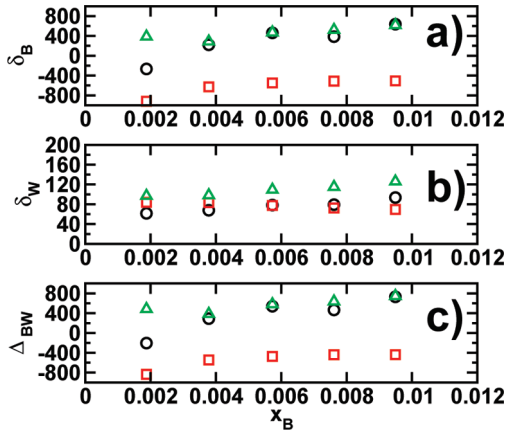


Figure 4. Preferential solvation parameters (a) $\delta_B = G_{BB} - G_{BW}$, (b) $\delta_W = G_{WW} - G_{BW}$, and (c) $\Delta_{BW} = G_{BB} + G_{WW} - 2G_{BW}$ in cm^3/mol. Results from the atomistic simulation in black, the CG(IBI) simulation in red, and the CG(PMF) simulation in green.

potentials are transferable to systems with higher benzene concentrations and reproduce the corresponding change of the benzene activity coefficient due to changes in the local solvent environment.

To further quantify how B−B and W−W affinities contribute to the observed trends in Figure 3, the preferential solvation parameter Δ_{BW} (eq 6) is written as the sum of two contributions, δ_B (characterizing the B−B affinity) and δ_W (characterizing the W−W affinity) defined as $\delta_B = G_{BB} - G_{BW}$ and $\delta_W = G_{WW} - G_{BW}$. As shown in Figure 4, we observed deviation from ideal solution behavior ($\Delta_{BW} \neq 0$) for all of the simulated mixtures, but Δ_{BW} is positive for the atomistic and CG(PMF) models while negative for the CG(IBI) model. Except for the lowest concentration, $\delta_B > \delta_W$ for the atomistic and CG(PMF) models. Hence, relative to benzene−water, the benzene−benzene affinity is stronger than the water−water affinity.

At a concentration of 0.6 M, the system starts to become unstable, as shown by the excess coordination numbers presented in Figure 5 as a function of the radius *R*. The values of the excess numbers at a long distance do not reach a plateau (as they do in a stable solution), but they diverge. With the atomistic model, the divergent trend is clearly observed for all three excess coordination numbers, and a first form of aggregation is observed as shown in the inset of Figure 5a. For the CG(PMF) model, only the benzene−water excess coordination number shows a tendency to diverge (red line in Figure 5c). The CG(IBI) model (Figure 5b) shows no indication of instability.

At a high concentration (9.5 M), all of the excess numbers obtained with the atomistic and CG(PMF) models clearly diverge, as shown in Figure 6a and c, respectively, while the excess numbers obtained with the CG(IBI) model oscillate. Phase separation between benzene and water is observed with the atomistic model after 5 ns of simulation, as shown in the inset in Figure 6a, where the two phases are easily discernible. Figure 6c clearly indicates thermodynamic instability, but phase separation cannot be easily detected visually, as shown in the inset in this figure.

6. Discussion

We have investigated two effective B−B potentials with respect to their ability to reproduce thermodynamic activities, which vary as a result of molecular association in benzene−water solutions at different mixture compositions. Because density fluctuations in aqueous solutions with hydrophobic solutes are large and strongly dependent on mixture composition, it is not a priori clear if an effective two-body potential can be found which is thermodynamically transferable, i.e., capable of describing changes in solution properties from one concentration to another. In simulations with (nonatomistic) coarse-grained models, transferability of the model cannot be guaranteed, and application of the potentials at state points away from where they were originally parametrized should always be carefully validated. Despite this obvious limitation of pair potentials used in coarse-grained models, ways exist to account for nonadditive (multibody) effects in effective pair potentials, provided that the physical nature of the nonadditivity is understood. Aqueous electrolyte solutions provide an example. For those, an implicit-solvent effective two-body potential has been previously reported that accounts for multibody contributions of the surrounding ions through their effect on the dielectric properties of the implicit solvent medium.[22]

For the present system, one may anticipate that the CG(IBI) (parametrized in the pure liquid) and CG(PMF) (parametrized at infinite dilution) potentials each are capable of describing one of the extreme cases of benzene/water mixtures. At concentrations close to instability, one might think that the CG(IBI) model should be the most appropriate choice because fluctuations are big, microscopic benzene clusters are formed, and B interacts predominantly with B. In the opposite limit, where the solution is sufficiently dilute to ignore occurrences of B−B interactions beyond two-body, we may argue that the CG(PMF) model is to be preferred. It is however not a priori clear which of the two coarse-grained potentials performs better at solution compositions located in between these two limits. A striking observation in Figure 3b is that the CG(PMF) potential is remarkably transferable over the entire range of stable-mixture compositions, while the CG(IBI) potential fails to reproduce the all-atom data of all mixtures.

A possible interpretation could be the following: Benzene aggregation in water results from hydrophobic interactions between the nonpolar benzene solutes. Small hydrophobic solutes like benzene reduce the volume of configuration space available for hydrogen bonding. The corresponding entropy loss can be minimized when solvent-separated nonpolar solutes are forced into contact. In our simulations, we use a single-site, coarse-grained water model, which, other than indirectly through the effect on the pair correlation function, does not account for hydrogen bonding. Figures 2b and 3b indeed illustrate that this water model is incapable of mediating hydrophobic benzene interactions, as evidenced by the results obtained with the CG(IBI) model, which favors benzene solvation by water rather than benzene association. The B−B interaction in this model, by construction, contains contributions of the B−B van der Waals interactions only. The CG(PMF) model, on the other hand, performs remark-

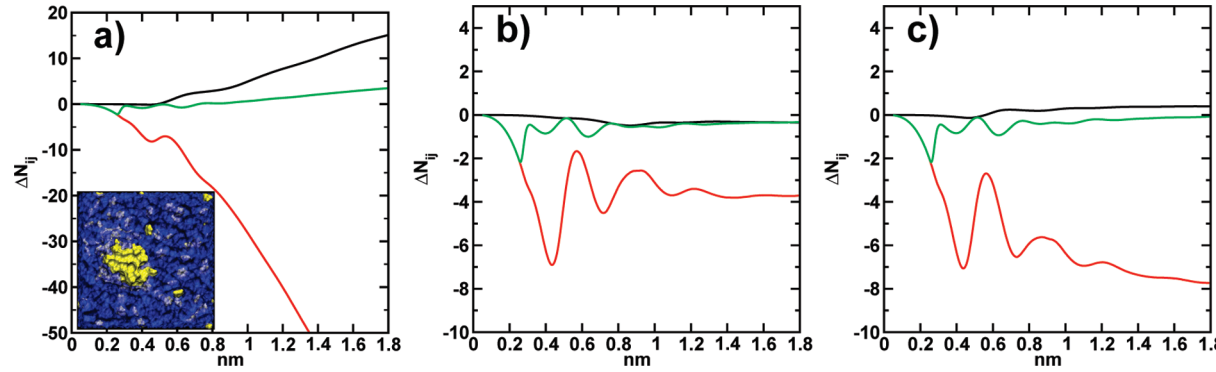


Figure 5. Benzene−benzene (black), benzene−water (red), and water−water (green) excess coordination numbers, ΔN_{ij}, as a function of the correlation distance R for a benzene−water mixture at a concentration of 0.6 M. Results from (a) atomistic, (b) CG(IBI), and (c) CG(PMF) models. The inset in part a shows an atomistic snapshot: benzene molecules in yellow and water molecules in blue.

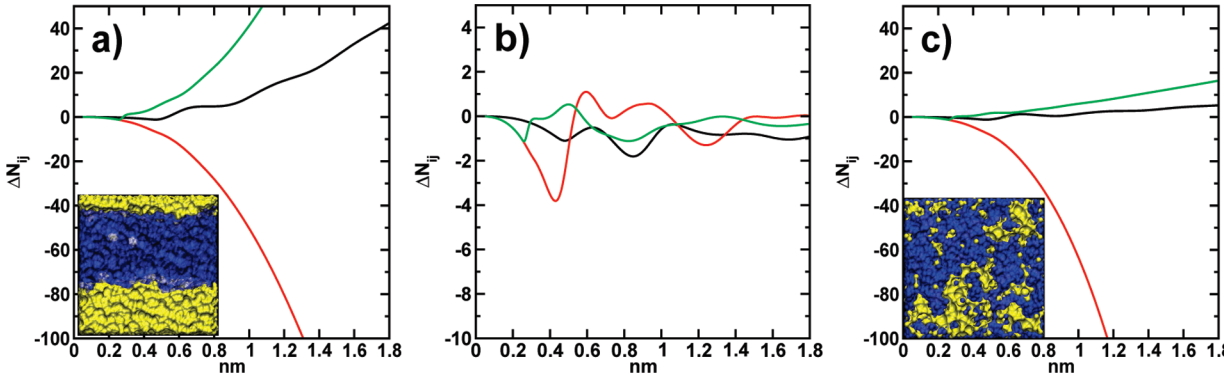


Figure 6. Benzene−benzene (black), benzene−water (red), and water−water (green) excess coordination number, ΔN_{ij}, for the benzene−water systems at a concentration of 9.5 M. Results from the (a) atomistic model, (b) CG(IBI) model, and (c) CG(PMF) model. Insets show an (a) atomistic and (c) CG(PMF) snapshot: benzene molecules in yellow and water molecules in blue.

ably well, despite the fact that it uses the same single-site water model. The CG(PMF) potential is obtained (see eq 3) by removing indirect correlations, mediated by the single-site coarse-grained water model, from the benzene−benzene pair potential of mean force in water. But since this pair potential of mean force has been obtained with a detailed atomistic model that describes water−water hydrogen bonding explicitly, the application of eq 3 yields an effective pair potential which retains the contribution of hydrophobic attraction between the solutes.

Hydrophobic association, observed with the detailed atomistic model, occurs on account of multibody correlations in the solvent (hydrogen bonding). The CG(PMF) model (eq 3) includes these multibody correlations in an *effective* pair potential, while the CG(IBI) model does not include them. The difference between these two potentials, shown in Figure 1c, indicates that the "hydrophobic contribution" has a short-range nature and a magnitude at least as large as the van der Waals interaction which the CG(IBI) model represents. This hydrophobic contribution for benzene association qualitatively compares well with the estimated strength of hydrophobic interaction between benzene solutes, as discussed by Pratt and Chandler.[65] Recent work by Wu and Prausnitz[66] on hydration free energies of short alkanes has indicated that hydrophobic interactions can be described with a pairwise additive potential; i.e., by considering the alkanes as hydrophobic associates of methylene units, it could be shown that only pairwise hydrophobic forces contribute to the free energy of hydrophobic association. The results by Wu and Prausnitz indicate that one pair potential is capable of describing hydrophobic associates with different chain lengths. In an aqueous solution of benzene, the B−B pair potential describing hydrophobic association at a low benzene concentration may therefore potentially also be used to describe hydrophobic association at larger benzene concentration. This idea is supported by the data shown in Figure 3b. This figure shows changes of the benzene activity coefficient obtained with the all-atom model that includes multibody correlations mediated by the solvent, as well as the results obtained with the CG(PMF) model in which these multibody contributions are effectively accounted for in the B−B pair potential. The observed transferability of the CG(PMF) model supports the notion that hydrophobic interactions are pairwise additive up to the concentration where the solution becomes unstable. Figure 7 shows CG and atomistic simulation snapshots of the solution at a benzene concentration just before the point where instability sets in and shows chainlike benzene clusters with up to three or four benzene molecules. The model however is not transferable to conditions where extended hydrophobic surfaces start to play a role. In such situations, where both small and large length scales are relevant, hydrophobic forces

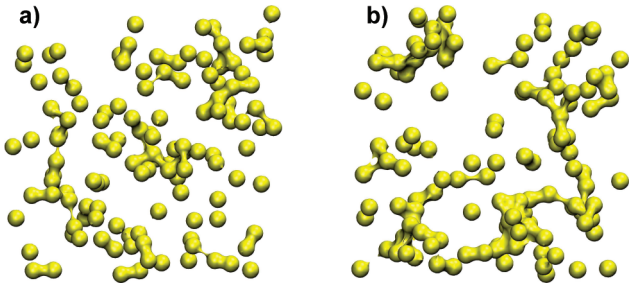


Figure 7. (a) Atomistic and (b) CG(PMF) snapshots for benzene solution at 0.5 M. For clarity, only benzene beads are shown. For the atomistic snapshots, the benzene molecules are described by the corresponding beads. The snapshots of cluster surfaces were obtained by VMD[68] using a rolling probe of 3 Å radius.

cease to be additive.[67] As evidenced by Figure 6, the CG(PMF) model does not form two well-defined phases separated by an interface under conditions where it should.

Figure 8 shows the RDFs for the AA, CG(IBI), and CG(PMF) models at two representative benzene concentrations. The WW and BW RDFs obtained with the CG(IBI) and CG(PMF) model agree very well with the AA model. The BB RDFs show some discrepancies. For example, the height of the first peak obtained with the CG(IBI) model differs clearly from that of the AA model. An important point to be made, however, is that the changes in the RDFs with concentration are fairly small. Therefore, little can be said about model transferability by investigating the RDFs. In contrast to the small differences observed in the RDFs, the KB analysis provides a much clearer picture and shows that the CG(PMF) model is perfectly transferable to other solution concentrations.

7. Conclusions

We have developed two CG models for binary mixtures of benzene and water. Both CG models are developed on the basis of atomistic reference data generated with molecular simulations, describe benzene and water molecules explicitly, but differ in the benzene−benzene interaction potential. In the CG(IBI) model, the benzene−benzene interaction potential is obtained by reproducing the structural properties of a pure liquid benzene system, while in the CG(PMF) model the parametrization has been targeted on two benzene molecules in aqueous solution at infinite dilution. Both the applied models are "structure-based" since they are derived from a benzene−benzene potential of mean force in liquid benzene or in water solution. We investigate the transferability of those CG potentials to solution compositions different from the one used in the parametrization. To this end, we use the Kirkwood−Buff theory of solutions which provides a link between the local structure of the solutions and the influence of global composition changes on chemical potentials and activity coefficients of the solution components.

We show that the CG(IBI) model is not well-suited to represent benzene in dilute aqueous solution. The model does not favor benzene aggregation as observed in detailed-atomistic simulations. We attribute this observation to the single-site CG water model, which lacks the hydrogen bonding properties needed to account for hydrophobicity, and the effective B−B pair potential, which contains only the contributions of van der Waals interactions. The CG(PMF) model correctly describes the benzene/water mixture at different benzene concentrations, as well as the structure of pure benzene solution in agreement with the atomistic model. The parametrization of this model has been based on all-atom simulations of two benzene solutes in water, which introduces hydrophobic interaction in the effective pair potential.

The Kirkwood−Buff analysis shows that the CG(PMF) potential can be used to describe benzene/water mixtures up to concentrations where the solution becomes unstable and phase separation occurs with the atomistic model. At all concentrations where the solution is stable, the agreement with the reference atomistic result is very good. At concentrations where phase separation occurs in the atomistic simulation, the solution simulated with the CG(PMF) model shows a tendency to become unstable as well, although it does not completely phase-separate, while with the CG(IBI) model the solution remains artificially stable.

The present work not only shows that carefully designed structure-based CG models are to some extent transferable

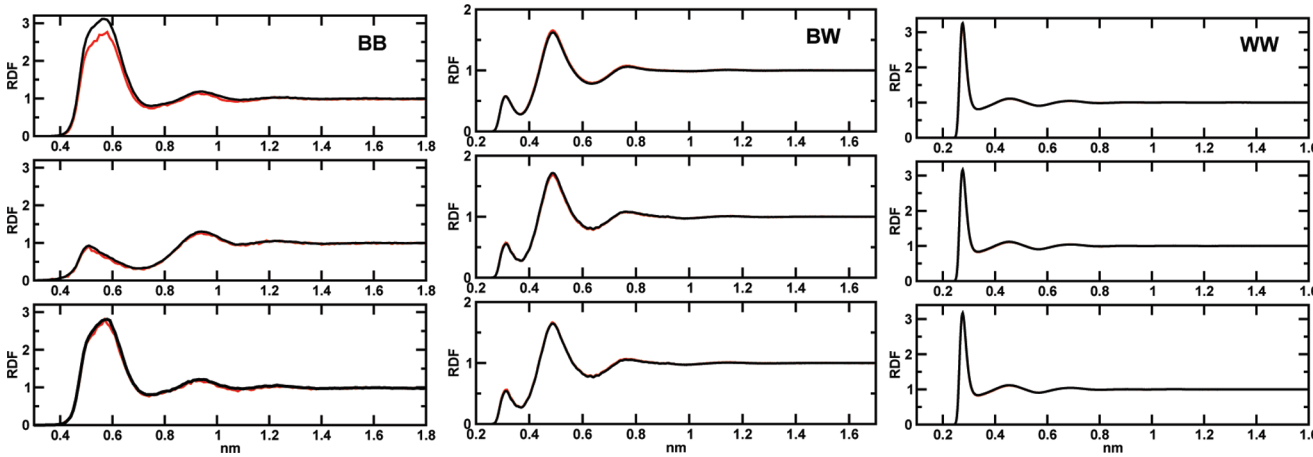


Figure 8. Molecular radial distribution function for a pair of molecules in a benzene−water mixture at different concentrations: 0.5 M (black line) and 0.2 M (red line). Results from atomistic (top), CG(IBI) (middle), and CG(PMF) (bottom) simulations.

to different concentrations, it also shows that models derived to reproduce structures of liquid mixtures or solutions are not by construction incapable of reproducing thermodynamic properties of those systems as well. By applying KB theory to our atomistic and CG simulations, we not only analyze the mixing properties of the different systems, we also illustrate in which way structural properties and thermodynamic properties of solutions are linked. KB theory provides a link between *local* solution structure and changes in the chemical potential and the activity coefficient that occur after a change in the *global* mixture composition. From a methodological perspective, KB theory, therefore, is ideally suited to analyze the transferability of CG potentials to different state points (global solution compositions). This transferability analysis works particularly well for structure-based CG potentials since in this case KB integrals of atomistic and CG simulation match at the reference state by construction.

The consistency of a CG model in terms of both structure and thermodynamics is essential if one studies biomolecular aggregation, where both structural arrangement and partitioning, hydrophobicity, or hydrophilicity are key properties. A certain level of transferability of the simulation model is also important since these processes are characterized by fluctuating and spatially varying concentrations.

Acknowledgment. We would like to thank Emiliano Brini and Jia-Wei Shen for carefully reading the manuscript. C.P. acknowledges financial support by the German Science Foundation within the Emmy Noether Programme (grant PE1625/1-1).

References

(1) Silbermann, J.; Klapp, S. H. L.; Schoen, M.; Chennamsetty, N.; Bock, H.; Gubbins, K. *J. Chem. Phys.* **2006**, *124*, 074105.

(2) Liu, P.; Izvekov, S.; Voth, G. A. *J. Phys. Chem. B* **2007**, *111*, 11566–11575.

(3) Johnson, M. E.; Head-Gordon, T.; Louis, A. A. *J. Chem. Phys.* **2007**, *126*, 144509.

(4) Qian, H.; Carbone, P.; Chen, X.; Karimi-Varzaneh, H.; Liew, C.; Muller-Plathe, F. *Macromolecules* **2008**, *41*, 9919.

(5) Allen, E. C.; Rutledge, G. C. *J. Chem. Phys.* **2008**, *128*, 154115.

(6) Fischer, J.; Paschek, D.; Geiger, A.; Sadowski, G. *J. Phys. Chem. B* **2008**, *112*, 13561–13571.

(7) Carbone, P.; Varzaneh, H. A. K.; Chen, X.; Muller-Plathe, F. *J. Chem. Phys.* **2008**, *128*, 064904.

(8) Mullinax, J.; Noid, W. *J. Chem. Phys.* **2009**, *131*, 104110.

(9) Sanghi, T.; Aluru, N. R. *J. Chem. Phys.* **2010**, *132*, 044703.

(10) Muller-Plathe, F. *ChemPhysChem* **2002**, *3*, 754–769.

(11) Tozzini, V. *Curr. Opin. Struct. Biol.* **2005**, *15*, 144–150.

(12) van der Vegt, N. F. A.; Peter, C.; Kremer, K. In *Coarse-Graining of Condensed Phase and Biomolecular Systems*; Chapman and Hall/CRC Press, Taylor and Francis Group: New York, 2008; Chapter: Structure-based coarse- and fine-graining in soft matter simulations.

(13) Murtola, T.; Bunker, A.; Vattulainen, I.; Deserno, M.; Karttunen, M. *Phys. Chem. Chem. Phys.* **2009**, *11*, 1869–1892.

(14) Peter, C.; Kremer, K. *Soft Matter* **2009**, *5*, 4357–4366.

(15) Rühle, V.; Junghans, C.; Lukyanov, A.; Kremer, K.; Andrienko, D. *J. Chem. Theory Comput.* **2009**, *5*, 3211–3223.

(16) Lyubartsev, A. P.; Laaksonen, A. *Phys. Rev. E* **1995**, *52*, 3730–3737.

(17) Soper, A. *Chem. Phys.* **1996**, *202*, 295–306.

(18) Reith, D.; Pütz, M.; Müller-Plathe, F. *J. Comput. Chem.* **2003**, *24*, 1626–1636.

(19) Peter, C.; Delle Site, L.; Kremer, K. *Soft Matter* **2008**, *4*, 859–869.

(20) McCoy, J. D.; Curro, J. C. *Macromolecules* **1998**, *31*, 9362–9368.

(21) Zacharopoulos, N.; Vergadou, N.; Theodorou, D. *J. Chem. Phys.* **2005**, *122*, 244111.

(22) Hess, B.; Holm, C.; van der Vegt, N. F. A. *Phys. Rev. Lett.* **2006**, *96*, 147801.

(23) Fritz, D.; Harmandaris, V. A.; Kremer, K.; van der Vegt, N. F. A. *Macromolecules* **2009**, *42*, 7579–7588.

(24) Shelley, J.; Shelley, M.; Reeder, R.; Bandyopadhyay, S.; Klein, M. *J. Phys. Chem. B* **2001**, *105*, 4464–4470.

(25) Nielsen, S. O.; Lopez, C. F.; Srinivas, G.; Klein, M. L. *J. Chem. Phys.* **2003**, *119*, 7043.

(26) Marrink, S. J.; de Vries, A. H.; Mark, A. E. *J. Phys. Chem. B* **2004**, *108*, 750–760.

(27) Bond, P. J.; Sansom, M. S. P. *J. Am. Chem. Soc.* **2006**, *128*, 2697–2704.

(28) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. *J. Phys. Chem. B* **2007**, *111*, 7812–7824.

(29) Shinoda, W.; Devane, R.; Klein, M. *Mol. Simul.* **2007**, *33*, 27.

(30) Michel, J.; Orsi, M.; Essex, J. W. *J. Phys. Chem. B* **2008**, *112*, 657–660.

(31) Derreumaux, P.; Mousseau, N. *J. Chem. Phys.* **2007**, *126*, 025101.

(32) Brown, S.; Fawzi, N. J.; Head-Gordon, T. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 10712–10717.

(33) Izvekov, S.; Voth, G. A. *J. Phys. Chem. B* **2005**, *109*, 2469–2473.

(34) Zhou, J.; Thorpe, I. F.; Izvekov, S.; Voth, G. A. *Biophys. J.* **2007**, *92*, 4289–4303.

(35) Allen, E. C.; Rutledge, G. C. *J. Chem. Phys.* **2009**, *130*, 034904.

(36) Mognetti, B. M.; Virnau, P.; Yelash, L.; Paul, W.; Binder, K.; Mueller, M.; MacDowell, L. G. *J. Chem. Phys.* **2009**, *130*, 044101.

(37) Murtola, T.; Karttunen, M.; Vattulainen, I. *J. Chem. Phys.* **2009**, *131*, 055101.

(38) Kirkwood, J.; Buff, F. *J. Chem. Phys.* **1951**, *19*, 774–777.

(39) Ben-Naim, A.; Navarro, A. M.; Leal, J. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2451–2460.

(40) Henderson, R. L. *Phys. Lett. A* **1974**, *49*, 197–198.

(41) Jain, S.; Garde, S.; Kumar, S. K. *Ind. Eng. Chem. Res.* **2006**, *45*, 5614–5618.

(42) Lyubartsev, A. P.; Mirzoev, A.; Chen, L.; Laaksonen, A. *Faraday Discuss.* **2010**, *144*, 43–56.

(43) Villa, A.; van der Vegt, N. F. A.; Peter, C. *Phys. Chem. Chem. Phys.* **2009**, *11*, 2077–2086.

(44) Ben-Naim, A. *Molecular Theory of Solutions*; Oxford Univ. Press: New York, 2006.

(45) Weerasinghe, S.; Smith, P. E. *J. Phys. Chem. B* **2003**, *107*, 3891–3898.

(46) Weerasinghe, S.; Smith, P. E. *J. Chem. Phys.* **2003**, *119*, 11342–11349.

(47) Lee, M. E.; van der Vegt, N. F. A. *J. Chem. Phys.* **2005**, *122*, 114509.

(48) Hess, B.; van der Vegt, N. F. A. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 13296–13300.

(49) Klasczyk, B.; Knecht, V. *J. Chem. Phys.* **2010**, *132*, 024109.

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

(51) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269–6271.

(52) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. *Comput. Phys. Commun.* **1995**, *91*, 43–56.

(53) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.

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

(55) Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981**, *52*, 7182–7190.

(56) Nosé, S. *Mol. Phys.* **1984**, *52*, 255–268.

(57) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695–1697.

(58) Miyamoto, S.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 952–962.

(59) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. *J. Comput. Chem.* **1997**, *18*, 1463–1472.

(60) Hess, B. *J. Chem. Theory Comput.* **2008**, *4*, 116.

(61) van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Simul.* **1988**, *1*, 173–185.

(62) Wang, H.; Junghans, C.; Kremer, K. *Eur. Phys. J.* **2009**, *28*, 221–229.

(63) *CRC Handbook of Chemistry and Physics*, 89th ed.; Lide, D., Ed.; CRC Press: Boca Raton, FL, 2008.

(64) Schravendijk, P.; van der Vegt, N. F. A. *J. Chem. Theory Comput.* **2005**, *1*, 643–652.

(65) Pratt, L. R.; Chandler, D. *J. Chem. Phys.* **1977**, *67*, 3683–3704.

(66) Wu, J.; Prausnitz, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 9512–9515.

(67) Chandler, D. *Nature* **2005**, *437*, 640–647.

(68) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33–38.

CT100228T

JCTC Journal of Chemical Theory and Computation

Theoretical Study of the Dual Fluorescence of 4-(*N*,*N*-Dimethylamino)benzonitrile in Solution

Ignacio Fdez. Galván,* M. Elena Martín, and Manuel A. Aguilar

Química Física, Edif. José María Viguera Lobo, Universidad de Extremadura, Avda. de Elvas s/n, 06071 Badajoz, Spain

Received December 15, 2009

Abstract: We have performed high-level quantum calculations of absorption and emission properties of 4-(*N*,*N*-dimethylamino)benzonitrile (DMABN) in gas phase and in polar solvents, including the solvent effects with an explicit mean field model. Two excited states of DMABN have been found and optimized, corresponding to the two excited states assumed by the generally accepted hypotheses for the dual fluorescence of this molecule: a locally excited (LE) state and a charge transfer (CT) state. The results show that, in the gas phase, the charge transfer state is severely distorted and higher in energy than the locally excited state, while in polar solvents, it becomes almost an ideal twisted intramolecular charge transfer state (TICT) and is stabilized with respect to the locally excited state and the Franck−Condon point. The relative free energies calculated for the two excited states in solution suggest that both states are accessible from initial Franck−Condon excitation, the charge transfer state being highly favored in aqueous solution and less so in tetrahydrofuran. The results support the validity of the TICT model in this system.

1. Introduction

Since the discovery of dual fluorescence in 4-(*N*,*N*-dimethylamino)benzonitrile (DMABN), 50 years ago,[1] a great number of publications have been devoted to the investigation and explanation of its properties, as well as to similar or analogous compounds. In gas phase and in nonpolar solvents, DMABN has a single normal fluorescence band. When placed in polar solvents, it exhibits two fluorescence bands, which is an uncommon feature in fluorescence spectra. The high-energy fluorescence band corresponds to the band found in nonpolar solvents, and a new red-shifted band appears in polar solvents. Lippert et al.[1] attributed the additional band to an intramolecular charge transfer state, stabilized by the polar solvent. Grabowski et al.[2] suggested that this charge transfer state would need a twist of the dimethylamino group with respect to the phenyl ring, leading to what is known as a twisted intramolecular charge transfer (TICT) state. The participation of a TICT state is probably the most widely accepted model for the observed dual fluorescence of DMABN, but it is not the only existing explanation, a summary of other hypotheses can be found in the review by Grabowski, Rotkiewicz, and Rettig.[3] In addition to the TICT model, the participation of a planar intramolecular charge transfer state (PICT), with the dimethylamino group in the same plane as the phenyl ring, and a rehybridized intramolecular charge transfer state (RICT), with a nonlinear nitrile group, have received the most attention in recent years.

Different experimental studies[4−10] have suggested the existence of an equilibrium between the two emitting excited states, the "normal" locally excited (LE) state and the charge transfer (CT) state. Theoretical calculations, in general, support the TICT model and the interconversion between the LE and CT states as well,[8,11−18] particularly calculations that include the solvent effects, such as the work of Mennucci et al. using continuum solvation[19] or that of Minezawa and Kato with reference interaction site model (RISM) solvation.[20]

Despite the abundance of theoretical and experimental studies, the exact nature of the CT state is still under controversy. The comparison of experimental data of DMABN

* To whom correspondence should be addressed. E-mail: jellby@unex.es.

10.1021/ct9006713 © 2010 American Chemical Society
Published on Web 07/01/2010

and analogous compounds, with different geometric constraints and substituents, usually leads to various conclusions regarding the structure of the CT state. Compounds, such as 3,5-dimethyl-4-(*N*,*N*-dimethylamino)benzonitrile, where the dimethylamino group is forced to be twisted, show only the CT fluorescence band, suggesting a TICT, while other compounds where the twisting is hindered (like 6-cyano-1,2,3,4-tetrahydroquinoline, NTC6) can also present dual fluorescence, suggesting a PICT. As mentioned above, theoretical calculations tend to favor the TICT model for the DMABN molecule. For instance, the already mentioned study of Minezawa and Kato[20] ruled out the participation of a PICT state based on the calculated fluorescence energies.

In this work, we have investigated the structure and absorption and emission properties of DMABN in gas phase and in polar solvents. We considered an aprotic solvent of moderate polarity, tetrahydrofuran (THF), and a highly polar and protic solvent, water. There is experimental data on both solvents, and they will let us examine the influence that a protic solvent may have on the absorption and emission properties of DMABN. High-level quantum calculations have been performed on the ground state and two excited states (LE and CT), and the solvent effect has been included with an explicit model implementing a mean field approximation. By studying the structures and energies of the electron states presumedly implied in the dual fluorescence phenomenon, we expect to cast further light on the validity of the TICT model for this system. When appropriate, we will also compare the results obtained in this work with those of a previous study on a similar system: *N*,*N*-dimethylaniline (DMA).[21]

The rest of this paper is structured as follows. In section 2, the computational methods and details are set out; section 3 presents and discusses the results obtained from the calculations, in gas phase and in solution, and finally, section 4 contains the main conclusions we can draw from this work.

2. Methods and Details

Solvent effects on the DMABN UV/vis spectra were calculated with the ASEP/MD (Average Solvent Electrostatic Potential from Molecular Dynamics) method. This is a sequential quantum mechanics/molecular mechanics (QM/MM) method implementing the mean field approximation. It combines, alternately, a high-level quantum mechanics (QM) description of the solute with a classical molecular mechanics (MM) description of the solvent. One of its main features is the fact that the solvent effect is introduced into the solute's wave function as an average perturbation. Details of the method have been described in previous papers,[22−24] so here we will only present a brief outline.

As mentioned above, ASEP/MD is a method combining QM and MM techniques, with the particularity that full QM and MD (molecular dynamics) calculations are alternated and not simultaneous. During the MD simulations, the intramolecular geometry and charge distribution of all molecules is considered as fixed. From the resulting simulation data, the average electrostatic potential generated by the solvent on the solute (ASEP) is obtained. This potential is introduced as a perturbation into the solute's quantum mechanical Hamiltonian, and by solving the associated Schrödinger equation, one gets a new charge distribution for the solute, which is used in the next MD simulation. This iterative process is repeated until the electron distribution of the solute and the solvent structure around it are mutually equilibrated.

The ASEP/MD framework can also be used to optimize the geometry of the solute molecule.[25] At each step of the ASEP/MD procedure, the gradient and Hessian on the system's free-energy surface (including the van der Waals contribution) can be obtained, and thus, they can be used to search for stationary points on this surface by some optimization method. After each MD simulation, the solute geometry is optimized within the fixed "average" solvent structure by using the free-energy derivatives. In the next MD simulation, the new solute geometry and charge distribution are used. This approach allows the optimization of the solute geometry in parallel to the solvent structure.

For calculating transition energies, the iterative process is performed on the initial state of the transition (the ground state for absorption, the excited state for emission), that is, the atomic charges for the MD and the energy derivatives for the geometry optimization of the solute are calculated with the initial state's wave function. Then, with a frozen solvent model, the transition energies between the different states are obtained. It is also possible to calculate transition energies with a polarizable solvent model. In this case, once the solute and solvent structure have been optimized for the initial state of the solute, each state energy and wave function is calculated with the same solvent structure, but where the solvent molecules' charges are replaced by gas-phase charges plus a molecular polarizability, and an additional iterative polarization process is performed.[26,27] In this work, we used a nonpolarizable solvent model in all cases, as previous test calculations with a similar system and polarizable solvent did not show an important enough influence to compensate for the increased computational effort required.

With the transition energies calculated in solution (ΔE) and in gas phase (ΔE^0), the solvent shift δ can be obtained as the difference:

$$\begin{aligned}\delta &= \Delta E - \Delta E^0 \\ &= (\langle\Psi_{ex}|\hat{H}_{QM} + \hat{V}|\Psi_{ex}\rangle - \langle\Psi_g|\hat{H}_{QM} + V|\Psi_g\rangle) - \\ &\quad (\langle\Psi^0_{ex}|\hat{H}^0_{QM}|\Psi^0_{ex}\rangle - \langle\Psi^0_g|\hat{H}^0_{QM}|\Psi^0_g\rangle) \\ &= (\langle\Psi_{ex}|\hat{H}_{QM} + \hat{V}|\Psi_{ex}\rangle - \langle\Psi^0_{ex}|\hat{H}^0_{QM}|\Psi^0_{ex}\rangle) - \\ &\quad (\langle\Psi_g|\hat{H}_{QM} + \hat{V}|\Psi_g\rangle - \langle\Psi^0_g|\hat{H}^0_{QM}|\Psi^0_g\rangle)\end{aligned} \quad (1)$$

where the subindices ex and g denote the excited and ground state state of the transition, $\hat{H}_{QM}$ is the QM Hamiltonian of the solute at the in-solution geometry, without the solute−solvent interaction, $\hat{V}$, and $\hat{H}^0_{QM}$ is the QM Hamiltonian at the gas-phase geometry; Ψ and Ψ^0 are the wave functions optimized in solution and in gas phase, respectively. This solvent shift can be partitioned in different contributions, namely a geometry contribution δ_{geo}, an electronic distortion contribution δ_{dist}, and an electrostatic solute−solvent contribution δ_{elec}. If we introduce Ψ' as the wave function

optimized for the $\hat{H}_{QM}$ Hamiltonian (the geometry in solution, but without solute−solvent interaction)

$$\begin{aligned}
\delta &= \delta_{geo} + \delta_{dist} + \delta_{elect} \\
\delta_{geo} &= (\langle\Psi'_{ex}|\hat{H}_{QM}|\Psi'_{ex}\rangle - \langle\Psi^0_{ex}|\hat{H}^0_{QM}|\Psi^0_{ex}\rangle) - \\
&\quad (\langle\Psi'_{g}|\hat{H}_{QM}|\Psi'_{g}\rangle - \langle\Psi^0_{g}|\hat{H}^0_{QM}|\Psi^0_{g}\rangle) \\
\delta_{dist} &= (\langle\Psi_{ex}|\hat{H}_{QM}|\Psi_{ex}\rangle - \langle\Psi'_{ex}|\hat{H}_{QM}|\Psi'_{ex}\rangle) - \\
&\quad (\langle\Psi_{g}|\hat{H}_{QM}|\Psi_{g}\rangle - \langle\Psi'_{g}|\hat{H}_{QM}|\Psi'_{g}\rangle) \\
\delta_{elect} &= \langle\Psi_{ex}|\hat{V}|\Psi_{ex}\rangle - \langle\Psi_{g}|\hat{V}|\Psi_{g}\rangle
\end{aligned} \tag{2}$$

Thus, δ_{geo} is the solvent shift resulting from the change in geometry between gas phase and solution, δ_{elec} corresponds to the difference in solute−solvent interaction energy between the excited and ground states, and δ_{dist} to the difference in the wave function distortion energy. Note that the van der Waals component of the interaction energy is not included in the above expressions, since we adopt the approximation of considering it constant for all electronic states of the solute, and therefore, it vanishes when vertical transition energies are considered.

Once the different solute electronic states and the solvent structure around them have been optimized and equilibrated, the free energy differences between those states can be calculated, within the ASEP/MD framework, making use of the free energy perturbation method.[28,29] The expression we use to calculate the free energy difference between two species in equilibrium in solution, ΔG is

$$\Delta G = \Delta E + \Delta G_{int} + \Delta V \tag{3}$$

where ΔE is the difference in the internal quantum energy of the solute between the two species, ΔG_{int} is the difference in the solute−solvent interaction energy, which is calculated classicaly with the free energy perturbation method, and ΔV is a term that includes the difference in the zero point energy (ZPE) and entropic contributions of the solute. The last term, ΔV, is normally evaluated by applying the harmonic approximation to the vibrational modes of the solute in solution, and it needs the information provided by the Hessian matrix. In this work, obtaining an accurate enough Hessian matrix required too large computational resources, and we decided to approximate the results by neglecting this term. It must be noted that this ΔV term refers only to the internal nuclear degrees of freedom of the solute, free energy contributions from the solvent around the solute are properly accounted for in the ΔG_{int} term. Minezawa and Kato[20] estimated ZPE contributions to be typically lower than 1 kcal/mol.

The quantum calculations of the solute molecule were done with the complete active space self-consistent field (CASSCF) method,[30] using the 6-311G** basis set. Gas phase calculations were also done with the cc-pVTZ basis set. The active orbitals were the six π and π^* orbitals of the phenyl ring, the four π and π^* of the nitrile group, and the nonbonded orbital of the nitrogen, and 12 electrons were included in these orbitals, for a (12,11) total active space. Geometry optimizations in gas phase and in solution were performed except when noted on pure roots (the ground state, S_0, or the first excited singlet state, S_1), but transition energies were always calculated with a state-average (SA) calculation of the first 3 singlet states at the given geometry, S_0, S_1, and S_2. To obtain accurate transition energies, it is known that the inclusion of dynamic correlation in the quantum calculations is necessary, which we did with the complete active space second-order perturbation (CASPT2) method,[31,32] using the SA-CASSCF(12,11) wave functions as reference. A new IPEA (ionization potential - electron affinity) shifted zeroth-order Hamiltonian has been recently proposed for CASPT2 calculations,[33] which is supposed to reduce systematic overstabilization errors in open-shell systems (as is the case of the excited states studied here). We did all CASPT2 with the proposed IPEA shift of 0.25 E_h (CASPT2(0.25)), as well as with no IPEA shift (CASPT2(0.00)). To minimize the appearance of intruder states, an additional imaginary shift of 0.1i E_h was used. No symmetry was assumed in any case.

The MD simulations were carried out with rigid molecules; THF and water were used as solvents. Lennard-Jones parameters and solvent atomic charges were taken from the OPLS-AA (optimized potentials for liquid simulations, all atoms) force field,[34] solute atomic charges were calculated from the quantum calculations with the CHELPG (charges from electrostatic potential in a grid) method.[35] The geometry of tetrahydrofuran was optimized with the Becke's three-parameter Lee−Yang−Parr density functional (B3LYP)/6-311G**, for water the TIP3P model was employed. A total of 215 THF molecules or 800 water molecules, and the solute were included at the experimental density of the solvent. Periodic boundary conditions were applied, and spherical cutoffs were used to truncate the interatomic interactions at 12 Å; long-range interactions were calculated using the Ewald sum technique. The temperature was fixed at 298 K by using the Nosé−Hoover thermostat. A time step of 0.5 fs was used during the simulations, and each one was run for 100 ps after 25 ps equilibriation.

At each step of the ASEP/MD procedure, 500 configurations evenly distributed from the MD run were used to calculate the ASEP, and a radius of 15 bohr for THF and 12 bohr for water was used for including explicit solvent charges. Each ASEP/MD run was continued until the energies and solute geometry and charges are stabilized for at least 5 iterations; results are reported as the average of these last 5 iterations.

For in solution calculations, the ASEP/MD software[23] was used. During the ASEP/MD runs, quantum calculations (CASSCF optimizations) were performed with the Gaussian 98 package.[36] The final SA-CASSCF and CASPT2 calculations were done with Molcas 6.4.[37] All MD simulations were performed using Moldy.[38]

3. Results and Discussion

3.1. Gas Phase. The geometry of DMABN was optimized in gas phase at CASSCF(12,11)/6-311G** level. The resulting structure, similar to the case of *N*,*N*-dimethylaniline (DMA),[21] features a pyramidal amine N atom, slightly above the plane of phenyl ring, with the CH_3 groups below the plane forming a CH_3-N-CH_3 angle of 115.84°, and a

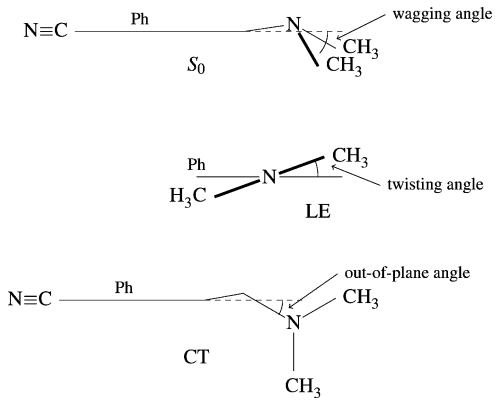


Figure 1. Scheme showing the wagging, twisting, and out-of plane angles in the ground and excited states of DMABN.

wagging angle (the angle between the CH_3-N-CH_3 plane and the phenyl ring plane) of 24.48°. This pyramidal structure and the angle values qualitatively agree with the experimental estimates[39] of 115.7° and 15°, respectively.

The first excited state, S_1, was also optimized in gas phase and, again analogously to DMA, a planar local structure on the amine N atom was obtained, with a CH_3-N-CH_3 angle of 116.91° and a twisting angle of the CH_3-N-CH_3 plane with respect to the phenyl ring of 15.53°. The experimental estimate of this angle is rather higher, 30°, but it includes significant uncertainty.[39] The bonds of the phenyl ring become longer, around 1.43 Å (they are typically 1.40 Å in the ground state), indicating a loss of aromatic character. Conversely, the bonds between the phenyl ring and the amine nitrogen and the nitrile carbon become 0.02 and 0.01 Å shorter.

Although there is no experimental evidence of emission from a CT state in gas phase, we could also optimize an excited state with high charge transfer character and perpendicular CH_3-N-CH_3 and ring planes. This state, according to the TICT model, is responsible for the dual fluorescence of DMABN observed in polar solvents. In gas phase, the torsion of the CH_3-N-CH_3 by 90° around the C−N bond is accompanied by a significant distortion of the ring, which is no longer planar (the C_4 atom moves out of the plane) and a rotation of the whole $N(CH_3)_2$ moiety, on its plane, around the C_4 atom, as displayed in Figure 1. This ring distortion agrees with the results of coupled cluster CC2 calculations.[16] The charge transfer character is reflected on the ring bond lengths, which becomes quinoidal (1.37 Å for the C_2-C_3 and C_5-C_6 bonds, 1.43 Å for the others); the phenyl−nitrogen bond is lengthened as well, from 1.37 to 1.46 Å (identical to the $N-CH_3$ bonds), while the phenyl−nitrile bond is shortened from 1.44 to 1.42 Å, which is further evidence of the decoupling between the dimethylamino and benzonitrile groups. At this distorted geometry, this state is the first excited state, and from now onward, to distinguish this state from the one in the previous paragraph, when referring to emission, we will call this state CT and the other state LE.

The absorption energies calculated at the optimized S_0 geometry are collected in Table 1. Transition energies to the first two excited states were obtained at three calculation levels (SA-CASSCF, CASPT2(0.25), and CASPT2(0.00)) and two basis sets (6-311G** and cc-pVTZ). It can be seen that SA-CASSCF values are significantly overestimated, especially for the $S_0 \rightarrow S_2$ transition, and hardly dependent on the basis set; the inclusion of dynamic correlation in CASPT2 is clearly needed to correctly reproduce even the qualitative energies of the first two absorption bands. Both CASPT2 variants, with the two triple-ζ basis sets tried, agree in providing a separation between the two absorption bands of around 0.3 eV, in good accord with the experimental values, and also in agreement with the experiments are the oscillator strengths, predicting a weak band for the almost forbidden $S_0 \rightarrow S_1$ transition and a much stronger absorption for $S_0 \rightarrow S_2$.

When comparing the two CASPT2 methods, we can observe that CASPT2(0.25) values are higher than CASPT2(0.00) ones, and all values slightly decrease when the basis set quality is improved. The same discrepancy and trend in the two methods was found in a previous study of DMA,[21] which has similar absorption properties. When the 0.25 E_h IPEA shift was proposed,[33] it was suggested that it would reduce systematic errors in open-shell systems. Our experience, however, is that CASPT2(0.00) transition energies may agree better with experiment, while CASPT2(0.25) overestimates them. We attribute this apparently poorer performance of CASPT2(0.25) to a cancellation of errors in the CASPT2(0.00) calculations: the limited basis sets and active spaces would tend to overestimate excited state energies, but this is compensated by the overstabilization of open-shell systems inherent to CASPT2(0.00). When CASPT2(0.25) is used in these cases, only the basis set and active space errors remain, and the result is an overestimation of the transition energies. In any case, for a given system and electron transition, the difference between the values of the two methods is quite constant, so that trends like solvatochromic shifts can be confidently evaluated from either of them.

In Table 2, we show the calculated band origins (energy difference between the S_0 and LE minima) and the emission energies from the two optimized excited states, LE and CT. Again, the inclusion of dynamical electron correlation improves the results, and CASPT2(0.00) gives consistently lower values for all transitions than CASPT2(0.25), while the latter overestimates the experimental data, and all values decrease with increasing basis set quality. The difference between the two transitions experimentally available is well reproduced (around 0.2 eV).

Unfortunately, there is no experimental evidence of a CT $\rightarrow S_0$ fluorescence in gas phase. The estimations in nonpolar solvents place the maximum at around 3.2 eV.[41] If this value could be transferred to the gas phase situation, it would make the CASPT2(0.00)/6-311G** result underestimated by about 0.4 eV; however, this value must be taken with caution because the above estimation is based on a very weak band that is completely buried into the more intense LE fluorescence and is, therefore, affected by a high uncertainty. Furthermore, the CC2 calculations of Köhn and Hättig,[16] where a similar geometry for the CT state was obtained, yielded also a low value for the fluorescence energy, 2.49 eV. The molecular geometries and transition energies are,

Table 1. Absorption Energies Calculated in Gas Phase (eV)[a]

	SA-CASSCF		CASPT2(0.25)		CASPT2(0.00)	
	$S_0 \rightarrow S_1$	$S_0 \rightarrow S_2$	$S_0 \rightarrow S_1$	$S_0 \rightarrow S_2$	$S_0 \rightarrow S_1$	$S_0 \rightarrow S_2$
6-311G**	4.74	6.24	4.58 (0.006)	4.90 (0.472)	4.17 (0.006)	4.46 (0.430)
cc-pVTZ	4.77	6.23	4.54 (0.006)	4.82 (0.465)	4.08 (0.005)	4.34 (0.418)
experimental[40]			4.25 (weak)	4.56 (strong)		

[a] Oscillator strength in parentheses.

Table 2. Band Origins and Fluorescence Energies Calculated in Gas Phase (eV)[a]

	SA-CASSCF			CASPT2(0.25)			CASPT2(0.00)		
	0−0	LE → S_0	CT → S_0	0−0	LE → S_0	CT → S_0	0−0	LE → S_0	CT → S_0
6-311G**	4.55	4.32	4.10	4.33	4.15(0.012)	3.25 (0.006)	3.91	3.75 (0.010)	2.77 (0.006)
cc-pVTZ	4.57	4.34	4.10	4.29	4.11 (0.012)	3.21 (0.006)	3.82	3.66 (0.011)	2.69 (0.006)
experimental[42]				4.00	3.76	3.2[b]			

[a] Oscillator strength in parentheses. [b] Estimation in cyclohexane.[41]

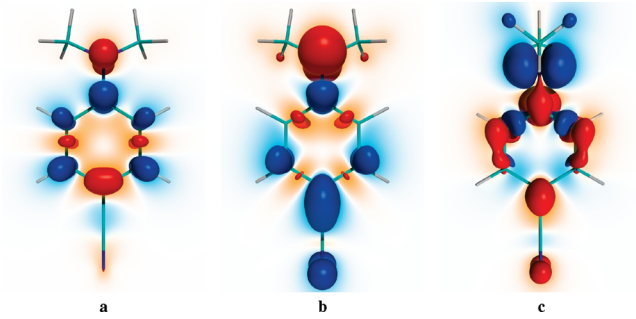


Figure 2. Electron density change in the $S_0 \rightarrow S_1$ transition (a), in the $S_0 \rightarrow S_2$ transition (b), and in the CT → S_0 transition (c). Isosurfaces for a change of 0.0032e/a_0^3: red for a decrease in density and blue for an increase. Densities calculated at SA-CASSCF/6-311G** level.

in general, in good agreement with previous ab initio calculations.[12,15,16,18,43]

To characterize the electron transitions studied in this work, we can examine differences in electron density and dipole moments. In Figure 2 we display three-dimensional maps of the electron density variation upon vertical transition, for $S_0 \rightarrow S_1$, $S_0 \rightarrow S_2$, and CT → S_0 (the map for LE → S_0 is quite similar, but inverted, to that of $S_0 \rightarrow S_1$). It can be observed that $S_0 \rightarrow S_1$ corresponds mainly to a local excitation in the phenyl ring, while $S_0 \rightarrow S_2$ has a marked charge transfer character, involving an important transfer of electron density from the amine N atom to the CN end of the molecule. This is also recognized from the main configuration state function (CSF) contributions in the electron states: from S_0, S_1 is well described by $\pi-\pi^*$ transitions local to the phenyl ring (particularly from the bonding to the antibonding orbitals of the C−C bonds parallel to the long axis of the molecule), S_2 corresponds mainly to a single excitation from the dimethylamino end to the rest of the molecule.[12,18] The CT → S_0 transition has, as expected, a significant charge transfer character, in the opposite direction from the $S_0 \rightarrow S_2$ transition.

The dipole moments of the different states confirm the above, as seen in Table 3. In the absorption processes (S_0 geometry), the dipole moments of the S_0 and S_1 states are almost the same, confirming the "local excitation" character of the transition. The same holds for the emission from the LE state, where the dipole moments of both excited and ground state are practically identical. The dipole moment of the S_2 state, however is almost twice the as large as that of the ground state, indicative of a significant charge transfer inside the molecule. This same observation can be made in the CT geometry, where the dipole moment of the molecule is reduced to half the initial value upon emission from the CT state to the S_0 ground state.

The optimized CT state is calculated to lie 4.27 kcal/mol above the optimized LE state. If both excited states are in thermodynamic equilibrium prior to radiation emission, as suggested by several studies,[4,5,7,12] the LE state would hold more than 99.9% of the excited state population. Another important factor is the significant geometry distortion that the DMABN molecule must suffer to reach this CT state; the structure of the LE state is much more similar to the optimized S_0 geometry that the CT is. Therefore, both thermodynamic and kinetic arguments support the absence of observed emission from the CT state in gas phase.

All the above results serve to identify the geometries and electronic states that will subsequently be studied in solution. The main features are the existence of absorption processes to low-lying states of different character, and the possibility of obtaining two different optimized excited state structures, each of them of a different character as well.

3.2. Solution. The three optimized structures of DMABN (S_0, LE, and CT) were solvated in water and tetrahydrofuran, and the geometries were reoptimized with the ASEP/MD method. The geometry optimization was performed, as in gas phase, at the CASSCF(12,11)/6-311G** level; only in the case of the CT structure, it was necessary to use state averaging for the first two roots, relative weights in 1:3 proportion were used for the S_0 and CT states, respectively. Transition energies were calculated applying CASPT2 on a SA-CASSCF reference wave function, now with three states and equal weights, in all cases. We report here only CASPT2(0.00) results, the corresponding CASPT2(0.25) transitions are always around 0.4 eV higher, as observed in the gas phase.

The obtained dipole moments of the different states and the characteristic angles of the optimized structures are shown in Table 3. There are several things to be noted. The geometrical changes in the S_0 and LE structures are small in

Table 3. Characteristic Angles (Wagging Angle for S_0, Twisting Angle for LE, and out-of-Plane Angle for CT, in Degrees, see Figure 1) and Dipole Moments (at CASPT2(0.00)/6-311G** Level, in D) for DMABN

	S_0 geometry				LE geometry			CT geometry		
	wag	$\mu(S_0)$	$\mu(S_1)$	$\mu(S_2)$	twist	$\mu(S_0)$	μ(LE)	o.o.p.	$\mu(S_0)$	μ(CT)
gas	24.4	6.45	6.48	12.53	15.5	6.88	6.94	45.7	4.18	8.76
THF	23.5	7.29	7.32	13.50	7.3	8.06	8.13	31.2	6.75	17.18
water	27.8	8.19	8.11	14.64	9.8	9.41	9.59	6.9	9.64	21.26

both solvents, the overall shape of the molecule is maintained. The dipole moments of the different states are enhanced, as expected in polar solvents, but for the S_0 and LE structures the differences between states remain approximately constant. The largest variations are seen for the CT structure. Here the phenyl ring becomes increasingly planar and the out-of-plane angle progressively vanishes as the solvent polarity is increased. Simultaneously, the dipole moments of the two electron states are considerably strengthened and the difference between them becomes very large, more than 10 D. In water the molecule can be considered almost in an ideal TICT state: with a planar phenyl ring, a perpendicularly twisted dimethylamino substituent, and a very large internal charge transfer character.

Among the bond lengths, the only significant changes are a shortening of the C−N bond in the amine end, in the LE state, from gas phase to aqueous solution (0.01 Å), and in the CT state, a lengthening of the C−C bonds of the phenyl ring closer to the C≡N (0.01 Å), a shortening of those closer to the $N(CH_3)_2$ (0.02 Å), a shortening of the C−C bond between the phenyl and the nitrile (0.03 Å), and a shortening of the C−N bond in the amine end (0.04 Å). Other differences in the bond lengths are lower than 0.01 Å.

To analyze the solvent distribution around the solute molecule we refer to Figure 3, which displays the regions in space with higher concentration of oxygen atoms when the DMABN molecule is put in aqueous solution. The solvent distributions around the S_0 and LE structures show the existence of hydrogen bonding with the amine N atom in the ground state and with the phenyl ring in the excited state. However, when compared to the solvent distributions around the analogous states of DMA,[21] it is clear that the solvent structure around DMABN is significantly weaker than around DMA. This fact explains why the anomalous solvatochromic shifts of DMA in water are not observed for DMABN and supports the strong hydrogen bonds as responsible for the aforementioned anomalous behavior. This difference can be traced back to a larger negative atomic charge on the N atom in DMA: −0.64 *e* vs −0.55 *e* in DMABN.

A very strong solvent structure, however, is observed around the CT state of DMABN (Figure 3c). There are two regions of high water concentration at either side of the phenyl ring, where the water molecules orient one of their hydrogens toward the solute. Around the positions perpendicular to the $N(CH_3)_2$ plane, there is a high water concentration too, but now it is the oxygen that is oriented toward the electron-depleted nitrogen. The nitrile end forms hydrogen bonds with the solvent as well, but because of the axial symmetry, it does not induce a clear spatial orientation in the water molecules, and only the regions of high hydrogen occupancy are visible in the figure. It can be noted that the preferred hydrogen bonds are not colinear with the C≡N bond, but form an approximately tetrahedral angle, so that 3 water molecules can be simultaneously hydrogen-bonded to the nitrile.

The radial distribution functions (rdf) of water oxygens around the solute nitrogens are displayed in Figures 4 and 5. Figure 4 clearly shows the peak at 3 Å corresponding to three water molecules around the nitrile N atom of DMABN. For comparison, the rdf of DMA is also drawn, and it shows a peak for the $N(CH_3)_2$ nitrogen that is just a shoulder in DMABN; this peak corresponds to the water molecule hidrogen-bonded to the N atom thanks to its enhanced negative charge in DMA. The peak at 4.7 Å corresponds to water molecules loosely bound around the methyl groups. The rdfs for the LE state (not shown) are essentially similar to those of the S_0 state, but lacking the peak or shoulder at 3 Å for the amine nitrogen. The rdfs for the CT state are displayed in Figure 5. The peak for the nitrile nitrogen is higher, showing the increased structure of the solvent, and both rdfs have a broader peak at 4 Å, which can be attributed

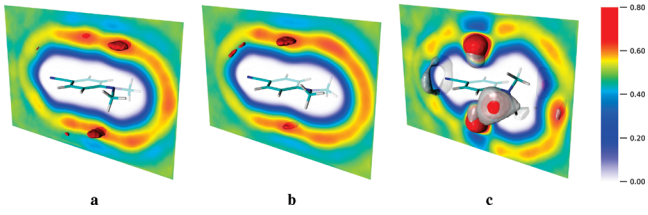


Figure 3. Occupancy maps of water oxygen atoms (considered as van der Waals spheres, as calculated by VMD[44]) around DMABN for (a) the optimized S_0 structure, (b) the optimized LE structure, and (c) the optimized CT structure. Solid red isosurfaces shown for values of 0.64 in a and b and 0.70 in c. The gray translucent surfaces in c show the occupancy of hydrogen atoms at a value of 0.42.

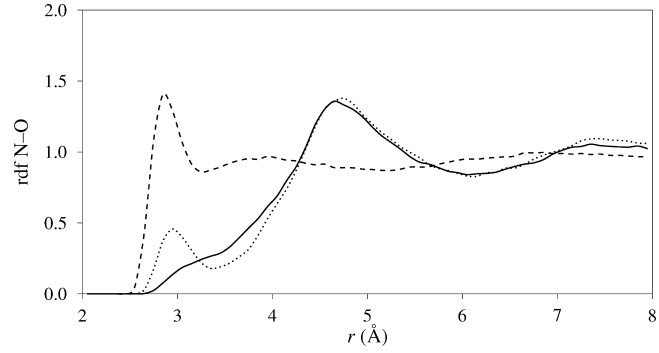


Figure 4. Radial distribution functions of water oxygen atoms around solute nitrogens for the S_0 state. Full line: $N(CH_3)_2$ nitrogen. Dashed line: C≡N nitrogen. Dotted line: DMA nitrogen (see ref 21).

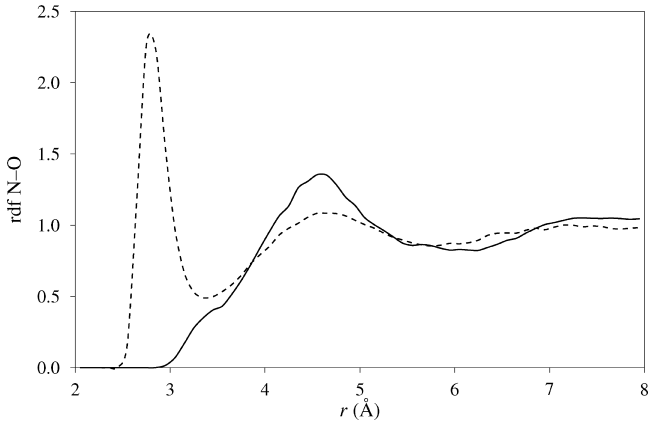


Figure 5. Radial distribution functions of water oxygen atoms around solute nitrogens for the CT state. Full line: $N(CH_3)_2$ nitrogen. Dashed line: C≡N nitrogen.

to the water molecules placed at the sides of the phenyl ring; the molecules at the sides of the $N(CH_3)_2$ group are visible as a shoulder at 3.5 Å.

Table 4 collects the calculated transition energies for the two absorption and two fluorescence bands, in gas phase and in the two studied solvents. We observe a small solvatochromic shift in the $S_0 \rightarrow S_1$ absorption band and a more significant red shift in the $S_0 \rightarrow S_2$ band, both effects being consistent with the dipole moment variations in those transitions. The energy difference between the two bands is reduced in solution, which can explain why the experimental spectra of DMABN in THF[7] and water[45] do not show any visible feature corresponding to the $S_0 \rightarrow S_1$ absorption, although in other nonpolar solvents[7] the $S_0 \rightarrow S_1$ absorption can be detected as a shoulder in the $S_0 \rightarrow S_2$ band. The calculated oscillator strengths (not shown) are very similar to the gas phase values, this supports the disappearance of the weak $S_0 \rightarrow S_1$ band into the stronger $S_0 \rightarrow S_2$ band in the absorption spectra in solution.

When comparing the current results with those obtained for DMA,[21] it can be seen that the apparent anomaly found for DMA in water (a blue shift in absorption bands, when other polar solvents induce a red shift) is not so strong in this case. There is indeed a small blue shift in the $S_0 \rightarrow S_1$ absorption, but the shift of the $S_0 \rightarrow S_2$ still occurs to the red side. This can be explained by weaker hydrogen bonds in DMABN compared to DMA as discussed above. From Table 4, it is clear that our results underestimate the red shift observed in the $S_0 \rightarrow S_2$ and LE $\rightarrow S_0$ bands when the DMABN molecule passes from gas phase to solution. For example, for the absorption band, we obtain −0.10 eV in THF and −0.15 eV in water, while the experimental values are −0.27 eV and −0.41 eV, respectively. This discrepancy can be mainly attributed to the neglect of the dispersion components of the solute−solvent interactions, as we proposed in the previous study of DMA, where we obtained a similar systematic error in solvent shifts. In DMA, the dispersion contribution was estimated in around −0.10 eV, and we can expect it to be somewhat higher, maybe up to −0.20 eV in DMABN because of the its larger size and therefore larger polarizability. The addition of this estimated error would bring the calculated solvent shift values much closer to experiment.

In the fluorescence from the LE state a small solvent effect is predicted, similarly to the $S_0 \rightarrow S_1$ absorption. The CT state is however strongly affected by the solvent, as seen in Table 3, and the corresponding fluorescence band suffers a large red shift in THF and water. Both results agree qualitatively with experimental observations,[5,7] which find that the position of the LE band is not much affected by the solvent nature, while the CT band, when present, is strongly red-shifted in polar solvents. Quantitatively, our results underestimate the red shift in the LE band, probably because of the neglect of the dispersion component, as noted above, but the shift in the fluorescence energies when passing from THF to water is correctly reproduced (−0.05 vs ±0.06 eV for LE, −0.50 vs −0.45 eV for CT). The energies for the LE fluorescence are overestimated and for the CT fluorescence are underestimated, which results in differences between the two fluorescence maxima that are too large by about 0.8 eV. This can be attributed to the apparent underestimation of the CT $\rightarrow S_0$ energy with the quantum method used, as discussed in the previous section, but also to the overstabilization of the CT state with the point-charge model used in the MD simulations, as it is known that point-charge models tend to emphasize solvent structure and charge separation. This overstabilization of the CT state further decreases the CT $\rightarrow S_0$ energy. Another reason for this problem could be the use of the same Lennard-Jones parameters for all electronic states of the solute; this can lead to unrealistic solvation for excited states when their electron distribution is significantly different from the ground state, as is the case of the CT state.

The solvatochromic shifts in the electron transitions can be partitioned as indicated it eq 2; the results are shown in Tables 5 and 6. It is evident that the electrostatic contribution δ_{elec} is dominant in both $S_0 \rightarrow S_2$ and CT $\rightarrow S_0$ transitions. The effect of the change of geometry, δ_{geo}, on the solvent shifts is small in all but the CT $\rightarrow S_0$ transition, but compared to the total δ, its relative size is more significant, showing that the effect of the solvent on the equilibrium geometries should not be neglected.

The relative free energies of the optimized LE and CT minima are given in Table 7, calculated in solution according to eq 3. In both water and THF, the CT state has a lower energy than the LE state, unlike in gas phase, where the LE state is more stable. The stabilization of the CT state is almost double in water than in THF and is dominated by the solute−solvent interaction component ΔG_{int}. These results explain qualitatively the experimental fluorescence spectra in the two solvents, with a CT band that is, in relation with the LE band, much more intense in water than in THF.[5,7]

Another quantity to consider is the relative energy of the CT state with respect to the Franck−Condon S_1 state, the results obtained are shown in Table 8 (the E_{exc} term corresponds to excitation energy $S_0 \rightarrow S_1$, calculated at the ground state geometry). In gas phase, the CT state is very close in energy to the Franck−Condon S_1 state; in water and THF the CT state lies significantly below the Franck−Condon state.

Table 4. Transition Energies (eV) Calculated in Solution at CASPT2(0.00)/6-311G** Level[a]

	absorption		fluorescence	
	$S_0 \rightarrow S_1$	$S_0 \rightarrow S_2$	LE $\rightarrow S_0$	CT $\rightarrow S_0$
gas (refs 40 and 42)	4.17 (4.25)	4.46 (4.56)	3.75 (3.76)	2.77
THF (ref 7)	4.15	4.36 (4.29)	3.74 (3.47)	2.33 (2.79)
water (refs 46 and 47)	4.21	4.31 (4.15)	3.69 (3.41−3.54)	1.83 (2.34)

[a] Experimental values in parentheses.

Table 5. Solvent Shifts and Their Components (kcal/mol) in Tetrahydrofuran Calculated at CASPT2(0.00)/6-311G** Level

	$\delta =$	$\delta_{geo} +$	$\delta_{dist} +$	δ_{elec}
$S_0 \rightarrow S_1$	−0.50	−0.14	−0.56	0.20
$S_0 \rightarrow S_2$	−2.14	−0.01	0.38	−2.59
LE $\rightarrow S_0$	−0.34	0.66	−1.04	0.04
CT $\rightarrow S_0$	−10.23	7.99	−0.36	−18.58

Table 6. Solvent Shifts and Their Components (kcal/mol) in Water Calculated at CASPT2(0.00)/6-311G** Level

	$\delta =$	$\delta_{geo} +$	$\delta_{dist} +$	δ_{elec}
$S_0 \rightarrow S_1$	0.98	0.26	0.11	0.65
$S_0 \rightarrow S_2$	−3.52	0.36	0.92	−4.77
LE $\rightarrow S_0$	−1.34	−0.12	−1.66	0.44
CT $\rightarrow S_0$	−21.65	9.44	7.40	−38.48

Table 7. Free Energy of the CT State Relative to the LE State (kcal/mol)

	$\Delta G =$	$\Delta E +$	ΔG_{int}
Gas		4.27	
THF	−5.36	4.83	−10.19
Water	−15.67	12.88	−28.55

Table 8. Free Energy of the CT State Relative to the Franck−Condon S_1 State (kcal/mol)

	$\Delta G =$	$\Delta E +$	$\Delta G_{int} -$	ΔE_{exc}
Gas	−1.67	94.55		96.21
THF	−11.46	94.23	−9.98	95.71
Water	−21.57	103.30	−27.68	97.19

The relative free energies of the different states and geometries are summarized in Figure 6. The CT state is above the LE state at both the S_0 and LE geometries in all cases. At the CT geometry, the LE state was not obtained, it is presumed to have a much higher energy. The most clear difference between the three diagrams is the great stabilization in solution of the CT state at its optimized geometry that is placed below the Franck−Condon and the optimized LE state, thus permitting the appearance of the corresponding fluorescence band.

4. Conclusions

In this work, we have shown high-level quantum calculations of absorption and emission properties of 4-(*N*,*N*-dimethylamino)benzonitrile (DMABN) in gas phase and in polar solvents: tetrahydrofuran and water. Two excited states of DMABN, with different optimum geometries, have been identified, corresponding one of them to a local excitation (LE) on the phenyl ring (S_1 at the ground state geometry) and the other to a intramolecular charge transfer (CT) from the $N(CH_3)_2$ moiety to the phenyl ring (S_2 at the ground state geometry). The CT state displays a twisted $N(CH_3)_2$ with respect to the ring, as suggested by the TICT model.

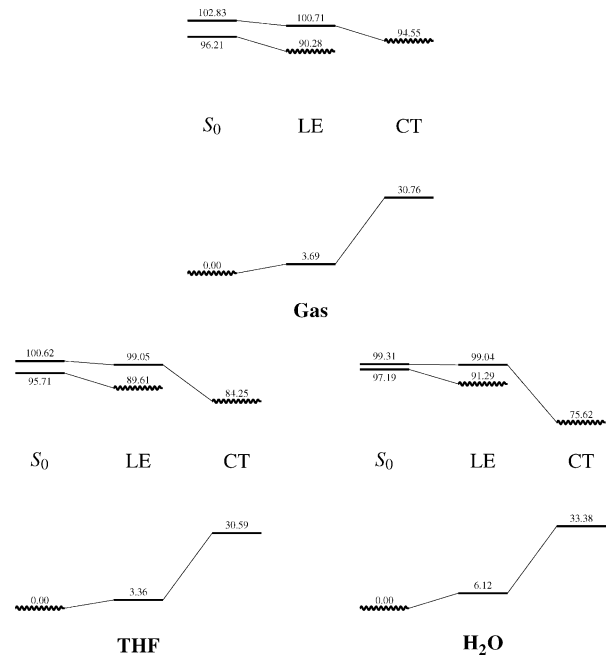


Figure 6. Relative free energies at CASPT2(0.00)//CASSCF(12,11)/6-311G** level in kcal/mol of the studied states of DMABN in gas phase and in the two solvents. The S_0, LE, and CT labels refer to the three optimized geometries obtained, and the state optimized for each geometry is marked with a wavy line.

The results for absorption show a weak S_1 band and a strong S_2 band. In solution, the S_2 band is significantly shifted to lower energies, overlapping and covering the S_1 band, which agrees with the experimental spectra.

In gas phase, the CT excited state has a higher energy than the LE state and the Franck−Condon S_1 state in absorption, and its geometry features a strongly distorted phenyl ring. This makes the CT state hardly accessible from excitation at the ground state geometry, and explains why no dual fluorescence is observed in gas phase.

In solution, the CT state acquires a more planar phenyl ring and is stabilized with respect to the LE and Franck−Condon states. In THF, all three states have a similar energy, while in water, the CT state is much more stable. The relative stabilities of the LE and CT states are in qualitative agreement with the experimental fluorescence intensities: in THF two bands of similar intensities are observed, in water the lower-energy (CT) band is much more intense.

In sum, the results of this work support the validity of the TICT model for the dual fluorescence of DMABN, since a

twisted intramolecular charge transfer state is found, it is stabilized in polar solvents, and its emission properties are in agreement with experimental observations.

Acknowledgment. This work was supported by the CTQ2008-06224 Project from the Ministerio de Educación y Ciencia of Spain and the PRI08A056 Project from the Consejería de Economía, Comercio e Innovación of the Junta de Extremadura. I.F.G. acknowledges the Junta de Extremadura and the European Social Fund for financial support.

References

(1) Lippert, E.; Lüder, W.; Moll, F.; Nägele, W.; Boos, H.; Prigge, H.; Seibold-Blankenstein, I. *Angew. Chem.* **1961**, *73*, 695–706.

(2) Rotkiewicz, K.; Grellmann, K. H.; Grabowski, Z. R. *Chem. Phys. Lett.* **1973**, *19*, 315–318.

(3) Grabowski, Z. R.; Rotkiewicz, K.; Rettig, W. *Chem. Rev.* **2003**, *103*, 3899–4032.

(4) Leinhos, U.; Kuehnle, W.; Zachariasse, K. A. *J. Phys. Chem.* **1991**, *95*, 2013–2021.

(5) Changenet, P.; Plaza, P.; Martin, M. M.; Meyer, Y. H. *J. Phys. Chem. A* **1997**, *101*, 8186–8194.

(6) Druzhinin, S. I.; Ernsting, N. P.; Kovalenko, S. A.; Pérez Lustres, L.; Senyushkina, T. A.; Zachariasse, K. A. *J. Phys. Chem. A* **2006**, *110*, 2955–2969.

(7) Galievsky, V. A.; Zachariasse, K. A. *Acta Phys. Pol., A* **2007**, *112*, S39–S56.

(8) Lee, J.-K.; Fujiwara, T.; Kofron, W. G.; Zgierski, M. Z.; Lim, E. C. *J. Chem. Phys.* **2008**, *128*, 164512.

(9) Gustavsson, T.; Coto, P. B.; Serrano-Andrés, L.; Fujiwara, T.; Lim, E. C. *J. Chem. Phys.* **2009**, *131*, 031101.

(10) Zachariasse, K. A.; Druzhinin, S. I.; Kovalenko, S. A.; Senyushkina, T. *J. Chem. Phys.* **2009**, *131*, 224313.

(11) Gorse, A.-D.; Pesquer, M. *J. Phys. Chem.* **1995**, *99*, 4039–4049.

(12) Serrano-Andrés, L.; Merchán, M.; Roos, B. O.; Lindh, R. *J. Am. Chem. Soc.* **1995**, *117*, 3189–3204.

(13) Hayashi, S.; Ando, K.; Kato, S. *J. Phys. Chem.* **1995**, *99*, 955–964.

(14) Sudholt, W.; Sobolewski, A. L.; Domcke, W. *Chem. Phys.* **1999**, *240*, 9–18.

(15) Rappoport, D.; Furche, F. *J. Am. Chem. Soc.* **2004**, *126*, 1277–1284.

(16) Köhn, A.; Hättig, C. *J. Am. Chem. Soc.* **2004**, *126*, 7399–7410.

(17) Gómez, I.; Reguero, M.; Boggio-Pasqua, M.; Robb, M. A. *J. Am. Chem. Soc.* **2005**, *127*, 7119–7129.

(18) Amatatsu, Y. *J. Phys. Chem. A* **2005**, *109*, 7225–7235.

(19) Mennucci, B.; Toniolo, A.; Tomasi, J. *J. Am. Chem. Soc.* **2000**, *122*, 10621–10630.

(20) Minezawa, N.; Kato, S. *J. Phys. Chem. A* **2005**, *109*, 5445–5453.

(21) Fdez. Galván, I.; Martín, M. E.; Muñoz Losa, A.; Aguilar, M. A. *J. Chem. Theory Comput.* **2009**, *5*, 341–349.

(22) Sánchez, M. L.; Aguilar, M. A.; Olivares del Valle, F. J. *J. Comput. Chem.* **1997**, *18*, 313–322.

(23) Fdez. Galván, I.; Sánchez, M. L.; Martín, M. E.; Olivares del Valle, F. J.; Aguilar, M. A. *Comput. Phys. Commun.* **2003**, *155*, 244–259.

(24) Aguilar, M. A.; Sánchez, M. L.; Martín, M. E.; Fdez. Galván, I. An Effective Hamiltonian Method from Simulations: ASEP/MD. In *Continuum Solvation Models in Chemical Physics*, 1st ed.; Mennucci, B., Cammi, R., Eds.; Wiley: West Sussex, England, 2007; Chapter 4.5, pp 580−592.

(25) Fdez. Galván, I.; Sánchez, M. L.; Martín, M. E.; Olivares del Valle, F. J.; Aguilar, M. A. *J. Chem. Phys.* **2003**, *118*, 255–263.

(26) Martín, M. E.; Muñoz Losa, A.; Fdez. Galván, I.; Aguilar, M. A. *J. Chem. Phys.* **2004**, *121*, 3710–3716.

(27) Muñoz Losa, A.; Fdez. Galván, I.; Aguilar, M. A.; Martín, M. E. *J. Phys. Chem. B* **2007**, *111*, 9864–9870.

(28) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420–1426.

(29) Fdez. Galván, I.; Aguilar, M. A.; Ruiz-López, M. F. *J. Phys. Chem. B* **2005**, *109*, 23024–23030.

(30) Roos, B. O.; Taylor, P. R.; Siegbahn, P. E. M. *Chem. Phys.* **1980**, *48*, 157–173.

(31) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. *J. Phys. Chem.* **1990**, *94*, 5483–5488.

(32) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218–1226.

(33) Ghigo, G.; Roos, B. O.; Malmqvist, P.-Å. *Chem. Phys. Lett.* **2004**, *396*, 142–149.

(34) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.

(35) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361–373.

(36) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, Ö.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C. S.; Adamo, C.; Clifford, S.; Ochterski, J. W.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Salvador, P.; Dannenberg, J. J.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T. A.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andrés, J. L.; González, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.11.3; Gaussian, Inc.: Pittsburgh, PA, 2001.

(37) Karlström, G.; Lindh, R.; Malmqvist, P.-Å.; Roos, B. O.; Ryde, U.; Veryazov, V.; Widmark, P.-O.; Cossi, M.; Schimmelpfennig, B.; Neogrády, P.; Seijo, L. *Comput. Mater. Sci.* **2003**, *28*, 222–239.

(38) Refson, K. *Comput. Phys. Commun.* **2000**, *126*, 310−329.

(39) Kajimoto, O.; Yokoyama, H.; Ooshima, Y.; Endo, Y. *Chem. Phys. Lett.* **1991**, *179*, 455–459.

(40) Bulliard, C.; Allan, M.; Wirtz, G.; Haselbach, E.; Zachariasse, K. A.; Detzer, N.; Grimme, S. *J. Phys. Chem. A* **1999**, *103*, 7766–7772.

(41) Schuddeboom, W.; Jonker, S. A.; Warman, J. M.; Leinhos, U.; Kuehnle, W.; Zachariasse, K. A. *J. Phys. Chem.* **1992**, *96*, 10809–10819.

(42) Lommatzsch, U.; Gerlach, A.; Lahmann, C.; Brutschy, B. *J. Phys. Chem. A* **1998**, *102*, 6421–6435.

(43) Gómez, I.; Mercier, Y.; Reguero, M. *J. Phys. Chem. A* **2006**, *110*, 11455–11461.

(44) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graph.* **1996**, *14*, 33–38.

(45) Monti, S.; Marconi, G.; Manoli, F.; Bortolus, P.; Mayer, B.; Grabner, G.; Köhler, G.; Boszczyk, W.; Rotkiewicz, K. *Phys. Chem. Chem. Phys.* **2003**, *5*, 1019–1026.

(46) Figueroa, I. D.; El Baraka, M.; Quinones, E.; Rosario, O. *Anal. Chem.* **1998**, *70*, 3974–3977.

(47) Saigusa, H.; Iwase, E.; Nishimura, M. *J. Phys. Chem. A* **2003**, *107*, 4989–4998.

CT9006713

JCTC Journal of Chemical Theory and Computation

Chemically Meaningful Atomic Charges That Reproduce the Electrostatic Potential in Periodic and Nonperiodic Materials

Thomas A. Manz* and David S. Sholl*

School of Chemical and Biomolecular Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0100

Received March 5, 2010

Abstract: Net atomic charges (NACs) can be used both to understand the chemical states of atoms in a material as well as to represent the electrostatic potential, *V*, of the material outside its electron distribution. However, many existing definitions of NACs have limitations that prevent them from adequately fulfilling this dual purpose. Some charge methods are not applicable to periodic materials or are inaccurate for systems containing buried atoms, while others work for both periodic and nonperiodic materials containing buried atoms but give NACS that do not accurately reproduce *V*. We present a new approach, density derived electrostatic and chemical (DDEC) charges, that overcomes these limitations by simultaneously optimizing the NACs to be chemically meaningful and to reproduce *V* outside the electron distribution. This atoms-in-molecule method partitions the total electron density among atoms and uses a distributed multipole expansion to formally reproduce *V* exactly outside the electron distribution. We compare different methods for computing NACs for a broad range of materials that are periodic in zero, one, two, and three dimensions. The DDEC method consistently performs well for systems with and without buried atoms, including molecules, nonporous solids, solid surfaces, and porous solids like metal organic frameworks.

Introduction

The electron density in a material, $\rho(\vec{r})$, controls almost all fundamental properties of the material. In many contexts, however, it is useful to replace this complicated function by the simpler concept of a single charge associated with each of the material's atoms. Atomistic simulations based on interatomic force fields, which provide a crucial tool in the multiscale modeling of complex materials,[1] often rely on charges of this kind for their definition. Atomic charges also provide a simple approach for characterizing charge transfer and similar properties. The task of assigning net atomic charges (NACs) to individual atoms from $\rho(\vec{r})$ obtained from a quantum chemistry calculation is a long-standing challenge that has no unique solution.[2] To be widely applicable, an approach to this task should satisfy several criteria. First, the method should be directly applicable to both periodic and nonperiodic systems; specifically, the method should not require the use of a nonperiodic cluster to assign charges for periodic materials. Second, to accurately describe electrostatic interactions between chemical species, the NACs should accurately reproduce the electrostatic potential, *V*, outside the electron distribution. Third, to be chemically meaningful, the assigned atomic distributions and NACs should resemble those of real atoms in appropriate reference states. Fourth, the method should be applicable to both porous and nonporous materials. Finally, the method should yield a well-defined set of NACs that are independent of the basis sets used to generate the electron density.

Common methods for computing NACs fail to satisfy one or more of the criteria above. Electrostatic potential (ESP) fitting methods assign NACs that minimize the RMS deviation in *V* over a set of grid points outside the van der Waals (vdW) surface of a material relative to *V* from the full electron density.[3,4] In systems containing buried atoms, many different combinations of charges on atoms far from

* E-mail: thomasamanz@gmail.com (T.A.M.), david.sholl@chbe.gatech.edu (D.S.S.).

10.1021/ct100125x © 2010 American Chemical Society
Published on Web 07/15/2010

the vdW surface can give almost the same V, giving ESP charges that are not chemically meaningful.[5] (An atom is buried if its minimum distance to the system's vdW surface is larger than the atom's vdW radius.) For large molecules, this problem can be addressed,[3,6] but for spatially extended materials, ESP charges are typically obtained from nonperiodic cluster models, where it is not straightforward to determine the optimal cluster size or truncation.[7] The recent REPEAT method overcomes these problems, allowing NACs to be fit to V in porous periodic materials without constructing a nonperiodic cluster model, but it has only limited applicability to nonporous solids.[8] The Hirshfeld (HD) method[9] is applicable to both periodic and nonperiodic systems but underestimates the magnitudes of NACs.[10,11] The recent iterative stockholder atom (ISA) and iterative Hirshfeld (IH) methods reproduce V and molecular dipoles,[10,12−16] but they have not yet been applied to compute NACs in periodic materials. An alternative widely applied to periodic systems is the Bader decomposition, which associates the electron density in cleverly constructed volumes to each atom.[17,18] Bader NACs are chemically meaningful and suitable for materials containing buried atoms, but they do not accurately reproduce V outside the electron distribution. The well-known Mulliken method has the key drawback of explicit basis set dependence.[19] Natural population analysis (NPA) solves this basis set problem,[19] but NPA is not readily available for periodic systems. This brief review indicates that no approach is currently available that meets all of the criteria listed above.

Below, we introduce a new approach, density derived electrostatic and chemical (DDEC) charges, to associating charges with atoms from a known electron density. DDEC charges satisfy all of the criteria defined above. We present results for materials that are periodic in zero, one, two, and three dimensions, for dense solids, for porous solids, for surfaces of solids, for small molecules, and for large molecules with buried atoms. These examples include many different chemical elements, emphasizing the broad applicability of the method. In each example, we compare the performance of the DDEC method to several existing charge calculation methods. The DDEC method is among the best performing methods for each system, and none of the other charge methods we tested were as accurate as DDEC for such a wide range of systems. The ESP (REPEAT) method is the most accurate for fitting V in nonperiodic (periodic) materials without buried atoms but requires the addition of fine-tuned constraints to accurately treat systems with buried atoms.[3,6,8] The DDEC method does not require such constraints to compute accurate charges in systems with buried atoms. We also find that the DDEC NACs exhibit good transferability between related chemical systems. These properties make the DDEC NACs well-suited for the construction of force fields used in atomistic simulations and for the interpretation of charge transfer during chemical reactions.

Theory

We begin by defining an arbitrary material as a set of atoms $\{A\}$ at positions $\{\vec{R}_A\}$, in a reference unit cell, **U**. For a nonperiodic system (e.g., a molecule), **U** is any parallelpiped enclosing the entire electron distribution. The total electrostatic potential, V, can be written as a sum of atomic electrostatic potentials:

$$V(\vec{r}) = \sum_{k_1} \sum_{k_2} \sum_{k_3} \sum_{A} V_A(\vec{r}_A) \quad (1)$$

Here, the reference unit cell has $k_1 = k_2 = k_3 = 0$ and summation over A means summation over all atoms in this unit cell. For a periodic direction, k_i ranges over all integers with the associated lattice vector $\vec{v}_i$. For a nonperiodic direction, $k_i = 0$ and $\vec{v}_i$ is the corresponding edge of **U**. In this sum

$$\vec{r}_A = \vec{r} - k_1\vec{v}_1 - k_2\vec{v}_2 - k_3\vec{v}_3 - \vec{R}_A \quad (2)$$

where $V_A(\vec{r}_A)$ is defined as the electrostatic potential due to the charge assigned to atom A

$$V_A(\vec{r}_A) = \frac{z_A}{r_A} - \oint \frac{\rho_A(\vec{r}'_A)\, d^3\vec{r}'_A}{|\vec{r}_A - \vec{r}'_A|} = \frac{q_A}{r_A} + B + C \quad (3)$$

where $\rho_A(\vec{r}_A)$ is the non-negative electron density assigned to atom A, z_A is the nuclear charge, q_A is the NAC, and $r_A = |\vec{r}_A|$. Since $V_A(\vec{r}_A)$ outside the charge distribution can be exactly represented by a multipole expansion,[20,21] we can rewrite $V_A(\vec{r}_A)$ as shown on the right-hand side of eq 3 where B and C are terms due to atomic multipoles (AMs) and penetration of the atom's electron density, respectively. AMs can be computed by applying the method of Laidig[21] or Kosov and Popelier[20] to $\rho_A(\vec{r}_A)$. The penetration term, C, decays approximately exponentially with increasing r_A and is essentially zero for r_A greater than a critical radius where ρ_A becomes negligible.[22] As in other atoms-in-molecule (AIM) methods,[9,10,14,17,23] we impose the constraint that the entire electron density is included in the collection of atomic densities, which requires $\Theta(\vec{r}) = 0$ where

$$\Theta(\vec{r}) = \rho(\vec{r}) - \sum_{k_1} \sum_{k_2} \sum_{k_3} \sum_{A} \rho_A(\vec{r}_A) \quad (4)$$

To be chemically meaningful, NACs should indicate the amount of charge transfer between atoms. Imagine building a chemical system from isolated atoms in a thought experiment. Noninteracting atoms are first placed at the appropriate atomic positions. The choice of noninteracting atoms for this prematerial is not unique: for NaCl, we could choose either neutral Na and neutral Cl or Na^+ and Cl^-. Because there is no charge transfer between the atoms yet, the computed NACs for this noninteracting prematerial should equal the isolated atom charges. The IH and ISA methods satisfy this prematerial charge criterion,[13] but the Bader method does not.[24] The HD method only satisfies this criterion if the prematerial atoms are neutral, but this is not possible for systems containing a net charge. When the interactions between atoms are included, charges are transferred between them to form the real system. We now show how an AIM method can be developed that simultaneously accounts for the different types of charge transfer in interacting systems.

For ionic charge transfer, the density of each atom in the interacting system, $\rho_A(\vec{r}_A)$, should approximately equal the density of that atom in a noninteracting reference state, $\rho_A^{ref}(r_A,n_A)$, having the same number of electrons, n_A. To achieve this, we minimize a distance measure between $\rho_A(\vec{r}_A)$ and $\rho_A^{ref}(r_A,n_A)$ subject to the constraint $\Theta(\vec{r}) = 0$. The IH method involves a suitable distance measure of this type:[10]

$$F_{chem} = \sum_A \oint \rho_A(\vec{r}_A) \ln\left(\frac{\rho_A(\vec{r}_A)}{\rho_A^{ref}(r_A,n_A)}\right) d^3\vec{r}_A \quad (5)$$

Setting $\partial F_{chem}/\partial\rho_A(\vec{r}_A) = 0$ leads to a proportional distribution of electrons between atoms, $\rho_A(\vec{r}_A)/\rho_A^{ref}(r_A,n_A) = \rho_B(\vec{r}_B)/\rho_B^{ref}(r_B,n_B)$, which is a type of Stockholder partitioning.[15] The reference state density is the linear interpolation

$$\rho_A^{ref}(r_A,n_A) = (1-f)\rho_A^{\tau}(r_A) + f\rho_A^{\tau+1}(r_A) \quad (6)$$

between the spherically averaged ground state densities of isolated atoms of the same element having the closest lower (τ) and higher ($\tau + 1$) integer number of electrons, where $f = n_A - \text{floor}(n_A) = n_A - \tau$.[10,12,13] Instead of being based on single atoms, the reference states could be based on chemical fragments containing several atoms, which would give fragment net charges rather than NACs. A modification of this type to the IH method has been used to study charge transfer between a graphene sheet and adsorbates.[25]

The second type of charge transfer is covalent bonding. This type of charge transfer is directional in nature and generates nonspherical distortions in the electron density of each atom. Multipole contributions sum to zero for spherically symmetric distributions, so eq 1 converges most rapidly if each $\rho_A(\vec{r}_A)$ is optimized to be as close to spherically symmetric as possible. Since a key objective is to compute NACs that accurately reproduce V outside the electron distribution, the atomic distributions should be optimized to be close to spherically symmetric. In the ISA method,[13−15] this is done by minimizing the information distance, F_{ESP}, between $\rho_A(\vec{r}_A)$ and its spherical average, $\rho_A^{avg}(r_A)$, subject to the constraint $\Theta(\vec{r}) = 0$, where

$$F_{ESP} = \sum_A \oint \rho_A(\vec{r}_A) \ln\left(\frac{\rho_A(\vec{r}_A)}{\rho_A^{avg}(r_A)}\right) d^3\vec{r}_A \quad (7)$$

Setting $\partial F_{ESP}/\partial\rho_A(\vec{r}_A) = 0$ gives $\rho_A(\vec{r}_A)/\rho_A^{avg}(r_A) = \rho_B(\vec{r}_B)/\rho_B^{avg}(r_B)$, which is a type of Stockholder partitioning where density is given to each atom in proportion to $\rho_A^{avg}(r_A)$. The effects on V of slight deviations of $\rho_A(\vec{r}_A)$ from spherical symmetry are accounted for by the AMs of $\rho_A(\vec{r}_A)$. Because AM contributions are minimized by making $\rho_A(\vec{r}_A)$ close to spherically symmetric, the ISA NACs are often very similar to those obtained with the ESP method.

The third type of charge transfer is a response to overall changes in the electric field. For example, the electron density of an isolated O^{2-} ion in a vacuum decays very slowly with increasing radius because of strong repulsion between the excess valence electrons. A crystal containing O^{2-} ions contains cations that provide charge compensation, and the dielectric constant of the crystal is larger than that of the vacuum. Charge compensation and dielectric screening reduce the strength of electrostatic repulsion between valence electrons in an O^{2-} ion, causing it to have a smaller effective radius in a crystal than in a vacuum. As shown in the Results and Discussion section below, the use of uncompensated reference states (i.e., vacuum phase ions) to construct $\rho_A^{ref}(r_A,n_A)$ leads to chemically unreasonable NACs in many cases, for example, charge magnitudes significantly greater than 2 in crystalline MgO. To solve this problem, charge-compensated ions should be used to construct $\rho_A^{ref}(r_A,n_A)$. One possible approach is to place all of the compensating charge on the surface of an atom-centered sphere having a radius chosen to approximately coincide with the anionic radius,[26] such as has been used in the interpretation of experimental electron density profiles for MgO and other ionic solids.[27] Our modification of this approach allows the anionic radius to be determined self-consistently by energy minimization. For cations, we developed a different approach involving a spherical charge compensation shell. As a conceptually simpler alternative, we also developed a periodic charge compensation scheme for both anions and cations that employs a uniform background charge. Details of these approaches are described in the Methods section below.

Table 1. Optimization of the Parameter χ

	h-BN crystal			BN sheet		
χ	iterations	AM_{max}	q_B	iterations	AM_{max}	q_B
0.00	422	0.27(q)	0.64	1550	0.11(q)	1.07
0.05	243	0.21(q)	1.00	171	0.16(q)	1.04
0.10	149	0.24(q)	0.96	142	0.23(q)	0.99
0.20	90	0.32(q)	0.90	86	0.36(q)	0.91
0.50	47	0.58(q)	0.79	46	0.66(q)	0.79
1.00	29	0.92(q)	0.70	28	1.03(q)	0.69

When all atoms are near the surface of a material, only one set of NACs approximately reproduces V. Because the ISA NACs are optimized to reproduce V, the optimization landscape for F_{ESP} is likewise steep when the system contains no buried atoms. In this case, the ideal force field NACs are those which approximately minimize F_{ESP} irrespective of F_{chem} and correspond to the well-determined ESP charges. For systems containing buried atoms, multiple sets of NACs can reproduce V with similar accuracy. Including F_{chem} in the optimization functional will ensure that the optimized NACs are chemically meaningful, because the atomic distributions are optimized to resemble those of real atoms in appropriate reference states. Thus, we propose an optimization functional of the form

$$G = \chi F_{chem} + (1-\chi)F_{ESP} + \int_U \lambda(\vec{r})\,\Theta(\vec{r})\,d^3\vec{r} \quad (8)$$

to conveniently handle systems with and without buried atoms, where the Lagrange multiplier $\lambda(\vec{r})$ enforces the constraint $\Theta(\vec{r}) = 0$. F_{chem} and F_{ESP} should be weighted nonnegatively, so $0 \le \chi \le 1$. For optimum performance, F_{chem} should be weighted much less than F_{ESP} so that F_{chem} only becomes important when F_{ESP} defines a shallow landscape (i.e., for systems containing buried atoms); thus, $0.01 \le \chi \le 0.4$. Table 1 shows the effects of changing χ on the charges and largest atomic multipole (AM_{max}) for a single sheet of

hexagonal BN and the three-dimensional solid form of h-BN formed by stacking these sheets (see Figure 1). Computational details are provided in the Methods section below. Since the sheets in h-BN are weakly bound to each other, an ideal charge method should give similar NACs for these two structures. When no reference density was used (i.e., $\chi = 0$), significantly different charges were obtained for these two structures, illustrating the poor transferability of the ISA method for systems with buried atoms. This problem was solved by including a reference density (i.e., $\chi > 0$), leading to similar charges for the two structures. (C1 charge compensated densities, defined in detail below, were used in this example.) The number of iterations required for convergence decreased monotonically as χ increased from 0 to 1. For h-BN, AM_{max} went through a minimum around $\chi = 0.05$. As described above, small AMs are desirable for the NACs to accurately reproduce V. In the following results, we used $\chi = 0.10$, which provided small AMs at reasonable computational cost. Our approach also makes it possible for the first time to compute ISA and IH NACs for spatially periodic materials, where ISA corresponds to $\chi = 0$ and IH corresponds to $\chi = 1$ using reference ions placed in a vacuum without charge compensation.

Charges are calculated using an iterative scheme. Setting $\partial G/\partial \rho_A(\vec{r}_A) = 0$ for constant ${\rho_A}^{ref}(r_A,n_A)$ gives

$$\frac{\partial G}{\partial \rho_A(\vec{r}_A)} = (1-\chi)\ln\left(\frac{\rho_A(\vec{r}_A)}{\rho_A^{avg}(r_A)}\right) + \chi \ln\left(\frac{\rho_A(\vec{r}_A)}{\rho_A^{ref}(r_A,n_A)}\right) + \chi - \lambda(\vec{r}) \quad (9)$$

Exponentiating eq 9 gives

$$e^{\lambda(\vec{r})-\chi} = \frac{\rho_A(\vec{r}_A)}{w_A(r_A)} = \frac{\rho_B(\vec{r}_B)}{w_B(r_B)} \quad (10)$$

where the relative weight factor for each atom is

$$w_A(r_A) = (\rho_A^{avg}(r_A))^{(1-\chi)}(\rho_A^{ref}(r_A,n_A))^{\chi} \quad (11)$$

Equation 10 can be more conveniently written by summing over B to give

$$\rho_A(\vec{r}_A) = w_A(r_A)\,\rho(\vec{r})/\sum_{k_1}\sum_{k_2}\sum_{k_3}\sum_{B} w_B(r_B) \quad (12)$$

Neutral atom densities, ${\rho_A}^{ref}(r_A,0)$, are used as initial estimates. In each iteration, the current estimate of $\{w_A(r_A)\}$ is used to compute $\{\rho_A(\vec{r}_A)\}$ by eq 12. After calculating

$$n_A = \oint \rho_A(\vec{r}_A)\, d^3\vec{r}_A = z_A - q_A \quad (13)$$

the new reference density ${\rho_A}^{ref}(r_A,n_A)$ is computed via eq 6. Finally, ${\rho_A}^{ref}(r_A,n_A)$ and the spherical average of $\rho_A(\vec{r}_A)$ are combined using eq 11 to generate an improved estimate of $\{w_A(r_A)\}$. The process is repeated until the NACs for consecutive iterations differ by less than a chosen tolerance. Since neutral atom densities are used for the initial guess, the initial iteration of this scheme gives the HD NACs. (The value of χ does not affect the HD NACs because χ first enters the calculation after the initial iteration is completed.) Because no charge compensation is involved for neutral atoms, the compensated and uncompensated reference density sets give the same HD NACs. Depending on the choice of χ and reference densities, the converged results give the DDEC, ISA, or IH NACs.

Methods

Except where otherwise noted, density functional theory (DFT) with the PW91 exchange-correlation functional was used to obtain $\rho(\vec{r})$. The PBE functional was used for the iron oxides introduced below. All periodic DFT calculations were performed in the Vienna Ab Initio Simulation Package[28,29] (VASP) using the projector augmented wave (PAW) method[30] and a 400 eV cutoff. For periodic systems, the product of the number of k points and the unit cell volume exceeded 4000 Å^3. Except where otherwise noted, nonperiodic DFT calculations were performed using Gaussian software[31] with 6-311++G** basis sets. LANL2DZ basis sets and the B3LYP functional were used for the Zr complexes introduced below. DDEC charges should be computed using an all-electron density. PAW is a frozen-core method that yields all-electron densities. Since LANL2DZ employs an effective core potential (ECP) for the 28 innermost electrons of Zr, we added the PAW frozen-core distribution of these 28 electrons to the LANL2DZ-derived SCF electron density to obtain the all-electron density of each Zr complex, which was then used to compute the DDEC charges. The PAW and ECP methods treat the core electrons relativistically, which is important for modeling heavy atoms. Experimental structures were used for MIL-53 (Al),[32] h-BN,[33] and IRMOF-1,[34] while all other geometries were optimized. Bader charges were computed using the software developed by Henkelman and co-workers.[35−37]

Electrostatic potentials in the periodic systems were computed using the Ewald summation of Smith that includes point charges and optionally point dipoles or quadrupoles.[38] Enough terms were included in the real and reciprocal space sums to converge $V(\vec{r})$ to better than 0.06 kcal/mol. In the

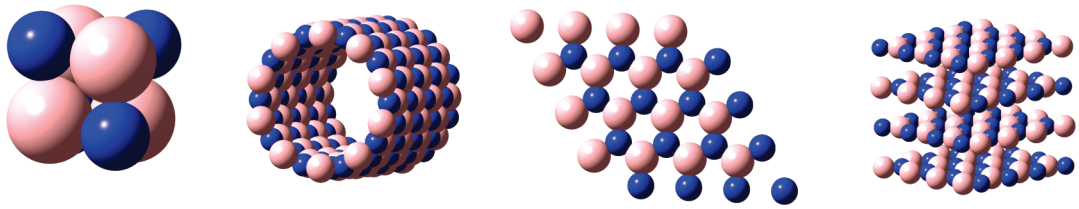


Figure 1. BN structures used in Table 2. From left to right: B_4N_4 cluster, (10,10) nanotube, hexagonal sheet, and hexagonal crystal (h-BN).

results below, ΔV_{NAC} is the RMS difference between the ab initio $V(\vec{r})$ and the potential due to the NACs. ΔV_{dip} is similarly defined, except it includes the potential due to both the NACs and the atomic dipoles. The grid points used to compute ΔV_{NAC} and ΔV_{dip} were uniformly distributed between the surfaces defined by 1.4× and 2.0× vdW radii, which is the region of interest for ESP fitting methods.[4] In periodic systems, V is only determined up to an arbitrary constant,[8] so the average V over these grid points was set equal for the two systems. We used previously reported vdW radii applicable to the first 96 elements.[39]

Since alternative methods can be used to generate charge compensated reference densities, one should include the reference density set label when reporting charges obtained by the DDEC method. In the results below, DDEC/c1 denotes the use of $\chi = 0.1$ with "c1" reference densities computed in VASP by the following method. The ion was placed in the center of a cubic unit cell having an edge length of 10 Å. The cutoff energy, PW91 functional, k-point spacing, and PAW method described in the first paragraph of this section were used. The total number of electrons, n_A, was specified to give the desired ion charge. The energy and electron density were then converged to give the spin state with the lowest energy. During this DFT calculation, the unit cell is made neutral by adding a uniform background charge. This uniform background charge provides charge compensation and only affects the electron density by modifying the net electric field. Moreover, the periodic array of atoms provides dielectric screening that further reduces the electric field strength. This simulates the charge compensation and dielectric screening that occurs in real materials, thereby providing more accurate reference densities than would be obtained from uncompensated ions. For comparison, we also performed charge analysis using uncompensated vacuum reference ions. In the results below, uncompensated (UC) denotes the use of $\chi = 0.1$ with reference densities computed in Gaussian using the PW91 functional and 6-311++G** basis sets (LANL2DZ on Zr and WTBS on Sr).

DDEC/c2 denotes the use of $\chi = 0.1$ with "c2" reference densities computed using a spherical surface of compensating charge having radius R_{comp}. Through trial and error, we found that a careful choice of R_{comp} is required to obtain accurate NACs. No compensating charge was used for the neutral atoms. Let q_1 be the net charge (including the nuclear charge) closer to the atomic center than R_{comp}, q_2 the compensating charge at R_{comp}, and $q_3 < 0$ the net charge due to electron density farther from the nucleus than R_{comp}. Since $q_3 < 0$, making the overall system neutral (i.e., $q_1 + q_2 + q_3 = 0$) gives $q_1 + q_2 > 0$, which produces a cation-like behavior for the spherical ball including R_{comp}. Thus, for cations, the surface charge q_2 was constrained to make the overall system neutral. Since $q_3 < 0$, for $q_1 + q_2 = 0$ the net charge enclosed in a ball of radius $r > R_{comp}$ is slightly negative, and this causes a weak Coulombic repulsion of electrons from the nucleus for $r > R_{comp}$. Because this is the correct behavior for a compensated anion, we infer that $q_1 + q_2 = 0$ is a suitable constraint for anions. By Gauss's law, the constraint $q_1 + q_2 = 0$ can be imposed by choosing a surface charge distribution that makes the electric field zero infinitesimally outside R_{comp}. For anions, R_{comp} was iteratively adjusted by increments of 0.05 Å until the system's total energy was minimized. Values of R_{comp} determined in this manner are similar to crystal-structure-derived anionic radii reported in the CRC Handbook of Chemistry and Physics; specifically, R_{comp} (crystal-structure-derived anionic radii[40]) values in Å are 1.25 (1.33) for F^-, 1.80 (1.81) for Cl^-, 1.95 (1.96) for Br^-, 2.20 (2.20) for I^-, 1.30 (1.21−1.42) for O^{2-}, 1.80 (1.84) for S^{2-}, 1.95 (1.98) for Se^{2-}, and 2.20 (2.21) for Te^{2-}.

For a cation, R_{comp} should be chosen so electrons are attracted less strongly to the nucleus than in the uncompensated cation, but more strongly than in the neutral atom. The following procedure ensures this. First, consider the mth moment of the radius

$$\langle r_i^{(m)} \rangle = \oint r^m \phi_i^*(\vec{r})\, \phi_i(\vec{r})\, d^3\vec{r} \qquad (14)$$

for each occupied Kohn−Sham spin−orbital in the neutral atom. Because the spin−orbital is normalized, the moment for $m = 0$ is 1, while $\langle r_i^{(1)} \rangle$ gives the average radius of the spin−orbital. $\langle r_i^{(-1)} \rangle$ is the average electrostatic potential magnitude exerted by the spin−orbital on the center of the nucleus. Because $m = 1$ weights points in proportion to r while $m = -1$ weights points in proportion to $1/r$, the average potential magnitude exerted on the nucleus is greater than if the electron's entire charge were placed at the average radius:

$$\langle r_i^{(-1)} \rangle > \langle r_i^{(1)} \rangle^{-1} \qquad (15)$$

Accordingly, if we choose R_{comp} as the average of the q largest $\langle r_i^{(1)} \rangle$ chosen from the z_A occupied spin−orbitals of the neutral atom, then the electrostatic potential magnitude exerted by the compensating charge on the nucleus will necessarily be less than for any q electron spin−orbitals present in the neutral atom but greater than if no compensating charge is used. This procedure was used to compute the R_{comp} of the cations, because it gives compensation in the required range between that of the neutral atom and the uncompensated $+q$ cation.

Close examination shows that the conductor-like polarizable continuum model (CPCM) and the c2 compensation method have the same mathematical form when the cavity radius in the CPCM model is set equal to R_{comp}. When no further constraints are added, the CPCM model chooses a polarization charge distribution to make the electric field zero infinitesimally outside the cavity wall, which is the appropriate constraint for c2 anions. The wave function and polarization charge distribution are determined self-consistently with each affecting the other. Analogous to the Hartree−Fock method in which the electron−electron potential operator is twice the electron−electron energy operator, the unconstrained CPCM potential operator is twice the CPCM energy operator.[41] The Gaussian output file lists the polarization charge, q_{CPCM}, corresponding to the potential operator, so $q_{CPCM} = 2q_2$ for anions. For example, direct integration of the enclosed charge for the c2 O^{2-} ion gave $q_1 = -0.865$, while the Gaussian output gave a polarization charge of $q_{CPCM} = 1.735$, which is within an integration tolerance of $2q_2 = -2 \times q_1 = 1.73$. For cations, q_{CPCM} is constrained (using

the Gaussian keyword IComp=1) to give a net charge of zero (i.e., $q_1 + q_{CPCM} + q_3 = 0$), which gives $q_{CPCM} = q_2$ for cations. In this case, q_{CPCM} has a constant value independent of the wave function so it behaves as an external potential with equal energy and potential operators. For the c2 Mg^{2+} ion, the Gaussian output listed a polarization charge of $q_{CPCM} = -1.998$ which is within an integration tolerance of $q_2 = -2$. The c2 reference densities were computed in Gaussian 09 using the CPCM smoothing method of York and Karplus.[41] Example input files are given in the Supporting Information. To ensure applicability to all elements of the periodic table, a universal Gaussian basis set (UGBS) based on the work of de Castro and Jorge[42] was used, and relativistic corrections were included using the fourth order Douglas−Kroll−Hess method with spin−orbit coupling and a finite nuclear model. In this UGBS, we used the same exponents and same spdf shells for all elements, giving 294 basis functions per atom. Having two different charge compensation methods (c1 and c2) makes it easier to compute a complete set of reference densities for the chemical elements by providing options in cases where convergence of one method is more difficult than the other. Because they do not involve periodic boundary conditions, the c2 reference densities converged more readily than the c1 reference densities and decayed exponentially with increasing radius.

Our code implementing the DDEC method is freely available from ddec.sourceforge.net. Our results were computed with this code as follows. Valence and core electron densities were output from VASP or Gaussian on a grid of $N_1 \times N_2 \times N_3$ points defined by translation vectors $\vec{v}_1/N_1$, $\vec{v}_2/N_2$, and $\vec{v}_3/N_3$. We chose density grids such that the summed valence density gave the exact number of valence electrons, n_{val}, to within the greater of 0.1 electrons or 0.03%, then normalized the valence density from the grid to give n_{val}. To ensure integration accuracy, our code did not allow any single grid point to contribute more than 0.03 valence electrons. The code ensured that the electron density of each grid point was not negative and used a zero tolerance to avoid division by zero in eqs 5, 7−12. ${\rho_A}^{\tau}(r_A)$ and ${\rho_A}^{avg}(r_A)$ were computed on 50 radial shells equally spaced between $r_A = 0$ and 3.0 Å. NACs were considered converged when successive iterations resulted in a change of less than 5×10^{-5} e for each atom. Upon convergence, the code generates AMs up to quadrupole order. We performed the following tests for MgO and NaF solids to ensure the accuracy of these integration parameters: (a) doubling N_1, N_2, and N_3 changed the DDEC/c1 (c2) cation charge by −0.006 (−0.007) e for MgO and −0.0002 (−0.0003) e for NaF and (b) using 100 radial shells equally spaced between $r_A = 0$ and 5.0 Å changed the DDEC/c1 (c2) cation charge by +0.042 (+0.002) e for MgO and +0.009 (+0.015) e for NaF.

To improve integration accuracy and reduce computational expense, we treat valence and core electron densities separately. Since $\rho(\vec{r})$ has a cusp near each nucleus, an extremely fine grid spacing would be required to accurately integrate ${\rho_A}^{core}(\vec{r}_A)$. We avoided this problem by using the (integer) number of core electrons, ${n_A}^{core}$, obtained from chemical knowledge. For each atom, the number of valence electrons, ${n_A}^{val}$, is obtained by integrating

$$\rho_A^{val}(\vec{r}_A) = \rho_A(\vec{r}_A) - \rho_A^{core}(\vec{r}_A) \quad (16)$$

which is used to compute ${n_A} = {n_A}^{core} + {n_A}^{val}$. ${\rho_A}^{core}(\vec{r}_A)$ is found by performing an ISA analysis on $\rho^{core}(\vec{r})$, which is accurate because core electron distributions for different atoms are spherically symmetric and barely overlap. Previous implementations of the IH and ISA methods did not use valence-core separation (VCS).[10,12−16] To validate this approach, NACs for h-BN and formamide were computed using the HD, IH, ISA, DDEC/c1, and UC methods. For each method, the RMS difference in NACs computed with and without VCS was only 0.01e, except that IH gave a RMS difference of 0.03e. Not using VCS led to a change in the DDEC/c1 (c2) cation charge of −0.007 (−0.007) e for MgO and +0.018 (+0.030) e for NaF solids relative to NACs computed with VCS. VCS is critical for calculating NACs in large systems, where a fine grid would be too computationally expensive. Using VCS, the DDEC and IH NACs for IRMOF-1, which contains 424 atoms in the unit cell, were computed in much less time than was required for the DFT calculation to generate the electron density. ISA took substantially more iterations and time to converge than DDEC and IH.

Results and Discussion

We now turn to a series of examples that illustrate the properties of DDEC charges and their advantages relative to other charge partitioning methods. In all of the results below, atomic charges and multipoles are given in Hartree atomic units, which use 1e and Bohr as the unit charge and length, respectively. All ΔV_{NAC} and ΔV_{dip} values are in kcal/mol for a test charge of 1e. AM_{max} denotes the largest atomic dipole magnitude (labeled d) or quadrupole component (labeled q) in atomic units. The dominant quadrupole component was

$$Q_{3z^2-r^2} = -\oint (3(\vec{r}_A \cdot \hat{z})^2 - r_A^2)\rho_A(\vec{r}_A)\, d^3\vec{r}_A \quad (17)$$

The z direction was perpendicular to the plane containing the formamide atoms, the BN sheet, the hexagonal planes in h-BN, and the BN nanotube axis.

To highlight the applicability of DDEC charges to both clusters and extended materials, we first examined the zero-, one-, two-, and three-dimensional BN structures shown in Figure 1. The atomic charges computed with a variety of AIM methods are summarized in Table 2, where it is clear that the choice of AIM method has a large effect on the computed NAC. For the 3D crystal, for example, q_B varied from 0.19 (HD) to 2.13 (Bader) with ISA, IH, DDEC, and UC giving intermediate values. Table 2 also includes information about the atomic multipoles (AMs) for each BN material. Including F_{ESP} in the calculation of atomic charges reduced the magnitudes of the AMs. For the DDEC, ISA, and UC methods, which include F_{ESP}, every atomic dipole and quadrupole component was ≤0.5 for all four materials. Bader, HD, and IH calculations, which do not include F_{ESP},

Table 2. Charge Analysis for the BN Structures Shown in Figure 1, with AIM Methods Listed in Order of Ability to Reproduce *V*

	B_4N_4 cluster			BN nanotube			BN sheet			h-BN crystal	
	q_B	AM_{max}	ΔV_{NAC}	q_B	AM_{max}	$\Delta V_{NAC}(\Delta V_{dip})$	q_B	AM_{max}	ΔV_{NAC}	q_B	AM_{max}
V-fit	0.78		0.4	0.47		0.89	0.86		0.5		
DDEC/c1	0.79	0.2d	0.6	1.00	0.2q	7.1(3.2)	0.99	0.2q	0.5	0.96	0.2q
DDEC/c2	0.817	0.2d	0.7	1.088	0.3q	8.2(3.7)	1.09	0.4q	0.5	1.07	0.4q
ISA	0.82	0.2d	0.8	1.09	0.1q	8.5(2.4)	1.07	0.1q	0.5	0.64	0.3q
UC	0.84	0.2d	0.9	1.14	0.4q	8.9(4.2)	1.14	0.5q	0.5	1.12	0.5q
IH	0.89	0.5d	1.3	1.12	2.6q	8.6(16.2)	1.10	2.8q	0.5	1.14	2.3q
HD	0.27	0.5q	4.3	0.20	1.8q	3.5(10.7)	0.20	1.8q	0.6	0.19	1.5q
Bader	0.92	0.7d	1.5	2.16	2.4d	21.0	2.13	2.7d	0.9	2.13	1.6d
no charge	0		6.7	0		5.9	0		0.7		

Table 3. Charge Analysis for Several Solid Materials and NaF Slab

	NaF slab q_{Na} for layer				Na F bulk	NaCl bulk	MgO bulk	MgH_2 bulk		$SrTiO_3$ bulk		
	top	2	3	4	q_{Na}	q_{Na}	q_{Mg}	q_{Mg}	q_H	q_{Sr}	q_{Ti}	q_O
Bader	0.82	0.80	0.80	0.80	0.80	0.82	1.71	1.58	−0.79	1.54	1.88	−1.14
DDEC/c1	1.00	1.03	1.02	1.02	1.02	0.97	2.02	1.54	−0.77	1.96	2.48	−1.48
DDEC/c2	1.00	1.00	1.00	1.00	1.00	0.94	2.05	1.62	−0.81	1.99	2.60	−1.53
HD	0.37	0.34	0.35	0.35	0.33	0.27	0.37	0.32	−0.16	0.66	0.45	−0.37
IH	1.05	1.07	1.08	1.09	1.06	1.01	2.23	2.06	−1.03	2.50	3.11	−1.87
ISA	0.94	0.68	0.70	0.62	0.51	0.41	0.63	0.62	−0.31	2.44	−0.22	−0.74
UC	1.01	1.03	1.02	1.02	1.02	0.98	2.13	1.58	−0.79	2.29	2.72	−1.67

had some atomic dipole or quadrupole components ≥0.5 for all structures. The overall accuracy of AIM methods for reproducing *V* in these materials was DDEC > ISA > UC > IH > HD > Bader. The c1 and c2 reference densities gave similar results. In agreement with eq 3, methods that minimized AMs were the best overall at reproducing *V*. Charges that minimize ΔV_{NAC} are designated *V*-fit in Table 2. For B_4N_4, these are the ESP charges. For the BN sheet and nanotube, these are the unrestrained REPEAT charges using a uniform grid between the surfaces defined by 1.4× and 2.0× vdW radii. The large difference in *V*-fit charges for the BN sheet ($q_B = 0.86$) and nanotube ($q_B = 0.47$) highlight their poor transferability between similar structures. The standard deviation of AIM charges between the four BN structures followed the trend HD < IH < DDEC < UC < ISA < Bader, which shows that AIM methods including a reference state (HD, IH, DDEC, UC) exhibit the best charge transferability. These results show that the DDEC NACs accurately reproduce *V* with reasonable charge transferability compared to other AIM methods.

To better characterize different AIM methods, it is helpful to use a structure that can be described from chemical knowledge. In crystalline NaF, which is highly ionic, a small spillover of electrons from F^- to Na^+ would lead to $0.6 < q_{Na} \approx 0.9 < 1$, and q_{Na} based on experimental diffraction data is 0.95.[43] For this material, the computed q_{Na}'s were 0.33 (HD), 0.51 (ISA), 0.80 (Bader), 1.00 (DDEC/c2), 1.02 (DDEC/c1 and UC), and 1.06 (IH). The HD charge magnitude was too small because it uses neutral instead of charged reference state atoms.[10] To better understand the ISA result, we examined a NaF(100) surface in a slab calculation of eight layers with the middle four layers fixed at bulk positions, the other layers relaxed, and a vacuum spacing of ∼10 Å. As shown in Table 3, q_{Na} for ISA decreased from 0.94 in the surface's top layer to 0.62 in the middle of the slab. This is an example of ISA giving problematic results for a material with buried atoms, an idea supported by the observation that ISA gave reasonable charges for the surface layers. ISA does not constrain the atomic distributions to decay like real atoms as r_A increases. HD, IH, UC, and DDEC do by using a reference density, and the Bader method is an extreme limit where atomic distributions do not overlap. As a result, each of these methods except ISA gave similar NACs for the top and middle layers of the NaF(100) slab.

Charges were also computed for bulk NaCl, $SrTiO_3$, MgO, and MgH_2. Like NaF, q_{Na} for NaCl is expected to be ∼0.9. Since $SrTiO_3$ is partially ionic, one expects the NACs to lie between zero and the oxidation states of +2 (Sr), +4 (Ti), and −2 (O). Likewise, the Mg charge is expected to lie between zero and the oxidation state of +2 in MgH_2. The IH and UC NACs did not fall within this range, indicating that these methods are unsuitable for computing NACs in ionic solids. For $SrTiO_3$, the IH (UC) method gave $q_{Sr} = 2.50$ (2.29), which is too large by ∼0.50 (∼0.29) *e*. The HD and ISA methods failed for the same reasons explained above for the NaF system. The HD charges were too small in magnitude, while the ISA method gave a negative instead of positive charge on Ti in $SrTiO_3$. For all four solids, the Bader method gave NACs less than the oxidation states. The DDEC methods showed that the MgO solid is completely ionic, which agrees with charges extracted from electron diffraction data.[27] The DDEC methods gave reasonable charges, except they often overestimated charge magnitudes by a few hundredths of a unit in completely ionic solids like NaF and MgO. It is hard to pinpoint the source of this small error. Since IH and UC use uncompensated reference densities, while DDEC uses compensated reference densities, these results show the importance of using compensated reference densities when computing charges in solids.

A slab calculation for $SrTiO_3$(100) showed that both DDEC and Bader methods gave chemically reasonable charges for all layers with charges for the middle layers

Table 4. Charge Analysis for Three Oxides Containing Iron in Different Oxidation States

				DDEC c1(c2)			Bader		
	phase	magnetism	Fe oxid. state	q_O	q_{Si}	q_{Fe}	q_O	q_{Si}	q_{Fe}
Fe_2O_3	hematite	antiferro	+3	−1.19 (−1.16)		1.78 (1.74)	−1.05		1.57
Fe_2SiO_4	spinel	antiferro	+2	−1.16 (−1.13)	2.23 (2.21)	1.20 (1.15)	−1.45	3.11	1.34
Fe_3O_4	magnetite	ferri	+3 (tet)	−1.21 (−1.20)		1.76 (1.73)	−1.10		1.46
			+2.5 (oct)			1.54 (1.53)			1.46

converging to the bulk values. This slab was symmetric with a total of four SrO layers and three TiO_2 layers. The geometry of the outer four layers was relaxed, and the inner three layers were fixed at bulk positions. For the top layer, the Sr (O) charges were 1.72 (−1.41) for DDEC/c1, 1.75(−1.46) for DDEC/c2, and 1.61 (−1.26) for Bader, giving a net layer charge of 0.31 (DDEC/c1), 0.29 (DDEC/c2), and 0.35 (Bader). For the next layer, the Ti (O) charges were 2.42 (−1.48) for DDEC/c1, 2.54 (−1.52) for DDEC/c2, and 1.82 (−1.14) for Bader, giving a net layer charge of −0.54 (DDEC/c1), −0.50 (DDEC/c2), and −0.46 (Bader). The subsequent layer charges were similar to the bulk values. Diffraction experiments show that the charge distribution around Sr in bulk $SrTiO_3$ is nearly spherically symmetric, while that between Ti and O is asymmetric; this suggests Sr does not participate in covalent bonding while Ti and O do.[44] Reflections from the spherically symmetric Sr +0 and +2 ions are indistinguishable within the resolution of existing diffraction experiments,[45] making experimental determination of the NACs problematic. Because the DDEC NAC of Sr is approximately an integer (+2) while those of Ti (∼2.5) and O(∼−1.5) are nonintegers, the DDEC results also suggest Sr does not participate in covalent bonding while Ti and O do.

We now consider a series of oxides containing iron in different oxidation states. The oxidation state of Fe in Fe_2SiO_4 is +2, while that in Fe_2O_3 is +3. Magnetite (Fe_3O_4) has an inverse spinel structure containing one Fe +3 oxidation atom in a tetrahedral site and two Fe +2.5 oxidation atoms in octahedral sites. O (Si) have +2 (+4) oxidation states in these compounds. For a given element, a core electron's binding energy measured by X-ray photoelectron spectroscopy (XPS) increases as the net charge or oxidation state of the atom becomes more positive. Yamashita and Hayes measured the Fe $2p_{3/2}$ binding energies as 709.0 ± 0.02 (fayalite phase of Fe_2SiO_4), 710.56 ± 0.05 (magnetite), and 711.0 ± 0.01 (hematite phase of Fe_2O_3) eV, and these binding energies follow the expected trend with increasing average oxidation state.[46] A corresponding trend does not exist for the O 1s binding energies, suggesting no significant differences in the O net charge or oxidation state between these compounds.[46] An unresolved shoulder in the magnetite Fe 3p peak indicates that two different types of Fe are present. Table 4 compares DDEC and Bader charge analysis for Fe_2O_3, Fe_2SiO_4, and Fe_3O_4. For O, the DDEC charges are approximately the same (∼−1.2) in all three materials, while the Bader charge is different (−1.45) for Fe_2SiO_4 than for the other two compounds (∼−1.1). Covalent bonding is suggested since the net charges are much lower in magnitude than the oxidation states. Both the DDEC and Bader methods give higher Fe charge for Fe_2O_3 than for Fe_2SiO_4, which follows the same trend as the Fe oxidation states. For magnetite, DDEC gives a higher net charge for Fe^{tet} than for Fe^{oct}, which follows the expected oxidation state trend and can potentially explain the two different types of Fe observed in the XPS spectrum. The Bader method gives a net charge of 1.46 for both Fe^{tet} and Fe^{oct}.

All of the examples of 3D-periodic materials discussed above were dense materials or surfaces of these solids. As examples of three-dimensional porous materials, we examined two metal−organic frameworks (MOFs). Assigning charges to the atoms in MOFs is crucial in force-field-based simulations of molecular adsorption in these materials,[47] and in this context it is of course desirable that these charges correctly reproduce the electrostatic potential in the MOF's pores. We chose materials for which NACs have been previously reported. Ramsahye et al. reported ESP charges from a nonperiodic cluster model of the MOF large-pore MIL-53 (Al) and Mulliken charges for the periodic crystal.[48] The c1 and c2 reference densities gave approximately equal NACs for this material. Figure 2 shows the structure of this MOF and compares the Bader and DDEC/c1 NACs to the previous results. The ESP and DDEC/c1 NACs were approximately the same. The Mulliken (Bader) NACs were strongly correlated to the DDEC/c1 NACs but had significantly lower (higher) magnitudes. Table 5 lists the ΔV_{NAC} and ΔV_{dip} values for several different charge methods applied to this MOF. Due to the effects of capping atoms in the nonperiodic cluster model, using Ramsahye et al.'s ESP results for the periodic unit cell gives a system whose NACs do not sum to zero; therefore, ΔV_{NAC} for these charges is arbitrarily large, illustrating the advantage of assigning charges in systems like this using the fully periodic system

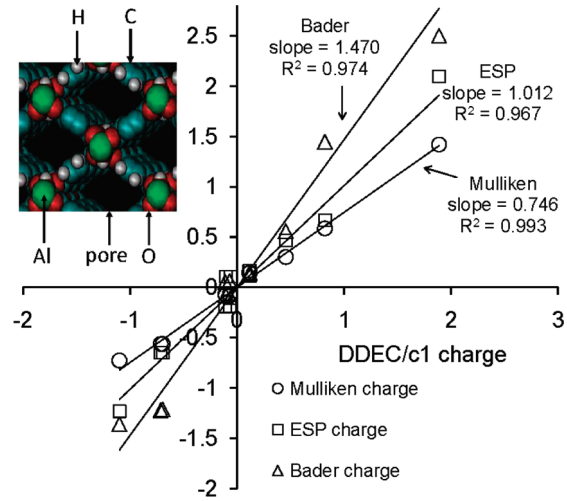


Figure 2. A comparison of ESP, Bader, and Mulliken atomic charges to the DDEC/c1 charges for the MOF MIL-53 (Al).

Table 5. Charge Analysis for Large-Pore MIL-53 (Al), with Methods Listed in Order of Ability to Reproduce *V*

Atom	ISA	DDEC c1(c2)	UC	HD	IH	no charges	Mulliken[48]	Bader	ESP for fragment[48]
ΔV_{NAC}	2.3	2.9 (2.6)	3.0	3.0	4.8	4.9	5.1	14.3	n.a.
ΔV_{dip}	0.9	1.2 (1.2)	1.2	3.2	3.1		n.a.	n.a.	n.a.
α	1	1.05 (1.01)	1.07	0.25	1.43		0.76	1.45	1.09
R^2	1	0.994 (0.997)	0.997	0.977	0.984		0.985	0.965	0.970

Table 6. Charge Analysis for IRMOF-1, with Methods Listed in Order of Ability to Reproduce *V*

atom	REPEAT method[8]	ISA	ESP for fragment[49]	UC	DDEC c1(c2)	ESP for fragment[50]	HD	no charges	Bader
Zn	1.28	1.27	1.26	1.32	1.21 (1.16)	1.31	0.42		1.30
O_1	−1.57	−1.59	−1.44	−1.68	−1.40 (−1.33)	−1.79	−0.34		−1.23
O_2	−0.61	−0.66	−0.67	−0.68	−0.66 (−0.62)	−0.63	−0.22		−1.13
C_1	0.52	0.73	0.68	0.79	0.76 (0.69)	0.62	0.18		1.37
C_2	0.14	−0.08	0.06	−0.16	−0.15 (−0.12)	0.05	−0.01		0.11
C_3	−0.18	−0.10	−0.16	−0.05	−0.06 (−0.06)	−0.12	−0.02		0.22
H	0.17	0.15	0.16	0.12	0.12 (0.12)	0.12	0.04		−0.16
ΔV_{NAC}	0.7	1.0	1.4	1.4	1.5 (1.3)	1.7	2.7	4.2	23.6
ΔV_{dip}	n.a.	0.7	n.a.	0.9	0.9 (0.9)	n.a.	4.0		n.a.
α	1	1.02	0.97	1.07	0.94 (0.90)	1.09	0.27		1.03
R^2	1	0.981	0.989	0.969	0.960 (0.967)	0.993	0.946		0.785

rather than a nonperiodic cluster. ΔV_{NAC} for the remaining methods followed the trend: ISA < DDEC < UC < HD < IH < no charges < Mulliken < Bader. For the ISA, DDEC, UC, and IH methods, ΔV_{dip} was less than ΔV_{NAC}, and the inclusion of atomic dipoles improves the accuracy of the force field for this material. NACs for different charge methods were fit to a line of the form $q = \alpha q_{best}$, where q_{best} was the set of NACs that most accurately reproduced *V* (i.e., the ISA charges). The DDEC, UC, and ESP for fragment methods had slopes and correlation coefficients (R^2) very close to 1, indicating that they give charges similar to the optimal values. On the other hand, the HD, IH, Mulliken, and Bader methods have slopes or correlation coefficients significantly different from 1, indicating that they do not give charges similar to the optimal values.

The recently developed REPEAT method provides electrostatic potential derived atomic charges using a modified error functional that eliminates the need to build a nonperiodic cluster model to obtain *V*-fit NACs for porous periodic systems.[8] Table 6 compares the NAC, ΔV_{NAC}, and ΔV_{dip} values of IRMOF-1 (also called MOF-5) for several AIM methods to those of the REPEAT method and those determined using the ESP method applied to two different nonperiodic clusters. In Table 6, the unique atoms are labeled according to Figure 6b of Campana et al.[8] ΔV_{NAC} for these methods followed the trend REPEAT < ISA < ESP fragment 1, UC, DDEC < ESP fragment 2 < HD < no charges < Bader. (The IH method failed, because it tried to assign a charge magnitude >3 for some of the oxygen atoms.) Thus, the REPEAT method was the most accurate for obtaining force-field charges in this porous periodic system. This implementation of the REPEAT method did not involve any constraints and thus would not be applicable to systems with buried atoms. Of the methods directly applicable to systems with buried atoms, the DDEC and UC methods were the most accurate at reproducing *V*. NACs for different charge methods were fit to a line of the form $q = \alpha q_{best}$, where q_{best} was the set of NACs that most accurately reproduced *V* (i.e., the REPEAT charges). The ISA, DDEC, UC, and ESP for fragment methods had slopes and correlation coefficients (R^2) close to 1, indicating that they give charges similar to the optimal values. The c1 and c2 reference densities gave similar results. On the other hand, the HD and Bader methods had slopes or correlation coefficients significantly different from 1, indicating that they do not give charges similar to the optimal values.

As an example of applying our methods to a large molecular system including heavy atoms, we studied the reaction of Stanciu et al.,[51] in which a 129-atom Zr complex with two bulky bidentate ligands reacts with two O_2 molecules to form a bisperoxy complex. It is important in examples of this kind to characterize charge transfer within the complex before and after reaction. The reaction scheme and contour surfaces for a density value of 0.1 are shown in Figure 3. These contour surfaces show that the charge density concentrates into regions surrounding each ligand, each O_2 molecule, and the metal. Table 7 summarizes the net charges of atoms in these regions computed using several different charge methods. NPA, Bader, DDEC, and UC give similar results and show a charge of ∼+2 residing on the Zr both before and after oxidation. These four methods show that a charge of ∼−1.2 is transferred from each ligand to each O_2

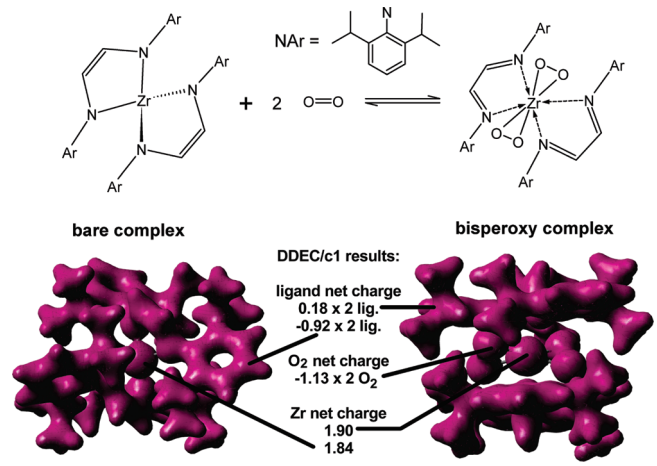


Figure 3. Charge analysis for an electron transfer reaction.

Table 7. Charges for Components of the Zr Complex of Figure 3 after and before Oxidation

	bisperoxy complex			bare complex	
	Zr	O_2	ligand	Zr	ligand
Bader	2.36	−1.24	0.06	2.20	−1.10
DDEC/c1	1.90	−1.13	0.18	1.84	−0.92
DDEC/c2	1.90	−1.09	0.14	1.96	−0.98
HD	0.74	−0.59	0.22	0.84	−0.42
IH	2.60	−1.27	−0.03	2.64	−1.32
ISA	1.66	−0.99	0.16	1.40	−0.70
Mulliken	1.20	−0.90	0.30	1.51	−0.65
NPA	2.00	−1.21	0.21	1.92	−0.96
UC	2.04	−1.13	0.11	2.10	−1.05

during oxidation. The HD, ISA, and Mulliken methods gave lower charge magnitudes, while the IH method gave higher charge magnitudes.

We now consider the accurate characterization of molecular multipoles for gas-phase molecules. Because V at large distances from a molecule is dominated by the lowest-order nonzero molecular multipole moment (MMM), atomistic simulations of gas-phase molecules should use force fields that accurately reproduce the leading MMM. Using formamide ($CHONH_2$) as an example, the molecular dipole magnitude (μ) from each method discussed above was compared to μ defined from the full electron density on the grid density used in the calculation of the NACs. As shown in Table 8, NACs from six of the methods gave small errors in μ: −0.01 (ESP), 0.02 (DDEC and UC), 0.03 (ISA), 0.09 (Mulliken), and 0.11 (IH). The remaining methods gave much larger μ errors: 0.34 (NPA), −0.43 (HD), and 0.84 (Bader). The molecular dipole vector is the sum of the dipole vector due to the NACs and the atomic dipole vectors, so AIM methods that give small atomic dipole magnitudes also give NACs that accurately reproduce μ. As noted above, including F_{ESP} in eq 8 makes the atomic dipole magnitudes small. This means DDEC, ISA, and UC will give more accurate molecular dipoles than other AIM methods for most molecules, not just formamide. For formamide, ΔV_{NAC} followed the trend ESP < ISA < UC, DDEC < IH < Mulliken < NPA < HD < Bader < no charges. ΔV_{NAC} for the best four methods (ESP, ISA, UC, and DDEC) differed by only 0.3 kcal/mol. The RMS difference between the DDEC/c1 and ESP NACs was 0.06e, which was greater than that for the ISA method (0.02), but less than that for the IH (0.07), NPA (0.09), Bader (0.35), Mulliken (0.37), and HD (0.43) methods. Similar NACs, dipole moments, and ΔV_{NAC} were obtained for the c1 and c2 reference densities. Repeating the DDEC/c1 calculation on $\rho(\vec{r})$ obtained from a DFT calculation using plane-waves instead of 6-311++G** basis sets gave only a 0.01e RMS change in NACs, which shows that the DDEC NACs are insensitive to the basis sets used to generate the electron density if the basis set size is adequate.

A useful feature of the DDEC approach is that the DDEC AMs are the multipole moments of $\rho_A^{val}(\vec{r}_A)$, because the core density for each atom does not contribute to the AMs by symmetry. The pth order MMM can be directly formed from the AMs up to order p.[20,21] Therefore, including up to pth order DDEC AMs in the force field exactly reproduces the molecular multipole moments of the valence density up to order p. Due to the finite grid spacing used in DDEC computation, these MMMs may differ by a small amount from those of the full wave function, but this can of course be reduced to an arbitrary level by refining the grid. A key drawback of the ESP method is that adding atomic dipoles requires a new optimization, which leads to new atomic charges. For example, the N charge in formamide changed from −0.89 to −1.47 when atomic dipoles were added to the ESP method; moreover, the atomic dipoles were large in magnitude ($AM_{max} = 1.5$). The DDEC method does not have this drawback, as the atomic charges are unchanged regardless of the number of AMs used and the AMs are optimized to be small in magnitude. For formamide, including atomic dipoles decreased the DDEC/c1 RMS V error from 1.6 to 1.1 kcal/mol and increased the ESP RMS V error from 1.3 to 11.8 kcal/mol. The increase in RMS V error when atomic dipoles are added to the ESP method is possible because the grid points used to evaluate the force field's RMS V error were not the same as those used to fit the atomic dipoles. The poor performance of the ESP method when atomic dipoles are included is due to the large AM_{max}. The DDEC approach does not have this problem and provides a useful framework for defining AMs in situations where extremely accurate approximation of the electrostatic potential outside the electron distribution is desired.

Finally, we consider the formal relationship between the DDEC method and computational chemistry methods. Like DFT, the DDEC method assumes electron transfer occurs

Table 8. Charge Analysis for Formamide, with Methods Listed in Order of Ability to Reproduce V

	net atomic charges						quantitative measures of V fitting			
	C	O	N	H[b]	H[c]	H[c]	$\Delta\mu$[a]	ΔV_{NAC}	ΔV_{dip}	AM_{max}
ESP	0.62	−0.53	−0.89	−0.01	0.42	0.38	−0.01	1.3	n.a.	n.a.
ISA	0.63	−0.53	−0.87	−0.01	0.40	0.38	0.03	1.4	1.0	0.4 (q)
DDEC/c2	0.58	−0.52	−0.79	0.00	0.37	0.36	0.02	1.5	1.0	0.4 (q)
UC	0.62	−0.54	−0.82	0.00	0.38	0.37	0.02	1.5	1.0	0.4 (q)
DDEC/c1	0.59	−0.53	−0.76	0.00	0.36	0.35	0.02	1.6	1.1	0.5 (q)
IH	0.50	−0.47	−0.81	0.04	0.37	0.36	0.11	2.1	2.3	0.9 (q)
Mulliken	0.00	−0.33	−0.29	0.09	0.27	0.25	0.09	3.9	n.a.	n.a.
NPA	0.47	−0.57	−0.79	0.10	0.40	0.39	0.34	4.3	n.a.	n.a.
HD	0.09	−0.28	−0.12	0.04	0.13	0.13	−0.43	5.4	1.9	1.0 (q)
Bader	1.27	−1.02	−1.12	0.04	0.43	0.40	0.84	12.5	4.9	2.9 (q)
ESP + dipoles	0.92	−1.00	−1.47	0.33	0.80	0.41	0.00	14.6	11.8	1.5 (d)
no charges							−1.56	16.2	n.a.	n.a.

[a] Error in μ compared to grid electron density. [b] H bonded to C. [c] H bonded to N.

Table 9. Charge Analysis for Formamide and Three Ozone States Using Different Exchange-Correlation Theories

				ISA		DDEC/c1		DDEC/c2	
	μ (au)	X−O (Å)	O−O−O (deg)	μ_{NAC}	q_O[a]	μ_{NAC}	q_O[a]	μ_{NAC}	q_O[a]
formamide ($q_O = -0.56 \pm 0.03$, $\mu_{NAC} = 1.60 \pm 0.02$)									
PW91	1.56	1.22		1.58	−0.53	1.58	−0.53	1.58	−0.52
B3LYP	1.59	1.21		1.63	−0.59	1.62	−0.57	1.62	−0.56
CCSD	1.54	1.21		1.62	−0.60	1.61	−0.59	1.60	−0.58
exp.	1.46[b]	1.24[b]							
ozone singlet ($q_O = 0.36 \pm 0.03$, $\mu_{NAC} = 0.45 \pm 0.03$)									
PW91	0.25	1.28	118.4	0.46	0.37	0.46	0.37	0.45	0.36
B3LYP	0.27	1.26	118.5	0.48	0.40	0.49	0.40	0.47	0.39
CCSD	0.25	1.24	117.8	0.46	0.37	0.46	0.38	0.45	0.37
CAS(12,9)	0.21	1.28	116.9	0.40	0.32	0.41	0.32	0.40	0.31
exp.	0.21[c]	1.27[d]	116.8[d]						
ozone triplet ($q_O = 0.21 \pm 0.03$, $\mu_{NAC} = 0.34 \pm 0.06$)									
PW91	0.29	1.35	99.3	0.35	0.21	0.34	0.21	0.33	0.20
B3LYP	0.32	1.32	99.2	0.38	0.23	0.37	0.23	0.37	0.22
CCSD	0.32	1.32	97.5	0.39	0.24	0.39	0.23	0.38	0.23
CAS(12,9)	0.22	1.36	99.3	0.21	0.16	0.26	0.16	0.26	0.16
exp.		1.35[e]	98.9[e]						
ozone +1 doublet ($q_O = 0.48 \pm 0.05$, $\mu_{NAC} = 0.14 \pm 0.03$)									
PW91	0.01	1.23	134.3	0.14	0.49	0.14	0.49	0.14	0.48
B3LYP	0.01	1.21	134.6	0.18	0.54	0.18	0.54	0.17	0.53
CCSD	0.01	1.18	133.4	0.14	0.50	0.15	0.50	0.14	0.49
CAS(11,9)	0.04	1.23	128.1	0.09	0.42	0.09	0.42	0.08	0.41
exp.		1.25[f]	131.5[f]						

[a] For ozone, the center O atom has charge q_O, and each outer O atom has charge $-q_O/2$ (singlet, triplet) or $(1 - q_O)/2$(+1 cation). [b] Ref 54. [c] Ref 55. [d] Ref 56. [e] Ref 57. [f] Ref 53.

to minimize a system's energy subject to optional constraints; therefore, the DDEC method is applicable to the same types of systems as DFT. Optional constraints include fixing the net charge, one or more atomic positions, and the number of up electrons minus down electrons. As described above, the DDEC charge distributions depend only on $\rho(\vec{r})$, the reference densities, and χ. Quantum chemistry methods such as DFT, coupled cluster theory, configuration interaction, and quantum Monte Carlo can formally reproduce the exact $\rho(\vec{r})$ in the limit of an exact exchange-correlation methodology with relativistic corrections and arbitrarily large basis sets. All these methods yield the same $\rho(\vec{r})$ and thus identical DDEC charges in this limit. Actual computations involve compromises that approximate $\rho(\vec{r})$, and this causes different computational chemistry methods to give slightly different DDEC charges.

A system is called single-reference if the configuration interaction expansion of its wave function has one Slater determinant whose coefficient is an order of magnitude greater than that of any other Slater determinant; otherwise, the system is called multireference. Examples in this article are single-reference systems, except for the magnetic systems (Fe_2O_3, Fe_2SiO_4, and Fe_3O_4) discussed above and the ozone states discussed below. (Ozone is a commonly studied multireference system.[52,53]) In antiferromagnetic systems like Fe_2O_3 and Fe_2SiO_4, each magnetic center (e.g., Fe atom) has a quantum superposition of spin-up and spin-down magnetism. When applying DFT to these antiferromagnetic systems, we used the common technique of breaking symmetry by assigning spin-up magnetism to one Fe atom and spin-down magnetism to the other Fe atom in the unit cell, and this approximation gave reasonable $\rho(\vec{r})$ and NACs. For small systems, the complete active space (CAS) and related methods can be used to find correlated multireference wave functions.

Table 9 shows dipole moments, select geometric parameters, and oxygen charges for formamide and three ozone states using different exchange-correlation theories. μ and μ_{NAC} are the dipole moments due to $\rho(\vec{r})$ and the NACs, respectively. AUG-cc-pVTZ basis sets were used for coupled-cluster theory with single and double excitations (CCSD) and CAS, and 6-311++G** basis sets were used for the DFT methods. CAS(n,m) denotes an active space containing n electrons and m orbitals. $m = 9$ and $n = 12$ (11 for +1 cation) are the same size as the 2p valence space for 3 O atoms and a reasonable choice for the active space. In the CAS calculations, molecular orbitals with the lowest energy eigenvalues were fully occupied, those with the highest eigenvalues were completely unoccupied, and those with intermediate eigenvalues were included in the active space. The ozone potential energy surface contains dozens of equilibrium geometries having various spin multiplicities and net charges.[52,53] Here, we consider three equilibrium structures that are similar to the experimentally observed global energy minima for the singlet, triplet, and +1 doublet states. Each of these structures has two equal O−O bond lengths. The similarity of the computed and experimental geometries verifies that the appropriate minima on the potential energy surface have been found. For each state, the average and standard deviation of q_O and μ_{NAC} are given in parentheses in Table 9. Each average and standard deviation included all the levels of theory (PW91, B3LYP, CCSD, and CAS for ozone states) and charge methods (ISA, DDEC/c1, and DDEC/c2). In each case, the standard deviation was small (≤ 0.06), because different levels of theory and charge

methods gave similar results. Most importantly, the standard deviation for a given molecular state was much smaller than the change observed between two different molecular states. For example, $q_O = 0.21 \pm 0.03$ for the ozone triplet state was decisively lower than $q_O = 0.36 \pm 0.03$ for the ozone singlet state. For a given molecular state, the maximum change in q_O between two different exchange-correlation theories was 0.13 even when changing between c1 and c2 reference densities. The overall message from Table 9 is that DDEC charges can be reliably computed for a variety of computational chemistry methods, including DFT, coupled-cluster theory, and multireference methods like CAS.

Additional tests showed that the DDEC NACs depend less on the exchange-correlation theory used to compute the reference densities than they do on the exchange-correlation theory used to compute $\rho(\vec{r})$. For the five bulk materials in Table 3, the RMS change in NAC (Δq_{rms}) for c1 ref densities computed with the Ceperley-Alder LSDA functional versus PW91 (geometry and $\rho(\vec{r})$ from LSDA in both cases) was $0.023e$. This was smaller than $\Delta q_{rms} = 0.026e$ when the geometry and $\rho(\vec{r})$ were obtained from LSDA versus PW91 (using PW91 c1 ref densities in both cases). When the geometries, $\rho(\vec{r})$, and c1 ref densities were all computed with LSDA versus PW91, the change was slightly larger ($\Delta q_{rms} = 0.031e$). For the three magnetic materials in Table 4, $\Delta q_{rms} = 0.034e$ when the c1 ref densities were recomputed with PBE. Similar insensitivity was also observed for the c2 ref densities. Specifically, for the four BN structures in Table 2, $\Delta q_{rms} = 0.006e$ when the c2 ref densities were recomputed with the hybrid functional MPW1PW91. For formamide using PW91 c2 ref densities, $\Delta q_{rms} = 0.037$ $(0.059)e$ for geometry and $\rho(\vec{r})$ from B3LYP (CCSD) versus PW1. When c2 ref densities were recomputed with MPW1PW91, all of the B3LYP, CCSD, and PW91 NACs for formamide changed by less than $0.005e$. These results show it is not necessary to compute the reference densities and $\rho(\vec{r})$ with the same exchange-correlation theory. The optimal exchange-correlation functional for computing the reference densities should correctly reproduce each element's ground state electron configuration. The PW91 functional did this for all of the elements we studied, so all of the NACs given in Tables 1−9 used PW91 reference densities.

Conclusions

A systematic comparison has been made between different methods for computing atomic charges, with particular emphasis on AIM methods. Although a unique definition for the charge associated with an atom in a material does not exist, we believe an ideal AIM method should satisfy several criteria. First, it should give NACs that reflect the amount of charge transfer between atoms. For example, it should give NACs that are approximately less than or equal to the formal oxidation state for ionic compounds. Second, it should give NACs that approximately reproduce V outside the electron distribution. Third, it should be applicable to a wide variety of materials, including periodic and nonperiodic materials with and without buried atoms. Fourth, it should give NACs with good transferability between closely related systems. We provided evidence that existing AIM methods fall short of these goals. Prior research has shown that the HD method almost always underestimates NACs,[10,11] and our results confirm this. Upon calculating IH and ISA NACs for the first time in periodic materials, we found that these methods do not work well for nonporous solids. The ISA method is problematic for systems containing buried atoms, while the IH method overestimates charge magnitudes in many extended systems. Because the ISA method does not require the calculation of reference densities, it is an appealing method for calculating NACs in systems that do not contain buried atoms, and we found that it accurately reproduced V in these systems. The Bader method consistently gave chemically reasonable charges for the nonporous solids we studied, but Bader NACs did not accurately reproduce V outside the electron distribution in porous materials.

We developed a new AIM method, called density derived electrostatic and chemical (DDEC) charges to address these problems. DDEC optimizes the NACs to be chemically meaningful and to reproduce V outside the electron distribution by simultaneously optimizing atomic electron density distributions to be close to spherically symmetric and to resemble reference state atoms. For accurate calculation of atomic charges in extended systems, we found that the reference densities must include the effects of charge compensation. These improvements allow the DDEC method to account for three different types of charge transfer that take place in materials: (a) ionic bonding, (b) covalent bonding, and (c) charge compensation and dielectric screening. Charge analysis was performed for a B_4N_4 cluster, a hexagonal BN sheet, a BN nanotube, nine nonporous solids (h-BN, NaF, NaCl, MgO, MgH_2, $SrTiO_3$, Fe_2O_3, Fe_2SiO_4, and Fe_3O_4), a NaF slab, a $SrTiO_3$ slab, the MIL-53 (Al) MOF, IRMOF-1, a $ZrN_4C_{52}H_{72}$ organometallic complex, a $ZrO_4N_4C_{52}H_{72}$ organometallic complex, formamide, and three ozone states (singlet, triplet, and +1 cation). These examples included systems with periodicity along zero, one, two, and three dimensions; porous and nonporous solids; solid surfaces; and both large and small molecules. The DDEC method consistently gave reasonable charges in a wide variety of periodic and nonperiodic materials with and without buried atoms. The DDEC NACs were consistent with experimental results for NaF, MgO, $SrTiO_3$, and a series of oxides containing iron, and for these systems showed slightly better correspondence than the Bader charges to the experimental data. Furthermore, the DDEC NACs were not sensitive to the basis sets used to generate the electron density. However, the DDEC method often overestimated charge magnitudes by a few hundredths of a unit for completely ionic solids.

Two methods were developed for calculating charge compensated reference densities. In the c1 method, each ion is placed in the center of a 10 Å × 10 Å × 10 Å periodic unit cell made neutral by a uniform background charge. The c2 method does not involve periodicity and places each ion in the center of a spherical shell of compensating charge. For anions, the compensating charge is chosen to make the electric field zero infinitesimally outside the shell radius which is varied to minimize the system's total energy. These

optimized radii were similar to crystal-structure-derived anionic radii from the literature (where available). For a $+q$ cation, the compensating charge is constrained to $-q$ and the shell radius is set equal to the average radius of the outermost q occupied Kohn−Sham spin−orbitals in the neutral atom. The DDEC/c1 and DDEC/c2 methods gave similar results for all of the systems we studied. The c1 and c2 reference densities computed with the PW91 functional gave good performance even when $\rho(\vec{r})$ for the system being studied was computed with a different level of theory (e.g., PBE, B3LYP, LSDA, CCSD, or CAS); thus, it is not critical to recompute the reference densities at different levels of theory. Between different reference density sets or exchange-correlation theories, the DDEC NACs usually differed by $\leq 0.1e$. The DDEC NACs were also within $0.1e$ of experimental charges for NaF and MgO. All of these results suggest that the DDEC NACs can be routinely computed with an absolute accuracy of approximately $0.1e$.

Tests were performed to assess the relative ability of NACs computed using different methods to reproduce V outside the electron distribution. The ESP method, and its recent extension to periodic materials called the REPEAT method, were found to be the most accurate methods for reproducing V in nonperiodic and periodic materials, respectively. This is expected, because these two methods specifically optimize the NACs to minimize ΔV_{NAC}. Unlike the ESP and REPEAT methods, the DDEC method does not require the addition of constraints to handle systems with buried atoms. Of the methods we tested that were applicable to systems with buried atoms, the DDEC method was found to be the most accurate for reproducing V outside the electron distribution. The Mulliken, HD, and Bader methods were found to be inaccurate for reproducing V. Because DDEC charges are optimized to resemble a chemical reference state, they exhibited better transferability than charges simply optimized to fit V. Because the AMs are optimized to be small, DDEC NACs accurately reproduce the dipole moments of molecules and can be used to quantify the amount of charge transfer between atoms during chemical reactions. We anticipate that DDEC charges will find application in the development of interatomic potentials for atomistic simulations of complex materials, particularly for spatially periodic materials containing buried atoms.

Acknowledgment. Financial support was provided by the National Science Foundation through grant CHE0651182 and the Department of Energy through contract DE-SC0001058. Supercomputing resources were provided by Georgia Institute of Technology and Teragrid grants TG-CHE090019 and TG-CTS100027.

Supporting Information Available: Coordinates of optimized geometries, more extensive results tables, and sample input files for computing the c2 reference densities. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Phillpot, S. R.; Sinnott, S. B. *Science* **2009**, *325*, 1634–1635.

(2) Wiberg, K. B.; Rablen, P. R. *J. Comput. Chem.* **1993**, *14*, 1504–1518.

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

(4) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129–145.

(5) Dognon, J.-P.; Durand, S.; Granucci, G.; Levy, B.; Millie, P.; Rabbe, C. *J. Mol. Struct.* **2000**, *507*, 17–23.

(6) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269–10280.

(7) Sherwood, P.; de Vries, A. H.; Collins, S. J.; Greatbanks, S. P.; Burton, N. A.; Vincent, M. A.; Hillier, I. H. *Faraday Disc.* **1997**, *106*, 79–92.

(8) Campana, C.; Mussard, B.; Woo, T. K. *J. Chem. Theory Comput.* **2009**, *5*, 2866–2878.

(9) Hirshfeld, F. L. *Theor. Chim. Acta* **1977**, *44*, 129–138.

(10) Bultinck, P.; Van Alsenoy, C.; Ayers, P. W.; Carbo-Dorca, R. *J. Chem. Phys.* **2007**, *126*, 144111.

(11) Davidson, E. R.; Chakravorty, S. *Theor. Chim. Acta* **1992**, *83*, 319–330.

(12) Bultinck, P.; Ayers, P. W.; Fias, S.; Tiels, K.; Van Alsenoy, C. *Chem. Phys. Lett.* **2007**, *444*, 205–208.

(13) Bultinck, P.; Cooper, D. L.; Van Neck, D. *Phys. Chem. Chem. Phys.* **2009**, *11*, 3424–3429.

(14) Lillestolen, T. C.; Wheatley, R. J. *Chem. Commun.* **2008**, 5909–5911.

(15) Lillestolen, T. C.; Wheatley, R. J. *J. Chem. Phys.* **2009**, *131*, 144101.

(16) Van Damme, S.; Bultinck, P.; Fias, S. *J. Chem. Theory Comput.* **2009**, *5*, 334–340.

(17) Bader, R. F. W.; Macdougall, P. J.; Lau, C. D. H. *J. Am. Chem. Soc.* **1984**, *106*, 1594–1605.

(18) Bader, R. F. W.; Matta, C. F. *J. Phys. Chem. A* **2004**, *108*, 8385–8394.

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

(20) Kosov, D. S.; Popelier, P. L. A. *J. Phys. Chem. A* **2000**, *104*, 7339–7345.

(21) Laidig, K. E. *J. Phys. Chem.* **1993**, *97*, 12760–12767.

(22) Freitag, M. A.; Gordon, M. S.; Jensen, J. H.; Stevens, W. J. *J. Chem. Phys.* **2000**, *112*, 7300–7306.

(23) Nalewajski, R. F.; Parr, R. G. *Proc. Nat. Sci. Acad. U.S.A.* **2000**, *97*, 8879–8882.

(24) Maslen, E. N.; Spackman, M. A. *Aust. J. Phys.* **1985**, *38*, 273–287.

(25) Leenaerts, O.; Partoens, B.; Peeters, F. M. *Appl. Phys. Lett.* **2008**, *92*, 243125.

(26) Watson, R. E. *Phys. Rev.* **1958**, *111*, 1108–1110.

(27) Zuo, J. M.; OKeeffe, M.; Rez, P.; Spence, J. C. H. *Phys. Rev. Lett.* **1997**, *78*, 4777–4780.

(28) Hafner, J. *J. Comput. Chem.* **2008**, *29*, 2044–2078.

(29) Kresse, G.; Furthmüller, J. *Phys. Rev. B* **1996**, *54*, 11169–11186.

(30) Kresse, G.; Joubert, D. *Phys. Rev. B* **1999**, *59*, 1758–1775.

(31) 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.; Nakasuji, 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.; Nakamjima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A. J.; 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, N. J.; 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, O.; Foresman, J. B.; Ortiz, J. V.; Ciolowski, J.; Fox, D. J. *Gaussian 09*, Revisions up to and including A.02; Gaussian, Inc.: Wallingford, CT, 2009.

(32) Loiseau, T.; Serre, C.; Huguenard, C.; Fink, G.; Taulelle, F.; Henry, M.; Bataille, T.; Ferey, G. *Chem.—Eur. J.* **2004**, *10*, 1373–1382.

(33) Wyckoff, R. W. G. *Crystal Structures*; John Wiley & Sons: New York, 1963; Vol. 1, 85−237.

(34) Li, H.; Eddaoudi, M.; O'Keeffe, M.; Yaghi, O. M. *Nature* **1999**, *402*, 276–279.

(35) Henkelman, G.; Arnaldsson, A.; Jonsson, H. *Comput. Mater. Sci.* **2006**, *36*, 354–360.

(36) Sanville, E.; Kenny, S. D.; Smith, R.; Henkelman, G. *J. Comput. Chem.* **2007**, *28*, 899–908.

(37) Tang, W.; Sanville, E.; Henkelman, G. *J. Phys. Condens. Matter* **2009**, *21*, 084204.

(38) Smith, W. *CCCP5 Newsletter* **1998**, *46*, 18–30.

(39) Manz, T. A.; Sholl, D. S. *J. Comput. Chem.* **2010**, *31*, 1528–1541.

(40) *CRC Handbook of Chemistry and Physics*; 90 ed.; Lide, D. R., Ed.; CRC Press: Baco Raton, FL, 2009; Sec. 14, pp 11−12.

(41) York, D. M.; Karplus, M. *J. Phys. Chem. A* **1999**, *103*, 11060–11079.

(42) de Castro, E. V. R.; Jorge, F. E. *J. Chem. Phys.* **1998**, *108*, 5225–5229.

(43) Su, Z. W.; Coppens, P. *Acta Crystallogr., Sect. A* **1995**, *51*, 27–32.

(44) Ikeda, T.; Kobayashi, T.; Takata, M.; Takayama, T.; Sakata, M. *Solid State Ionics* **1998**, *108*, 151–157.

(45) Jauch, W.; Reehuis, M. *Acta Crystallogr., Sect. A* **2005**, *61*, 411–417.

(46) Yamashita, T.; Hayes, P. *Appl. Surf. Sci.* **2008**, *254*, 2441–2449.

(47) Keskin, S.; Liu, J.; Rankin, R. B.; Johnson, J. K.; Sholl, D. S. *Ind. Eng. Chem. Res.* **2009**, *48*, 2355–2371.

(48) Ramsahye, N. A.; Maurin, G.; Bourrelly, S.; Llewellyn, P.; Loiseau, T.; Ferey, G. *Phys. Chem. Chem. Phys.* **2007**, *9*, 1059–1063.

(49) Amirjalayer, S.; Tafipolsky, M.; Schmid, R. *Angew. Chem., Int. Ed.* **2007**, *46*, 463–466.

(50) Sagara, T.; Klassen, J.; Ganz, E. *J. Chem. Phys.* **2004**, *121*, 12543–12547.

(51) Stanciu, C.; Jones, M. E.; Fanwick, P. E.; Abu-Omar, M. M. *J. Am. Chem. Soc.* **2007**, *129*, 12400–12401.

(52) Kalemos, A.; Mavridis, A. *J. Chem. Phys.* **2008**, *129*, 054312.

(53) Willitsch, S.; Innocenti, F.; Dyke, J. M.; Merkt, F. *J. Chem. Phys.* **2005**, *122*, 024311.

(54) Kurland, R. J.; Wilson, E. B. *J. Chem. Phys.* **1957**, *27*, 585–590.

(55) Mack, K. M.; Muenter, J. S. *J. Chem. Phys.* **1977**, *66*, 5278–5283.

(56) Tyuterev, V. G.; Tashkun, S.; Jensen, P.; Barbe, A.; Cours, T. *J. Mol. Spectrosc.* **1999**, *198*, 57–76.

(57) Bouvier, A. J.; Inard, D.; Veyret, V.; Bussery, B.; Bacis, R.; Churassy, S.; Brion, J.; Malicet, J.; Judge, R. H. *J. Mol. Spectrosc.* **1998**, *190*, 189–197.

CT100125X

JCTC Journal of Chemical Theory and Computation

On the Interfragment Exchange in the X-Pol Method

Alessandro Cembran, Peng Bao, Yingjie Wang, Lingchun Song,* Donald G. Truhlar,* and Jiali Gao*

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

Received May 21, 2010

Abstract: The inclusion of exchange repulsion terms in the explicit polarization (X-Pol) model is examined by antisymmetrizing the X-Pol Hartree product wave function; this yields X-Pol with full eXchange, called X-Pol-X. When the monomers are treated by Hartree−Fock theory, this calculation can be accomplished by using the formalism of block-localized wave functions (BLW) that has been used in a variety of applications. In this case the block-localized structure in the X-Pol-X wave function allows for decomposition of the full Fock matrix of a dimension of *M* blocks into *M* smaller Fock matrices. The method is illustrated by considering two trimer structures of water clusters, and it is found that the total exchange repulsion energies in these hydrogen-bonding test cases are adequately treated and—to a good approximation—are pairwise additive. We also present a formalism to yield a simplified Fock matrix by making use of the neglect of interfragment differential overlap (NIDO) approximation, which is less severe than the neglect of diatomic differential overlap (NDDO) approximation.

1. Introduction

The explicit polarization (X-Pol) method,[1] initially called molecular orbital derived potentials for liquids (MODEL), was introduced in 1997[2] for the treatment of macromolecular systems in statistical mechanical Monte Carlo and molecular dynamics simulations using electronic structure theory.[3−5] It is based on the partition of a whole system into subsystems, which are called fragments or blocks, each of which is represented by an antisymmetrized wave function.[2,3] The system wave function is then a Hartree product of fragment wave functions, which are optimized by the self-consistent field (SCF) method in the presence of the external field of all other blocks until the energy or electron density of the entire system is converged.[2,3,6−8] The individual wave functions of the subsystems can be obtained at any level of theory: ab initio Hartree−Fock (HF), semiempirical molecular orbital theory, correlated wave function theory, or Kohn−Sham (KS) density functional theory (DFT).[9] In general, the X-Pol method can be applied to post-SCF calculations such as MP2 and multiconfigurational methods, and a post-SCF treatment can be used to estimate dispersion interaction energies between different fragments. In contrast to the methods of Stoll et al.[10,11] (whose emphasis was on the localization of orbitals on atomic centers with neighboring bonding partners and lone-pairs) and Yang[12] (whose emphasis was on full calculations for large systems by a divide-and-conquer approach), X-Pol is not designed to reproduce the full HF or KS energy; rather, it is intended to establish a theoretical framework for economical and accurate applications to macromolecular systems and liquids.[3,5] In the present article we apply it to clusters.

The Hartree product approximation of the X-Pol wave function implies that interfragment exchange and correlation interactions are neglected.[1−3,6,8] The most noticeable contributions to the interfragment exchange term are from the repulsion interactions at geometries where orbitals on different blocks have significant overlap (short-range), whereas the dominant contributions to interfragment correlation energy are dispersion interactions at geometries where they have negligible overlap (long-range). To achieve computational efficiency, we have proposed to use empirical Lennard-Jones terms to account for these interactions in the spirit of developing a quantal force field for macromolecular simulations.[1−3,5,6] Alternatively, interfragment exchange can be included in the X-Pol method by fully antisymmetrizing the block-localized molecular orbitals (BLMOs) to form a

* Corresponding authors: e-mail lcsong2007@gmail.com (L.S.), truhlar@umn.edu (D.G.T.), and gao@jialigao.org (J.G.).

10.1021/ct100268p © 2010 American Chemical Society
Published on Web 07/20/2010

determinant wave function, and this is equivalent to the block-localized wave function (BLW) method described by Mo and Peyerimhoff,[13] which has been applied to a variety of problems.[9,14−25] In the present context, this approach will be called X-Pol with full eXchange or X-Pol-X. In such a wave function the orbital coefficient matrix is block-diagonal, which has the property that orbitals within each block are orthogonal, and orbitals between different subsystems are nonorthogonal.[11,13,14,26] The latter is characteristic of valence bond theory.[9,16,23,27] Methods for treating nonorthogonal determinantal wave functions have been described a long time ago and used in various applications, including the work of King et al.,[28] Gallup,[29] Stoll and co-workers,[11,30] Gianinetti and co-workers,[31,32] Mo et al.,[13−16,27] Head-Gordon and co-workers,[33] and Cembran et al.[9] The BLW method has been used to define valence bond-like states in the mixed molecular orbital and valence bond (MOVB) theory[16,22,23,27] and in the multistate density functional theory (MSDFT) based on valence bond or VBDFT.[9] Here, we use the X-Pol-X (i.e., BLW) method to elucidate the magnitude of the exchange repulsion term that is required in the X-Pol method.[15] In the Appendix we derive a simplified X-Pol-X method by invoking neglect of interfragment differential overlap (NIDO).

It should be noted that, in 1999, Kitaura et al.[34,35] used the wave functions of dimeric fragments in a post-SCF correction to estimate the pairwise exchange repulsion. The expansion of the monomeric subspace into two fragments also includes pairwise charge transfer effects that are important for describing short-range interactions.[10] This approach was called the fragment molecular orbital (FMO) method, and it has been used in applications with various electronic structural models.[36] However, aside from the dimer correction term, first described by Stoll and Preuss,[10] the construction of block-localized (fragment) orbitals and the SCF computational algorithm (see page 321 of ref 34) of the FMO method[34,35] are essentially the same as that (see page 660 of ref 2) in X-Pol theory.[2,3] Nevertheless, these studies illustrate the usefulness of block localization of molecular orbitals for macromolecular systems.

Section 2 presents the X-Pol method in a general way and then describes the X-Pol-X extension of the X-Pol method to include the exchange repulsion energy in SCF calculations. Section 3 describes an illustrative application at the ab initio Hartree−Fock level. Section 4 gives the results and discusses them, with special emphasis on comparison of X-Pol to X-Pol-X. Section 5 summarizes the conclusions. The NIDO simplification procedure, which may be very useful in future work, is presented in the Appendix.

2. Method

We first describe the X-Pol method. Then, we present the SCF procedure for optimization of an X-Pol-X wave function. It is especially interesting to consider the case where the fragment wave functions are Hartree−Fock wave functions, obtained either by ab initio theory or by semiempirical molecular orbital theory. Then the fragment wave functions are optimized by the self-consistent field (SCF) method in the presence of the external field of all other blocks.[2,3,6−8] This is sometimes called the double SCF (DSCF) procedure because one must achieve self-consistency both within each fragment and among all blocks in the system. As a consequence of the interfragment self-consistency, the fragment wave functions are polarized.

2.1. Block Localization and the X-Pol Wave Function. We partition a large molecular system into M blocks (which have been equivalently and interchangeably called subgroups, subsystems, molecules, residues, monomers, fragments, or bodies). The term block is more general than some other terms, such as fragment and body, in that orbitals on the same atom can be divided into different subsystems without the physical constraint of atomic structure fragmentation.[14] Block a contains k_a basis functions and n_a electrons. Thus, the total number of primitive basis functions, K, and the total number of electrons, N, in the system are

$$K = \sum_{a=1}^{M} k_a \quad \text{and} \quad N = \sum_{a=1}^{M} n_a \tag{1}$$

The molecular orbitals φ_j^a in each block a are written as linear combinations of the primitive basis functions $\{\chi_\mu^a;\ \mu = 1, ..., k_a\}$ in that specific subspace:

$$\varphi_j^a = \sum_{\mu=1}^{k_a} c_{j\mu}^a \chi_\mu^a \tag{2}$$

Thus, by construction, the molecular orbitals are strictly localized within the subspace of each block.

The X-Pol wave function is approximated as a product of the individual, antisymmetric fragment wave functions of all M blocks:

$$\Psi^{\text{X-Pol}} = \prod_{a=1}^{M} \hat{A}\Phi_a \tag{3}$$

where $\hat{A}$ is an antisymmetrization operator, and Φ_a is a product of the occupied spin orbitals in block a (eq 2):

$$\Phi_a = \varphi_1^a \varphi_2^a ... \varphi_{n_a}^a \tag{4}$$

The Hartree product wave function in eq 3 includes Coulomb interactions, but the exchange interactions are ignored between different blocks. If we make a further approximation by representing the Coulomb potentials by a classical multipole expansion or by the potential due to distributed monopoles calculated by Mulliken population analysis, the computation of the Fock matrix for block a, in the presence of the external field of the rest of the system, is simplified to

$$\mathbf{F}^a = \mathbf{H}^a + \sum_i \left(\mathbf{J}_i^a - \frac{1}{2}\mathbf{K}_i^a \right) + \sum_{b \neq a} \mathbf{V}_b^a \tag{5}$$

where $\mathbf{H}^a$ is the one-electron Hamiltonian matrix, $\mathbf{J}_i^a$ and $\mathbf{K}_i^a$ are the Coulomb and exchange integral matrices associated with molecular orbital i in block a, and the last summation term is the Coulomb interaction matrix due to the nuclei and wave functions of all other blocks in the system. The interaction between blocks is at a level comparable to a force field for macromolecular systems.[1,2,6,8] The Fock matrix $\mathbf{F}^a$ depends on the instantaneous polarization of the wave functions of other

blocks through $\mathbf{V}_b^a$, which in turn depends on $\mathbf{F}^a$. The electronic energy of the full system is obtained by the DSCF method. We have presented a nonvariational version of the Fock operator to illustrate the main ideas as simply as possible, which is sufficiently accurate for use in statistical mechanical Monte Carlo simulations without the need for computing analytical gradients,[2−4] and this approach has been adopted in other applications.[34−36] A variational version that can be used to calculate analytic gradients, suitable for geometry optimization and molecular dynamics simulations,[5] is presented elsewhere.[7,8]

We note that the electrostatic potential $\mathbf{V}_b^a$, which is derived from the wave function of block b, can be calculated in more than one way.[2] For example, it can be determined by integration over the BLMOs in that block, it can be represented by fitted atomic densities, or it can be approximated by electrostatic-potential-fitted charges or by a multipole expansion.[2,6,37] These choices may be dictated by the context in which X-Pol is used. In previous applications, we have used charges obtained by Mulliken population analysis to approximate the Coulomb potentials of each of the monomers.[1,2,5,6,8] In the present work, we use the explicit evaluation of all two-electron Coulomb integrals to account for interblock electrostatic interactions. Therefore, the calculation of Coulomb interactions between different monomer blocks does not introduce approximations here, and this makes it easier to evaluate the new theoretical approach to exchange repulsion.

The short-range exchange repulsion energies between different blocks can be approximated empirically by Lennard-Jones terms, $\varepsilon_{ab}^{\mathrm{LJ}}$, which also account for long-range dispersion interactions.[2,3] Thus, the total X-Pol energy of the system is

$$E^{\text{X-Pol}} = E^{\mathrm{SCF}} + \frac{1}{2}\sum_a \sum_{b\neq a} \varepsilon_{ab}^{\mathrm{LJ}} \tag{6}$$

2.2. X-Pol-X: Block-Localized Wave Function That Accounts for Exchange Interactions. The use of the R^{-12} repulsive terms in the Lennard-Jones potential (where R is a distance between atoms in different monomers) to approximate the short-range exchange repulsion significantly reduces the computational cost needed to solve the Hartree−Fock−Roothaan equations for the entire block-localized (i.e., fragmental) system. Alternatively, the exchange energies can be determined without the use of an empirical potential by antisymmetrizing the X-Pol wave function, keeping the block-localized structure of the molecular orbitals:[13−16,27]

$$\Psi^{\text{X-Pol-X}} = \hat{A}\prod_{a=1}^{M}\Phi_a \tag{7}$$

The coefficient matrix of the corresponding orbitals has the following block-diagonal form:

$$\mathbf{C} = \begin{pmatrix} \mathbf{C^1} & 0 & ... & 0 \\ 0 & \mathbf{C^2} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \mathbf{C^k} \end{pmatrix} \tag{8}$$

where $\mathbf{C}^a$ is a $k_a \times k_a$ matrix, with the orbital coefficients of eq 2 arranged as column vectors. As mentioned in the Introduction, this is equivalent to constructing a block-localized wave function (BLW) by use of the block-localized molecular orbitals (BLMO),[13−15] and the computational procedure has been described by several authors in various contexts,[9,13−16,27,32,33,38] dating back to the work of Stoll et al.[11]

An alternative to the expression of eq 7 is to use the antisymmetrized function of antisymmetrized block-localized wave functions of eq 3. However, in the present case, the molecular orbitals are strictly block-localized within each fragment space, and the difference from eq 7 is likely to be minimal (antisymmetrizing determinant states is useful if the block-localized fragments are used to form multiconfigurational states).[39] In addition, the present test results show that the exchange repulsion energy is adequately treated, suggesting that eq 7 in the current form is a good choice for the purpose of this study.

The molecular orbitals in eq 7 are subject to the constraint that orbitals within the same subgroup are orthogonal, whereas orbitals in different subgroups are nonorthogonal:

$$\langle \varphi_i^a | \varphi_j^b \rangle = \begin{cases} \delta_{ij}^{aa}, a = b \\ S_{ij}^{ab}, a \neq b \end{cases} \tag{9}$$

where S_{ij}^{ab} is the overlap integral between molecular orbitals i and j from blocks a and b, respectively, and is given by

$$S_{ij}^{ab} = (\mathbf{c}_i^a)^{\mathrm{T}} \mathbf{R}^{ab} \mathbf{c}_j^b \tag{10}$$

where $\{\mathbf{c}_i^a; i = 1, ..., n_a; a = 1, ..., k_a\}$ are the columns of $\mathbf{C}^a$ in eq 8 (the components of these vectors are $\{c_{i\mu}^a; \mu = 1, ..., k_a\}$), where T denotes a transpose, and where $\mathbf{R}^{ab}$ is the block of the overlap matrix $\mathbf{R}$ corresponding to basis functions in fragments a and b:

$$R_{\mu\nu}^{ab} = \langle \chi_\mu^a | \chi_\nu^b \rangle \tag{11}$$

The orthogonality constraint on orbitals in the same block does not affect the total energy of the system because the energy is invariant to a unitary transformation of the orbitals within a block.

The density matrix $\mathbf{D}$ is defined in terms of a matrix $\tilde{\mathbf{C}}$ formed from the matrix $\mathbf{C}$ of eq 8 by keeping only the columns corresponding to occupied orbitals. Thus $\mathbf{C}$ is $K \times K$, but $\tilde{\mathbf{C}}$ is $K \times N$, and

$$\mathbf{D} = \tilde{\mathbf{C}}(\tilde{\mathbf{C}}^{\mathrm{T}}\mathbf{R}\tilde{\mathbf{C}})^{-1}\tilde{\mathbf{C}}^{\mathrm{T}} \tag{12}$$

The Fock matrix, $\mathbf{F}$, has an identical form to that in Hartree−Fock theory, in which an element, in the atomic basis$\{\chi_\mu\}$, is given for a closed-shell system as[9,16,27]

$$F_{\mu\nu} = H_{\mu\nu} + \sum_{\lambda\sigma}^{K} D_{\lambda\sigma}\left\{(\chi_\mu\chi_\nu|\chi_\lambda\chi_\sigma) - \frac{1}{2}(\chi_\mu\chi_\lambda|\chi_\nu\chi_\sigma)\right\} \tag{13}$$

where $H_{\mu\nu}$ is an element of the one-electron matrix representing kinetic energy and nuclear attraction, and the Coulomb and exchange integrals are shown in the Mulliken notation. The total X-Pol-X energy is

$$E^{\text{X-Pol-X}} = \sum_{\mu\nu}^{K} D_{\mu\nu}\{H_{\mu\nu} + F_{\mu\nu}\} \tag{14}$$

In comparison with the X-Pol SCF energy E^{SCF} (eq 6), the exchange energy included in the X-Pol-X expression (eq 14) originates from the orthogonalization constraint $(\mathbf{C}^T\mathbf{RC})^{-1}$ in eq 12.[15,40]

Since the orbital coefficient matrix is block-diagonal (eq 8), the $K \times K$ Fock matrix can be reduced to smaller sizes, one for each block ($k_a \times k_a$) by transforming the original Fock matrix into a block-diagonal form using a projection operator.[9,11] Thus, for block a

$$\mathscr{T}^{aa} = \mathbf{F}^{aa} + \mathbf{V}_2 + \mathbf{V}_3 + \mathbf{V}_4 \tag{15}$$

where

$$\mathbf{V}_2 = -\sum_{b\neq a}^{M} \mathbf{F}^{ab}\left(\sum_{c\neq a}^{M} \mathbf{D}^{bc}\mathbf{R}^{ca}\right) \tag{16}$$

$$\mathbf{V}_3 = -\sum_{b\neq a}^{M} \left(\sum_{c\neq a}^{M} \mathbf{D}^{bc}\mathbf{R}^{ca}\right)^T \mathbf{F}^{ba} \tag{17}$$

and

$$\mathbf{V}_4 = \sum_{d\neq a}^{M}\sum_{b\neq a}^{M}\left(\sum_{c\neq a}^{M} \mathbf{D}^{bc}\mathbf{R}^{ca}\right)^T \mathbf{F}^{bd}\left(\sum_{c\neq a}^{M} \mathbf{D}^{dc}\mathbf{R}^{ca}\right) \tag{18}$$

In the above equations the dimension of a matrix $\mathbf{X}^{ab}$ is $k_a \times k_b$ (see eq 1). Note that eq 15 is general in that it can be used in both Hartree−Fock[13−16,27] and DFT calculations.[9] In the latter case, the exchange potential in Hartree−Fock theory is replaced by the exchange−correlation potential.[9]

3. Computational Details

All computations have been performed by use of a locally modified GAMESS program[41] and the Xiamen University Valence Bond (XMVB) program.[42] The minimally augmented polarized valence double-ζ basis set 6-31+G(d) was used for all calculations. All X-Pol-X calculations are fully self-consistent; that is, the orbitals are optimized in the presence of interfragment exchange.

We present illustrative calculations for two examples, both of which are structures of water trimer. First, we chose the minimum energy configuration of a cyclic water trimer structure, **c-W$_3$**, in which each water molecule donates and accepts one hydrogen bond from one of the other two water monomers.[43] All three pairwise interactions are stabilizing. For the second structure, we constructed a trimer complex structure that has a significantly Coulomb-repulsive interaction as follows. We first optimized the dimer complex structure, and then we placed the third water molecule at the C_2 image position of the donor water molecule about the bisection axis of the acceptor water molecule. This complex is denoted as the symmetric trimer, **s-W$_3$**. Both structures are depicted in Figure 1.

In the following discussion, each water monomer in the trimer complexes is treated as a single block or fragment, and the internal geometries of the monomers are kept the same as in the fully optimized configuration at the HF/6-31+G(d) level for **c-W$_3$** and the same as constructed above for **s-W$_3$**. The X-Pol-X wave function in the three-block

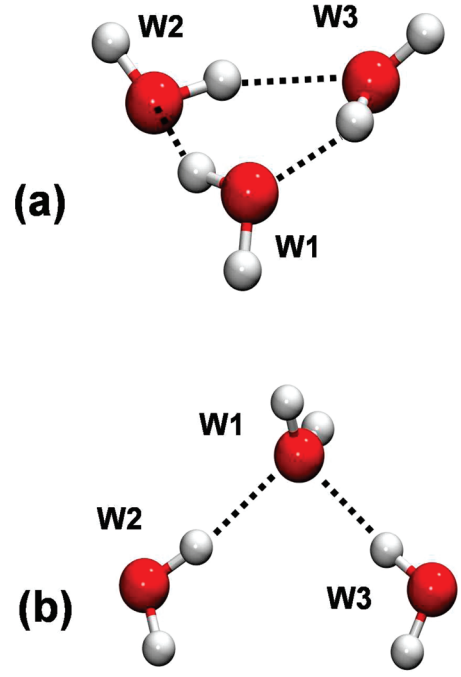


Figure 1. Illustration of (a) the minimum water trimer structure (**c-W$_3$**) and (b) a symmetric configuration (**s-W$_3$**).

system is $\hat{A}\psi_a\psi_b\psi_c$. The energy and energy components of the X-Pol-X wave function will be compared to those for the unrelaxed wave function $\hat{A}\psi_a^0\psi_b^0\psi_c^0$, where ψ_w^0 is the wave function of a monomer in the gas phase (unpolarized wave function) at the geometry in the trimer complex.

The energy components we consider are Coulomb energy, exchange energy, polarization energy, and charge transfer energy. The Coulomb and exchange components are the same in all energy decomposition analysis (EDA) methods, including the approach of Kitaura and Morokuma,[44] whereas the polarization and charge transfer terms are defined in the block-localized wave function energy decomposition (BLW-ED) method developed by Mo et al.[15]

4. Results and Discussion

The main goal of the present paper is to show that the theory presented in section 2 can provide a reasonable estimate of the exchange repulsion energy in the X-Pol method.

Table S1 (see Supporting Information) lists the total energies and energy components determined for the fully separated monomers and for the trimer complexes as calculated from $\hat{A}\psi_a\psi_b\psi_c$ and from $\hat{A}\psi_a^0\psi_b^0\psi_c^0$ and also by full Hartree−Fock calculations. In this discussion we shall focus on interaction energies, defined as the energy of the trimer minus the sum of the energies of the three separated monomers, each with the same geometry that it has in the trimer. We shall also discuss the components of the interactions. The interaction energies and their components are given in Table 1; all values in Table 1 were computed from the absolute energies in Table S1 (see Supporting Information).

Row 1 of Table 1 gives the X-Pol-X interaction energy, and the following rows of the table report quantities that help us dissect this interaction energy. Row 2 is calculated from $\psi_a^0\psi_b^0\psi_c^0$, and row 3 is calculated from $\hat{A}\psi_a^0\psi_b^0\psi_c^0$. Row 2 gives the unpolarized Coulomb contribution to the X-Pol

Table 1. Computed Relative Energies for the Cyclic Water Trimer Minimum Structure and for a Symmetric Trimer Geometry[a]

		computed relative energy (kcal/mol)	
row	quantity[a]	**c-W$_3$**	**s-W$_3$**
1	$\Delta E^{\text{X-Pol-X}}$	−12.5	−6.7
2	Coulomb energy with unpolarized monomer orbitals	−25.6	−16.0
3	exchange energy with monomer orbitals	16.3	10.8
4	total energy for monomer orbitals	−9.3	−5.2
5	X-Pol-X distortion energy	3.6	1.6
6	X-Pol-X Coulomb energy	−31.8	−18.3
7	X-Pol-X exchange eneergy	15.7	10.1
8	X-Pol-X polarization energy	−3.2	−1.5
9	charge transfer energy	−3.0	−2.1
10	full Hartree−Fock	−15.5	−8.8

[a] All energies are determined by use of the 6-31+G(d) basis set and are relative to the sum of the unpolarized monomer energies.

Table 2. Additivity of Dimer Contributions to the Coulomb and Exchange Repulsion Interaction Energies in the Cyclic Water Trimer Minimum Structure and in a Symmetric Trimer Geometry[a]

	c-W$_3$		**s-W$_3$**	
	Coulomb (kcal/mol)	exchange (kcal/mol)	Coulomb (kcal/mol)	exchange (kcal/mol)
dimer *12*	−10.92	5.43	−10.15	5.02
dimer *23*	−9.86	4.75	1.97	0.02
dimer *13*	−11.02	5.56	−10.15	5.02
sum of dimers	−31.80	15.74	−18.34	10.07
X-Pol-X	−31.80	15.71	−18.34	10.09
nonadditivity	0.00	−0.03	0.00	0.02

[a] All energies are determined by use of the 6-31+G(d) basis set.

or X-Pol-X binding energy. The results in Table 2 show that the Coulomb interaction contribution to the total binding energy is quite large in both water trimer complexes. In particular, with the monomers unpolarized, the total Coulomb interaction energies are −25.6 and −16.0 kcal/mol for **c-W$_3$** and **s-W$_3$** complexes, whereas row 3 shows that the exchange repulsion energies are 16.3 and 10.8 kcal/mol, respectively. The sum of Coulomb and exchange energies is often called the total electrostatic energy; in row 4 it is given as computed from the charge densities of separated, unpolarized monomers in the gas phase, and it amounts to −9.3 and −5.2 kcal/mol, respectively, for **c-W$_3$** and **s-W$_3$**. The water trimer has been extensively studied, but all but one of these studies are not discussed here since our goal is not to investigate the water clusters. In one energy decomposition calculation using the reduced variational space (RVS) method,[45] which is analogous to the BLW-ED process, Chen and Gordon[46] found that the total electrostatic interaction energy is −9.0 kcal/mol using the 6-31++G(d,p) basis set at the minimum-energy configuration.

Next we consider the analogous breakdown of the X-Pol-X interaction energies and interaction energy components. For these calculations, the wave function is $\hat{A}\psi_a\psi_b\psi_c$ where the antisymmetrization is performed before the orbitals are optimized, so that they are optimized in the presence of interfragment exchange. Now, as compared to rows 2−4 in Table 1, there is an additional component of the total interaction energy, namely, the increase in internal energy of the monomers due to polarization of the orbitals. This is given in row 5, the Coulomb energy is in row 6, and the exchange energy is in row 7. The sum of rows 5−7 gives the X-Pol-X interaction energy in row 1. Note that row 4 is the total energy of the antisymmetrized fragment product wave function where the orbitals are the unpolarized monomer orbitals, whereas row 1 is the total energy for the X-Pol-X wave function, in which the mutual charge and exchange polarizations of the monomer charge density are included.

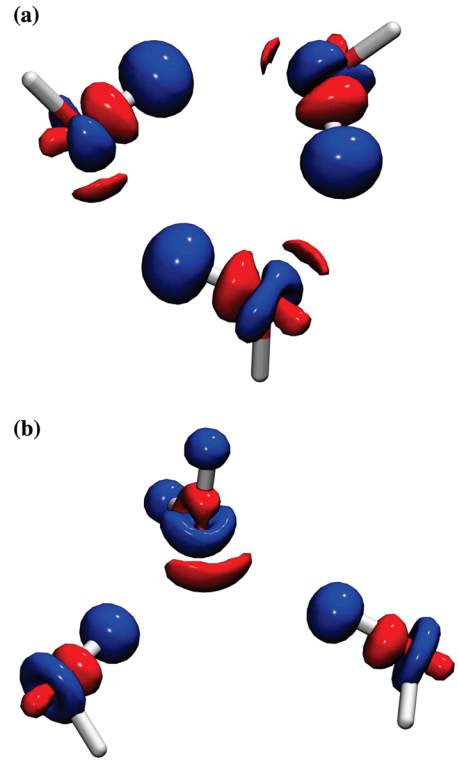


Figure 2. Electron density difference contour between the polarized and unpolarized systems, $\Delta\rho = \rho(\hat{A}\psi_a\psi_b\psi_c) - \rho^0(\psi_a^0\psi_b^0\psi_c^0)$, for (a) the minimum-energy structure (**c-W$_3$**) and (b) the symmetric configurations (**s-W$_3$**). The structure orientations are identical to those shown in Figure 1 and the contour levels are at 0.003 au, with red contours representing gain in electron density and blue for charge depletion.

Mutual polarization of the block-localized monomer wave functions in the presence of other monomers enhances Coulomb interactions by −6.2 and −2.3 kcal/mol (difference between rows 6 and 2) and the exchange repulsion slightly. Taking into account of the energy costs for polarizing the monomer wave functions in the trimer complex and the small decrease in exchange repulsion due to polarization effects, we see that the overall polarization, calculated at this level, lowers the **c-W$_3$** interaction energy from −9.3 to −12.5 kcal/mol and the **s-W$_3$** interaction energy from −5.2 to −6.7 kcal/mol. The differences between these quantities are listed in row 8, which shows a net X-Pol polarization energy of −3.2 and −1.5 kcal/mol, or 26% and 22%, for the two complexes, respectively. The electronic polarization effects in these two water trimer complexes are illustrated in Figure 2, which depicts electron density shifting away for the hydrogen atoms that donate hydrogen bonds to the acceptor oxygen atoms.

Concomitantly, electron density is attracted toward regions where oxygen atoms bind the proton donor. It is of interest to notice that, in the **s-W$_3$** structure (Figure 2b), electron density is reduced from both hydrogen atoms in the W1 water molecule due to block localization that disallows charge transfer between different water molecules, although they are not directly donating any hydrogen bonds. In the study of Chen and Gordon,[46] polarization effects were found to contribute −2.7 kcal/mol with the slightly larger 6-31++G(d,p) basis set for **c-W$_3$**.

Row 10 in Table 1 lists the interaction energy from a full (delocalized) Hartree−Fock calculation. The energy difference between the Hartree−Fock SCF result and the X-Pol-X energy for each trimer complex is defined as the charge transfer energy, which is not included in the X-Pol-X method and will be addressed in a separate study. This charge transfer energy (row 9) is −3.0 and −2.1 kcal/mol for the 6-31+G(d) basis functions in the cyclic and symmetric complexes, respectively. For comparison, Chen and Gordon[46] found a charge transfer energy of −1.7 kcal/mol using the RVS approach, smaller than the present value. The BLW energy decomposition analysis (BLW-ED) method has been shown to be relatively stable in computed polarization and charge transfer energies with increased size of basis set.[15,18,47]

It is interesting to further analyze the exchange repulsion energy of the X-Pol-X method. The numerical results in Table 1 show that the exchange repulsion is significant, reducing the Coulomb interaction energies by more than 50% in both cases. Table 2 shows the pairwise decomposition of the total Coulomb (row 6 in Table 1) and exchange (row 7 in Table 1) energies. The Coulomb and exchange energies for each dimer were obtained from a calculation with the optimized monomer wave functions in the trimer complex described by the X-Pol-X method ($\hat{A}\psi_1\psi_2\psi_3$). For example, the first row is determined using the Hartree product ($\psi_1\psi_2$) and antisymmetrized ($\hat{A}\psi_1\psi_2$) functions, respectively. The variation of Coulomb energies for the dimer interactions depicted in the first three rows in Table 2 suggests that stronger Coulomb attractions are accompanied by greater charge overlap and exchange repulsions. The Coulomb interaction energy is additive by definition in the energy decomposition analysis,[15,40,44] but the exchange repulsion term is not necessarily additive due to orthogonalization of the product of nonorthogonal block-localized orbitals. Nevertheless, Table 2 reveals that the exchange repulsion energies are additive to within 0.03 kcal/mol in both trimer complexes. This result holds even though one complex (**c-W$_3$**) has three significant pairwise exchange repulsion interactions, while the other (**s-W$_3$**) has only two. The finding of high additivity for exchange repulsion interactions suggests that it is sufficient to construct an X-Pol-X wave function that involves only pairwise antisymmetrized blocks, which can greatly reduce the computational cost of the fully antisymmetrized system.

The dependence of exchange repulsion on intermolecular separation is depicted in Figure 3 for a water dimer as a function of the hydrogen-bond distance between the donor hydrogen and acceptor oxygen atoms. As is well-known, the exchange repulsion decreases exponentially to a diminishing value as the hydrogen-bond distance is elongated by more than 1 Å from the optimal geometry, well within the first solvation layer in liquid water. Figure 3 shows that, for all practical purposes, one needs to consider only the explicit exchange repulsion for monomers within the first solvation shell, and there is no need to construct a fully antisymmetric X-Pol determinant wave function for the entire system.[47] Further simplifications of the X-Pol-X method can be made by imposing the NIDO approximation presented in the Appendix.

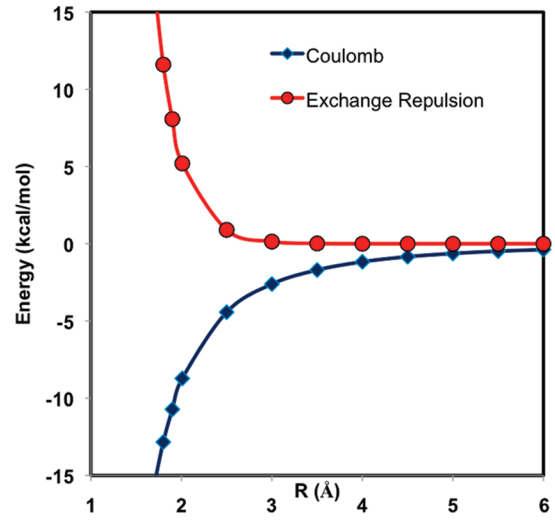


Figure 3. Computed Coulomb (X-Pol) and exchange repulsion (X-Pol-X) interaction energies as a function of the hydrogen-bond distance between donor hydrogen and acceptor oxygen atoms in the water dimer complex with C_s symmetry. Energies are computed by use of the 6-31+G(d) basis at the optimized geometries under the C_s symmetry constraint at the HF/6-31+G(d) level.

Several topics suggest themselves for further work. First, it would be useful to examine more systems. Second, the method can be used to account for charge transfer effects.[9,16,23,27,47] Third, the method may be examined by treating the monomers with density functional theory.[6,9] Fourth, it would be interesting to examine the effect of computing the interfragment exchange terms of X-Pol-X in a post-SCF fashion, that is, optimizing the orbitals by X-Pol using a constant interfragment exchange based on the unpolarized fragmental wave functions and then using the optimized BLMOs to compute the exchange repulsion. Fifth, the NIDO approximation presented in the Appendix can be used either with full X-Pol-X calculations, including the use of full NDDO approximation in effective-Hamiltonian MOVB method,[25] or with post-SCF interfragment exchange. Finally, it would be interesting to use the interfragment exchange terms of X-Pol-X calculations as a basis for parametrizing more accurate analytical pairwise exchange repulsion terms.

5. Conclusions

In this article, we have examined the inclusion of exchange repulsion energies directly into the explicit polarization (X-Pol) model by antisymmetrizing the Hartree product wave function in a way that is equivalent to the block-localized wave function (BLW) approach. In general, this may be called X-Pol with full exchange or X-Pol-X. The block-localized structure in the X-Pol-X wave function allows for

decomposition of the full Fock matrix of M blocks into M smaller Fock matrices, reducing the computational costs from $Q[(KM)^3]$ scaling to $Q[M(K)^3]$. Furthermore, the complexity of the Fock matrix can be reduced by making use of the neglect of fragmental diatomic overlap approximation, a modification of the traditional neglect of diatomic differential overlap (NDDO) method. Another way to reduce the cost would be to compute the exchange repulsion by a set of pairwise X-Pol-X dimer calculations, including the contributions of the nearby pairs. The computation of the exchange repulsion energy is illustrated by considering two trimer structures of water, and it was found that the total exchange repulsion energies in these simple but general and strongly hydrogen-bonding cases are pairwise additive to an excellent approximation.

Appendix: X-Pol-X with the Neglect of Interfragment Differential Overlap Approximation

The neglect of diatomic differential overlap (NDDO)[48,49] approximation is adopted in most modern semiempirical quantum mechanical methods such as MNDO,[50] AM1,[51] PM3,[52] and RM1.[53] In this method, electronic integrals involving the charge density $\chi_\mu^A\chi_\nu^B$ are assumed to be zero when the basis functions χ_μ^A and χ_ν^B are located on different atoms, $A \neq B$, with the exception that the two-center one-electron nuclear attraction integrals are retained due to their importance for chemical bonding. Here we introduce a less severe approximation called the neglect of interfragment differential overlap (NIDO). In this approximation, we retain all electronic integrals within the same monomer, but electronic integrals involving differential overlap between basis functions belonging to different fragments are neglected. Applying the NIDO approximation greatly simplifies eq 15, as explained in the rest of this appendix. Below, we present the explicit expression for each term in eqs 16−18 by applying the NIDO approximation.

First, the block diagonal terms of the unprojected matrix are reduced to

$$F_{\mu\nu}^{aa} = \left[H_{\mu\nu}^{aa} + \sum_{\lambda\sigma}^{k_a} D_{\lambda\sigma}^{aa}\left\{(\chi_\mu^a\chi_\nu^a|\chi_\lambda^a\chi_\sigma^a) - \frac{1}{2}(\chi_\mu^a\chi_\lambda^a|\chi_\nu^a\chi_\sigma^a)\right\}\right] + \left[\sum_{b\neq a}\sum_{\lambda\sigma}^{k_b} D_{\lambda\sigma}^{bb}(\chi_\mu^a\chi_\nu^a|\chi_\lambda^b\chi_\sigma^b)\right] \quad \text{(A1)}$$

Note that the "exchange" terms from orbitals on other fragments, $D_{\lambda\sigma}^{ab}(\chi_\mu^a\chi_\lambda^b|\chi_\nu^a\chi_\sigma^c)$, are all zero because of the NIDO approximation. Thus, the unprojected Fock matrix element for orbitals on fragment a can be written as

$$F_{\mu\nu}^{aa} = (F_{\mu\nu}^{aa})^0 + \left\langle \chi_\mu^a \middle| \sum_{b\neq a} \mathbf{V}_b^a \middle| \chi_\nu^a \right\rangle \quad \text{(A2)}$$

Equation A2 includes contributions only from the Coulomb integrals between different fragments, which is similar to a combined quantum mechanical and molecular mechanical (QM/MM) method. In the original X-Pol potential,[2,3] in which the Hartree product wave function is used, the rest of the density projection terms in eq 15 are not needed.

The $\mathbf{V}_2$ term of eq 16 simplifies under the NIDO approximation to

$$\begin{aligned}(\mathbf{V}_2)_{\mu\nu}^{aa} &= \sum_{b\neq a}^{M}\sum_{\beta}^{k_b}\left[H_{\mu\beta}^{ab} + \sum_{c,d}^{M}\sum_{\lambda\sigma}^{k_c,k_d} D_{\lambda\sigma}^{cd}\{(\chi_\mu^a\chi_\beta^b|\chi_\lambda^c\chi_\sigma^d) - \frac{1}{2}(\chi_\mu^a\chi_\lambda^c|\chi_\beta^b\chi_\sigma^d)\}\right]P_{\beta\nu}^{ba} \\ &= -\frac{1}{2}\sum_{b\neq a}^{M}\sum_{\beta}^{k_b}\left[\sum_{\lambda}^{k_a}\sum_{\sigma}^{k_b} D_{\lambda\sigma}^{ab}(\chi_\mu^a\chi_\lambda^a|\chi_\beta^b\chi_\sigma^b)\right]P_{\beta\nu}^{ba}\end{aligned} \quad \text{(A3)}$$

in which only the exchange terms having $a = c$ and $b = d$ remain nonzero.

The $\mathbf{V}_3$ term becomes

$$\begin{aligned}(\mathbf{V}_3)_{\mu\nu}^{aa} &= \sum_{b\neq a}^{M}\sum_{\beta}^{k_b}(P_{\beta\mu}^{ba})^{\mathrm{T}}\left[H_{\beta\nu}^{ba} + \sum_{c,d}^{M}\sum_{\lambda\sigma}^{k_c,k_d} D_{\lambda\sigma}^{cd}\left\{(\chi_\beta^b\chi_\nu^a|\chi_\lambda^c\chi_\sigma^d) - \frac{1}{2}(\chi_\beta^b\chi_\lambda^c|\chi_\nu^a\chi_\sigma^d)\right\}\right] \\ &= \sum_{b\neq a}^{M}\left[-\frac{1}{2}\sum_{\beta}^{k_b}(P_{\beta\mu}^{ba})^{\mathrm{T}}\sum_{\lambda}^{k_b}\sum_{\sigma}^{k_a} D_{\lambda\sigma}^{ba}(\chi_\beta^b\chi_\lambda^b|\chi_\nu^a\chi_\sigma^a)\right]\end{aligned} \quad \text{(A4)}$$

Similarly, the nonzero terms are in the exchange part for orbitals that are on fragments such that $b = c$ and $a = d$ in the double summation over fragments in parentheses.

Finally, the elements of the last term of of eq 15 become

$$\begin{aligned}(\mathbf{V}_4)_{\mu\nu}^{aa} &= \sum_{d\neq a}^{M}\sum_{\delta}^{k_d}\left\{\sum_{b\neq a}^{M}\sum_{\beta}^{k_b}(P_{\beta\mu}^{ba})^{\mathrm{T}}\left[H_{\beta\delta}^{bd} + \sum_{c,f}^{M}\sum_{\lambda\sigma}^{k_c,k_f} D_{\lambda\sigma}^{cf}\left\{(\chi_\beta^b\chi_\delta^d|\chi_\lambda^c\chi_\sigma^f) - \frac{1}{2}(\chi_\beta^b\chi_\lambda^c|\chi_\delta^d\chi_\sigma^f)\right\}\right]\right\}P_{\delta\nu}^{da} \\ &= \sum_{b\neq a}^{M}\sum_{\beta}^{k_b}\sum_{\delta}^{k_b}(P_{\beta\mu}^{ba})^{\mathrm{T}}\left[H_{\beta\delta}^{bb} + \sum_{c}^{M}\sum_{\lambda\sigma}^{k_c} D_{\lambda\sigma}^{cc}(\chi_\beta^b\chi_\delta^b|\chi_\lambda^c\chi_\sigma^c)\right]P_{\delta\nu}^{ba} \\ &\quad - \frac{1}{2}\sum_{d\neq a}^{M}\sum_{\delta}^{k_d}\sum_{b\neq a}^{M}\sum_{\beta}^{k_b}(P_{\beta\mu}^{ba})^{\mathrm{T}}\left[\sum_{\lambda}^{k_b}\sum_{\sigma}^{k_d} D_{\lambda\sigma}^{bd}(\chi_\beta^b\chi_\lambda^b|\chi_\delta^d\chi_\sigma^d)\right]P_{\delta\nu}^{da}\end{aligned} \quad \text{(A5)}$$

We presented the NIDO approximation here for a general case of X-Pol-X; that is, the monomers can be treated by ab initio Hartree−Fock, by post-Hartree−Fock correlation methods, or by density functional theory. However, if desired, the monomers could also be treated by semiempirical molecular theory employing the NDDO approximation within the monomers. This procedure has been implemented into the program CHARMM for defining diabatic states in the effective Hamiltonian mixed molecular orbital and valence bond (EH-MOVB) method.[25]

Acknowledgment. We are grateful to Peng Zhang, Hannah R. Leverentz, and Carlos Sosa for valuable discussions. This work has been partially supported by the National Institutes of Health (Grant GM46736 for the development of effective Hamiltonian methods for enzymatic processes and Grant RC1-GM091445 for the development of a quantal force field for biomolecular simulations), by the National Science Foundation (Grant CHE09-57162 for the development of mixed molecular orbital and valence bond theory and Grant CHE09-56776 for incorporating quantum mechanics in the treatment of complex reactive dynamical systems).

Supporting Information Available: One table showing absolute energies and energy components in hartrees for trimer calculations. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Xie, W.; Gao, J. *J. Chem. Theory Comput.* **2007**, *3*, 1890.

(2) Gao, J. *J. Phys. Chem. B* **1997**, *101*, 657.

(3) Gao, J. *J. Chem. Phys.* **1998**, *109*, 2346.

(4) Wierzchowski, S. J.; Kofke, D. A.; Gao, J. *J. Chem. Phys.* **2003**, *119*, 7365.

(5) Xie, W.; Orozco, M.; Truhlar, D. G.; Gao, J. *J. Chem. Theory Comput.* **2009**, *5*, 459.

(6) Song, L.; Han, J.; Lin, Y. L.; Xie, W.; Gao, J. *J. Phys. Chem. A* **2009**, *113*, 11656.

(7) Xie, W.; Song, L.; Truhlar, D. G.; Gao, J. *J. Phys. Chem. B* **2008**, *112*, 14124.

(8) Xie, W.; Song, L.; Truhlar, D. G.; Gao, J. *J. Chem. Phys.* **2008**, *128*, 234108/1.

(9) Cembran, A.; Song, L.; Mo, Y.; Gao, J. *J. Chem. Theory Comput.* **2009**, *5*, 2702.

(10) Stoll, H.; Preuss, H. *Theor. Chem. Acc.* **1977**, *46*, 12.

(11) Stoll, H.; Wagenblast, G.; Preuss, H. *Theor. Chim. Acta* **1980**, *57*, 169.

(12) Yang, W. *Phys. Rev. Lett.* **1991**, *66*, 1438.

(13) Mo, Y.; Peyerimhoff, S. D. *J. Chem. Phys.* **1998**, *109*, 1687.

(14) Mo, Y.; Zhang, Y.; Gao, J. *J. Am. Chem. Soc.* **1999**, *121*, 5737.

(15) Mo, Y.; Gao, J.; Peyerimhoff, S. D. *J. Chem. Phys.* **2000**, *112*, 5530.

(16) Mo, Y.; Gao, J. *J. Comput. Chem.* **2000**, *21*, 1458.

(17) Gao, J.; Mo, Y. *Prog. Theor. Chem. Phys.* **2000**, *5*, 247.

(18) Mo, Y.; Gao, J. *J. Phys. Chem. A* **2001**, *105*, 6530.

(19) Gao, J.; Garcia-Viloca, M.; Poulsen, T. D.; Mo, Y. *Adv. Phys. Org. Chem.* **2003**, *38*, 161.

(20) Brauer, C. S.; Craddock, M. B.; Kilian, J.; Grumstrup, E. M.; Orilall, M. C.; Mo, Y.; Gao, J.; Leopold, K. R. *J. Phys. Chem. A* **2006**, *110*, 10025.

(21) Mo, Y.; Gao, J. *Acc. Chem. Res.* **2007**, *40*, 113.

(22) Mo, Y. *J. Chem. Phys.* **2007**, *126*, 224104.

(23) Song, L.; Gao, J. *J. Phys. Chem. A* **2008**, *112*, 12925.

(24) Valero, R.; Song, L.; Gao, J.; Truhlar, D. G. *J. Chem. Theory Comput.* **2009**, *5*, 1.

(25) Cembran, A.; Payaka, A.; Lin, Y. L.; Xie, W.; Song, L.; Mo, Y.; Gao, J. *J. Chem. Theory Comput.* [Online early access]. DOI: 10.1021/ct1001686. Published Online: June 3, 2010.

(26) Hunt, W. J.; Hay, P. J.; Goddard, W. A. I. *J. Chem. Phys.* **1972**, *57*, 749.

(27) Mo, Y.; Gao, J. *J. Phys. Chem. A* **2000**, *104*, 3012.

(28) King, H. F.; Staton, R. E.; Kim, H.; Wyatt, R. E.; Parr, R. G. *J. Chem. Phys.* **1967**, *47*, 1936.

(29) Gallup, G. A. *Int. J. Quantum Chem.* **1972**, *6*, 899.

(30) Shukla, A.; Dolg, M.; Stoll, H.; Fulde, P. *Chem. Phys. Lett.* **1996**, *262*, 213.

(31) Raimondi, M.; Gianinetti, E. *J. Phys. Chem.* **1988**, *92*, 899.

(32) Gianinetti, E.; Raimondi, M.; Tornaghi, E. *Int. J. Quantum Chem.* **1996**, *60*, 157.

(33) Khaliullin, R. Z.; Head-Gordon, M.; Bell, A. T. *J. Chem. Phys.* **2006**, *124*, 204105/1.

(34) Kitaura, K.; Ikeo, E.; Asada, T.; Nakano, T.; Uebayasi, M. *Chem. Phys. Lett.* **1999**, *313*, 701.

(35) Kitaura, K.; Sawai, T.; Asada, T.; Nakano, T.; Uebayasi, M. *Chem. Phys. Lett.* **1999**, *312*, 319.

(36) Fedorov, D. G.; Kitaura, K. *J. Phys. Chem. A* **2007**, *111*, 6904.

(37) Leverentz, H. R.; Gao, J.; Truhlar D. G. *Theor. Chem. Acc.* **2010** (submitted for publication).

(38) Nagata, T.; Takahashi, O.; Saito, K.; Iwata, S. *J. Chem. Phys.* **2001**, *115*, 3553.

(39) McWeeny, R. *Proc. R. Soc. London, Ser. A* **1959**, *253*, 242.

(40) Su, P. F.; Li, H. *J. Chem. Phys.* **2009**, *131*, 014102.

(41) 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. S. *J. Comput. Chem.* **1993**, *14*, 1347.

(42) Song, L.; Mo, Y.; Zhang, Q.; Wu, W. *J. Comput. Chem.* **2005**, *26*, 514.

(43) Schutz, M.; Burgi, T.; Leutwyler, S.; Burgi, H. B. *J. Chem. Phys.* **1993**, *99*, 5228.

(44) Kitaura, K.; Morokuma, K. *Int. J. Quantum Chem.* **1976**, *10*, 325.

(45) Stevens, W. J.; Fink, W. H. *Chem. Phys. Lett.* **1987**, *139*, 15.

(46) Chen, W.; Gordon, M. S. *J. Phys. Chem.* **1996**, *100*, 14316.

(47) Mo, Y.; Gao, J. *J. Phys. Chem. B* **2006**, *110*, 2976.

(48) Pople, J. A.; Santry, D. P.; Segal, G. A. *J. Chem. Phys.* **1965**, *43*, S129.

(49) Pople, J. A.; Segal, G. A. *J. Chem. Phys.* **1965**, *43*, S136.

(50) Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4899.

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

(52) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209.

(53) Rocha, G. B.; Freire, R. O.; Simas, A. M.; Stewart, J. J. P. *J. Comput. Chem.* **2006**, *27*, 1101.

CT100268P

JCTC Journal of Chemical Theory and Computation

Protein Folding Simulations Combining Self-Guided Langevin Dynamics and Temperature-Based Replica Exchange

Michael S. Lee*,†,‡,§ and Mark A. Olson§

Computational Sciences and Engineering Branch, U.S. Army Research Laboratory, Aberdeen Proving Ground, Maryland 21005, Biotechnology High Performance Computing Software Applications Institute, U.S. Army Medical Research and Materiel Command, Frederick, Maryland 21702, and Department of Cell Biology and Biochemistry, U.S. Army Medical Research Institute of Infectious Diseases, Frederick, Maryland 21702

Received February 2, 2010

Abstract: Computer simulations are increasingly being used to predict thermodynamic observables for folding small proteins. Key to continued progress in this area is the development of algorithms that accelerate conformational sampling. Temperature-based replica exchange (ReX) is a commonly used protocol whereby simulations at several temperatures are simultaneously performed and temperatures are exchanged between simulations via a Metropolis criterion. Another method, self-guided Langevin dynamics (SGLD), expedites conformational sampling by accelerating low-frequency, large-scale motions through the addition of an ad hoc momentum memory term. In this work, we combined these two complementary techniques and compared the results against conventional ReX formulations of molecular dynamics (MD) and Langevin dynamics (LD) simulations for the prediction of thermodynamic folding observables of the Trp-cage mini-protein. All simulations were performed with CHARMM using the PARAM22+CMAP force field and the generalized Born molecular volume implicit solvent model. While SGLD-ReX does not fold up the protein significantly faster than the two conventional ReX approaches, there is some evidence that the method improves sampling convergence by reducing topological folding barriers between energetically similar near-native states. Unlike MD-ReX and LD-ReX, SGLD-ReX predicts melting temperatures, heat capacity curves, and folding free energies that are closer in agreement to the experimental observations. However, this favorable result may be due to distortions of the relative free energies of the folded and unfolded conformational basins caused by the ad hoc force term in the SGLD model.

Introduction

Molecular dynamics (MD) simulations of small proteins provide insight into the mechanisms and thermodynamics of protein folding. The most traditional protocol is to simulate at a fixed temperature the folding/unfolding of a protein immersed in an explicit solvent. So far, only the smallest proteins have been folded in this way.[1] Sampling at the all-atom level is slow due to the presence of multiple small minima on the energy landscape, namely, kinetic traps, which lead to a "mountainous" and/or "pebbly" energy surface.[2] To combat this problem, several enhanced sampling methods have been developed and tested over the past decade, including replica exchange,[3] accelerated molecular dynamics,[4] self-guided molecular dynamics,[5] potential smoothing,[6] locally enhanced sampling,[7] and resolution reduction (e.g., implicit solvent models and lattice models).[8,9]

Temperature is a commonly used parameter to accelerate conformational motions in proteins.[3,10] One of the most popular approaches in recent years is temperature-based

* Corresponding author e-mail: michael.scott.lee@us.army.mil.
† U.S. Army Research Laboratory.
‡ U.S. Army Medical Research and Materiel Command.
§ U.S. Army Medical Research Institute of Infectious Diseases.

10.1021/ct100062b © 2010 American Chemical Society
Published on Web 07/26/2010

replica exchange (T-ReX),[3,11] which involves simultaneously simulating multiple fixed-temperature windows. The temperature values are assigned at exponentially spaced or adaptively spaced[12] intervals between room temperature and a user-specified higher temperature. At regular time intervals, the simulation temperatures are exchanged between neighboring pairs of thermal windows according to a Metropolis criteron.[13] The method effectively percolates conformations up and down a ladder of temperatures depending on their relative energies.

Thinking beyond sampling at various temperatures, the key to characterizing the thermodynamics of folding is to sample the multiple major conformational basins that exist between the unfolded and native states. Major conformational basins are often separated by low-frequency modes of the protein. By definition, traditional dynamics methods require a relatively long time to traverse these low-frequency modes. As a potential solution to this dilemma, the recently developed self-guided Langevin dynamics (SGLD) accelerates the lowest frequency modes of a system by the addition of an ad hoc atomic force term proportional to the running average of individual atomic momenta over a short time interval (e.g., 0.1−1 ps).[5] In this force term, high-frequency motions tend to cancel out, while low-frequency motions, which are typically unidirectional over the short averaging time, tend to be additive, and thus have the effect of adding an external boost along the low-frequency degrees freedom. Excess energy generated by the SGLD force term is removed via a fixed-energy constraint term. While the addition of an ad hoc force term may cause deviations from canonical ensemble behavior, tests to-date indicate many thermodynamic observables are not significantly altered.[5,14]

In the original paper describing SGLD, an α-helix peptide, $(AAQAA)_3$, was folded up using a simple distance-dependent dielectric electrostatic function.[5] Impressively, the SGLD simulation was estimated to be at least 65 times more efficient than standard LD in reaching the apparent lowest energy state of the α-helical conformation. Using a similar method, self-guided molecular dynamics, Wen et al. reported improved sampling versus conventional MD for the folding of a nontraditional peptide, $\beta\beta\alpha 1$, and the villin headpiece using the AMBER force field, parm94, with an analytical Poisson implicit solvation model.[15] SGLD has also been applied to various biophysical problems including ionization equilibria[16,17] and water content in the interior of proteins.[18]

In this work, we combined the merits of two disparate techniques, SGLD and temperature-based ReX, and asked if their synergy could provide further sampling enhancements versus traditional MD-ReX and LD-ReX in the prediction of thermodynamic observables. Our test system is the commonly studied 20-residue Trp-cage mini-protein "5b".[19−26] Besides its small size, the Trp-cage mini-protein is an exemplary model system because it contains many key structural elements found in larger proteins. Our evaluation criteria of MD-Rex, LD-ReX, and SGLD-ReX included how quickly the native basin was reached starting from the unfolded state and a comparison of melting temperatures, heat capacity curves, and folding free energies to recent experimental observations. In addition, we compared the free-energy landscapes generated by each method to better understand the differences in the predicted observables. Finally, we examined several criteria to determine to what extent the three ReX methods deviated from theoretical canonical ensemble behavior.

Methods

Self-Guided Langevin Dynamics. The self-guided Langevin dynamics method, developed by Wu and Brooks,[5] enhances conformational sampling by accelerating low-frequency modes through the use of an ad hoc time-averaged momentum term. The algorithm was preceded by the self-guided molecular dynamics method which, instead, used a time-averaged force term.[27] The running average of the momentum over a short-period simulation time is added back as an external force to the simulation system. This term has the effect of accelerating low-frequency motions, because the modes that are slower than the averaging time are expected to be additive. In principle, compared to MD and Langevin dynamics, SGLD should increase the rate of hopping between conformational basins which might include the lowest energy topology and/or the experimentally observed native conformation. The main drawbacks to this method, however, are that SGLD no longer provides rigorous canonical ensemble sampling and that kinetics predictions are no longer comparable to true observables.[5] Furthermore, it is not yet clear how one would recover rigorous ensemble averages by reweighting the population densities[28] resulting from a SGLD simulation.

As outlined and derived in the original work, the SGLD method uses the following equation of motion:

$$\dot{\mathbf{p}}_i = \mathbf{f}_i - \gamma_i \mathbf{p}_i + \mathbf{R}_i + \lambda \mathbf{g}_i \tag{1}$$

where the rate of change of the momentum of a particle i, $\dot{\mathbf{p}}_i$, is a function of the force on the particle, $\mathbf{f}_i$, a friction constant, γ_i, the momentum itself,$\mathbf{p}_i$, a random force, $\mathbf{R}_i$, and a memory function, $\mathbf{g}_i$, which is scaled by a guiding factor, λ. The memory function, $\mathbf{g}_i$, is defined by the moving average of the momentum of the system over an interval of time, L:

$$\mathbf{g}_i = \gamma_i \langle \mathbf{p}_i \rangle_L \tag{2}$$

Inevitably, the addition of a memory term $\mathbf{g}_i$ to the atomic forces will result in a lack of energy conservation, typically heating up the system. Therefore, an energy conservation term is applied to the equation of motion which leads to uniform scaling of the atomic velocities at each time step (see Wu and Brooks for details[5]). The result of this conservation mechanism is that while the diffusivities of the lower frequency modes are enhanced due to the memory function, the opposite effect occurs for the higher frequency modes.[5]

Temperature-Based Replica Exchange. The replica-exchange protocol involves performing simultaneous simulations over a range of temperatures and/or biasing potentials.[11] In this work, only a range of simulation temperatures is considered. Each simulation, a, exchanges its temperature with another simulation, b, if $\Delta_{ab} < 0$, or $\exp(-\Delta_{ab})$ is greater

than a random number with uniform distribution, $r \in (0,1)$, where Δ_{ab} is defined as

$$\Delta_{ab} = (U_b - U_a)(1/k_B T_a - 1/k_B T_b) \quad (3)$$

where U_i is a defined energy measure of the replica-exchange simulation client i corresponding to temperature T_i and k_B is Boltzmann's constant. The ReX procedure can be thought of as an autonomous heating and cooling procedure, whereby, to a large extent, lower energy conformations settle into lower temperature windows and vice versa.

Simulation Setup. In this work, we used the PARAM22[29] force field with the CMAP backbone dihedral cross-term extension[30] and the generalized Born (GB) implicit solvent model, GBMV2.[31] GBMV2 is one of the more accurate implicit solvent models currently available, as it correctly mimics the Poisson solvation energy using a molecular surface-based dielectric boundary.[32] Implicit solvent models greatly reduce the number of simulation degrees of freedom compared to explicit solvent. This has two key benefits. First, in the context of ReX simulations, a reduced magnitude of potential energies leads to a smaller number of temperature clients necessary to span the desired temperature range to ensure frequent Metropolis exchanges of among clients. Also, implicit solvent accelerates conformational sampling in its own right by eliminating the diffusive reorientation of solvent molecules upon changes in protein conformation.

An integration time step of 2 fs was used as the SHAKE algorithm[33] is applied to fix all covalent bonds with hydrogen atoms. Nonbonded electrostatics and van der Waals interactions were truncated smoothly from 12 to 14 Å. For the MD simulations, a Nose-Hoover thermostat was used with a temperature coupling constant of 50 kcal/s^2. For the Langevin dynamics and self-guided Langevin dynamics simulations, the friction constant, γ, was set to 1 ps^{-1} for all heavy atoms. The SGLD guiding factor, λ, was set to 1, while the averaging time, τ, was set to 1 ps. These two values were arrived at on the basis of a compromise between sampling efficiency and preserving the backbone $\varphi-\psi$ free energy landscape of the alanine tripeptide (results not shown). In particular, smaller values of λ lead to diminished sampling benefits of the self-guided formalism, while larger values lead to distortions in the $\varphi-\psi$ free energy map.

Replica exchange was performed using the MMTSB[34] script *aarex.pl* which is a front-end to the CHARMM molecular dynamics package (version c33b2).[35] The replica exchange protocol utilized 16 simulation clients, with temperatures exponentially spaced between 298 and 500 K. Default GBMV2 parameters were used with the exception of the β value, which was set to -12 to improve energy conservation.[36] A surface tension value of 0.00542 kcal/(mol·Å^2) was used for the solvent accessible surface area nonpolar solvation term.[37]

The original Trp-cage protein (dubbed by its inventors, "5b") was used in this work. It has the sequence "NLYIQWLKDGGPSSGRPPS" and was structurally determined by NMR (PDB ID: 1L2Y).[19] We performed six simulations in total: MD-ReX, LD-ReX, and SGLD-ReX each starting from the native (NMR conformer 1) and unfolded trans conformations. The simulations are heretofore labeled as method/starting structure, e.g., SGLD-ReX/native. The trans simulations were each run for 100 ns. The native simulations were each run for 200 ns.

Evaluation Metrics. We compared predicted structural and thermodynamic properties of protein folding between the three simulation methods and experimental observation. Our measure of model quality was root-mean-squared deviation of the α-carbon trace to the native NMR conformer 1 structure (C_α rmsd). We predicted the heat capacity as a function of temperature, melting temperature, and folding free energy of the Trp-cage, all of which can be compared directly to experiment.[38]

Heat capacity was calculated in two ways. The "instantaneous" heat capacity, C_v^1, was computed as

$$C_v^1(T) = \frac{{\sigma_U}^2}{k_B T^2} = \frac{\langle U^2 \rangle - \langle U \rangle^2}{k_B T^2} \quad (4)$$

where U is the potential energy derived from WHAM analysis (described below) and T is the temperature. The distinction of constant volume (subscript v) vs constant pressure (subscript p) is meaningless in the context of implicit solvent simulations. Useful for canonical ensemble tests (described below), an alternative heat capacity formula does not require WHAM analysis, but is only defined at the midpoints of adjacent replica-exchange client temperatures. It is computed as a finite difference of the potentials of two contiguous temperature clients:

$$C_v^2\left(\frac{T_a + T_b}{2}\right) = \frac{\langle U_b \rangle - \langle U_a \rangle}{T_b - T_a} \quad (5)$$

We estimated the melting temperature, T_m, using three approaches. First, T_m^1 is the location of the maximum value of the computed $C_v^1(T)$ function. Second, T_m^2 is computed as the point where the derivative of the average rmsd as a function of temperature is maximum, which is roughly the inflection point. Finally, T_m^3 is found by iteratively searching for the temperature at which the free-energy difference between the native and unfolded state, $\Delta\Delta G_{fold}(T)$, is approximately equal to zero. Free energy of folding as a function of temperature, $\Delta\Delta G_{fold}(T)$, was computed as

$$\begin{aligned} \Delta\Delta G_{fold}(T) &= -k_B T \ln\left(\frac{\rho_{fold}(T)}{\rho_{unfold}(T)}\right) \\ &= -k_B T \ln\left(\frac{\int_{rmsd<rmsd_{fold}} \rho(T)\, dV}{\int_{rmsd>rmsd_{fold}} \rho(T)\, dV}\right) \end{aligned} \quad (6)$$

where the population density, ρ, as a function of C_α rmsd is integrated over the rmsd-delineated "native" and "unfolded" domains and dV is the volume element. The boundary dividing the folded and unfolded regions, $rmsd_{fold}$, is somewhat arbitrary given that experimental rmsd values cannot be observed. Nonetheless, $rmsd_{fold}$ can be deduced as the point where the population density is equal in both domains (i.e., $\Delta\Delta G_{fold}(T) = 0$) when computed at a predicted melting temperature, T_m, using either the T_m^1 or T_m^2 definitions. On the basis of this analysis and inspection of the free-energy

landscapes of all simulation methods as a function of rmsd, $rmsd_{fold}$ was chosen to be 3.4 Å. This definition liberally includes not only the native-like basin (rmsd ~ 1.0 Å) but several near-native compactly folded basins.

The heat capacity, C_v^1, folding free energies, and free-energy landscapes were derived from the multidimensional temperature-based-weighted histogram method WHAM[13] algorithm applied to the ReX simulation data.[39] The following dimensions were binned in the WHAM algorithm: potential energy, C_α rmsd, and radius of gyration, R_g. Structural representatives on the rmsd vs R_g landscape were selected visually and verified by rmsd analysis against the entire set of structures at 270 K.

Because of the ad hoc force term, SGLD may not strictly adhere to rigorous statistical mechanics and thus fail to produce a canonical ensemble.[5] We wanted to ascertain to what extent, if any, does SGLD deviate from canonical behavior, especially in combination with T-ReX. There are several measures one can use to assess departure from theoretical canonical ensemble behavior.[40] We evaluated four criteria in this work. First, we compared the actual average temperatures of the replica-exchange client simulations vs the temperatures specified in the input. Second, the ratio of heat capacity values, g, derived from a single temperature vs two temperatures was calculated as

$$g(T) = C_v^2(T)/C_v^1(T) \tag{7}$$

In a hypothetical canonical ensemble of a single state, the value of g should be 1.[40] The skew, S, of the potential energy distribution for a fixed temperature client is defined as

$$S = \frac{\langle (U - \langle U \rangle)^3 \rangle}{\langle (U - \langle U \rangle)^2 \rangle^{3/2}} \tag{8}$$

The skew for a perfect canonical ensemble of a single state should be zero.[40] In other words, the canonical energy distribution far away from a transition temperature should be strictly Gaussian. Our final metric was the kinetic energies or "temperatures" of the individual normal modes of the system. This measure, as far we are aware, has not been reported elsewhere. The question addressed in this metric is whether SGLD-ReX overheats low-frequency modes and cools high-frequency modes to retain the correct total macroscopic kinetic energy (i.e., temperature). If we define A_{ni} to be the ith Cartesian degree of freedom for the nth normal mode vector, and v_i is the velocity of the ith degree of freedom, then the temperature of the nth normal mode, T_n is

$$T_n = \frac{1}{2k_B}\left(\sum_i A_{ni}\sqrt{m_i}v_i\right)^2 \tag{9}$$

Velocities were obtained at 0.1 ps intervals from independent 1 ns simulations using the MD, LD, and SGLD protocols. The Cartesian normal mode matrix, **A**, was computed with the *vibran* module in CHARMM after optimizing NMR conformer 1 with 2000 steps of adopted-basis Newton−Raphson minimization.

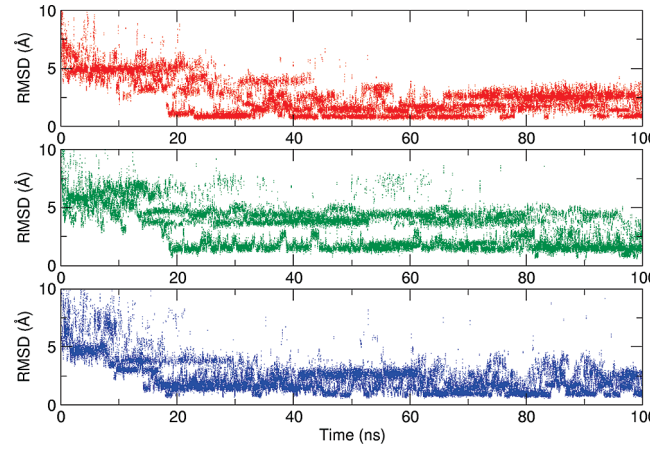


Figure 1. C_α rmsd of the lowest temperature ReX window (270 K) as a function of simulation time (smoothed by a 1 ns running average) for the folding simulations starting from the unfolded trans structure. Legend: red, MD-ReX; green, LD-ReX; blue, SGLD-ReX.

We also looked at three model systems to investigate possible deviations between SGLD and LD using one-dimensional potentials with two wells. The first model potential was a symmetric double well:

$$U_1(x) = x^2(x - 2)^2 \tag{10}$$

The second potential was a double well with one minimum higher than the other:

$$U_2(x) = x^2(x - 2)^2 + 2x \tag{11}$$

The third potential had two wells of different curvatures, narrow and broad:

$$U_3(x) = 0.499 + 0.7372x + 12.51x^2 - 23.883x^3 + 16.659x^4 - 5.1411x^5 + 0.59399x^6 \tag{12}$$

A fictitious particle was propagated along the model potential in CHARMM with a mass of 2171 amu. The mass was chosen to reproduce that of the Trp-cage. The y and z dimensions were restrained by a 500 (kcal/mol)/Å^2 harmonic potential. LD and SGLD parameters were the same as those used in the Trp-cage simulations. The time step was set to 10 fs, and the total simulation times were 10 μs. Trajectories of the x-coordinates were saved every picosecond and placed into histograms that were binned in 0.01 Å increments.

Results

First, we assessed how quickly the three simulation models fold up the Trp-cage protein starting with the trans conformation. In Figure 1, the C_α rmsd of the lowest temperature ReX window (270 K) is plotted as a function of simulation time. Because of the Metropolis criterion, low-temperature selection is a reasonable measure of energy-based structure detection. The SGLD-ReX, LD-ReX, and MD-ReX simulations first detect structures below 1.5 Å rmsd at 14.1, 18.5, and 18.5 ns, respectively. These results suggest only a modest speedup for SGLD-ReX in approaching the native basin starting from the unfolded state. This result is in stark

Table 1. Predicted Melting Temperatures and Folding Free Energies[a]

simulation	simulation time (ns)	T_m^1 (K) (max C_p)	T_m^2 (K) (rmsd)	T_m^3 (K) ($\Delta\Delta G_{fold} = 0$)	$\Delta\Delta G_{fold}$ ($T = 298$ K)
MD-ReX/trans	50−100	360	362	351	−1.8
MD-ReX/native	50−100	366	367	355	−1.6
MD-ReX/native	150−200	369	369	348	−1.7
LD-ReX/trans	50−100	375	352	290	0.2
LD-ReX/native	50−100	339	345	335	−1.9
LD-ReX/native	150−200	372	370	354	−1.4
SGLD-ReX/trans	50−100	315	315	311	−0.8
SGLD-ReX/native	50−100	339	338	331	−1.4
SGLD-ReX/native	150−200	324	320	306	−0.4
experiment	n/a	324	316[b]	317	−0.76 (0.05)

[a] Experimental results are taken from Streicher and Makhatadze.[38] [b] Derived from midpoint of "fraction unfolded" graph.[38]

comparison to the 65-fold speedup of helix formation reported in the original SGLD paper.[5] The likely explanation is that ReX already provides sufficient sampling enhancement for MD and LD to overcome the unfolded/folded transition barrier to fold up the Trp-cage. For all methods, the lowest C_α-rmsd basin appears to reside around 0.9 Å from the NMR conformer 1. This result compares favorably to the fact that the 37 other NMR conformers in PDB entry 1L2Y[19] are also, on average, 0.9 Å (rmsd) from conformer 1.

Comparison of calculated and experimental folding free energies requires defining the native basin. The resolution limit of our force-field/implicit solvent model energy function compelled us to use a liberal definition of the native basin. By evaluating the population distribution as a function of rmsd at the melting temperature of each simulation, we found that there were three compact basins that resided below 3.4 Å. Therefore, the dividing line between native/nonnative was set to 3.4 Å rmsd. Using this definition and data from the last 50 ns of the 200 ns native simulations, the free energies of folding predicted by SGLD-ReX, LD-ReX, and MD-ReX are −0.4, −1.4, and −1.7 kcal/mol, respectively, as reported in Table 1. The SGLD-ReX result is closest to the experimental folding free energy of −0.76 kcal/mol.

Also, in Table 1, predicted melting temperatures are compared. Melting temperatures derived from heat capacity (T_m^1) and rmsd (T_m^2) are nearly the same except for LD-ReX (see the PMF analysis below.) In contrast, melting temperatures derived from the transition point ($\Delta\Delta G_{fold} = 0$), T_m^3, are 10−20 K lower than T_m^1, which suggests the rmsd = 3.4 Å unfolded/folded dividing line is somewhat imprecise. SGLD-ReX simulations predict a melting temperature coinciding with experimental results. However, MD-ReX and LD-ReX predict melting temperatures ∼40−50 K higher than experiment. The large discrepancies in predicted melting temperatures between SGLD-ReX and traditional ReX approaches cannot be explained by the uncertainty of the calculations, which is ∼10 K.

To better understand the melting transition, the predicted heat capacity curves were compared against the experimental results in Figure 2. SGLD-ReX, as discussed earlier, predicts the closest temperature peak to experiment. Predictions from MD and LD are fairly consistent between the two simulation models in the temperature peak and profiles. The magnitudes of the heat capacities were not expected to correspond to experiment because implicit solvent models incorporate the free energy associated with the solvent degrees of freedom as part of the total potential energy, U, or "enthalpy" of the system.[41] This fact can distort calculations of observables on the basis of the enthalpy (e.g., heat capacity). However, the width of the heat capacity profile as a function of temperature was not expected to be affected to the same degree.

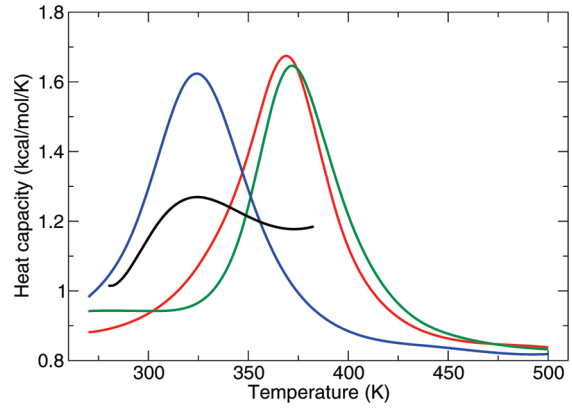


Figure 2. Predicted heat capacity as a function of temperature for the three simulation models (starting from the native structure) using the last 50 ns of 200 ns simulation data compared to experiment.[38] Legend: red, MD-ReX; green, LD-ReX; blue, SGLD-ReX; black, experiment.

To comprehend why the three different simulation models predict dissimilar observables even though the underlying potential energy function is identical, we investigated the potentials of mean force (PMFs) along various coordinate dimensions and performed several tests of canonical ensemble behavior. The PMFs along C_α rmsd to native and R_g coordinates in Figure 3 provide insight into the various conformational basins explored by the simulations. For the simulations starting from the unfolded state, the LD-ReX simulations generated three distinct intermediate folded states and only one near-native state when the PMF is mapped onto rmsd and R_g coordinates. In contrast, the SGLD-ReX simulations populated three near-native basins, while the MD-ReX simulations obtained similar results, further splitting the middle basin.

The nearest-to-native basin incorporates conformations with most of the correct features compared to the NMR structure as seen in a structure representative of that basin (Figure 3, model R1). The only defects are that the tryptophan side chain has a slight twist compared to the native, and the elusive 3_{10}-helix (as gauged by DSSP[42]) is

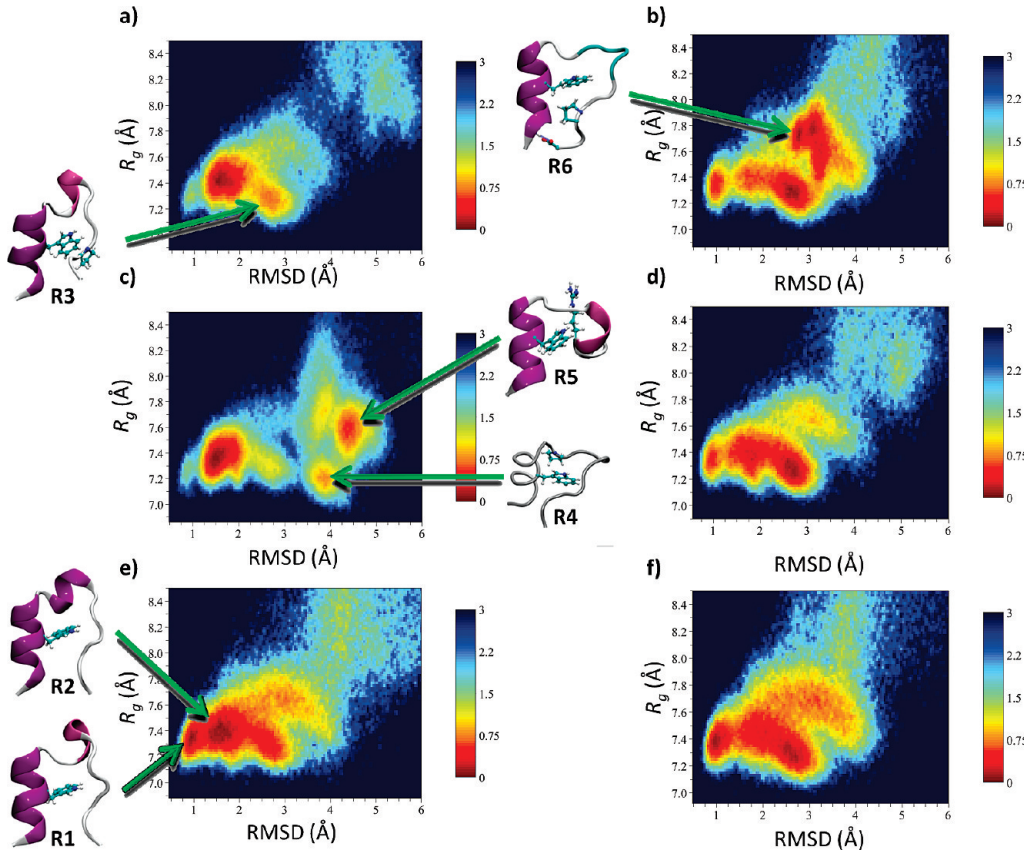


Figure 3. Free-energy landscapes at $T = 298$ K as a function of C_α rmsd to native and radius of gyration, R_g, using the last 50 ns of data from the 100 ns simulations starting from the folded conformation (right plots) and the unfolded trans conformation (left plots): (a, b) MD-REX; (c, d) LD-REX, and (e, f) SGLD-REX. The free energies are represented by a range of colors and are arbitrarily capped at 3 kcal/mol.

only sampled a small percentage of the time. In comparison, the main flaw of the representative structure of the 1.5 Å basin (Figure 3, R2) is incorrect folding in the turn region even though the Trp indole ring is stacked parallel to the correct proline (residue 18). The ∼2.7 Å rmsd basin representative (Figure 3, R3) has even larger flaws in the turn region despite the fact that the tryptophan is once again packed against Pro18. Because this basin is slightly more compact than the basins that are nearer to native, nonspecific hydrophobic collapse from the implicit solvent surface area term is a likely culprit. Representative compact structures of the LD-ReX/trans simulation that are even farther from the native (Figure 3, R4 and R5) show nonnative packing of the indole ring against residues Arg16 (R4) and Pro12 (R5). The MD simulation starting from the native produced a stable basin of structures (R6) in which the carboxyl terminus of Ser20 hydrogen bonds to the amide nitrogen of N-terminus residue Leu2, while Pro18 stacks perpendicularly to the tryptophan ring. In contrast, SGLD sees that same part of conformational space in rmsd/R_g dimensions but does not dwell there for any significant length of time, probably due to the ad hoc force term which favors increased global sampling vs local sampling.

From PMFs in the dimensions of potential energy and C_α rmsd to native at the transition temperature ($\Delta G_{fold} = 0$), further insights into the differences among the protocols can be gleaned as seen in Figure 4. All three methods sample the nearest-to-native basin (∼1 Å) at their respective transition temperatures, with SGLD-ReX having the most density there. Since the nearest-to-native basin does not appear to be the lowest in free energy, this could be due to the fact that SGLD-ReX performs the most excursions among basins in a given simulation time. Another positive feature of SGLD-ReX (Figures 4g−i) is how similar the PMFs are among the different starting conformations and data collection times. This suggests that, of the three methods, SGLD-ReX is the most self-consistent and arguably the most converged, at least in the conformational space of compact folds. The 150−200 ns data windows of MD-ReX and LD-ReX do have qualitative agreement with SGLD-ReX, suggesting that longer equilibration times can bring these three methods into better agreement. In particular, the sampled rmsd basins are similar among three methods at 150−200 ns. The main drawback with SGLD-ReX is that the computed transition temperature and resultant basin energies are significantly lower, suggesting a distortion of the relative free energy among folded and unfolded states. Finally, note that the rmsd in the range of 3.2−3.5 Å appears to be a sensible dividing line between folded/unfolded states, as, in many cases, the next basin above this line is higher in energy.

In Figure 5, we computed several properties of the simulation protocols to determine if there were any major deviations from theoretical canonical ensemble behavior. In Figure 5a, only MD-ReX produced temperatures precisely in line with those specified by the user, thanks to strict temperature control with a Nose-Hoover thermostat. The

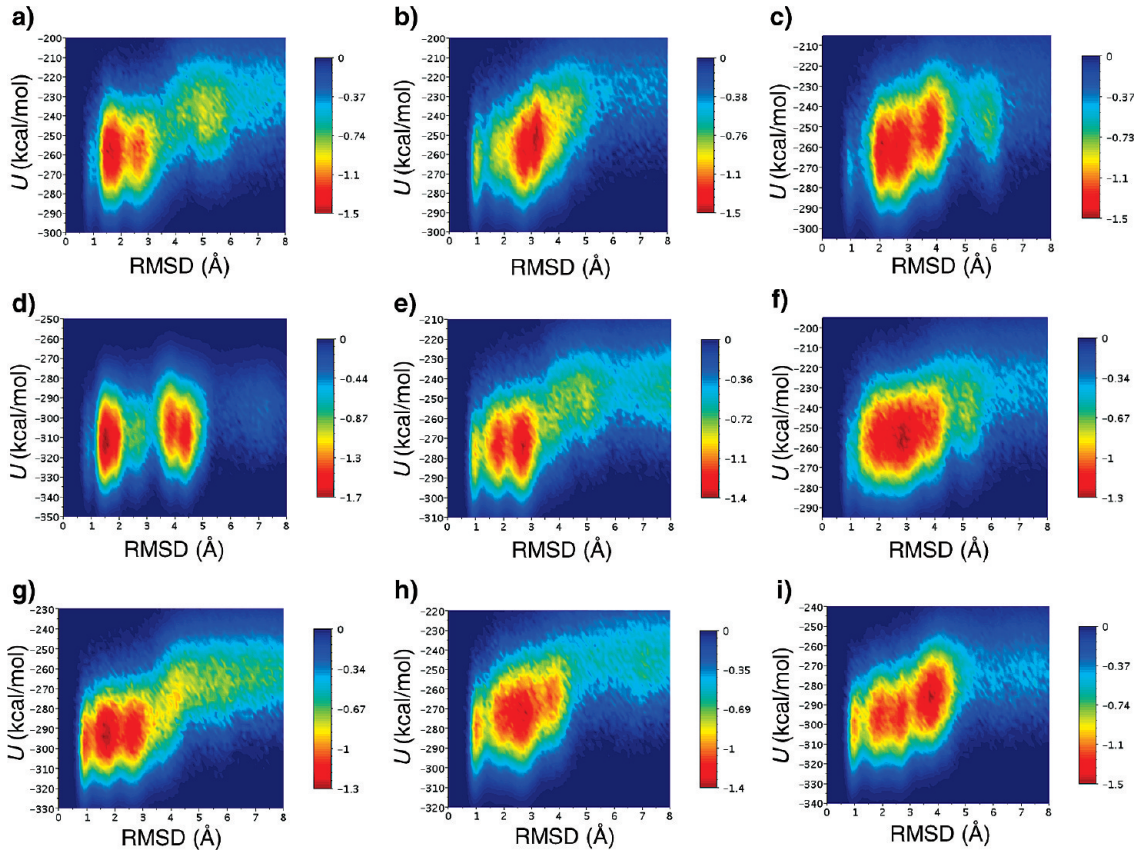


Figure 4. Free-energy landscapes at respective melting temperatures ($\Delta G_{fold} = 0$) of individual simulations ⟨*method/starting structure/simulation data*⟩ in the coordinates of potential energy, U, and C_α rmsd to native: (a) MD-ReX/trans/50−100 ns (T = 351.3 K), (b) MD-ReX/native/50−100 ns (T = 354.6 K), (c) MD-ReX/native/150−200 ns (T = 348.2 K), (d) LD-ReX/trans/50−100 ns (T = 290.1 K), (e) LD-ReX/native/50−100 ns (T = 335.1 K), (f) LD-ReX/native/150−200 ns (T = 353.9 K), (g) SGLD-ReX/trans/50−100 ns (T = 311.2 K), (h) SGLD-ReX/native/50−100 ns (T = 331.5 K), and (i) SGLD-ReX/native/150−200 ns (T = 306.4 K).

deviations for SGLD-ReX are the most substantial, up to 3 K too low for the highest temperature window. Fortunately, in Figure 5b, the Metropolis exchange factors implied by the simulation temperatures for SGLD-ReX are still consistent with the ones actually used for exchange. In Figure 5c, the ratio of heat capacities calculated by instantaneous and finite difference methods are all close to the theoretical value of 1 away from the transition state. In all cases, the ratio deviates from 1, as expected for a two-state superposition around the predicted melting temperatures of the respective methods. Figure 5d paints a similar story, whereby the potential energy histograms of all three methods have zero skew except for the transition temperature where two Gaussians are expected.

Next, the average potential energies of replica-exchange clients over the last 50 ns of simulation data elucidates how the client simulations sampled the potential energy surface starting from the unfolded structure (Figure 5e). The profile of the average energy over different client temperatures mirrors the heat capacity curve in that different transition temperatures (seen here as deviations from linearity) can be observed for the various protocols. In addition, among the three protocols at the lowest temperature, MD-ReX digs deepest into the potential energy surface, providing the lowest energy structures by several kilocalories per mole. This small discrepancy could partially explain the free energy differences between the three methods.

Another test stems from a concern that SGLD may overheat low-frequency modes and cool high-frequency modes to compensate for total energy conservation and maintenance of the user-specified system temperature. Fortunately, as seen in Figure 5f, all methods have roughly the same profile of normal mode temperature vs index of normal mode (going from slowest to fastest). Interestingly, the individual mode kinetic energies of all three methods tend to start out high and go lower with increasing mode frequency. This trend can be attributed as an artifact of SHAKE. When SHAKE is not run, the kinetic energy profiles are virtually flat along the entire range of modes with the correct value associated with the simulation temperature (results not shown).

Finally, we compare LD and SGLD for generating free-energy profiles of three double-well potentials (Figure 6). In the first potential model, Figure 6a, the potential is a symmetric double well. The deviation from ideal behavior is quite small for both LD and SGLD ($\tau = 0.1$ ps), namely, ∼0.005 kcal/mol. However, SGLD accelerates sampling by crossing between wells 4883 times, while LD only crosses 1117 times in the 10 μs simulation. The increased crossing frequency results in a lowered free-energy barrier between

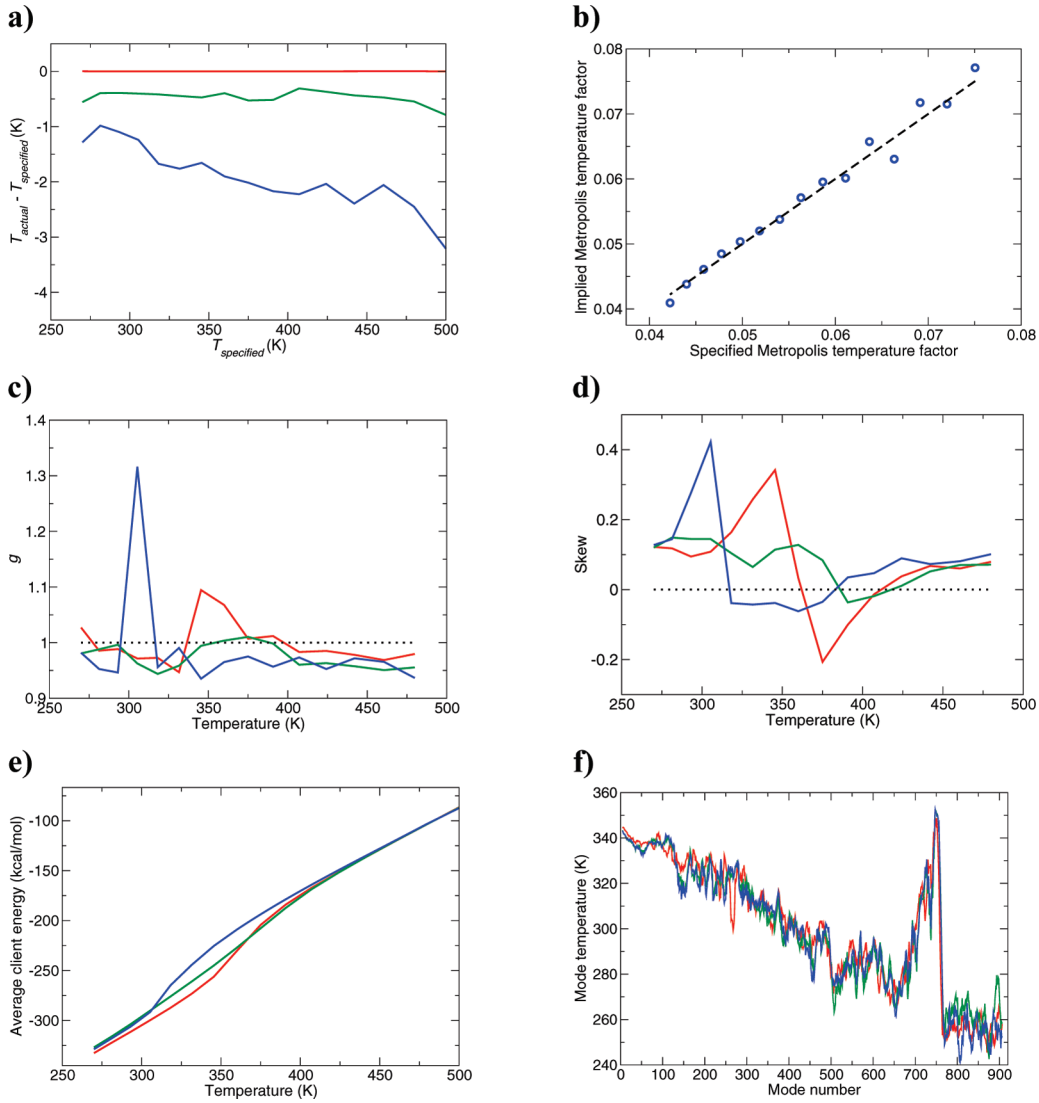


Figure 5. Various metrics of the ensembles generated at each replica-exchange temperature window for the three simulation protocols: (a) deviations from specified temperature (trans starting structures), (b) Metropolis temperature factors implied by simulation temperatures vs factors actually used for exchange, (c) *g*, ratio of heat capacities calculated by fluctuations and finite difference, (d) skew of potential energy distribution, (e) average potential energy of temperature windows over the last 50 ns of simulations starting from the unfolded trans conformation, and (f) temperatures of normal modes for each simulation protocol. Legend: red, MD-ReX; green, LD-ReX; blue, SGLD-ReX. Dashed lines represent optimal values.

the two wells of about 1 kcal/mol. Next, for a heterogeneous model system with one well higher in energy than the other as in Figure 6b, SGLD shifts the higher energy well down by ~0.3 kcal/mol. Finally, for a different heterogeneous system where the curvature of the higher well is broader than the lower well (Figure 6c) mimicking a folded/unfolded peptide landscape, the higher energy well is actually stabilized by ~0.3 kcal/mol. In both heterogeneous model systems, the free-energy barriers are reduced as expected. Unfortunately, the heterogeneous model systems indicate that SGLD distorts the relative free energy of minima with different energies and curvatures. How this result directly translates to a real protein system such as Trp-cage with a complex topological landscape is unclear.

Discussion

Self-guided Langevin dynamics was originally devised to accelerate low-frequency motions in order to enhance sampling. Our results for simulating folding−unfolding of the Trp-cage indicate that, compared to MD and LD, the topological free-energy barriers among major conformational basins were effectively reduced, but overall folding times were not significantly improved. The melting temperature predicted by SGLD is noticeably lower than those produced by the MD and LD approaches. This result could be a sign of actual changes in the effective population density sampled by the SGLD simulations or errors in canonical ensemble behavior that might be expected from the use of an ad hoc force term. Investigating the latter possibility, several canonical ensemble tests indicate that any error introduced by the SGLD force term can be ruled out. At worst, the actual simulation temperatures do drift moderately (~3 K) from their specifications, which could partly be due to the use of a high-energy derivative formalism such as GBMV[36] which, in turn, may compound errors associated with the total energy correction term in SGLD. As shown previously,[5] the diffu-

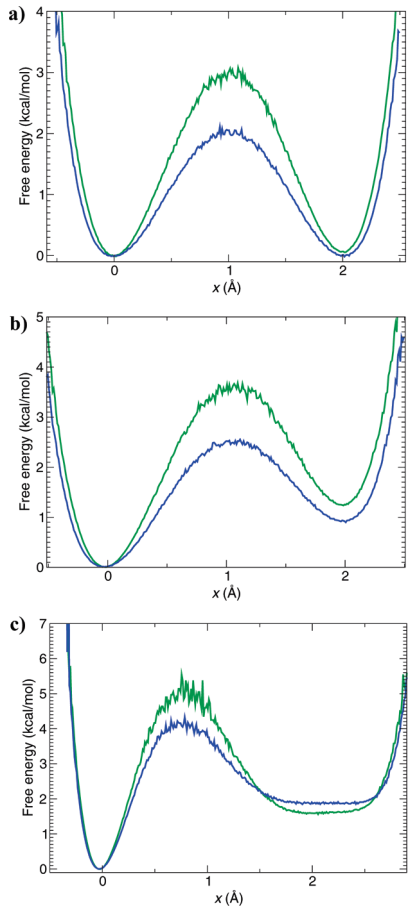


Figure 6. Potentials of mean force for three one-dimensional model systems (eqs 10, 11, and 12) using LD and SGLD integrators: (a) symmetric double well, (b) double well with different energies, and (c) narrow well/broad well combination. Legend: green, LD; blue, SGLD.

sivity of an SGLD simulation is greatly increased compared to LD. However, an analysis of the kinetic energies of the normal modes ruled out the possibility that the low-frequency modes in SGLD were actually "hotter" compared to MD and LD.

While our measured deviations from canonical ensemble behavior were not significant, predicted thermodynamic observables and the PMF landscapes in Figures 3 and 4 do differ substantially. Therefore, we cannot rule out the possibility that SGLD may have lowered the melting temperature by smoothing the basins of folded and unfolded states, thereby raising the relative free energy of the folded state. This could happen if the ad hoc force term in SGLD somehow propels the system out of narrow basins such as the native state more often than the broad unfolded regions. Another possibility is that SGLD slightly scales down the entire free-energy surface, reducing not just the free-energy barriers but the energetic difference between free-energy minima. To definitively test this prospect, we ran three model one-dimensional potentials. The SGLD simulation of the symmetric double well agreed with LD for the relative free energy of the minima but had a reduced barrier that increased crossing between states by a factor of 4. However, the SGLD simulation of the asymmetric narrow well potential yielded a free energy difference too small by 0.3 kcal/mol. Interestingly, SGLD simulation of narrow/broad well combination yielded the opposite effect: the magnitude of ΔG was 0.3 kcal/mol too large. This result confirms that SGLD does distort relative free energies of minima, albeit less than its favorable reduction of transition barriers. In any case, these test systems cannot be quantifiably translated to a complex system such as Trp-cage. Thermodynamic distortions could be definitively resolved by reweighting the simulation data. Unfortunately, such a reweighting scheme has not been devised yet.

While SGLD induces distortions in the simulations of simple model systems, molecular dynamics with a strict thermostat such as Nose-Hoover is not without fault. It should be granted that MD is better overall at locally sampling protein conformations compared to SGLD and LD. The evidence for this assertion is that in both the native and trans simulations, MD samples the lowest average energy of conformations at room temperature. In addition, the conformational basins as viewed by the PMF landscapes of MD are tighter than in LD and SGLD. The drawback of local sampling is that MD simulations can get trapped in certain conformational basins while insufficiently sampling others. In fact, it has been shown that in the limit of systems with a small number of degrees of freedom, Nose-Hoover MD simulations with only a single fictitious coordinate can become nonergodic, i.e., repeat the same trajectory ad infinitum.[43] Therefore, MD sampling may be incomplete or at least slower than SGLD at visiting all of the relevant compact conformational basins.

Thermodynamic issues aside, SGLD-ReX has a few salient properties. First, SGLD-ReX seems to show improved sampling convergence as the PMF native results at 50−100 ns visually match up quite well to the 150−200 ns data. This could be due to the ad hoc force term which increases transition probabilities among neighboring conformational basins. Better global sampling could partially explain the improved agreement of SGLD to experiment as previous studies have shown that all dominant basins need to be sampled to obtain accurate free-energy estimates.[41] Next, of the three approaches, SGLD-ReX has the highest sampling density of the nearest-native basin (<1 Å rmsd). This result may be due to the fact that SGLD-ReX does a better job of skimming the potential surface, revisiting the nearest-native basin more often. Finally, unlike coarse-grained methods, which smooth the potential energy surface by reducing spatial resolution, the self-guided formalism does not introduce any distortions into the generated conformations. This opens up the possibility of applying SGLD-ReX to the protein structure refinement problem.[44]

There are several caveats in this study. First and foremost, an implicit solvent model was used rather than explicit water molecules.[45] This choice enabled straightforward application of temperature-based replica exchange because of reduced degrees of freedom. Nonetheless, implicit solvent models have reduced conformational resolution due to artifacts such as too strong (or too weak) salt bridges[46] and hydrogen bonds, and missing noncovalent attractions between protein and solvent.[47] Another issue in this work is that we used a fixed-charge force field rather than a flexible-charge one

which can be problematic in studies such as this one where large conformational changes are expected[48] (i.e., unfolded to folded.) Also, while the simulation time of 100 ns was sufficient to permit ab initio folding of the Trp-cage protein followed by roughly 50 ns of production data, sampling convergence was not achieved in some cases. For example, the MD-ReX and LD-ReX simulations, starting from either native or unfolded trans conformations, produced quite different rmsd vs R_g PMF landscapes. Moreover, to fold up proteins larger than the Trp-cage, longer simulation times may be required. Finally, while fixed simulation temperatures for replica exchange was sufficient for this work, adaptive changes to the temperature set[12] may be necessary for larger proteins, especially ones with a sharper peak in their heat capacity profile (i.e., sharper energetic transition between the folded and unfolded state).[39]

Our study shows that combining SGLD and ReX produces a sampling method that has both advantages and disadvantages compared to Nose-Hoover-based MD-ReX and LD-ReX. Enhanced sampling convergence for SGLD-ReX is seen in the free-energy landscapes and is likely due to reducing the transition barriers between unfolded and folded states. Nonetheless, SGLD-ReX lacks direct application to studying kinetics of protein folding due to possibly modifying the folding pathway and its degrees of free-energy frustration. In addition, SGLD-ReX may produce lower predicted free energies of folding and melting temperatures by shifting the relative heights of free energy minima in the same way that it reduces transition barriers between conformational basins. If a proper reweighting scheme were devised, these problems would be alleviated. All in all, with appropriate selection of parameters and acknowledgment of some distortions due its ad hoc nature, SGLD-ReX should find application in the calculation of thermodynamics of protein folding−unfolding and protein−ligand association. The method is also currently being evaluated in the emerging field of comparative protein model refinement.

Acknowledgment. We would like to thank Dr. I.-C. Yeh for supplying his WHAM code. Funding for this research was provided by the U.S. Department of Defense Threat Reduction Agency Grants 3.10010_06_RD_B and TMTI0004_09_BH_T, and the Department of Defense Biotechnology High Performance Computing Software Applications Institute. Computational time was provided, in part, by the U.S. Army Research Laboratory Major Shared Resource Center and Maui High Performance Computing Center. The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the U.S. Army or of the U.S. Department of Defense. This paper has been approved for public release with unlimited distribution.

References

(1) Freddolino, P. L.; Schulten, K. *Biophys. J.* **2009**, *97*, 2338.

(2) Zuckerman, D. M.; Lyman, E. *J. Chem. Theory Comput.* **2006**, *2*, 1200.

(3) Ishikawa, Y.; Sugita, Y.; Nishikawa, T.; Okamoto, Y. *Chem. Phys. Lett.* **2001**, *333*, 199.

(4) Hamelberg, D.; Mongan, J.; McCammon, J. A. *J. Chem. Phys.* **2004**, *120*, 11919.

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

(6) Pappu, R. V.; Marshall, G. R.; Ponder, J. W. *Nat. Struct. Biol.* **1999**, *6*, 50.

(7) Hornak, V.; Simmerling, C. *Proteins* **2003**, *51*, 577.

(8) Skolnick, J.; Zhang, Y.; Arakaki, A. K.; Kolinski, A.; Boniecki, M.; Szilagyi, A.; Kihara, D. *Proteins* **2003**, *53* (Suppl. 6), 469.

(9) Zhang, Y.; Arakaki, A. K.; Skolnick, J. *Proteins* **2005**, *61*, Suppl. 7, 91−8.

(10) Kirkpatrick, S.; Gelatt, C. D., Jr.; Vecchi, M. P. *Science* **1983**, *220*, 671.

(11) Paschek, D.; Nymeyer, H.; Garcia, A. E. *J. Struct. Biol.* **2007**, *157*, 524.

(12) Trebst, S.; Troyer, M.; Hansmann, U. H. *J. Chem. Phys.* **2006**, *124*, 174903.

(13) Gallicchio, E.; Andrec, M.; Felts, A. K.; Levy, R. M. *J. Phys. Chem. B* **2005**, *109*, 6722.

(14) Wu, X.; Brooks, B. R. *Biophys. J.* **2004**, *86*, 1946.

(15) Wen, E. Z.; Hsieh, M. J.; Kollman, P. A.; Luo, R. *J. Mol. Graphics Modell.* **2004**, *22*, 415.

(16) Damjanovic, A.; Garcia-Moreno, E. B.; Brooks, B. R. *Proteins* **2009**, *76*, 1007.

(17) Damjanovic, A.; Wu, X.; Garcia-Moreno, E. B.; Brooks, B. R. *Biophys. J.* **2008**, *95*, 4091.

(18) Damjanovic, A.; Miller, B. T.; Wenaus, T. J.; Maksimovic, P.; Garcia-Moreno, E. B.; Brooks, B. R. *J. Chem. Inf. Model.* **2008**, *48*, 2021.

(19) Neidigh, J. W.; Fesinmeyer, R. M.; Andersen, N. H. *Nat. Struct. Biol.* **2002**, *9*, 425.

(20) Paschek, D.; Hempel, S.; Garcia, A. E. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 17754.

(21) Kannan, S.; Zacharias, M. *Proteins* **2009**, *76*, 448.

(22) Snow, C. D.; Zagrovic, B.; Pande, V. S. *J. Am. Chem. Soc.* **2002**, *124*, 14548.

(23) Zhou, R. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 13280.

(24) Steinbach, P. J. *Proteins* **2004**, *57*, 665.

(25) Juraszek, J.; Bolhuis, P. G. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 15859.

(26) Gattin, Z.; Riniker, S.; Hore, P. J.; Mok, K. H.; van Gunsteren, W. F. *Protein Sci.* **2009**, *18*, 2090.

(27) Wu, X.; Wang, S. *J. Chem. Phys.* **1999**, *110*, 9401.

(28) Andricioaei, I.; Dinner, A. R.; Karplus, M. *J. Chem. Phys.* **2003**, *118*, 1074.

(29) Mackerell, A. D., Jr.; Bashford, D.; Bellott, D. M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, I. W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586.

(30) MacKerell, A. D., Jr.; Feig, M.; Brooks, C. L., 3rd. *J. Am. Chem. Soc.* **2004**, *126*, 698.

(31) Lee, M. S.; Feig, M.; Salsbury, F. R., Jr.; Brooks, C. L., 3rd. *J. Comput. Chem.* **2003**, *24*, 1348.

(32) Feig, M.; Onufriev, A.; Lee, M. S.; Im, W.; Case, D. A.; Brooks, C. L., 3rd. *J. Comput. Chem.* **2004**, *25*, 265.

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

(34) Feig, M.; Karanicolas, J.; Brooks, C. L., 3rd. *J. Mol. Graphics Modell.* **2004**, *22*, 377.

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

(36) Chocholousova, J.; Feig, M. *J. Comput. Chem.* **2006**, *27*, 719.

(37) Feig, M.; Brooks, C. L., 3rd. *Proteins* **2002**, *49*, 232.

(38) Streicher, W. W.; Makhatadze, G. I. *Biochemistry* **2007**, *46*, 2876.

(39) Yeh, I. C.; Lee, M. S.; Olson, M. A. *J. Phys. Chem. B* **2008**, *112*, 15064.

(40) Rosta, E.; Buchete, N.-V.; Hummer, G. *J. Chem. Theory Comput.* **2009**, *5*, 1393.

(41) Chang, C. E.; Gilson, M. K. *J. Am. Chem. Soc.* **2004**, *126*, 13156.

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

(43) Cooke, B.; Schmidler, S. C. *J. Chem. Phys.* **2008**, *129*, 164112.

(44) Lee, M. S.; Olson, M. A. *J. Chem. Theory Comput.* **2007**, *3*, 312.

(45) Duan, Y.; Kollman, P. A. *Science* **1998**, *282*, 740.

(46) Zhou, R.; Berne, B. J. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 12777.

(47) Swanson, J. M.; Mongan, J.; McCammon, J. A. *J. Phys. Chem. B* **2005**, *109*, 14769.

(48) Patel, S.; Mackerell, A. D., Jr.; Brooks, C. L., 3rd. *J. Comput. Chem.* **2004**, *25*, 1504.

CT100062B

JCTC Journal of Chemical Theory and Computation

Simulating the Transition between Gel and Liquid-Crystal Phases of Lipid Bilayers: Dependence of the Transition Temperature on the Hydration Level

Bruno A. C. Horta,*,† Alex H. de Vries,‡ and Philippe H. Hünenberger*,†

Laboratory of Physical Chemistry, ETH Zürich, CH-8093 Zürich, Switzerland, and University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands

Received April 14, 2010

Abstract: Explicit-solvent molecular dynamics (MD) simulations of the monoglyceride glycerol-1-monopalmitin (GMP; bilayer patch of 2 × 6 × 6 lipids) at different hydration levels (full, half, or quarter hydration) and at different temperatures (318 to 338 K) are reported. The 40 ns simulations (some extended to 200 ns) are initiated from structures appropriate for the gel (GL) or liquid-crystal (LC) phases, with the goal of investigating whether atomistic MD simulations on this time scale can be used to monitor GL ↔ LC transitions in lipid bilayers, and to evaluate the corresponding transition temperatures T_m (as well as the influence of hydration on T_m) in a reliable fashion. The main conclusions are as follows: (*i*) The GL → LC transition of GMP can be observed on the 40 ns time scale. (*ii*) The LC → GL transition is comparatively slower, requiring simulations on the 200 ns time scale. (*iii*) A set of simulations initiated from a structure appropriate for the GL phase and carried out at slightly different temperatures permits the determination of a reliable value for T_m. (*iv*) The calculated T_m values reproduce the experimentally observed increase in this transition temperature upon decreasing the bilayer hydration. (*v*) The T_m values calculated at the three hydration levels considered are in essentially quantitative agreement with the experimental phase diagram of GMP. To our knowledge, this study represents the first accurate determination of the T_m of a lipid *via* atomistic simulations of the (reversible) GL ↔ LC phase transition, as well as the first direct simulation evidence for the increase in the transition temperature upon dehydration. The possible direct determination of T_m and the characterization of environmental effects on this quantity by simulation opens up promising perspectives in the contexts of force field refinement and the investigation of dehydration-induced damages in living cells (and bioprotection by cosolutes).

I. Introduction

Lipid bilayers are known to exist in several different lyotropic phases,[1,2] the relative stabilities of which depend on the temperature and pressure conditions, on the type of lipid molecules, on the identity and concentration of the solvent, and on the nature and concentration of possible cosolutes.[1,2] The most important and best studied of these phases are the crystal phases, the gel phases, and the liquid-crystal phase.[1] Different crystal and gel phases may be further distinguished, depending on the specific (regular) arrangement of the lipid molecules.

Crystal (CR) phases can be viewed as resulting from the periodic stacking of nearly dehydrated bilayers (except for the possible presence of a limited number of cocrystallization water molecules), in which the aliphatic lipid tails are arranged regularly in *all-trans* conformations.[1,2] These phases have a limited biological relevance, but polymorphism and crystallization phenomena are particularly important for the

* Corresponding authors. Phone: +41 44 632 5503. Fax: + 41 44 632 1039. E-mail: bruno.horta@gmail.com (B.A.C.H.), phil@igc.phys.chem.ethz.ch (P.H.H.).

† Laboratory of Physical Chemistry.

‡ University of Groningen.

10.1021/ct100200w © 2010 American Chemical Society
Published on Web 07/20/2010

food and pharmaceutical industries.[2] The crystal structures of numerous types of lipids have been solved by solid-state X-ray diffraction and have revealed interesting symmetry, enantiomerism, and polymorphism properties.[1−10]

Gel (GL) phases (in multilamellar systems) can be viewed as resulting from the regular (but not strictly periodic) stacking of significantly hydrated bilayers, in which the aliphatic lipid tails are arranged in nearly *all-trans* conformations and in orientations that are generally tilted with respect to the bilayer normal.[1] Due to the separation of successive bilayers by solvation shells, there is no or very little correlation between the arrangement of the lipids (e.g., tilting direction or undulations) in successive bilayers. Although the GL phase represents a highly ordered bilayer state, it is undoubtedly much more "dynamical" than the CR phase. This phase is not predominant in biological systems, although some evidence points toward the existence of small gel-like domains in cell membranes.[1,2]

The liquid-crystal (LC) phase (sometimes called fluid phase) also consists of significantly hydrated bilayers, but in which the aliphatic tails are conformationally disordered. Other structural and dynamic properties differentiating the LC from the GL phase are a broader distribution of the headgroup atoms along the bilayer normal, a larger area per lipid, a smaller bilayer thickness, and faster translational and rotational diffusion of the lipid molecules. Due to the highly disordered and "dynamical" (fluid-like) structure, many processes of biological importance can take place within a bilayer in the LC phase (e.g., insertion of macromolecules, transport, fusion, signaling). However, for the same reason, it is much more difficult to accurately measure structural properties of the LC phase compared to the more ordered CR and GL phases.[11]

The transition of a lipid bilayer from the LC phase to a more ordered (GL or CR) phase can be promoted by a temperature decrease (at constant water content). The temperature T_m at which this transition occurs (for a given water content) is referred to as the main transition temperature. Alternatively, the same transition can typically be promoted by dehydration (at constant temperature). In practice, however, the phase diagrams of lipid−water systems are generally complicated by the existence of other phases and by the slow kinetics associated with specific phase transitions, rendering sometimes difficult the clear-cut distinction between stable and metastable phases.

From the biological point of view, the LC ↔ GL phase transition is particularly important. Although the biologically relevant phase of cell membranes is considered to be the LC phase, the cholesterol-rich regions of this membrane may form small organized domains, in which the ordering approaches that of a GL phase.[12,13] This process is related to the formation of membrane rafts, which has been suggested to play a major role in cell signaling,[14−16] protein transport and association,[15] and other functions such as toxin binding.[17] Under stressful environmental conditions (e.g., low temperature or dehydration),[18,19] biological membranes can suffer undesirable phase transitions (e.g., LC → GL transition of a large portion of the membrane), leading to the loss of their biological function and, in more drastic situations, to the disruption of the membrane (upon reheating or rehydration). Many organisms have developed molecular strategies to prevent this transition, typically by adjusting the concentration of specific cosolutes (e.g., sugars) in the cell.[19−26] The lyotropic phase transitions are also important for many industrial processes such as food production and conservation and cosmetics and pharmaceutical formulations, as well as nanomaterials and synthetic applications.[1,27,28] Due to the great practical relevance of the problem, many experimental techniques have been developed for the accurate identification of lipid phases and the characterization of the corresponding transitions.[29−31]

Computer simulations have also greatly contributed to the characterization and understanding of the structure, dynamics, and thermodynamics of lipid bilayers. Among these methods, molecular dynamics (MD) simulation has proved to be particularly useful, because it provides information at a spatial (atomic level) and temporal (femtosecond) resolution inaccessible to experimentation, concerning system sizes (∼10 nm scale) and time scales (∼10−100 ns) still relevant for the evaluation of thermodynamic properties (statistical mechanics) and comparison with experimental data. Due to its biological relevance and to historical controversies concerning derived experimental data (e.g., value of the area per lipid,[11] possible presence of a surface tension[32,33]), many simulation studies have focused on the characterization of the LC phase.[34−39] However, some studies have also considered GL phases,[35,40−43] as well as other less common (but not less interesting) phases (e.g., ripple phase,[44] cubic phase[45,46]). On the other hand, although statistical models[47] and coarse-grained simulations[13] have been applied to characterize the GL ↔ LC transition, atomistic MD simulations that systematically investigate this process are scarce. The main reason is the relatively long time scale associated with this phase transition, which requires undertaking very long simulations. However, nowadays, time scales of 10−100 ns have become accessible for atomistic simulations considering systems with several thousand atoms. It is thus possible to simulate phase transitions of lipid−water systems in favorable cases, i.e., for systems where the equilibration of the individual phases occurs within time scales on this order. In the case of phospholipids, the time scale required for obtaining a fully equilibrated structure for the LC phase, which is a lower bound for the LC ↔ GL transition time scale, is already on the order of several nanoseconds.[48] For this reason, it may be more appropriate to study less complex bilayer systems such as surfactant or monoglyceride lipid systems. Since these have only one aliphatic tail per headgroup, their conformational and orientational relaxation is significantly faster, and convergence can be reached in comparatively shorter simulations.

In a recent study, Debnath et al.[49] reported MD simulations of a charged surfactant−cosurfactant system investigating the influence of the bilayer composition on the GL → LC transition. Different hydration levels were considered (50 and 80% w/w water concentrations), but no significant differences in the transition behavior were observed (in agreement with experimental results, given the high water content considered).

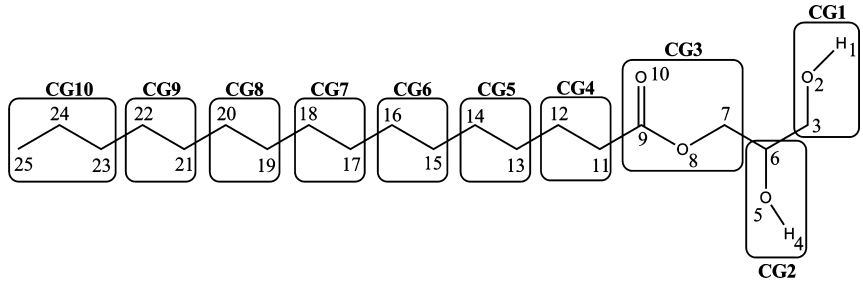


Figure 1. Chemical structure of the monoglyceride considered in the present study, glycerol-1-monopalmitate (GMP). The numbering refers to the GROMOS molecular topology. The charge group (CG) definitions are also indicated (all CGs are overall neutral). The methylene (11−24) and methyl (25) groups of the acyl chain are represented by united atoms with zero partial charge. The partial charges used for the headgroup atoms are reported in Table 1.

In the present study, the effect of the hydration level on the GL ↔ LC transition is investigated systematically, in the context of a saturated monoglyceride system. A monoglyceride is the molecule resulting from the esterification of a hydroxyl group of glycerol with a fatty acid. Monoglycerides present a number of key advantages, compared to, e.g., dipalmitoylphosphatidylcholine (DPPC), for a computational investigation of the GL ↔ LC phase transition: (*i*) the presence of only one aliphatic tail per headgroup, leading to a faster conformational and rotational relaxation; (*ii*) the limited role of long-range electrostatic interactions (uncharged, nonzwitterionic, and moderately polar headgroups), also leading to a faster relaxation and largely avoiding artifacts related to the approximate treatment of electrostatic interactions and the limited system size in simulations;[50−59] (*iii*) the absence of a ripple phase[44] as an intermediate state between the GL and the LC phases; and (*iv*) the availability of experimental structural and thermodynamic data.[60−69] More specifically, glycerol-1-monopalmitate (GMP; Figure 1) was chosen here as a model lipid, due to the availability of an experimental phase diagram.[69]

The phase diagram of the racemic GMP−water system[69] (as a function of hydration level and temperature, at pressure $P = 1$ atm) is reproduced in Figure 2. If a sample of GMP in the CR phase is heated in the presence of water at a concentration of about 30% w/w, the melting of the crystals and the formation of a LC phase is observed at about 55 °C. If the system is then cooled down relatively fast, a GL phase is formed at a temperature of about 50 °C. The GL phase is metastable and converts into a stable coagel (biphasic system consisting of CR and water phases), although this transformation only occurs over a period of a few hours (or even days). Note that this process is further complicated by the presence of polymorphism in the CR phase.

In this article, a total of 24 atomistic MD simulations (40−200 ns) of a racemic GMP bilayer patch performed at different temperatures and hydration levels, and starting from different initial conditions, are reported and compared. The main goal of this work is to investigate whether such simulations can account for (*i*) the GL → LC transition, the associated temperature T_m, and the increase of T_m upon decreasing the hydration level below about 20% w/w (Figure 2); (*ii*) the corresponding LC → GL transition, *a priori* more difficult to capture (disorder → order transition); and (*iii*) the main structural and dynamic properties of the GL and LC phases.

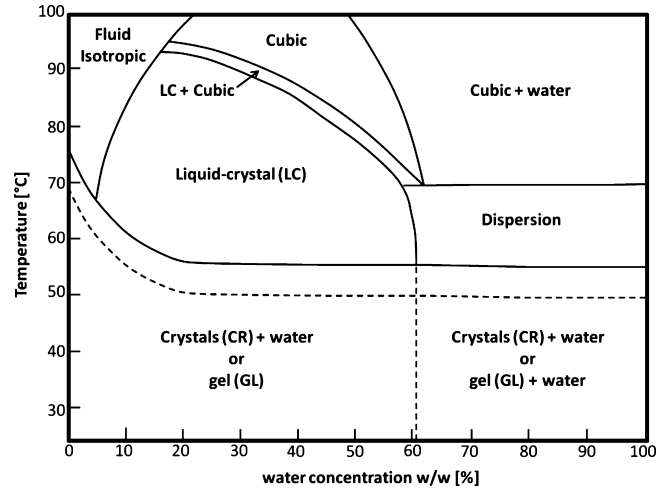


Figure 2. Phase diagram of the binary system glycerol-1-monopalmitate (GMP)−water. The horizontal axis represents the water concentration (in percent, weight-to-weight) and the vertical axis the temperature. Solid lines indicate equilibrium phase transitions, while the dashed line indicates the transitions involving the metastable GL phase (GL → LC and LC → GL). The phase diagram was adapted with permission from ref 69, copyright 1968, Elsevier.

II. Computational Details

II.1. Molecular Dynamics Simulations. All MD simulations were performed using the GROMOS96 program[70,71] together with a slightly modified version of the GROMOS 53A5 force field[72] for the lipids and the simple point charges (SPC) water model.[73] The only adjustment made to the original 53A5 force field concerns the partial charges of the headgroup atoms, which are reported in Table 1. These charges originated from an intermediate step in the parametrization of the new $53A6_{OXY}$ force field,[74] calibrated to reproduce experimental thermodynamic properties of oxygen compounds (alcohols, ethers, aldehydes, ketones, carboxylic acids, and esters). These charges were further slightly modified in an independent set of simulations (data not shown) to bring the equilibrium area per lipid as close as possible to the experimental values for both the GL and LC phases of GMP. The simulations were carried out under periodic boundary conditions based on rectangular boxes containing a hydrated GMP lipid bilayer patch of $2 \times 6 \times 6$ lipid molecules (in the *xy* plane, i.e., normal to the *z* axis). Newtons equations of motion were integrated using the

Table 1. Comparison between the Original 53A5 Charges for the Headgroup of GMP and the Modified Charges Used in the Present Study

charge group[a]	chemical function	atom label[a]	atom type[b]	IAC[b]	partical charges [*e*] original 53a5	partical charges [*e*] this work
1	alcohol	1	H	21	0.403	0.410
		2	OA	3	−0.611	−0.642
		3	CH2	15	0.208	0.232
2	alcohol	4	H	21	0.403	0.410
		5	OA	3	−0.611	−0.642
		6	CH1	14	0.208	0.232
3	ester	7	CH2	15	0.16	0.24
		8	OE	4	−0.36	−0.36
		9	C	12	0.58	0.66
		10	O	1	−0.38	−0.54

[a] Charge groups and atom labels are shown graphically in Figure 1. [b] Atom types and integer atom codes (IAC) refer to the 53A5 force field article.[72]

leapfrog scheme[75] with a time step of 2 fs. All bond lengths were constrained using the SHAKE procedure[76] with a relative geometric tolerance of 10^{-4}. The center of mass motion was removed every 100 ps. The simulations were performed in the *NPT* ensemble (constant number of particles N, pressure P, and temperature T) with a reference pressure $P = 1$ bar, and reference temperatures T ranging from 318 to 338 K. The temperature was maintained by weakly coupling the solute and solvent degrees of freedom separately to a temperature bath[77] at temperature T, with a relaxation time of 0.1 ps. The pressure was maintained by weakly coupling the particle coordinates and box dimensions in the xy plane and along the z axis separately[78] to a pressure bath at pressure P, with a relaxation time of 0.5 ps and an isothermal compressibility of 4.575×10^{-4} (kJ mol^{-1} nm^{-3})$^{-1}$ as appropriate for water.[70] The nonbonded interactions were computed using a twin-range scheme,[70,79] with short- and long-range cutoff distances set to 0.8 and 1.4 nm, respectively, and a frequency of five time steps for the update of the short-range pairlist and intermediate-range interactions. A reaction-field correction[80,81] was applied to account for the mean effect of electrostatic interactions beyond the long-range cutoff distance, using a relative dielectric permittivity of 61 as appropriate for the SPC water model.[82] Configurations were saved every 5 ps for analysis.

II.2. Simulated Systems. All simulations were carried out starting from two different initial (equilibrated) bilayer structures at full hydration (see below), one corresponding to a GL phase (I_{GL}) and the other one to a LC phase (I_{LC}). These structures were generated from exisiting snapshots of double-tailed DPPC phospholipids by reordering the atoms and deleting superfluous headgroup atoms, so as to map the two tail-to-tail DPPC molecules onto two tail-to-tail GMP molecules. The molecules were then placed in a box of lateral dimensions 0.6×0.6 nm and periodically replicated on a 6×6 lateral lattice to form a bilayer structure with 36 lipids per monolayer. Water molecules were added to the system by random insertion. The reference improper-dihedral controlling the chirality of the substituted glycerol moiety was set to opposite values for half of the lipid molecules in each leaflet, so as to generate a racemic bilayer (18 *R* and 18 *S* lipid molecules per leaflet). Water molecules initially inserted in the bilayer interior region were removed. Energy minimization and gentle equilibration provided the starting structure representative of the LC systems. The starting structure for the GL systems was obtained from a run of a LC system at a low temperature (308 K) during which the gel-like structure formed spontaneously. The run was continued at 318 K for 50 ns to further equilibrate this structure. Three different hydration levels were considered, which will be referred to as "full", "half", and "quarter" hydration. In the three cases, the computational box contained 72 lipid molecules along with 480 (full; 36.4% w/w), 240 (half; 18.2% w/w), or 120 (quarter; 9.1% w/w) water molecules. The simulations at full hydration were started directly from systems I_{GL} or I_{LC}. The initial coordinates of the systems at half hydration (I^{H}_{GL} and I^{H}_{LC}) were generated from configurations I_{GL} and I_{LC} *via* deletion of an appropriate number of water molecules (most remote from the bilayer midplane) and re-equilibration by 4 ns MD with positionally constrained lipid atoms (at constant normal pressure $P_z = 1$ bar and temperature $T = 338$ K). The same procedure was applied to generate the initial coordinates for the systems at quarter hydration (I^{Q}_{GL} and I^{Q}_{LC}) from those at half hydration (I^{H}_{GL} and I^{H}_{LC}).

Starting from the six initial configurations (I_{GL}, I_{LC}, I^{H}_{GL}, I^{H}_{LC}, I^{Q}_{GL}, and I^{Q}_{LC}), production simulations were preceded by energy minimization and a second MD equilibration (0.5 ns) to heat up the systems to the appropriate temperatures (318 K for the gel phase and 338 K for the LC phase) in the presence of progressively decreasing position restraints on the lipid atoms. From this point on, eight 40 ns production simulations were performed, at each of the three hydration levels: 3×6 simulations were started from the GL phase configurations and carried out at temperatures $T = 318$, 322, 326, 330, 334, or 338 K; 3×2 simulations were started from the LC phase configurations and carried out at temperatures $T = 318$ or 338 K. Each of the 24 production simulations will be referred to by a unique label characterizing the hydration level, initial bilayer structure, and reference simulation temperature. The labels, simulated systems, and simulation conditions are summarized in Table 2. Note that the average temperature in the simulated systems ($\langle T \rangle$ in Table 2) may slightly differ from the corresponding reference (thermostat) temperatures T. However, since the differences are of at most 1.8 K, they can be neglected, and the discussion will be performed in terms of T for simplicity.

II.3. Trajectory Analysis. The simulations were analyzed in terms of area per lipid a_{xy}, order parameters $S_{CH}(C_n)$ of the 14 methylene groups (C_n, with $n = 2-15$; atom numbers in Figure 1 minus nine), average number of *gauche* conformations per chain (n_G), average numbers (N) of intermolecular hydrogen bonds (H-bonds) between the different species, lipid lateral diffusion coefficients (D_{xy}), and lipid rotational (R) or wobbling (W) relaxation times τ and residual correlations c (τ_1^R, c_1^R, τ_2^R, c_2^R, τ_1^W, c_1^W, τ_2^W, c_2^W; distinguishing fast and slow decays by the indexes 1 and 2).

The area per lipid a_{xy} was calculated as the surface of the computational box in the xy plane divided by 36 and monitored as a function of time (indicator of a phase

Table 2. Simulated Systems and Simulation Conditions[a]

	hydration level					
label	hydration	c_{H_2O} [% w/w]	n_{H_2O}	starting configuration	T [K]	$\langle T \rangle$ [K]
$F_{GL}318$	full	36.4	480	GL	318	317.2
$F_{GL}322$	full	36.4	480	GL	322	321.1
$F_{GL}326$	full	36.4	480	GL	326	325.3
$F_{GL}330$	full	36.4	480	GL	330	329.3
$F_{GL}334$	full	36.4	480	GL	334	333.3
$F_{GL}338$	full	36.4	480	GL	338	337.3
$F_{LC}318$	full	36.4	480	LC	318	317.1
$F_{LC}338$	full	36.4	480	LC	338	337.3
$H_{GL}318$	half	18.2	240	GL	318	316.4
$H_{GL}322$	half	18.2	240	GL	322	320.4
$H_{GL}326$	half	18.2	240	GL	326	324.5
$H_{GL}330$	half	18.2	240	GL	330	328.6
$H_{GL}334$	half	18.2	240	GL	334	332.7
$H_{GL}338$	half	18.2	240	GL	338	336.6
$H_{LC}318$	half	18.2	240	LC	318	316.7
$H_{LC}338$	half	18.2	240	LC	338	336.6
$Q_{GL}318$	quarter	9.1	120	GL	318	316.3
$Q_{GL}322$	quarter	9.1	120	GL	322	320.3
$Q_{GL}326$	quarter	9.1	120	GL	326	324.3
$Q_{GL}330$	quarter	9.1	120	GL	330	328.2
$Q_{GL}334$	quarter	9.1	120	GL	334	332.2
$Q_{GL}338$	quarter	9.1	120	GL	338	336.3
$Q_{LC}318$	quarter	9.1	120	LC	318	316.4
$Q_{LC}338$	quarter	9.1	120	LC	338	336.3

[a] For each of the 24 simulations, the simulation label, hydration extent (relative to full hydration), water concentration (c_{H_2O}), number of water molecules in the computational box (n_{H_2O}), starting configuration (GL, gel; LC, liquid-crystal), reference (thermostat) temperature (T), and average (overall simulation) temperature ($\langle T \rangle$) are reported. In all cases, the computational box contains 2 × 6 × 6 lipid molecules, and the simulation is carried out for 40 ns at a reference pressure $P = 1$ bar.

transition) or averaged over the last 10 ns of the simulations (structural characteristic of the final states).

The order parameters S_{CH} of the hydrocarbon chains were calculated for the 14 methylene groups of the acyl chains of all GMP molecules, by computing the correlation functions describing the reorientation of the carbon−hydrogen vectors as described elsewhere[83] (see also ref 35). The values of S_{CH} were averaged over the 72 GMP molecules and monitored as a function of time (average S_{chn} over the 14 methylene groups as calculated from 1 ns time windows; indicator for a phase transition) or averaged over the last 10 ns of the simulations ($S_{CH}(C_n)$ or chain averaged value S_{chn}^{end}; structural characteristic of the final state).

The number n_G of *gauche* conformations per chain was calculated on the basis of the time-series of the 12 corresponding dihedral angles and averaged over the 72 lipid molecules and over time (last 10 ns). Because the distributions of the three staggered conformers in the hydrocarbon chains of lipids within bilayers are typically broad,[48,84] all dihedral angles in the ranges 0−120° (*gauche*+) and 240−360° (*gauche*−) were considered to be *gauche*.

The presence of a H bond was defined by a maximal hydrogen−oxygen distance of 0.25 nm and a minimal oxygen−hydrogen−oxygen angle of 135°. The occurring intermolecular H bonds were classified according to the different pairs of species present in the simulations (intralayer lipid−lipid N_{LL}^{intra}, interlayer lipid−lipid N_{LL}^{inter}, lipid−water N_{LW}, and water−water N_{WW}). Average numbers of H bonds were calculated considering the last 10 ns of the simulations.

The lipid lateral diffusion coefficients D_{xy} were calculated from the (average) mean-square displacement curves

$$\Delta(t) = \frac{1}{n_L}\sum_{i=1}^{n_L}\langle[\mathbf{r}_i(t+\tau) - \mathbf{r}_i(t)]^2\rangle_{\tau \le t_{av}-t}$$

where n_L is the number of lipid molecules (72), $\mathbf{r}_i$ stands for the coordinate vector of the center of geometry of lipid molecule i in the xy plane (two-dimensional vector; including displacements accross periodic boundaries), and the angular brackets indicate averaging over trajectory configurations considering all possible time origins τ (with the condition $\tau \le t_{av} - t$, where t_{av} is the duration of the averaging block). According to the Einstein model,[85] $\Delta(t)$ should be approximately linear, with a slope equal to $4D_{xy}$ (two-dimensional motion). In practice, deviations from linearity are observed at short times (cage effect) and at long times (statistical error). In the present study, the diffusion coefficients (along with error estimates) were evaluated as proposed by Essmann and Berkowitz.[86] The last 12 ns of each trajectory was divided into four 3 ns blocks, for which the diffusion coefficients were calculated by least-squares fit (LSF) to $\Delta(t)$ considering only the interval between 0.3 and 0.5 ns (where the function is close to linear). The D_{xy} value was obtained by averaging the four estimates, and the associated error was estimated by the corresponding standard deviation.

The rotational (R) and wobbling (W) motions of the lipid molecules were analyzed on the basis of the (average) autocorrelation function

$$C(t) = \frac{1}{n_L}\sum_{i=1}^{n_L}\langle \mathbf{v}_i(t+\tau)\cdot\mathbf{v}_i(\tau)\rangle_{\tau \le t_{av}-t}$$

where $\mathbf{v}_i$ is a unit vector characterizing the specific orientational property of lipid molecule i. The rotational motion of the lipids around their main axis was analyzed by considering the unit vector $\mathbf{v}_i$ along the C_9-O_{10} bond (Figure 1; carbonyl group). The wobbling motion was analyzed by considering the unit vector $\mathbf{v}_i$ along the C_9-C_{19} bond (Figure 1; along the aliphatic tail). The relaxation times (τ) and residual correlations (c) of these two modes were obtained by a LSF analysis of the logarithm of the corresponding $C(t)$ function. Both a fast and a slow decay were observed. The fast decay (τ_1^R, c_1^R, τ_1^W, c_1^W) was characterized by $C(t)$ in the interval t = 0−250 ps. The slow decay (τ_2^R, c_2^R, τ_2^W, c_2^W) was characterized by $C(t)$ in the interval t = 250 and 1500 ps. The analysis was performed by considering the last 20 ns of the simulations.

III. Results and Discussion

III.1. Characterization of the GL and LC Phases. Since the present study reports (to our knowledge) the first investigation of GMP bilayers by simulation, it is useful to first briefly discuss the main structural and dynamic properties of the GL and LC phases of this monoglyceride. To this purpose, the simulations $F_{GL}318$ and $F_{LC}338$ corresponding

Table 3. Structural Properties Averaged over the Last 10 ns of the Simulations[a]

label	phase (after 20 ns)	a_{xy} [nm^2]	S_{chn}^{end}	number of *gauche*/chain	H bonds			
					N_{LL}^{intra}	N_{LL}^{inter}	N_{LW}	N_{WW}
$F_{GL}318$	GL	0.248	0.276	1.7	143	0	156	1431
$F_{GL}322$	GL	0.249	0.277	1.8	142	0	155	1421
$F_{GL}326$	LC	0.286	0.210	2.9	120	0	183	1372
$F_{GL}330$	LC	0.291	0.204	3.0	120	0	182	1361
$F_{GL}334$	LC	0.294	0.199	3.0	117	0	185	1347
$F_{GL}338$	LC	0.293	0.195	3.1	117	0	182	1340
$F_{LC}318$	GL	0.251	0.261	1.5	141	0	159	1426
$F_{LC}338$	LC	0.296	0.194	3.1	118	0	181	1340
$H_{GL}318$	GL	0.240	0.278	1.4	148	0	147	640
$H_{GL}322$	GL	0.241	0.278	1.5	149	0	145	637
$H_{GL}326$	GL	0.243	0.276	1.6	144	0	150	626
$H_{GL}330$	GL	0.247	0.267	1.9	141	0	153	617
$H_{GL}334$	LC	0.297	0.191	3.0	118	3	181	571
$H_{GL}338$	LC	0.303	0.182	3.1	115	4	182	564
$H_{LC}318$	Int	0.275	0.229	2.7	126	2	178	599
$H_{LC}338$	LC	0.309	0.173	3.1	113	4	184	560
$Q_{GL}318$	GL	0.234	0.271	1.0	157	2	134	255
$Q_{GL}322$	GL	0.241	0.283	1.5	152	5	137	245
$Q_{GL}326$	GL	0.246	0.268	1.8	152	6	135	245
$Q_{GL}330$	GL	0.238	0.266	1.3	154	4	132	249
$Q_{GL}334$	GL	0.238	0.271	1.4	153	4	131	247
$Q_{GL}338$	LC	0.325	0.154	3.2	125	17	155	204
$Q_{LC}318$	Int	0.279	0.225	2.7	139	12	151	227
$Q_{LC}338$	LC	0.309	0.170	3.1	128	15	154	207

[a] The successive entries include the simulation label (Table 2), the phase after 20 ns (GL, gel; LC, liquid-crystal; Int, intermediate), the area per lipid a_{xy}, the chain-averaged order parameter S_{chn}^{end}, the average number of *gauche* conformations per chain n_G, the numbers of intralayer (N_{LL}^{intra}) and interlayer (N_{LL}^{inter}) lipid−lipid H bonds, and the numbers of lipid−water N_{LW} and water−water N_{WW} H bonds.

to systems at full hydration will be considered. These simulations were initiated from structures characteristic of the GL and LC phases, respectively, and carried out at temperatures where the considered phase is thermodynamically stable (experimentally, but also within the present force field, as will be discussed in sections III.2 and III.3); i.e., these represent equilibrium simulations for the two phases considered. The average values of key properties calculated over the last 10 ns of the trajectories are shown in Tables 3 (structural properties) and 4 (dynamic properties).

The average area per lipid a_{xy} in the GL phase ($F_{GL}318$), 0.248 nm^2, is slightly higher than the experimental estimate of 0.22 nm^2 reported by Pezron et al.[61] On the other hand, the corresponding average in the LC phase ($F_{LC}338$), 0.296 nm^2, is slightly lower than the corresponding experimental estimate[61] of 0.31 nm^2. However, the small differences observed between calculated and experimental values are well within the error associated with the experimental data. Although no error bar was provided for these specific measurements,[61] similar determinations considering the area per lipid of the LC phase of DPPC have led to a range of estimates as wide as[11] 0.48−0.73 nm^2 (the most commonly accepted value being[11] 0.64 nm^2). The present simulations are thus essentially in agreement with experimental results. As expected, the chain-averaged order parameters (S_{chn}^{end}) also indicate a higher extent of ordering of the aliphatic tails in the GL phase (0.276) compared to the LC phase (0.194). The average number n_G of *gauche* conformations per chain is also higher for the LC phase (3.1) compared to the GL phase (1.6). For the LC phase, the corresponding proportion of *trans* conformers (about 75%) also agrees well with available simulation and experimental data.[48]

Due to the thickness of the hydration layer in the systems at full hydration, no interlayer lipid−lipid H bonds were observed during these simulations ($N_{LL}^{inter} = 0$). Considering the available X-ray structures of monoglycerides in the CR phase,[6−8] both of the free glycerol hydroxyl groups are involved in two lipid−lipid H bonds, one interlayer and one intralayer (no cocrystallization water molecules were observed in these crystals). Assuming that the interlayer H bonds observed in the CR phase are replaced by lipid−water H bonds in the GL phase on a one-to-one basis (as suggested by Pezron et al.[61]), the calculated numbers $N_{LL}^{intra} = 143$ and $N_{LW} = 156$ for the GL phase (i.e., almost exactly two H bonds of either type per lipid molecule) are in excellent agreement with the crystallographic observations. The number of intralayer lipid−lipid H bonds N_{LL}^{intra} is about 20% lower for the LC phase compared to the GL phase. Such a decrease is expected to result from both the temperature increase (enhanced thermal fluctuations) and the higher inherent extent of disorder in the LC phase (suboptimal headgroup orientations for H-bonding). However, concomitantly, the number of lipid−water H bonds N_{LW} is about 20% higher in the LC phase. Such a (nearly quantitative) compensation effect between changes in the number of lipid−lipid and lipid-environment H bonds upon altering the phase or environment of a bilayer (i.e., the conservation of the headgroup H-bond saturation) has been observed previously in different contexts.[87−90]

As expected, the LC phase also evidences a more pronounced fluidity compared to the GL phase. The calculated lipid lateral diffusion coefficient D_{xy} is about twice as high for the LC phase (9.4×10^{-11} m^2•s^{-1}) compared to the GL phase (4.6×10^{-11} m^2•s^{-1}). These numbers should probably be viewed as upper bounds, due to finite-size effects in the microscopic simulated systems.[55] To our knowledge, the corresponding experimental values are not known for GMP. A comparison with the experimental[91] and simulation[55] values of 1.25 and 0.95×10^{-11} m^2•s^{-1}, respectively, for DPPC indicates that the lateral diffusion of GMP molecules (single acyl chain, low polarity headgroup) is much faster than that of DPPC (two acyl chains, zwitterionic headgroup). The rotational and wobbling motions of the lipids also take place on different time scales in the two phases. In the GL phase, the rotational motion appears to involve two distinct regimes (i.e., the corresponding autocorrelation function follows a biexponential form). The first regime involves a fast decay ($\tau_1^R = 0.27$ ns), leading to a residual correlation $c_1^R = 0.12$. After a second slower decay ($\tau_2^R = 2.6$ ns), the correlation essentially reaches zero ($c_2^R = 0.03$). In contrast, for the LC phase, a single fast decay is observed ($\tau_1^R = 0.19$ ns), leading to an essentially vanishing residual correlation ($c_1^R = 0.05$). The fast motion (time scale ~0.2 ns) can be attributed to the rotation of a single lipid molecule around its axis. The slow motion (time scale ~ 2.5 ns) observed in the GL phase can be attributed to

Table 4. Dynamic Properties Calculated over the Last 10 ns (12 ns for D_{xy}) of the Simulations[a]

		rotational motion					wobbling motion			
simulation code	phase (after 20 ns)	D_{xy} [10^{-11} m^2 s^{-1}]	τ_1^R [10^2 ps]	c_1^R	τ_2^R [10^3 ps]	c_2^R	τ_1^W [10^3 ps]	c_1^W	τ_2^W [10^4 ps]	c_2^W
$F_{GL}318$	GL	4.6 ± 2.0	2.7	0.12	2.6	0.03	19.0	0.96	7.7	0.93
$F_{GL}322$	GL	5.9 ± 3.4	2.5	0.10	4.0	0.04	15.6	0.95	5.7	0.91
$F_{GL}326$	LC	8.8 ± 2.0	2.1	0.07	9.7	0.04	3.1	0.83	2.6	0.74
$F_{GL}330$	LC	10.5 ± 1.8	1.9	0.06	12.6	0.03	2.9	0.82	2.7	0.73
$F_{GL}334$	LC	10.7 ± 1.3	1.9	0.05	4.1	0.04	2.7	0.80	3.2	0.73
$F_{GL}338$	LC	13.4 ± 1.6	1.9	0.05	8.5	0.04	2.6	0.80	3.5	0.73
$F_{LC}318$	GL	2.7 ± 1.2	2.7	0.12	2.7	0.03	25.4	0.97	18.5	0.95
$F_{LC}338$	LC	9.4 ± 1.2	1.8	0.05	13.0	0.03	2.7	0.80	4.4	0.74
$H_{GL}318$	GL	3.0 ± 1.5	2.6	0.11	3.2	0.04	28.4	0.97	15.2	0.96
$H_{GL}322$	GL	2.1 ± 0.4	2.3	0.09	2.9	0.03	21.5	0.96	9.5	0.94
$H_{GL}326$	GL	4.3 ± 2.5	2.2	0.08	3.1	0.02	16.6	0.96	8.7	0.93
$H_{GL}330$	GL	4.1 ± 2.1	2.1	0.07	5.4	0.04	15.0	0.95	14.2	0.93
$H_{GL}334$	LC	6.9 ± 1.5	2.0	0.06	5.8	0.03	2.8	0.82	2.1	0.72
$H_{GL}338$	LC	7.8 ± 1.3	1.9	0.06	187.1	0.04	2.3	0.78	2.5	0.69
$H_{LC}318$	Int	4.8 ± 0.7	2.6	0.11	3.2	0.03	3.9	0.86	2.2	0.76
$H_{LC}338$	LC	8.4 ± 1.4	1.8	0.05	36.7	0.04	2.2	0.78	2.1	0.68
$Q_{GL}318$	GL	1.6 ± 0.1	2.8	0.13	2.1	0.03	73.2	0.99	394.4	0.98
$Q_{GL}322$	GL	1.8 ± 0.3	2.8	0.13	2.4	0.04	21.6	0.96	8.2	0.93
$Q_{GL}326$	GL	1.5 ± 0.3	2.5	0.11	1.9	0.03	16.8	0.96	11.8	0.93
$Q_{GL}330$	GL	1.4 ± 0.4	2.4	0.10	1.6	0.01	36.0	0.97	18.4	0.96
$Q_{GL}334$	GL	1.5 ± 0.3	2.0	0.06	3.2	0.02	60.9	0.98	831.9	0.98
$Q_{GL}338$	LC	5.6 ± 0.3	2.2	0.08	9.5	0.03	2.6	0.80	1.5	0.67
$Q_{LC}318$	Int	2.9 ± 0.1	3.3	0.17	1.9	0.04	4.3	0.87	2.4	0.78
$Q_{LC}338$	LC	5.6 ± 0.9	2.3	0.08	11.5	0.05	2.4	0.79	1.8	0.68

[a] The successive entries include the simulation label (Table 2), the phase after 20 ns (GL, gel; LC, liquid-crystal; Int, intermediate), the lipid lateral diffusion coefficient D_{xy}, the first (τ_1^R) and second (τ_2^R) relaxation times corresponding to the two observed decays in the rotational motion with corresponding residual correlations (c_1^R and c_2^R), and the first (τ_1^W) and second (τ_2^W) relaxation times corresponding to the two observed decays in the wobbling motion with corresponding residual correlations (c_1^W and c_2^W).

correlated motions within a bilayer leaflet, such as the reorientation (precession) of the lipid tilting angle within the leaflet. The wobbling motion exhibits two decay regimes with high residual correlations, as expected for a vector that does not sample the accessible space isotropically (i.e., the aliphatic tails always point preferentially from the headgroup plane toward the membrane midplane). For the GL phase, the two decays are slow ($\tau_1^W = 19$ ns, $\tau_2^W = 77$ ns) and result in a very high residual correlation ($c_2^W = 0.96$). For the LC phase, the relaxation times for both decays are shorter ($\tau_1^W = 2.7$ ns, $\tau_2^W = 44$ ns), and the final residual correlation is comparatively smaller ($c_2^W = 0.74$), again indicative of a more limited tail ordering in this phase.

III.2. GL ↔ LC Transition: Reversibility. The reversibility of the GL ↔ LC phase transition (at the three levels of hydration) will be discussed considering the simulations initiated from structures appropriate for the GL or LC phases and performed at two different temperatures, 318 and 338 K (i.e., corresponding, respectively, to 10 K below and 10 K above the experimental $T_m = 328$ K at full hydration;[69] Figure 2). The presence of a transition can be assessed by monitoring the time series of the area per lipid a_{xy} and of the chain-averaged order parameter S_{chn}. These time series are displayed in Figure 3 for the simulations at full hydration, half hydration, and quarter hydration. The average values of key properties calculated over the last 10 ns of the trajectories are also reported in Tables 3 (structural properties) and 4 (dynamic properties).

The time evolutions of a_{xy} and of S_{chn} are clearly anticorrelated, both in terms of net changes (for systems undergoing a GL ↔ LC phase transition) and in terms of fluctuations (within time periods corresponding to a single phase). In other words, an increase (decrease) in the area per lipid always occurs concomitantly with a decrease (increase) in the ordering of the lipid tails. All simulations initiated from structures appropriate for the GL or LC phase and carried out within the temperature range in which the corresponding phase is expected to be stable (318 K for GL and 338 K for LC) do not evidence transitions, which justifies the equilibrium analysis of section III.1. In contrast, complete or incomplete transitions are observed in the simulations initiated from structures appropriate for the GL or LC phase and carried out at a temperature where the alternative phase is expected to be more stable (GL → LC at 338 K and LC → GL at 318 K; for simulations $H_{LC}318$ and $Q_{LC}318$, the transitions are incomplete on the 40 ns time scale, see below).

For the systems at full hydration (Figure 3a), the GL phase is stable at 318 K ($F_{GL}318$) with very limited fluctuations of a_{xy} and S_{chn}. The corresponding average values (last 10 ns) are 0.248 nm^2 and 0.276, respectively (Table 3). The LC phase is stable at 338 K ($F_{LC}338$), but a_{xy} and S_{chn} undergo larger fluctuations due to the enhanced thermal fluctuations. The corresponding average values (last 10 ns) are 0.296 nm^2 and 0.194. The systems $F_{GL}338$ and $F_{LC}318$ are not at equilibrium in their starting phase and undergo a transition, as expected. The simulation initiated with a structure appropriate for the GL phase and carried out above the phase transition temperature ($F_{GL}338$) evidences changes in a_{xy} from about 0.25 to about 0.30 nm^2 and in S_{chn} from about 0.27 to about 0.20 within the first 10 ns. Conversely, the simulation initiated with a structure appropriate for the LC phase and carried out below the phase transition temperature ($F_{LC}318$)

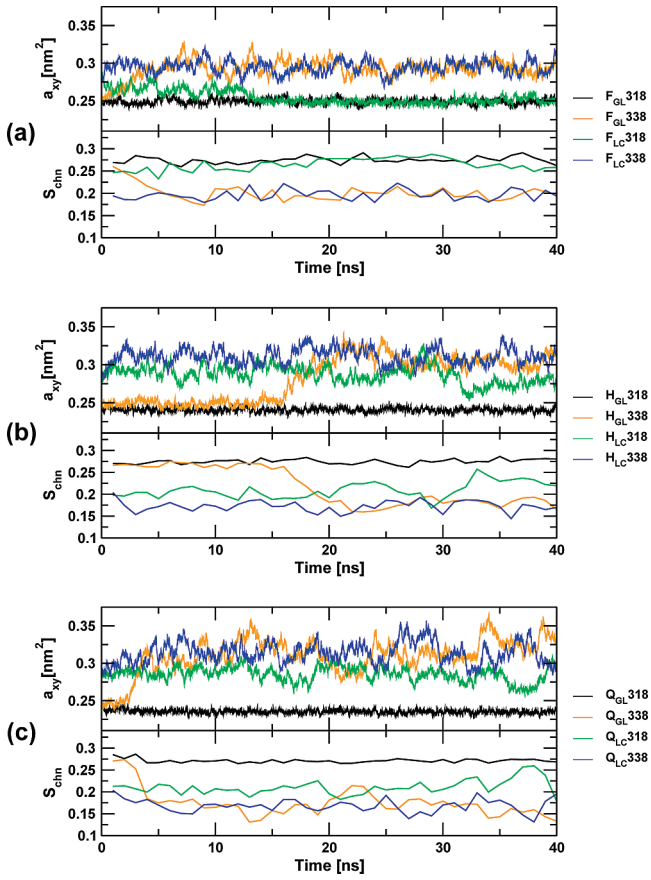


Figure 3. Time series of the area per lipid a_{xy} and of the chain-averaged order parameters S_{chn} for the simulations initiated from a structure appropriate for the GL or the LC phase and carried out at temperatures of 318 or 338 K. The three panels correspond to different hydration levels: (a) full hydration, (b) half hydration, and (c) quarter hydration. See Table 2 for the simulation labels.

evidences changes in a_{xy} from about 0.30 to about 0.25 nm^2 and in S_{chn} from about 0.20 to about 0.27 within the first 15 ns.

For the systems at half hydration (Figure 3b), the GL phase is stable at 318 K ($H_{GL}318$), and the LC phase is stable at 338 K ($H_{LC}338$). These two simulations are very similar to the corresponding simulations at full hydration ($F_{GL}318$ and $F_{LC}338$) in terms of all simulated properties (Figure 3, Tables 3 and 4). The simulation initiated from a structure appropriate for the GL phase and carried out above the phase transition temperature ($H_{GL}338$) evidences changes in a_{xy} from about 0.25 to about 0.30 nm^2 and in S_{chn} from about 0.27 to about 0.20 in the interval 15−20 ns. In contrast to the situation at full hydration, however, the simulation initiated from a structure appropriate for the LC phase and carried out below the phase transition temperature ($H_{LC}318$) evidences a progressive change in a_{xy} and S_{chn} in the direction of the GL phase values but does not reach converged values within the simulation time scale of 40 ns. As a result, at the end of the simulation, an intermediate state (the properties of which lie between those of the GL and LC phases) is reached (Tables 3 and 4). To verify that the incompleteness of the transition is due to kinetic (and not thermodynamic) factors, simulation $H_{LC}318$ was further extended to 200 ns (data not shown). The transition was found to occur after about 94 ns.

For the systems at quarter hydration (Figure 3c), the GL phase is stable at 318 K ($Q_{GL}318$), and the LC phase is stable at 338 K ($Q_{LC}338$). The transition from the GL to the LC phase ($Q_{GL}338$) also occurs, this time at the very beginning of the simulation. However, as for the corresponding simulation at half hydration, the simulation from the LC to the GL phase ($Q_{LC}318$) does not entirely converge within 40 ns and reaches an intermediate state. Here also, simulation $Q_{LC}318$ was extended to 200 ns (data not shown), and the LC → GL transition was found to occur after about 144 ns.

In summary, simulations on the 40 ns time scale are able to capture the reversible interconversion between the two phases and the correct transition temperature under conditions of full hydration. At lower hydration, the GL → LC transition can be appropriately simulated on this time scale, but the LC → GL transition appears to require significantly longer simulations (by about a factor five). The dependence of the transition temperature on the hydration level will be discussed in section III.3.

The present simulations suggest that the GL → LC transition tends to occur on a shorter time scale compared to the LC → GL transition. This difference is not visible at full hydration but is clearly observed for the systems at lower hydration. Two effects may be invoked to explain this difference. First, the LC phase is stable at higher temperatures compared to the GL phase. Thus, in order to promote a GL → LC transition, a system with a GL structure is heated up to a temperature above the phase transition temperature. In contrast, to promote a LC → GL transition, a system with a LC structure is cooled down to a temperature below the phase transition temperature. The second process is likely to be slower, as a direct consequence of the lower (average) atomic velocities in the simulation. Second, the GL phase is a more ordered phase compared to the LC phase. As a result, the LC → GL transition is expected to involve a high entropic barrier; i.e., the lipid molecules must sample a large extent of conformational space before reaching the regular arrangement characteristic of the GL phase. This process is also slowed down by the reduction of the area per lipid (packing constraints) and by the concomitant decrease in the lipid translational and rotational diffusion. (Table 4). In contrast, the GL → LC transition is expected to involve a comparatively much lower entropic barrier, due to the entropy increase of the overall process. The observation that the time scale associated with the LC → GL transition tends to increase when the hydration level is decreased is certainly related to the associated increase in the viscosity of the system. For example, the lipid lateral diffusion coefficient systematically decreases upon reducing the water content of the system, for both the GL and LC phases (Table 4).

Note, finally, that even for the simulations where a complete GL → LC ($F_{GL}338$, $H_{GL}338$, $Q_{GL}338$) or LC → GL ($F_{LC}318$) transition is observed, some properties might not be entirely converged after the transition given the 40 ns time scale of the simulations, as evidenced by comparison with the corresponding equilibrium simulations ($F_{LC}338$, $H_{LC}338$, $Q_{LC}338$, and $F_{GL}318$), see Tables 3 and 4.

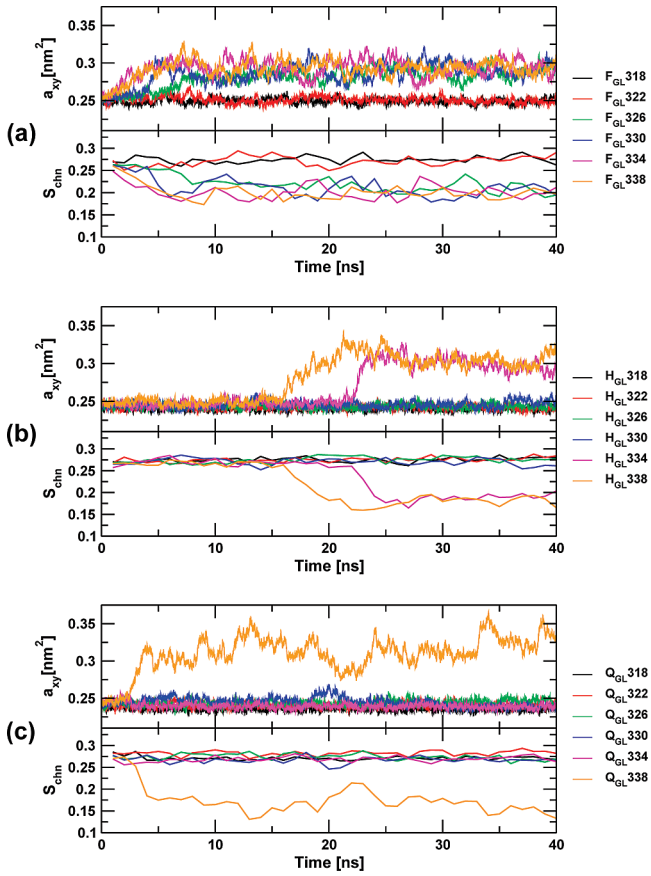


Figure 4. Time series of the area per lipid a_{xy} and of the chain-averaged order parameters S_{chn} for the simulations initiated from a structure appropriate for the GL phase and carried out at temperatures of 318, 322, 326, 330, 334, or 338 K. The three panels correspond to different hydration levels: (a) full hydration, (b) half hydration, and (c) quarter hydration. See Table 2 for the simulation labels.

III.3. GL → LC Transition: Temperature vs Hydration. Considering the above results, the determination of the lowest temperature for which a GL → LC transition is observed on the 40 ns time scale appears to provide a reliable method for the evaluation of the phase transition temperature T_m. Of course, a bracketing using the highest temperature for which a LC → GL transition is observed would be desirable but remains at present computationally too demanding (requiring simulations longer by at least a factor five).

The dependence of the transition temperature T_m on the level of hydration will thus be investigated considering the 18 simulations initiated from a structure appropriate for the GL phase at the three different hydration levels and carried out at distinct temperatures with a 4 K interval in the range 318−338 K. The corresponding time series of the area per lipid a_{xy} and of the chain-averaged order parameter S_{chn} are displayed in Figure 4.

At full hydration (Figure 4a), the simulations $F_{GL}318$ and $F_{GL}322$ remain stable in the GL phase during the entire trajectory, while the simulations $F_{GL}326$−$F_{GL}338$ evidence a transition to the LC phase within the first 10 ns. After 10 ns, the two sets of curves are clearly separated, suggesting that they both correspond to systems having reached their equilibrium phase at the given temperature. At half hydration (Figure 4b), the simulations $H_{GL}318$−$H_{GL}330$ remain stable in the GL phase during the entire trajectory, while the simulations $H_{GL}334$ and $H_{GL}338$ evidence a transition to the LC phase after about 23 and 20 ns, respectively. Here also, the two sets of curves are well separated after the first 25 ns. Finally, at quarter hydration (Figure 4c), the simulations $Q_{GL}318$−$Q_{GL}334$ remain stable in the GL phase and the only simulation evidencing a transition to the LC phase after about 4 ns is $Q_{GL}338$. Here again, the latter curve remains well separated from the five others after the first 5 ns.

The order parameters $S_{CH}(C_n)$ of the 14 methylene groups (C_n, $n = 2$−15) within the aliphatic tails calculated over the last 10 ns of the trajectories are shown in Figure 5. The average values of key properties calculated over the same interval are reported in Tables 3 (structural properties) and 4 (dynamic properties).

Overall, on the basis of these final properties, the simulations can unambiguously be classified into two distinct groups, depending on whether the GL → LC transition has occurred or not during the simulation. Residual differences between the different simulations within the two groups are due to the slightly different simulation temperatures, to the limited averaging time, and in the cases where a GL → LC transition has occurred, to insufficient equilibration after the transition (as mentioned in section III.2). These residual differences are least pronounced for simulations $F_{GL}318$ and $F_{GL}322$ (full hydration, no transition), except for a slightly higher diffusivity (lateral, rotational, wobbling) of the lipids in the latter case. They are slightly more important for simulations $F_{GL}326$−$F_{GL}338$ (full hydration, GL → LC transition), where the order parameters systematically decrease and the diffusivity of the lipids slightly increases upon increasing the temperature. Simulations $H_{GL}318$−$H_{GL}330$ (half hydration, no transition) show limited differences, except for noticeably lower order parameters at 330 K and a slight systematic increase in the diffusivity of the lipids upon increasing the temperature. Simulations $H_{GL}334$ and $H_{GL}338$ (half hydration, GL → LC transition) show similar trends to those of the corresponding simulations at full hydration ($F_{GL}326$−$F_{GL}338$). Simulations $Q_{GL}318$−$Q_{GL}334$ (quarter hydration, no transition) show similar trends to those of the corresponding simulations at half hydration ($H_{GL}318$−$H_{GL}330$). Finally, simulation $Q_{GL}338$ is the only one undergoing a phase transition at this hydration level.

Comparing the GL and the LC phases (see also section III.1), the former is characterized by significantly higher order parameters and a significantly lower diffusivity (lateral, rotational, wobbling) of the lipids. In general, a decrease in the hydration level tends to promote (*i*) a slight decrease in the order parameters for the GL phase, (*ii*) a somewhat more pronounced decrease in the order parameters for the LC phase, (*iii*) a slight decrease in the number of *gauche* conformations per chain for the GL phase, (*iv*) a significant decrease in the diffusivity (lateral, rotational, wobbling) of the lipids. By decreasing the hydration level, the thickness of the water layer between successive (periodic) bilayers decreases, and the H-bonded interaction of one leaflet with a periodic image of the other leaflet becomes possible. As a

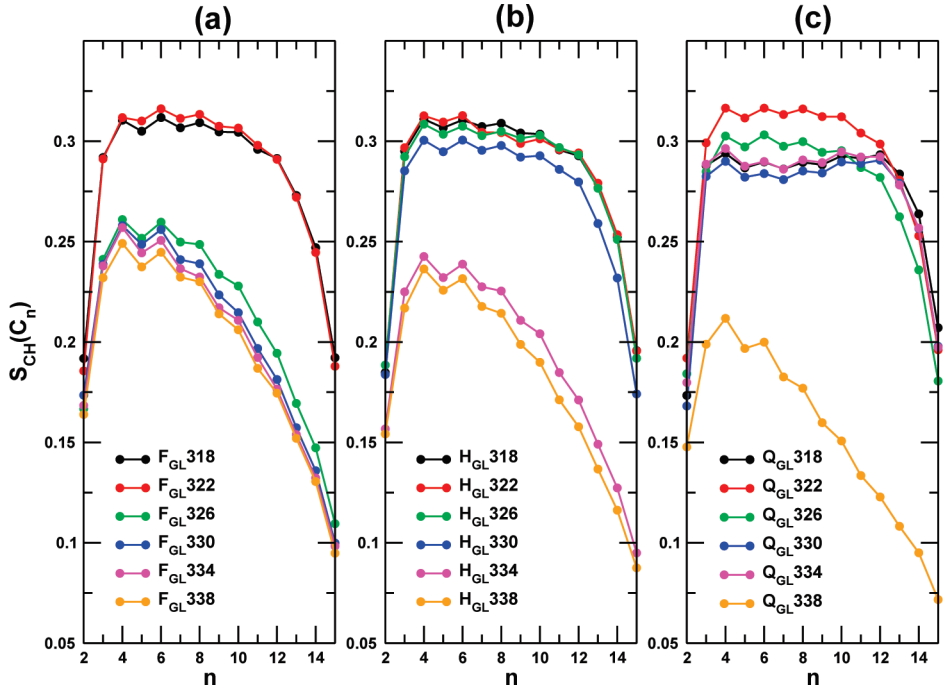


Figure 5. Order parameters $S_{CH}(C_n)$ of the 14 methylene groups (C_n, $n = 2-15$) calculated considering the last 10 ns of the simulations initiated from a structure appropriate for the GL phase and carried out at temperatures of 318, 322, 326, 330, 334, or 338 K. The three panels correspond to different hydration levels: (a) full hydration, (b) half hydration, and (c) quarter hydration. See Table 2 for the simulation labels.

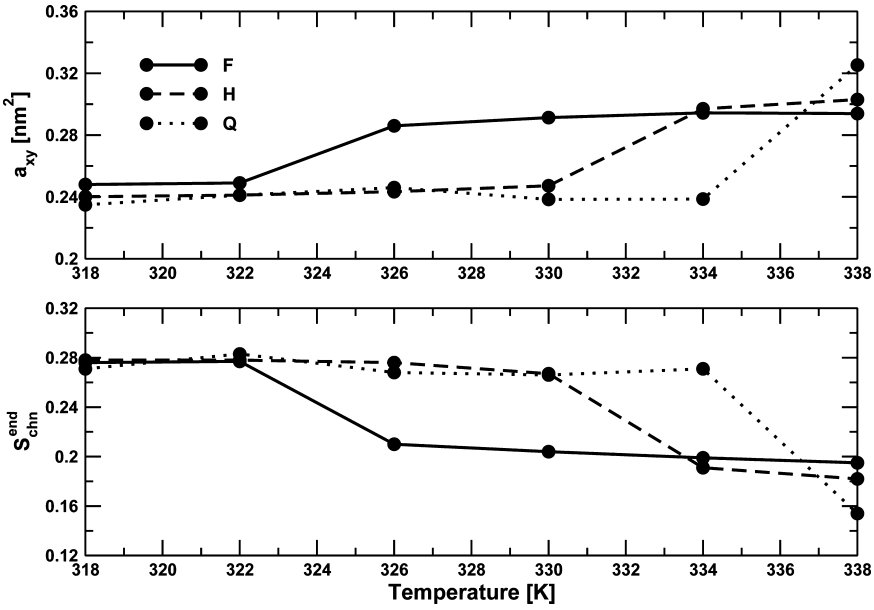


Figure 6. Average area per lipid a_{xy} (top panel) and chain-averaged order parameter S^{end}_{chn} (bottom panel) calculated considering the last 10 ns of the simulations initiated from a structure appropriate for the GL phase and carried out at different temperatures. The different linestyles correspond to the three considered hydration levels: full (F), half (H), and quarter (Q) hydration.

consequence, interlayer H bonds (N^{inter}_{LL}) are observed at half and, more markedly, quarter hydration. The presence of these H bonds represents an artifact of the simulation setup when taking the single-bilayer experiment as a reference. However, they are expected to occur in multilamellar systems. Finally, the H-bond compensation effect discussed in section III.1 (conservation of $N^{intra}_{LL} + N^{inter}_{LL} + N_{LW}$) is observed for all of the simulations evidencing a GL → LC transition.

On the basis of the above observations, the final values of the area per lipid a_{xy} and of the chain-averaged order parameters S^{end}_{chn} in the simulations may serve as an indicator for the most stable phase at a given temperature T. The corresponding results are displayed graphically in Figure 6. They suggest ranges for the transition temperature T_m of 322−326, 330−334, and 334−338 K in the systems at full, half, and quarter hydration, respectively. Considering the corresponding hydration levels (30.5, 18.2, and 9.1% w/w, respectively) these estimated transition temperatures (midpoints at 51, 59, and 63 °C, respectively) agree qualitatively very well with the experimental phase diagram[69] of GMP

(Figure 2; experimental T_m of about 50, 53, and 58 °C), and even essentially quantitatively given an estimated error of 2 °C on the simulation values and of 4 °C on the experimental ones (considering that the independent studies of refs 69 and 92 report T_m at full hydration of 50 and 54, respectively).

IV. Conclusion

The goal of the present study was to investigate whether atomistic MD simulations on the 10−100 ns time scale could be used to monitor GL ↔ LC transitions in lipid bilayers, and to evaluate the corresponding transition temperatures T_m (as well as environmental influences on T_m) in a reliable fashion. This time scale is possibly still too short in the context of lipids such as DPPC (two acyl tails, zwitterionic headgroup), in view of the slow translational and rotational relaxation of the lipid molecules in this case. For this reason, the simpler monoglyceride lipid GMP (one acyl tail, lower polarity headgroup) was considered instead in the present work. To our knowledge, this study represents the first investigation of GMP bilayers by simulation.

The main results of this work can be summarized as follows. First, simulations of 40 ns are able to account for the GL → LC transition and for the presence of a threshold temperature associated with this transition; i.e., they can provide a lower bound for T_m. Thus, by performing a set of simulations initiated from a structure appropriate for the GL phase and carried out at slightly different temperatures, it is possible to evaluate T_m accurately (here, within an interval of 4 °C). Simulations on longer time scales (about 200 ns) appear to be required for the inverse LC → GL transition, especially at water contents lower than full hydration. Although such simulations were not performed here in a systematic fashion, they could in principle provide another threshold temperature, i.e., an upper bound for T_m. The observation that the LC → GL transition is slower than the GL → LC transition is reasonable, considering that the former transition is a disorder → order transition (high entropy barrier). Second, an investigation of the influence of the hydration level (full, half, or quarter hydration) on the calculated T_m value could evidence the experimentally expected increase in this transition temperature upon decreasing the bilayer hydration. Another interesting observation is that in general, a decrease in the hydration level tends to promote a decrease in the order parameters as well as in the diffusivity (lateral, rotational, wobbling) of the lipids for both the GL and LC phases (the order parameters are themselves higher for the GL phase while the diffusivity is higher for the LC phase). Third, the T_m values calculated at the three hydration levels considered are in remarkable (qualitative and even essentially quantitative) agreement with the experimental phase diagram of GMP.[69] Qualitative agreement indicates that the model underlying the simulations (force field) is able to capture the essential physics of the system. Quantitative agreement, although pleasant, is probably largely fortuitous considering the approximations involved in this model and the empirical nature of its parameters.

To our knowledge, this study represents the first accurate determination of the T_m of a lipid *via* atomistic simulations of the (reversible) GL ↔ LC phase transition, as well as the first direct simulation evidence for the increase in the transition temperature upon dehydration. The direct determination of T_m by simulation is of high relevance in the context of force field development. The reason is that T_m typically represents one of the most accurately known experimental properties of the lipid−water system for a given lipid type. It would usefully complement the few other quantities available for the parametrization of a force field (transverse electron density profile and area per lipid, the latter being sometimes affected by a considerable experimental uncertainty). The investigation of the influence of hydration on the T_m value is also of great interest in the context of dehydration-induced damages in living cells and bioprotection by cosolutes.[19−26]

Important extensions of the present approach include (*i*) investigation of the feasibility of monitoring GL ↔ LC transitions in more complex (and biologically relevant) lipids such as DPPC (involving simulations on significantly longer time scales) and (*ii*) investigation of the influence of other environmental factors (besides hydration) on T_m (e.g., concentration of biologically relevant cosolutes[90] such as electrolytes, alcohols, anesthetics, or sugars). A particularly interesting example of a cosolute is the bioprotective sugar trehalose, which is known to depress the T_m of lipid bilayers at low hydration, thereby counteracting the potentially damaging T_m increase observed above.[87] Further extensions of this work along the two above lines are in progress.

Acknowledgment. Financial support from the Swiss National Science Fundation (Grant 21-121895) is gratefully acknowledged.

References

(1) Seddon, J. M.; Cevc, G. Lipid polymorphism: Structure and stability of lyotropic mesophases of phospholipids. In *Phospholipids Handbook*; Cevc, G., Ed.; Marcel Dekker, Inc.: New York, 1993; pp 403−454.

(2) Foubert, I.; Dewettinck, K.; Van de Walle, D.; Dijkstra, A. J.; Quinn, P. J. Physical Properties: Structural and Physical Characteristics. In *The Lipid Handbook*, 3rd ed.; Gunstone, F. D., Harwood, J. L., Dijkstra, A. J., Eds.;CRC Press, Div. of Taylor, Francis Group: Boca Raton, FL, 2007; pp 471−534.

(3) Jensen, L. H.; Mabis, A. J. *Acta Crystallogr.* **1966**, *21*, 770–781.

(4) Pascher, I.; Lundmark, M.; Nyholm, P.-G.; Sundell, S. *Biochim. Biophys. Acta* **1992**, *1113*, 339–373.

(5) Pascher, I. *Curr. Opin. Struct. Biol.* **1996**, *6*, 439–448.

(6) Larsson, K. *Acta Crystallogr.* **1966**, *21*, 267–272.

(7) Goto, M.; Kozo, K. *Bull. Chem. Soc. Jpn.* **1988**, *61*, 1434–1436.

(8) Goto, M.; Takiguchi, T. *Bull. Chem. Soc. Jpn.* **1985**, *58*, 1319–1320.

(9) Larsson, K. *Ark. Kemi* **1964**, *23*, 35–56.

(10) Dorset, D. L.; Pangborn, W. A.; Hancock, A. J.; van Soest, T. C. *Z. Naturforsch.* **1978**, *33c*, 50–55.

(11) Nagle, J. F.; Tristram-Nagle, S. *Biochim. Biophys. Acta* **2000**, *1469*, 159–195.

(12) Risselada, H. J.; Marrink, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 17367–17372.

(13) Marrink, S. J.; Risselada, J.; Mark, A. E. *Chem. Phys. Lipids* **2005**, *135*, 223–244.

(14) Brown, D. A.; London, E. *Annu. Rev. Cell Dev. Biol.* **1998**, *14*, 111–136.

(15) Simons, K. *Nature* **1997**, *387*, 569–572.

(16) Simons, K.; Toomre, D. *Nat. Rev. Mol. Cell. Biol.* **2000**, *1*, 31–40.

(17) Herreros, J.; Ng, T.; Schiavo, G. *Mol. Biol. Cell* **2001**, *12*, 2947–2960.

(18) Keilin, D. *Proc. R. Soc. London, Sect. B* **1959**, *150*, 149–191.

(19) Clegg, J. S. *Comp. Biochem. Physiol.* **2001**, *128B*, 613–624.

(20) Feovilova, E. P. *Appl. Biochem. Microbiol.* **2003**, *39*, 1–18.

(21) Zhmakin, A. I. *Physics—Uspekhi* **2008**, *51*, 231–252.

(22) Crowe, J. H.; Hoekstra, F. A.; Crowe, L. M. *Annu. Rev. Physiol.* **1992**, *54*, 579–599.

(23) Guppy, M.; Withers, P. *Biol. Rev.* **1999**, *74*, 1–40.

(24) Crowe, J. H.; Crowe, L. M.; Wolkers, W. F.; Oliver, A. E.; Ma, X.; Auh, J.-H.; Tang, M.; Zhu, S.; Norris, J.; Tablin, F. *Integr. Comp. Biol.* **2005**, *45*, 810–820.

(25) Hengherr, S.; Heyer, A. G.; Koehler, H. *FEBS J.* **2008**, *275*, 281–288.

(26) Crowe, L. M. *Comp. Biochem. Physiol.* **2002**, *131A*, 505–513.

(27) Treanor, R. L.; Weiss, R. G. *J. Am. Chem. Soc.* **1988**, *110*, 2170–2177.

(28) Ariga, K; Nakanishi, T.; Kawanami, S.-I.; Kosaka, T.; Kikuchi, *J. Nanosci. Nanotechnol.* **2006**, *6*, 1718–1730.

(29) Horváth, R.; Fricsovszky, G.; Papp, E. *Biosens. Bioelectron.* **2003**, *18*, 415–428.

(30) Liu, J.; Conboy, J. C. *J. Am. Chem. Soc.* **2004**, *126*, 8894–8895.

(31) Alaouie, A. M.; Smirnov, A. I. *J. Magn. Reson.* **2006**, *182*, 229–238.

(32) Jähnig, F. *Biophys. J.* **1996**, *71*, 1348–1349.

(33) Roux, B. *Biophys. J.* **1996**, *71*, 1346–1347.

(34) Tieleman, D. P.; Berendsen, H. J. C. *J. Chem. Phys.* **1996**, *105*, 4871–4880.

(35) Tieleman, D. P.; Marrink, S. J.; Berendsen, H. J. C. *Biochim. Biophys. Acta* **1997**, *1331*, 235–270.

(36) Cantor, R. S. *Biophys. J.* **1999**, *76*, 2625–2639.

(37) Feller, S. E.; Pastor, R. W. *J. Chem. Phys.* **1999**, *111*, 1281–1287.

(38) Feller, S. E. *Curr. Opin. Colloid Interface Sci.* **2000**, *5*, 217–223.

(39) Chandrasekhar, I.; Bakowies, D.; Glättli, A.; Hünenberger, P. H.; Pereira, C.; van Gunsteren, W. F. *Mol. Simul.* **2005**, *31*, 543–548.

(40) Heller, H.; Schaefer, M.; Schulten, K. *J. Phys. Chem.* **1993**, *97*, 8343–8360.

(41) Essmann, U.; Perera, L.; Berkowitz, M. L. *Langmuir* **1995**, *11*, 4519–4531.

(42) Venable, R. M.; Brooks, B. R.; Pastor, R. W. *J. Chem. Phys.* **2000**, *112*, 4822–4832.

(43) Tobias, D. J.; Tu, K.; Klein, M. L. *Curr. Opin. Colloid Interface Sci.* **1997**, *2*, 15–26.

(44) de Vries, A. H.; Yefimov, S.; Mark, A. E.; Marrink, S. J. *Proc. Natl. Acad. Sci.* **2005**, *102*, 5392–5396.

(45) Marrink, S. J.; Tieleman, D. P. *J. Am. Chem. Soc.* **2001**, *123*, 12383–12391.

(46) Fuhrmans, M.; Knecht, V.; Marrink, S. J. *J. Am. Chem. Soc.* **2009**, *131*, 9166–9167.

(47) Leermakers, F. A. M.; Scheutjens, J. M. H. M. *J. Chem. Phys.* **1988**, *89*, 6912–6924.

(48) Takaoka, Y.; Pasenkievicz-Gierula, M.; Miyagawa, H.; Kitamura, K.; Tamura, Y.; Kusumi, A. *Biophys. J.* **2000**, *79*, 3118–3138.

(49) Debnath, A.; Ayappa, K. G.; Kumaran, V.; Maiti, P. K. *J. Phys. Chem. B* **2009**, *113*, 10660–10668.

(50) Reif, M. M.; Krutler, V.; Kastenholz, M. A.; Daura, X.; Hünenberger, P. H. *J. Phys. Chem. B* **2009**, *113*, 3112–3128.

(51) Anézo, C.; de Vries, A. H.; Höltje, H.-D.; Tieleman, D. P.; Marrink, S.-J. *J. Phys. Chem. B* **2003**, *107*, 9424–9433.

(52) Baştuğ, T. P.; Kuyucak, S. M. *Chem. Phys. Lett.* **2006**, *425*, 320–323.

(53) Cordomi, A.; Edholm, O.; Perez, J. *J. Comput. Chem.* **2007**, *28*, 1017–1030.

(54) Herce, D. H.; Garcia, A. E. *J. Chem. Phys.* **2006**, *125*, 224711/1−224711/13.

(55) Klauda, J. B.; Brooks, B. R.; Pastor, R. W. *J. Chem. Phys.* **2006**, *125*, 144710/1−144710/8.

(56) Patra, M.; Karttunen, M.; Hyvnen, M. T.; Falck, E.; Lindqvist, P.; Vattulainen, I. *Biophys. J.* **2003**, *84*, 3636–3645.

(57) Patra, M.; Karttunen, M.; Hyvönen, M. T.; Falck, E.; Vattulainen, I. *J. Phys. Chem. B* **2004**, *108*, 4485–4494.

(58) Tieleman, D. P.; Hess, B.; Sansom, M. S. P. *Biophys. J.* **2002**, *83*, 2393–2407.

(59) de Vries, A. H.; Chandrasekhar, I.; van Gunsteren, W. F.; Hünenberger, P. H. *J. Phys. Chem. B* **2005**, *109*, 11643–11652.

(60) Krog, N.; Borup, A. P. *J. Sci. Food. Agric.* **1973**, *24*, 691–701.

(61) Pezron, I.; Pezron, E.; Bergenstahl, B. A.; Claesson, P. M. *J. Phys. Chem.* **1990**, *94*, 8255–8261.

(62) Pezron, I.; Pezron, E.; Claesson, P. M.; Bergenstahl, B. A. *J. Colloid Interface Sci.* **1991**, *144*, 449–457.

(63) Morley, W. G.; Tiddy, G. J. T. *J. Chem. Soc. Faraday Trans.* **1993**, *89*, 2823–2831.

(64) Cassin, G.; de Costa, C.; van Duynhoven, J. P. M.; Agterof, W. G. M. *Langmuir* **1998**, *14*, 5757–5763.

(65) Chupin, V.; Boots, J.-W. P.; Killian, J. A.; Demel, R. A.; de Kruijff, B. *Chem. Phys. Lipids* **2001**, *109*, 15–28.

(66) Sein, A.; Verheij, J. A.; Agterof, W. G. M. *J. Colloid Interface Sci.* **2002**, *249*, 412–422.

(67) van Duynhoven, J. P. M.; Broekmann, I.; Sein, A.; van Kempen, G. M. P.; Goudappel, G.-J. W.; Veeman, W. S. *J. Colloid Interface Sci.* **2005**, *285*, 703–710.

(68) Alberola, C.; Blümich, B.; Emeis, D.; Wittern, K.-P. *Colloids Surf., A* **2006**, *290*, 247–255.

(69) Krog, N.; Larsson, K. *Chem. Phys. Lipids* **1968**, *2*, 129–143.

(70) 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*; Verlag der Fachvereine: Zürich, Switzerland, 1996.

(71) Scott, W. R. P.; Hünenberger, P. H.; Tironi, I. G.; Mark, A. E.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Krüger, P.; van Gunsteren, W. F. *J. Phys. Chem. A* **1999**, *103*, 3596–3607.

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

(73) 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, The Netherlands, 1981; pp 331−342.

(74) Horta, B. A. C.; van Gunsteren, W. F.; Hünenberger, P. H. To be submitted to *J. Chem. Theory Comput.* **2010**.

(75) Hockney, R. W. *Methods Comput. Phys.* **1970**, *9*, 136–211.

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

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

(78) Klauda, J. B.; Venable, R. M.; MacKerell, A. D., Jr.; Pastor, R. W. *Curr. Top. Membr.* **2008**, *60*, 1–48.

(79) van Gunsteren, W. F.; Berendsen, H. J. C. *Angew. Chem., Int. Ed.* **1990**, *29*, 9921023.

(80) Barker, J. A.; Watts, R. O. *Mol. Phys.* **1973**, *26*, 789–792.

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

(82) Heinz, T. N.; van Gunsteren, W. F.; Hünenberger, P. H. *J. Chem. Phys.* **2001**, *115*, 1125–1136.

(83) Horta, B. A. C.; Perić-Hassler, L., Hünenberger, P. H. *J. Mol. Graph. Model.* **2010**, submitted.

(84) 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.

(85) Berendsen, H. J. C. *Simulating the Physical World*; Cambridge University Press: Cambridge, U.K., 2007.

(86) Essmann, U.; Berkowitz, M. L. *Biophys. J.* **1999**, *76*, 2081–2089.

(87) Pereira, C. S.; Lins, R. D.; Chandrasekhar, I.; Freitas, L. C. G.; Hünenberger, P. H. *Biophys. J.* **2004**, *86*, 2273–2285.

(88) Pereira, C. S.; Hünenberger, P. H. *J. Phys. Chem. B* **2006**, *110*, 15572–15581.

(89) Pereira, C. S.; Hünenberger, P. H. *Biophys. J.* **2008**, *95*, 3525–3534.

(90) Pereira, C. S.; Hünenberger, P. H. *Mol. Simul.* **2008**, *34*, 403–420.

(91) Vaz, W. L. C.; Clegg, R. M.; Hallmann, D. *Biochemistry* **1985**, *24*, 781–786.

(92) Larsson, K. *Chem. Phys. Lipids* **1972**, *10*, 177–180.

CT100200W

JCTC Journal of Chemical Theory and Computation

Human Inducible Hsp70: Structures, Dynamics, and Interdomain Communication from All-Atom Molecular Dynamics Simulations

Adrien Nicolaï,[†] Patrick Senet,*[,†] Patrice Delarue,[†] and Daniel R. Ripoll[‡,§]

[†] *Laboratoire Interdisciplinaire Carnot de Bourgogne, UMR 5209 CNRS-Université de Bourgogne, 9 Av. A. Savary, BP 47 870, F-21078 Dijon Cedex, France*

[‡] *Computational Biology Service Unit, Cornell Theory Center, Cornell University, Ithaca, New York 14853*

Received April 23, 2010

Abstract: The 70 kDa human heat shock protein is a major molecular chaperone involved in de novo folding of proteins *in vivo* and refolding of proteins under stress conditions. Hsp70 is related to several "misfolding diseases" and other major pathologies, such as cancer, and is a target for new therapies. Hsp70 is comprised of two main domains: an N-terminal nucleotide binding domain (NBD) and a C-terminal substrate protein binding domain (SBD). The chaperone function of Hsp70 is based on an allosteric mechanism. Binding of ATP in NBD decreases the affinity of the substrate for SBD, and hydrolysis of ATP is promoted by binding of polypeptide segments in the SBD. No complete structure of human Hsp70 is known. Here, we report two models of human Hsp70, constructed by homology with *Saccharomyces cerevisiae* cochaperone protein Hsp110 (open model) and with *Escherichia coli* 70 kDa DnaK (closed model) and relaxed for several tens to hundreds of nanoseconds by using all-atom molecular dynamics simulations in explicit solvent. We obtain two stable states, Hsp70 with SBD open and SBD closed, which agree with experimental and structural information for ATP-Hsp70 and ADP-Hsp70, respectively. The dynamics of the transition from the open to closed states is investigated with a coarse-grained model and normal-mode analysis. The results show that the conformational change between the two states can be represented by a relatively small number of collective modes which involved major conformational changes in the two domains. These modes provide a mechanistic representation of the communication between NBD and SBD and allow us to identify subdomains and residues that appear to have a critical role in the conformational change mechanism that guides the chaperoning cycle of Hsp70.

1. Introduction

The folding of many proteins *in vivo* is assisted by molecular chaperones.[1−4] The chaperones increase the efficiency of protein folding and inhibit the interactions of nascent proteins extruded from the ribosome with other proteins within the intracellular medium.[1−4] The 70 kDa heat shock protein (Hsp70) is one of the major molecular chaperones involved in de novo folding of proteins *in vivo*. Hsp70 molecular chaperones are ubiquitous and found in eukaryotes and prokaryotes.[1−4] Hsp70 is overexpressed under cellular stress conditions, such as, for example, heat shock and pathologies,[1−8] and this molecular chaperone has been found associated with pathological proteins in several "protein misfolding diseases" in humans.[9] An increase of the expression of Hsp70 in

* Corresponding author phone: +33 3 80 39 59 22; fax: +33 3 80 39 60 24; e-mail: psenet@u-bourgogne.fr.

§ Present address: Biotechnology HPC Software Applications Institute, Telemedicine and Advanced Technology Research Center, U.S. Army Medical Research and Materiel Command, 2405 Whittier Drive, Suite 200, Frederick, Maryland 21702

10.1021/ct1002169 © 2010 American Chemical Society
Published on Web 07/12/2010

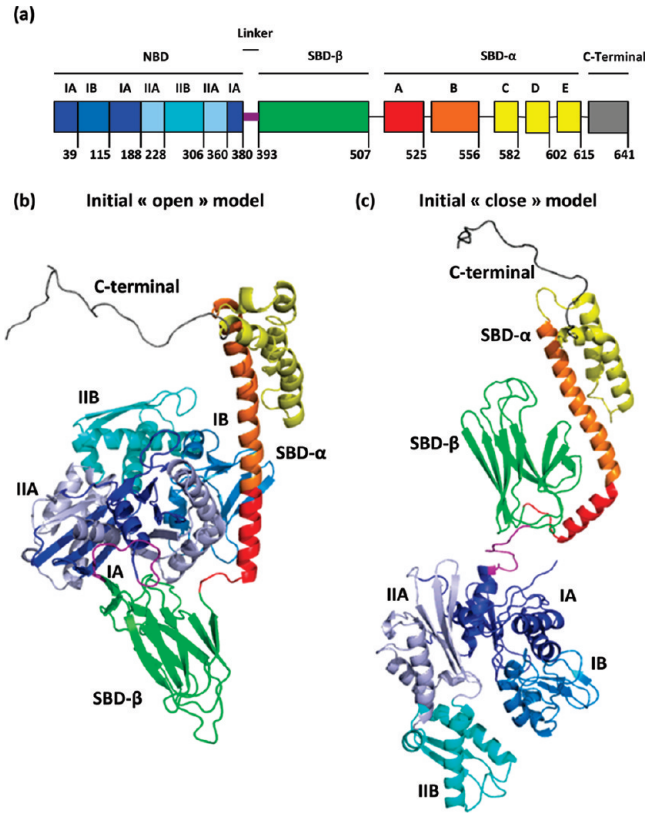


Figure 1. (a) Definition of the secondary structures of hHsp70. The color code is the following for the NBD (residues 1−380): subdomain IA (residues 1−39, 116−188, and 361−380; blue), subdomain IB (residues 40−115; marine), subdomain IIA (residues 189−228 and 307−360; light-blue), and subdomain IIB (residues 229−306; cyan). The SBD (residues 394−615) is formed by a β sandwich (SBD-β) and a helix bundle (SBD-α). The SBD-α has two helices, A (residues 512−525; red) and B (residues 526−556; orange), which are in contact with the SBD-β (residues 394−507; green) in the experimental structure of DnaK+peptide (PDB ID: 2KHO) and three other helices, C (residues 563−582; yellow), D (residues 589−602; yellow), and E (residues 605−615; yellow), forming a so-called hydrophobic core. The SBD is connected to the NBD by a short peptide (residues 381−393; purple) referred to as a linker. The C-terminal part (residues 616−641; gray) is a quite unstructured chain of amino acids terminated by the EEVD chaperone motif (residues 638−641; gray). (b) Ribbon diagram of the initial open model built by homology from the Hsp110 structure (PDB ID: 3C7N chain A). The color code is similar to the one in panel a. (c) Ribbon diagram of the initial closed model built by homology from the DnaK structure (PDB ID: 2KHO). The color code is similar to the one in panel a. This figure was prepared with PyMOL [http://www.pymol.org].

models of misfolding diseases has been shown to result in a decrease of the toxicity of disease proteins and/or to inhibit disease protein aggregation.[9−13] In addition, Hsp70 interacts with hundreds of client proteins of the intracellular environment.[1,5,8,14,15]

Hsp70 proteins are comprised of a nucleotide binding domain (NBD) and a substrate binding domain (SBD). The NBD is divided into two rather symmetrical lobes, I and II, each divided into two subdomains, A and B (Figure 1a). The SBD has a peptide binding pocket (this subdomain is referred to as SBD-β) and an α-helical "lid" (this subdomain is referred to as SBD-β; Figure 1a) which is believed to control the access of the substrate to the binding cavity.[1−5] The position of SBD-α relative to SBD-β defines two main conformational states of Hsp70 chaperone:[1−5] either the lid is open and the peptide can access the hydrophobic pocket within SBD-β, named the "open" conformation of the chaperone (as in the initial model shown in Figure 1b), or the lid is closed and the peptide is trapped in the pocket, refered to as the "closed" conformation (as the initial model shown in Figure 1c).

Hsp70 assists the folding of other proteins, through cycles of binding and release of unfolded polypeptide chains in the SBD by binding selectively short peptide stretches within the chain.[1−3] The binding and release of peptides is governed by ATP hydrolysis in the NBD and by nucleotide exchange[16,17] and is promoted by the action of cochaperone proteins.[18−20] The communication between the SBD and NBD, which regulates the binding and release of peptides, is governed by a linker (Figure 1a) through a mechanism which is not fully understood.[21,22] In ATP-Hsp70, the affinity for a substrate is low,[21] and therefore the molecular chaperone is expected to be in an open state (as in the initial model shown in Figure 1b). On the contrary, in ADP-Hsp70, the affinity for a substrate is high,[21] and the chaperone is expected to be thus in a closed state (as in the initial model shown in Figure 1c).

Because of their critical roles in many cellular processes and based on evidence that their mechanisms involve cycles of conformational changes, there has been strong interest in elucidating the structures of Hsp70′s. Crystal structures from bovine and human NBDs and NMR structures and crystal structures of SBDs from rat Hsc70 and *E. coli* Hsp70 (DnaK) have been solved during the past few years.[23−30] More recently, structures comprising both the NBD and SBD have become available; a truncated structure of the bovine Hsc70 chaperone (residues 1−554) without a nucleotide and substrate[31] and a truncated structure of *E. coli* DnaK (residues 4−603), complexed with ADP nucleotide and a short peptide (NRLLLTG), have been reported.[32] Furthermore, a crystal structure of *Saccharomyces cerevisiae* Hsp110 (residues 2−659), which is an Hsp70 homologue, was recently elucidated.[20] In this structure, Hsp110 was crystallized in an ATP state with the SBD open. Hsp110 is a cochaperon of Hsp70. It is a nucleotide-exchange factor (NEF) which binds to the NBD of the target chaperons[18−20] and induces ADP release from the NBD.

There is no structure of an Hsp70 locked in the ATP state (except the yeast Hsp110 homologue[20]). In addition, there are no structural data reporting the structure of the SBD of human Hsp70. Consequently, the alternation of conformations between the ATP and ADP states of a full length (human) Hsp70 has not been described at the atomic level. Prompted by the absence of a detailed atomic experimental structure, we carried out the task of generating computer models of human Hsp70 (hHsp70) in two conformational states: with the lid open (Figure 1b) and with the lid closed (Figure 1c). The sequence of human inducible Hsp70 (UniProt ID: P08107) was chosen for modeling purposes

because it has been reported as a new target for cancer therapies.[15,33−35]

Here, we first report a realistic atomic model of the complete structure of the inducible Hsp70 (residues 1−641) in an open conformation that was built from the yeast Hsp110 structure[18] (PDB ID: 3C7NA) using homology modeling (Figure 1b) and relaxed through several hundred nanoseconds by using all-atom molecular dynamics (MD) simulations in explicit solvent. These simulations resulted in a structural model for the open conformation of Hsp70, which agrees with the ATP state of hHsp70 when compared to the low-resolution small-angle X-ray scattering (SAXS) data.[36,37] We next report a second model of the human Hsp70 chaperone in a closed conformation that was built using homology modeling (Figure 1c) from the experimental structure of *E. coli* DnaK[32] (PDB ID: 2KHO). This second model was relaxed through hundreds of nanoseconds of all-atom MD and converged to a stable structural model of hHsp70. The latter model agrees with the ADP state of hHsp70 when compared to the low-resolution SAXS data.[36,37]

The transition between the open and closed conformations of hHsp70 was analyzed by computing the collective vibrational modes[38,39] of the two atomic models obtained by MD representatives of the open and closed conformations, respectively, using a coarse-grained model.[40−46] We show conclusively that a few collective modes of Hsp70 contribute to the interpolated conformational "pathway" from ATP-Hsp70 to ADP-Hsp70. In addition, these collective modes provide a possible mechanism to propagate information between the NBD and the SBD at a distance as large as 100 Å.

2. Methods

2.1. Generation of the Initial 3D Models of Hsp70. Three dimensional models of the human Hsp70 protein were generated by using homology modeling techniques. A BLAST[47] search using the sequence of Hsp70 against all sequences in the Protein Data Bank [i.e., sequences for which three-dimensional (3D) information has been collected] identifies few highly homologous sequences. Among these possible templates with the largest degree of residue identity (RI) detected by BLAST were 3C7N monomer B, 1YUW monomer A, and 2VZ7 monomer A. However, the actual 3D data from these structural determinations do not fully cover the whole sequence range of our protein of interest [e.g., for 3C7N_B from a total of 641 residues in Hsp70, only 554 are aligned (∼86%)]. On the other hand, BLAST was able to detect few other sequences that, while sharing lower RI, provided better coverage (i.e., a larger number of aligned residues). Among these structures are (a) 2KHO monomer A, for which the alignment provides coverage for 584 of 641 residues in the Hsp70 sequence (91%), with RI = 310/611 (50%), 402/611 (65%) residues with positive scores, and 26/611 (4%) gaps; and (b) 3C7N monomer A, for which the alignment provides coverage for 598 residues out of 641 (or 93%), with RI = 183/646 (28%), 315/646 (48%) residues with positive scores, and 49/646 (7%) gaps. On the basis of these results, we decided to use the experimentally determined structure of the *E. coli* Hsp70 (DnaK) chaperone[32] (PDB ID: 2KHO monomer A) and the structure of the heat shock protein Hsp110 homologue Sse1[18] (PDB ID: 3C7N monomer A) as templates to produce a set of 3D models for the human Hsp70 sequence.

To carry out this task, we made use of the program MODELLER[48−50] (version 9.3), which is capable of converting the information from (a) pairwise alignments of the query sequence with its templates and (b) the structural data provided by these templates into a set of distance and dihedral-angle restraints. By minimizing the violations to these derived restrains, MODELLER[48−50] is able to generate the 3D models for the sequence of interest. Initial models were energy minimized using 500 to 1000 steps of a steepest-descent algorithm provided in the DS modeling program (Accelrys, Inc.). The optimized models showing the closest C^{α} RMSD (root-mean-square deviation) with respect to the templates from which they were derived were used as starting points for MD simulations. In particular, the structural superposition of the open model of the APO-Hsp70 used as an initial conformation of one of the MD runs (APO-1 run), and its template, the structure of Hsp110, shows very small differences with a C^{α} RMSD of 0.8 Å for an optimal superposition of 609 C^{α} atoms. Similarly, the structural superposition of the MD initial conformation for the closed model of Hsp70 (APO run) and its template, the structure of DnaK, led to a C^{α} RMSD of 1.6 Å for an optimal superposition of 590 C^{α} atoms. Structure alignments were carried out by using the combinatorial extension method.[51]

2.2. Molecular Dynamics Simulations. All-atom molecular dynamics simulations in explicit water of Hsp70 have been carried out with the GROMACS software package[52] using the simple point charge (SPC) water model and the GROMOS96 ffG43a1 force field.[52] The time step used in all simulations was 0.001 ps, and the list of neighbors was updated every 0.005 ps with the "grid" method and a cutoff radius of 1 nm. The coordinates of all the atoms in the simulation box were saved every 2 ps. The initial velocities were chosen randomly. We used the NPT ensemble with a cubic box. The temperature and pressure were kept to the desired value by using the Berendsen method and an isotropic coupling for the pressure (T = 300 K, τ_T = 0.1 ps; P_0 = 1 bar, coupling time τ_P = 1 ps). The electrostatic term was computed by using the particle mesh Ewald (PME) algorithm (with a radius of 1 nm) with the fast Fourier transform optimization (on 117 points for each axis for the open model and 160 points for the closed model and an order equal to 4 for the interpolation). The cutoff algorithm was applied for the noncoulomb potentials with a radius of 1 nm. The system was warmed up for 40 ps and equilibrated for 600 ps with lower restraints, finishing with no restraints at 300 K. We performed five runs, four for the open model and one for the closed model.

Open Model. APO-1 Run. The initial structure of APO-Hsp70 (open model, run APO-1) was built by using the structure of Hsp110 as a template.[18] The APO-Hsp70 protein has been solvated in a cubic box with 82 065 SPC water molecules keeping a minimum distance of 0.9 nm between the solute and each face of the box. We used the periodic

boundary conditions, and the initial value of the box side is 13.683 nm. The charge of APO-Hsp70 has been neutralized by adding 11 Na^+ counterions. The energy of the model was first optimized with the "steepest descent minimization" algorithm and then by using the "conjugate gradients" algorithm. The production period was 400 ns.

ADP-1 Run. The ADP-Hsp70 model was built using a structure of Hsp70 relaxed after a 3 ns period of production of run APO-1. We added one ADP molecule with PRODRG[53] and one Ca^{2+} ion (as in the experimental structure of the human Hsp70 NBD, PDB ID: 1HJO[26]) to this APO structure. The protein was solvated with 82 089 water molecules in a cubic box of the same dimensions as in the APO-1 run, and 12 Na^+ ions were added to neutralize the system. The production period was 480 ns.

APO-2 Run. A second initial structure of APO-Hsp70 was built as follows. We observed that the initial conformation of ADP-Hsp70 was stable until 320 ns. The "A+B" helix did not break into two helices before 320 ns. To test the effect of the nucleotide and to produce different initial conditions for the APO form, we selected the structure of ADP-Hsp70 at 200 ns (before any conformational change). We removed the ADP molecule, the Ca^{2+} ion, and one Na^+ ion, and we added eight water molecules (with initial velocities equal to zero; the velocities of all the other atoms in the ADP-Hsp70 structure at 200 ns were conserved). This new APO structure was the initial structure of the APO-2 run. The production period was 184 ns.

ADP-2 Run. A second initial structure of ADP-Hsp70 was built similarly to the first trajectory ADP-1 but with different initial velocities. The production period was 192 ns.

Closed Model. APO Run. The initial structure of Hsp70 (closed model, APO run) was built by using the structure of DnaK as a template.[32] The closed model Hsp70 protein was solvated in a cubic box with 212 632 SPC water molecules keeping a minimum distance of 0.9 nm between the solute and each face of the box. We used periodic boundary conditions with an initial value of the box side of 18.664 nm. The charge of the system in the closed model was neutralized by adding 11 Na^+ counterions. The energy of the system was first optimized with the "steepest descent minimization" algorithm and then by using a "conjugate gradient" algorithm. The production period was 130 ns.

2.3. The Anisotropic Network Model (ANM). An important issue in the analysis of protein dynamics is the identification of the mechanism of slow, large-amplitude motions, also termed collective motions. The collective motions are believed to be essential for a protein to function. A widely used technique for studying collective motions of large proteins is the anisotropic network model[41−46] (ANM). In ANM, a protein in its folded state is represented by a three-dimensional elastic network of nodes, which are located at the C^α position of each residue. The interactions between residues in close proximity (distance < cutoff distance r_c) are replaced by harmonic springs.[41] The ANM predicts the relative sizes of structural fluctuations of a protein in its native state and their directionalities and permits the decomposition of the molecular motions into a series of $3N - 6$ modes ($N = 641$ for the human Hsp70).

In ANM, the intramolecular free energy of a protein of N residues in its native state is expressed as a series expansion in the fluctuations $\Delta\vec{R}_i$ of individual residue positions ($1 \le i \le N$)

$$F - F(0) = \sum_{i=1}^{N}\left(\frac{\partial F}{\partial\overrightarrow{\Delta R_i}}\right)_0.\overrightarrow{\Delta R_i} + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{\partial F^2}{\partial\overrightarrow{\Delta R_i}\partial\overrightarrow{\Delta R_j}}\right)_0:\overrightarrow{\Delta R_i}\overrightarrow{\Delta R_j} = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\Phi_{ij}:\overrightarrow{\Delta R_i}\overrightarrow{\Delta R_j} \quad (1)$$

where the first term on the right-hand side of the first line equality is zero because, in the native state

$$\left(\frac{\partial F}{\partial\overrightarrow{\Delta R_i}}\right)_0 = 0, \forall i$$

The $3N \times 3N$ Φ matrix in eq 1 is defined as follows

$$\Phi_{ij}^{\alpha\beta} = \left(\frac{\partial F^2}{\partial\overrightarrow{\Delta R_i}\partial\overrightarrow{\Delta R_j}}\right) = -A\frac{(R_i^\alpha - R_j^\alpha)(R_i^\beta - R_j^\beta)}{R_{ij}^2}\Gamma_{ij} \quad (2)$$

where R_i^α is the Cartesian component α (= x, y, or z) of the position $\vec{R}_i$ of the ith C^α atom and A and Γ_{ij} are parameters of the ANM model, which depend on the protein studied. Equation 2 is the simplest form of Φ, which is invariant by global translation and rotation of the molecule.

The elements Γ_{ij} of the matrix in eq 2 are dimensionless and take the value 1 if the inter-residue distance R_{ij} remains below the cutoff distance r_c and is 0 otherwise. The parameter A is the average force constant between two connected residues. The cross-correlations between residue fluctuations are found from

$$\sigma_{ij}^{\alpha\beta} = \langle\Delta R_i^\alpha\cdot\Delta R_j^\beta\rangle = kT[\Phi^{-1}]_{ij}^{\alpha\beta} \quad (3)$$

where $[\Phi^{-1}]_{ij}^{\alpha\beta}$ is the $\alpha\beta$ element of the ijth 3×3 bloc matrix of the $3N \times 3N$ inverse matrix Φ^{-1}.

The mean square fluctuation of the bonds R_{ij} is computed from eq 3, with $i = j$, i.e.,

$$\langle(\Delta R_i)^2\rangle = kT\sum_\alpha[\Phi^{-1}]_{ii}^{\alpha\alpha} \quad (4)$$

Equation 3 will be used here to calculate the B factors (B) of each residue:

$$B_i = \frac{8\pi^2}{3}\langle(\Delta R_i)^2\rangle \quad (5)$$

The matrix Φ is rewritten as the product of three matrices, the diagonal matrix Λ of its nonzero eigenvalues λ_m ($1 \le m \le 3N - 6$), the matrix $\mathbf{E}$ of the corresponding eigenvectors $\vec{e}_m$, and the transpose of $\mathbf{E}$:

$$\Phi = \mathbf{E}\Lambda\mathbf{E}^{-1} = [e_1e_2...e_{3N-6}]\,\mathrm{diag}(\lambda_1\lambda_2...\lambda_{3N-6})[e_1e_2...e_{3N-6}]^T \quad (6)$$

where $\mathbf{E}$ is an orthonormal matrix. The inverse of Φ can be readily found from eq 6 and is given by

$$\Phi^{-1} = \mathbf{E}\Lambda^{-1}\mathbf{E}^{-1} = \sum_m \left[\frac{\vec{e}_m.\vec{e}_m^{\mathrm{T}}}{\lambda_m}\right] \quad (7)$$

The eigenvalues of Φ are ordered as $\lambda_1 < \lambda_2 < \lambda_3 ... < \lambda_n$ where λ_1 is the first nonzero eigenvalue. These eigenvalues λ have the physical dimension of a restoring force but are also called frequencies by analogy with the normal modes analysis. The collective modes correspond to low λ. The number of modes of nonzero frequency for hHsp70 is 1917, and the slow modes here are defined as the first hundred modes.

The covariance of the structural fluctuations for the pair of residues $i-j$ is related to the spectrum of Φ by

$$\langle \overrightarrow{\Delta R_i} \cdot \overrightarrow{\Delta R_j} \rangle = kT \sum_m \left[\frac{\vec{e}_m \cdot \vec{e}_m^{\mathrm{T}}}{\lambda_m}\right]_{ij} \quad (8)$$

The quantity $1/\lambda_m$ is the reciprocal of the mth nonzero eigenvalue λ_m of Φ. According to eqs 4 and 8, the slowest mode (with the lowest frequency λ_1) has the most dominant contribution to the structural fluctuations of the protein and corresponds to displacements along $\vec{e}_1$. The slowest modes are thermally excited with a large amplitude and generally describe functional motions that are of great biological interest.

The ANM requires specification of two parameters: the force constant A that is a measure of the strength of the restoring forces of the $C^\alpha - C^\alpha$ virtual bonds of the native fold (eq 2) and the cutoff distance r_c that defines whether two given nodes are linked or not. The values of A and r_c in ANM were fitted in order to reproduce the B factors extracted from the crystallographic data of the truncated bovine Hsc70[31] (bHsc70) (PDB ID: 1YUW; data not shown). The value of the cutoff distance r_c was obtained as follows. We noted that there were more than six zero eigenvalues in the ANM model for $r_c \leq 10$ Å for both the bHsc70 structure and for the model average structures (computed between $t = 100$ ns and $t = 400$ ns for the open model and computed between $t = 30$ ns and $t = 130$ ns for the closed model). This implied that some deformation of the structure did not cost internal energy: for these deformations, the structure was unstable. To remove such physically unrealistic behavior, we adopted a larger cutoff distance, which was evaluated by studying the variation of the spectrum of Φ for the bHsc70 structure for different values of r_c. In particular, we computed the density of states, $g(\lambda)$, of the low frequency modes. The distribution $g(\lambda)$ was found to increase linearly with the frequency[54] for the slow modes (defined here as the first 100 modes which contribute to about 70% of the total fluctuations), implying that the cumulative density of modes, $G(\lambda) = \int_0^\lambda du\, g(u) \propto \lambda^2$, deviated from the Debye model of elastic solids where $G(\lambda) \propto \lambda^3$. The anomalous spectral dimension $d_s \approx 2$ of proteins, defined as the power of λ, i.e., $G(\lambda) \propto \lambda^{ds}$, was observed by inelastic neutron scattering for lysozyme,[55] for which $d_s \approx 1.4$. The spectral dimension of bHsc70 for each r_c value between 11 and 20 Å was computed from a power law that best fits the cumulative density of modes for the first 100 modes. The results showed that the value of d_s is nearly constant (≈ 1.9), and the shape of the density of modes $g(\lambda)$ does not change for $11 \leq r_c \leq 14$ Å. Therefore, we have chosen a $r_c = 11$ Å. The best value of A reproducing the B factors of bHsc70[31] for $r_c = 11$ Å was 0.4−0.5 kcal/mol/Å^2.

Another parametrization of ANM was tested on the basis of the so-called parameter-free ANM[45] (pfANM), published during the course of the present work. In pfANM, there is no cutoff distance r_c, and the force constant A decays as an inverse power n of the distance between a pair of C^α's. A value of $n \geq 6$ is recommended to reproduce the collective modes of a protein. The collective modes computed with pfANM were compared to the collective modes computed with ANM and $r_c = 11$ Å (data not shown). The collective modes were very similar to each other. The overlap of the eigenvectors of the first 10 nonzero frequency modes was greater than 0.93, both for the open and for the closed models of Hsp70 (data not shown).

2.4. Involvement Coefficient. In order to evaluate the contribution of a given mode to the transition between the open and closed states, individual and cumulative involvement coefficients[56] were calculated as follows.

A transition pathway was determined by linearly interpolating the transition end-points (i.e., the initial, $\vec{X}_A$, and final, $\vec{X}_B$, structures of a conformational transition) upon optimal superposition of all the C^α atoms. The resulting "displacement vector" $\vec{X}_A - \vec{X}_B$ was expanded in a linear combination of normal modes $\vec{e}_k$ of the initial state. Then, the projection L_k of the normalized displacement vector on the kth normal mode vector was computed:

$$I_k = \frac{\overrightarrow{X_A} - \overrightarrow{X_B}}{|\overrightarrow{X_A} - \overrightarrow{X_B}|} \cdot \vec{e}_k \quad (9)$$

The involvement coefficient is defined as $I_k = |I_k|$. It describes the degree of involvement of the kth mode in the conformational transition $A \rightarrow B$. Thus, the individual involvement coefficient indicates in a semiquantitative way which collective motions are important for a given conformational change. A complementary quantity which indicates the involvement of a set of collective modes in the conformational change is the cumulative involvement coefficient CI_k, which is computed as

$$CI_k = \sum_{k=1}^{N} I_k^2 \quad (10)$$

3. Results and Discussion

3.1. Structures of hHsp70 Relaxed by MD in an Open State. *Initial Model Built by Homology.* The initial model of the human inducible Hsp70 structure in an open state (named the open model hereafter) was built by homology using the experimental structure of Hsp110[18] as a template in the software MODELLER[48−50] (see Methods). There is a main relevant structural difference between Hsp110 and all of the known Hsp70 truncated structures: in the SBD-α of Hsp110,[18] the helices A and B (Figure 1a) form a unique long helix (Figure 1b) that we call "A+B".

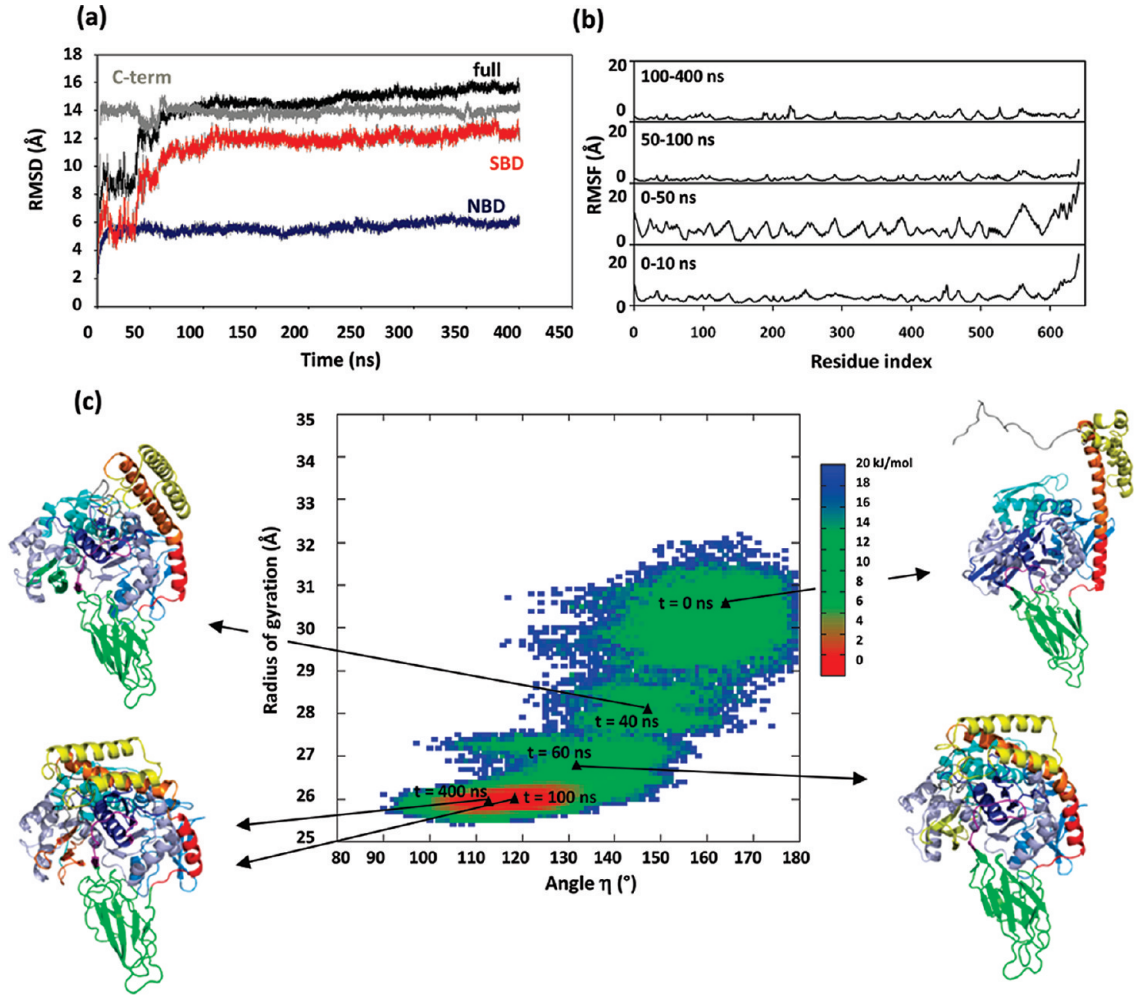


Figure 2. Time evolution of structural parameters of the open model as a function of time (run APO-1 is shown). (a) The C^{α} RMSD (for the full structure, the NBD, the SBD, and the C-terminal part) computed with respect to the initial open model built by homology modeling. (b) The C^{α} RMSF ($\sqrt{\langle \Delta R_i^2 \rangle}$, see eq 4 in Methods) computed for the open model in different time windows. (c) Energy surface (in kJ/mol) for the relaxation of the open model studied by MD. The energy surface is defined by $E = -kT \ln[P(R_g,\eta)/\mathrm{Max}\{P(R_g,\eta)\}]$ where $P(R_g,\eta)$ is the probability density of the radii of gyration R_g and of the angle η defined in section 3.1 and computed from MD. E is zero at the maximum value of $P(R_g,\eta)$ (red basin).

To the best of our knowledge, there is no structure of isolated Hsp110, and its long helix might be formed when Hsp110 interacts with its target protein[18] (Hsp70) or with itself, as shown in a crystal of the Hsp110 dimer[20] (PDB ID: 2QXL).

MD of the Open Model without Nucleotide. The initial model without nucleotide (APO-hHsp70) was relaxed by all-atom MD at 300 K (see Methods). The structural changes and the stability of the structure simulated by MD have been monitored by computing the C^{α} RMSD with respect to the initial model (both for the full structure and for the SBD, NBD, and C-terminal domains separately) as a function of time for two trajectories of hHsp70 without a nucleotide (named APO-1 and APO-2, see Methods), as illustrated in Figure 2a for the APO-1 trajectory. The structure of the NBD simulated by MD converges to a stable conformation in less than 10 ns, as shown by the convergence of the C^{α} RMSD with respect to the initial model in Figure 2a. The relatively huge RMSD value between the relaxed open model and the initial model found at 400 ns (16 Å; Figure 2a) is due to the reorganization of the SBD and of the C-terminal part and to the reorientation of the SBD relative to the NBD. Analysis of the C^{α} RMSF (root-mean-square fluctuations) along the primary sequence (Figure 2b) showed that the SBD structure and the C-terminal part of the initial open model built by homology were not stable. Indeed, we observed huge fluctuations of the SBD-α and of the C-terminal loops (residues index ≥ 525 in Figure 2b), reflecting a reorganization of the structure up to about 50 ns of simulation time in the MD trajectory APO-1 (Figure 2b). The relaxation of the C-terminal part of the initial model of hHsp70 occurred quickly (in few nanoseconds; Figure 2a) and induced unfolding of the helix E of its SBD-α (Figure 3a). The main structural changes that occurred at about 50 ns (shown as a sharp increase of the C^{α} RMSD at 50 ns in Figure 2a for the SBD and for the full structure) were the break of the initial fused helix "A+B" of the SBD-α[18,20] into two helices A and B and the binding of the SBD-α on the NBD, as shown in Figure 3a. There was still a slow structural evolution of the model between 100 and 400 ns (as shown by a slow evolution of the C^{α} RMSD of the full structure in Figure 2a), but the structures (not shown) computed between 100 and 400 ns were very similar.

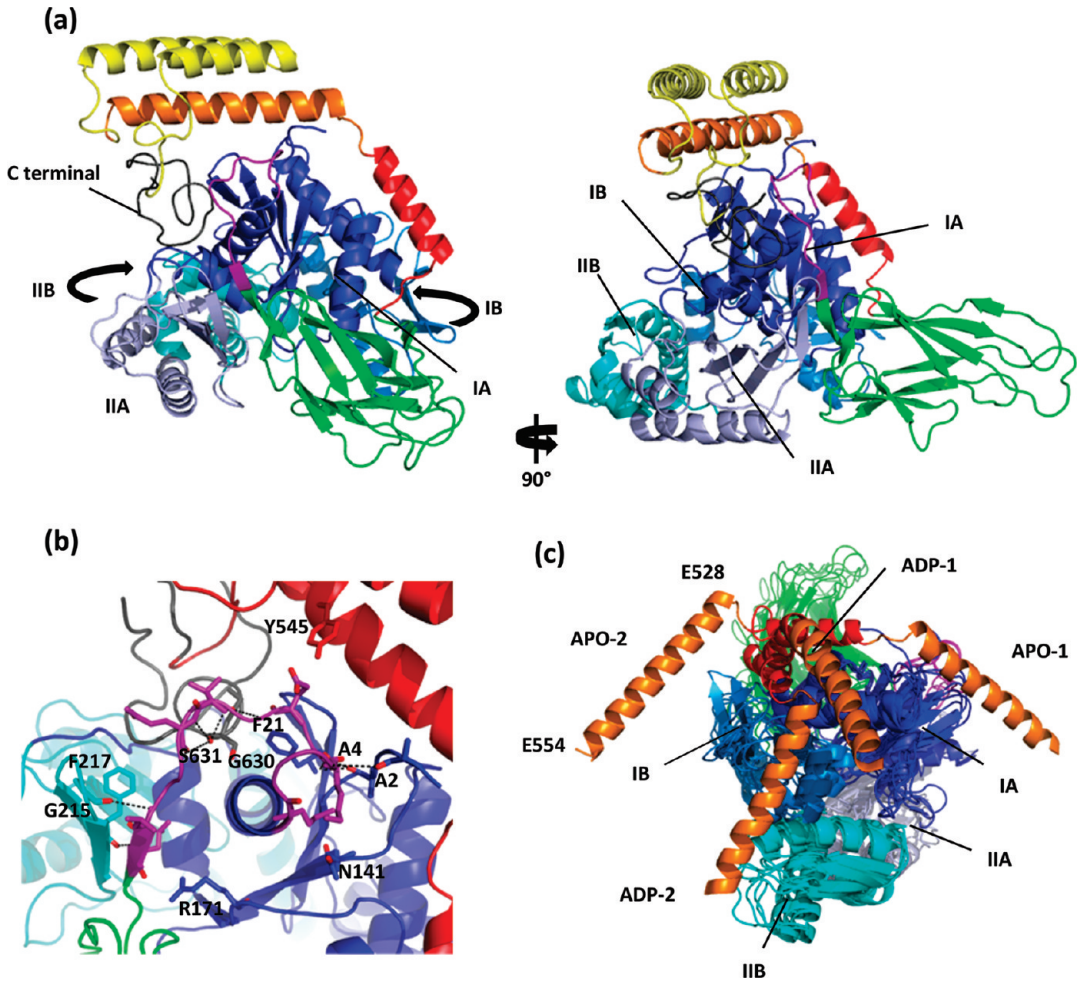


Figure 3. Typical structures of the open model relaxed by MD simulations. The color code is similar to the one in Figure 1a. Ribbon diagram of the average open model (run APO-1) in two views rotated by 90°. (b) Close-up view of the contacts between the linker and the other domains. (c) Comparison of the average structures of the open model relaxed by MD: run APO-1, run APO-2, run ADP-1, and run ADP-2. The structures are truncated after residue 554 to show clearly the position of the SBD-α. The superposition of the structures was done by minimizing the RMSD of the C^{α} of the NBD. This figure was prepared with PyMOL [http://www.pymol.org].

The reorganization of the structure of the initial open model was best represented by the energy surface $E(R_g,\eta)$; Figure 2c) defined as an effective potential of mean field: $E(R_g,\eta) \equiv -kT \ln[P(R_g,\eta)/\mathrm{Max}\{P(R_g,\eta)\}]$, where k is the Boltzmann constant, T is the temperature, and $P(R_g,\eta)$ is the 2-D residential probability density[57] of the radius of gyration of the protein (R_g) and of the angle measuring the local curvature of the "A+B" helix (η) in the region where the helix was found bent in the relaxed structure. The angle η was calculated as the angle between the vector joining the C^{α} of the residues 512 and the C^{α} of the residue 516 at the N-terminal part of the "A+B" helix and the vector joining the C^{α} of the residues 530 and the C^{α} of the residue 534 of the helix B. The four-residue spacing used to define the $C^{\alpha}-C^{\alpha}$ vectors was chosen because in a perfect α helix a rotation of over 360° corresponds to a 3.6 residue spacing; in this way, the two vectors chosen are almost parallel to the helical axis with an angle of almost 180° in the initial model. The sequence of the major events of the reorganization of the model studied by MD is shown in Figure 2c: the initial break of the helix occurred at about 40 ns, and between 40 and 60 ns, the helix curvature increased (the angle η decreased) while simultaneously the SBD-α rotated, implying a stronger interaction between the SBD and the NBD. Binding of the helix on the NBD is observed around 100 ns. The structure found at 100 ns is closely similar to the one at 400 ns (the C^{α} RMSD between the structures at 100 and 400 ns is 3.2 Å). The energy scale computed from canonical MD simulations, such as the one shown in Figure 2c, is approximate because the conformational sampling of canonical MD simulations is necessarily limited for a system of the size of Hsp70, and because the energy barriers in a 2-D energy surface may not represent the actual energy barriers in the multidimensional conformational space of the protein.

In the stable average structure computed between 100 and 400 ns, the helices A and B interact with the lobe IA of the NBD (Figure 3a). The interdomain linker (residues 381−393) is bound to a cleft between subdomains IA and IIA of the NBD (Figure 3b). The residues M381, K384, E386, and L393 are in close contact with the residues N141, A2 and A4, F21, and R171 of the subdomain IA, respectively, whereas the residue L392 is in close contact with residues G215 and F217 of the subdomain IIA (Figure 3b). In a second trajectory

(APO-2), the initial open model adopted a stable conformation after 70 ns of simulation time, as shown by the time evolution of the C^{α} RMSD relative to the initial model, which converged to a constant value (data not shown). Similar events were observed in trajectories APO-1 and APO-2: a break of the initial fused helices A+B, a folding of the SBD-α on the NBD, and the binding of the linker between subdomains IA and IIA. However, in the APO-2 trajectory, helices A and B were found respectively in contact with lobes IA and IB of the NBD, as shown in Figure 3c.

MD of the Open Model with ADP. The stable typical structures of the initial open model with ADP (see Methods) relaxed by all-atom MD at 300 K are presented in Figure 3c. In the two trajectories of hHsp70 (ADP-1 and ADP-2, see Methods), we observed a break of the initial fused helix A+B of the initial open model (Figure 1b) and the displacement of the helix bundle toward the NBD (Figure 3c) after 320 ns in ADP-1 and 100 ns in ADP-2, as reflected by the sharp variations of the C^{α} RMSD with respect to the initial model at these times (data not shown). The stable positions of the α-helical part of the SBD relative to the NBD are different from the ones obtained for the APO-hHsp70 simulations (Figure 3c). In ADP-1, helices A and B are bound to lobe IA, and in ADP-2, helix A is bound to lobe IA and helix B is bound to lobe IB (Figure 3c).

Relatively long all-atom MD simulations are needed to relax the initial model of hHsp70 in an open state (Figure 1b). The model undergoes a major structural relaxation before converging to a stable structure (Figure 3) after several tens of nanoseconds in MD simulations. The structural changes and energy surfaces (not shown) are similar in all trajectories: the fused A+B initial helix breaks into two helices A and B, and the SBD helix bundle moves toward the NBD. The stable average structures of the NBD and of the SBD-β of the open model simulated by MD are similar in the four MD runs: the C^{α} RMSD between the four NBD structures varies between 1.9 Å and 4.9 Å, and the C^{α} RMSD between the four SBD-β structures varies between 2.5 Å and 3.4 Å. The C^{α} RMSD between the SBD-α structures of the four runs varies between 6.4 Å and 10.6 Å because the helices A and B of the SBD-α interact with the NBD, and their precise positions relative to the NBD differ in the different trajectories (Figure 3c). With the NBD of hHsp70 being rather symmetrical, it might be possible for the open state of hHsp70 to exist in different stable conformations involving different positions of the SBD-α relative to the NBD. The structure of the open state of hHsp70, which should correspond to ATP-hHsp70, has not been solved at the atomic level so far.

The hydrolysis of ATP is known to promote substrate binding and facilitate the closure of SBD.[16,17] In the homologous truncated structure solved experimentally,[32] ADP-Hsp70 was found in a closed state (as in the initial model shown in Figure 1c) with the linker exposed to solvent.[32] However, the sole binding of ADP to the initial open model does not provoke the closure of the SBD in our MD simulations (Figure 3c). This might be due to the existence of an activation barrier between the open and closed states that the system cannot cross spontaneously within the limited time-scale sampled (400 ns) by our MD simulations. This activation barrier may correspond to the extrusion of the linker from its strongly bound position on a hydrophobic cleft between subdomains IA and IIA[22] of the NBD, as found in the initial open model (Figure 1b) and in the stable relaxed APO and ADP structures (Figure 3c). *In vivo*, the transition of hHsp70 between open and closed states involved the hydrolysis of ATP[16,17] and is assisted by cochaperones such as Hsp40 and Hsp110[18−20] not considered in the present study. Simulations of the conformational dynamics of hHsp70 on a longer time-scale ($\gg\mu$s) and with cochaperones are beyond the reach of all-atom simulations and will require the application of anharmonic coarse-grained models like UNRES.[58]

3.2. Structure of Hsp70 Relaxed by MD in a Closed State. *Initial Model Built by Homology.* The initial model of hHsp70 in a closed state (named the closed model here) was built by homology[48−50] using the recent complete (unfortunately still truncated by 30 residues) structure of DnaK in solution (PDB ID: 2KHO;[32] see Methods). The experimental structure of the bacteria homologue DnaK is locked in the ADP state with a short peptide (NRLLLTG) bound within the SBD-β (PDB ID: 2KHO).[32] The coordinates of the nucleotide and of the peptide substrate were not resolved in the data deposited in the PDB. Compared with the initial open model, the SBD-α of this second model has a kink between the two helices A and B (Figure 1c), as in all known Hsp70 truncated structures.

MD of the Closed Model without a Nucleotide. We considered one trajectory of the initial closed model (Figure 1c) without a nucleotide and without a substrate (APO; see Methods). The APO closed model adopted a stable conformation after about 30 ns, as shown by the time evolution of the C^{α} RMSD of the full structure with respect to the initial model (Figure 4a) and by the time evolution of the C^{α} RMSF of the structure along the sequence (Figure 4b). In Figure 4a, the C^{α} RMSD of the full structure and of the SBD converges to a constant value after 30 ns, whereas the C^{α} RMSD of the NBD converges after only 10 ns. In Figure 4b, the C^{α} RMSF computed between 30 and 60 ns and that between 30 and 130 ns were very similar. Analysis of the C^{α} RMSF along the primary sequence in the first 30 ns (Figure 4b) showed that the largest structural fluctuations of the initial model were mainly located in the C-terminal part of the model (residue index ≥ 616 in Figure 4b). The C-terminal part of hHp70 (residues 616−641) was absent in the experimental DnaK template (2KHO) and added in the homology modeling procedure. The largest fluctuations observed in Figure 4b in the first 30 ns were due to the folding of this initial unstructured C-terminal part of the protein. The folded C-terminal part moved between the SBD-β and helix B of the SBD-α, inducing the unfolding of helix E of the initial model (Figure 5a). In addition, the kink between helix A and helix B in SBD-α is more elongated in the relaxed structure than in the initial model, and its SBD-β is rotated (Figure 5a). No other major structural changes were observed after 30 ns up to the end of the MD trajectory (130 ns). The SBD is always closed (Figure 5a) and adopted similar conformations to that

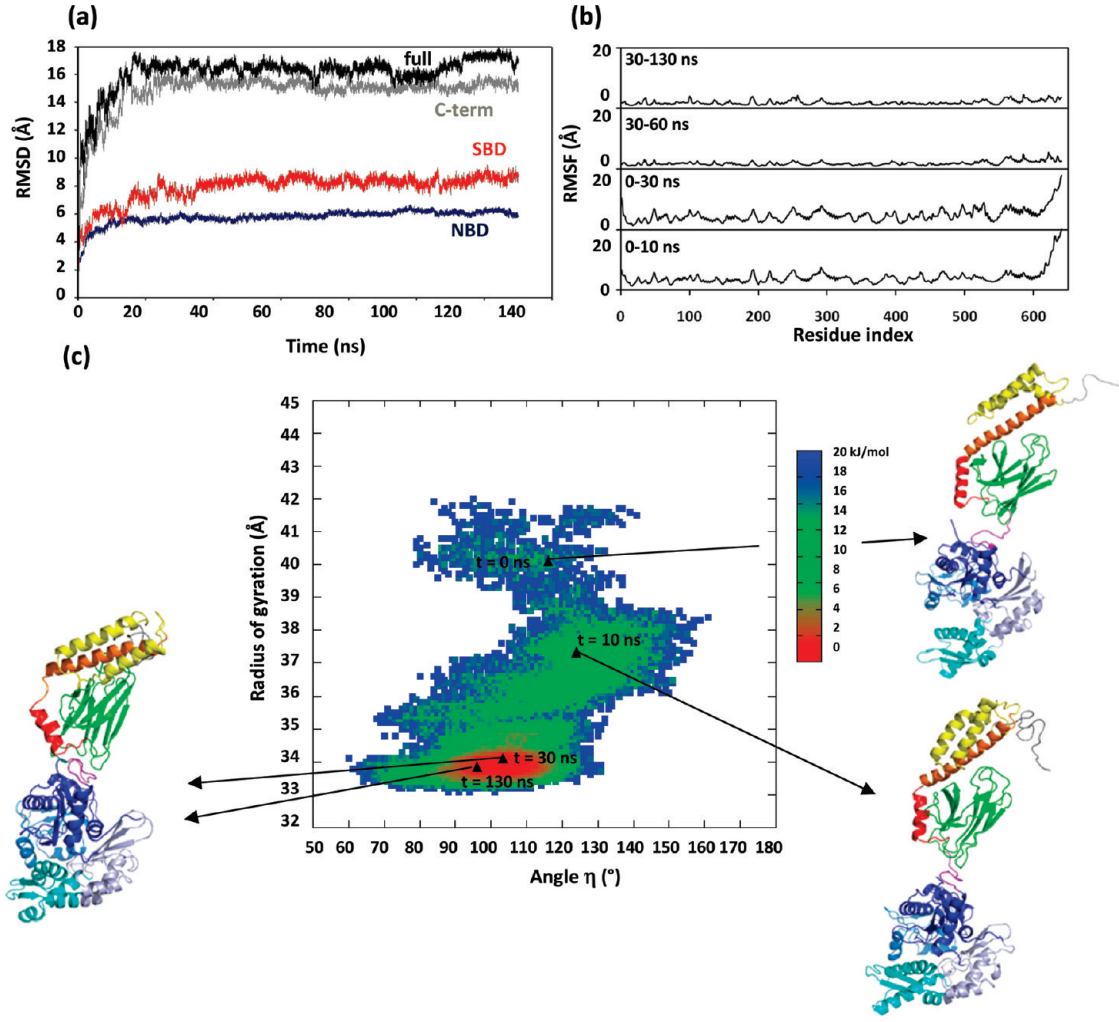


Figure 4. Time evolution of structural parameters of the closed model as a function of time. (a) The C^{α} RMSD (for the full structure, the NBD, the SBD, and the C-terminal part) computed with respect to the initial closed model built by homology modeling. (b) The C^{α} RMSF ($\sqrt{\langle \bar{\Delta} R_i^2 \rangle}$, see eq 4 in Methods) computed for the closed model in different time windows. (c) Energy surface (in kJ/mol) for the relaxation of the closed model studied by MD. The energy surface is defined by $E = -kT \ln[P(R_g,\eta)/\mathrm{Max}\{P(R_g,\eta)\}]$ where $P(R_g,\eta)$ is the probability density of the radii of gyration R_g and of the angle η defined in section 3.1 and computed from MD. E is zero at the maximum value of $P(R_g,\eta)$ (red basin).

observed in ADP-DnaK.[29,32] On the other hand, preliminary results (unpublished data) for two all-atom MD runs (150 ns) of DnaK in the closed state showed that the bacterial structure adopted stable conformations between 20 and 50 ns which are very similar to that of human Hsp70 in the closed state (Figure 5a). Because the model of the closed state converged quickly to a stable conformation and because the simulations of the closed model of hHsp70 are computationally demanding (the simulation cell contains about 645 000 atoms) and our computational resources are limited, we did not perform additional MD runs.

The sequence of major events of the relaxation of the closed model simulated by MD is shown on the energy surface $E(R_g,\eta)$ (Figure 4c) where R_g and η are the radius of gyration of the protein and the angle measuring the curvature of the "A+B" helix as defined in section 3.1. In the first 10 ns, folding of the C-terminal part of the protein was observed. After 30 ns, the SBD and NBD have moved toward each other with the SBD bound to the NBD in a stable conformation, which does not change significantly until the end of the simulation (130 ns; the RMSD between the structure at 30 ns and the one at 130 ns is 3.4 Å).

The interdomain linker of hHsp70 relaxed by MD in the closed state is not bound in the same hydrophobic cleft between subdomains IA and IIA (Figure 5b), as was the case in the open state of hHsp70 (Figure 3b). However, in the closed state of hHsp70, the linker makes close contacts with the subdomain IA of the NBD; namely, the residues M381, V388, Q389, and L390 of the linker are respectively in close interactions with the residues N141, R171, V169, and N168 of domain IA. In addition, the linker contacts SBD-β: the residues E386, V388, Q389, and L393 of the linker are in close contact with the residues N505, N483, A482 and I420, and P421 of the SBD-β, respectively.

3.3. Comparison with Experimental Structural Information. Several structures of two-domain Hsp70 NBD-SBD[31,32] constructs have been published in the past few years. However, all of these structures were truncated, so a comparison of our full-length models with the experimental structures solved by XRD (X-ray diffraction) or NMR

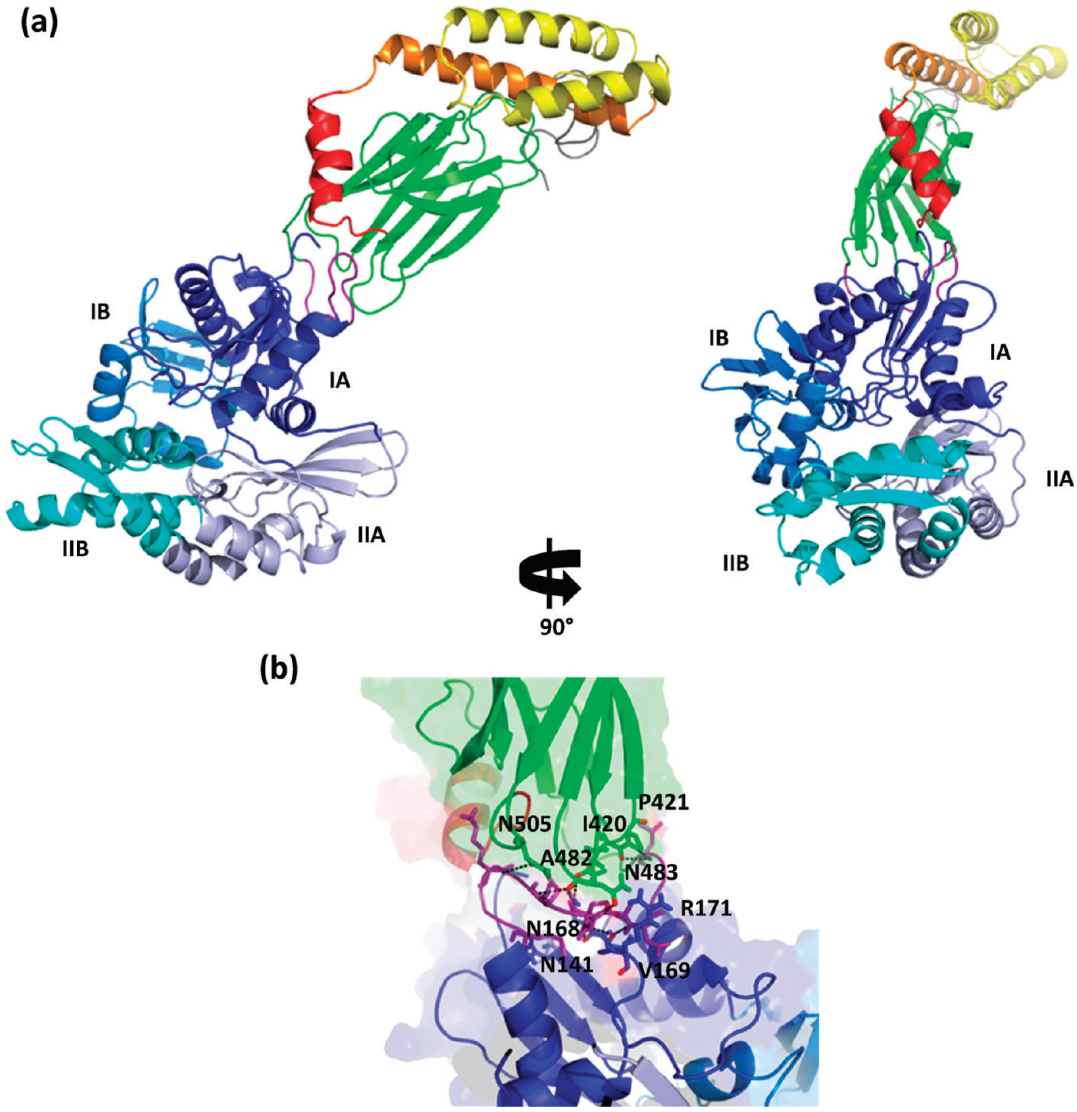


Figure 5. Typical structure of the closed model after relaxation by MD simulations. The color code is similar to the one in Figure 1a. (a) Ribbon diagram of the average structure of the closed model in two views rotated by 90°. (b) Close-up view of the contacts between the linker and the other domains. This figure was prepared with PyMOL [http://www.pymol.org].

Table 1. Experimental Structural Information for Hsp70 Chaperone and Their Homologues

protein	Uniprot ID	sequence length (aa)	domains	sequence of the constructs (aa)	PDB ID	ligand	exptl. techniques
hHsp70	**P08107**	641	NBD	3−382	1HJO[26]	ADP	XRD
bHsc70	**P19120**	650	NBD+SBDβ	1−554	1YUW[31]	ATP	XRD
			NBD	1−386	1NGA[24]	ADP	XRD
			NBD	1−386	1NGG[24]	ATP	XRD
DnaK	**P0A6Y8**	638	NBD+SBDβ+α	4−603	2KHO[32]	ADP	NMR
Sse1 (Hsp110)	**P32589**	693	NBD+SBDβ+α	1−649	3C7NA[18]	ATP	XRD

(nuclear magnetic resonance) is not fully appropriate (Table 1). The only experimental data available for a complete bovine Hsc70[36] (650 amino acids) and *E. coli* DnaK[37] (638 amino acids) Hsp70 chaperones are the low-resolution SAXS data.[36,37] Two quantities are typically extracted from SAXS: the radius of gyration R_g (Table 2) and the probability distribution $P(r)$ of the distances r between the heavy atoms (Figure 6).

Radius of Gyration R_g. In refs 36 and 37, the radii of gyration R_g was evaluated from the SAXS data by using both the Guinier approximation [$\ln (I(Q)/I(0) = -\exp(-R_g^2Q^2/3)$, where $I(Q)$ is the intensity collected at the (small) scattering wave-vector Q] and by computing numerically the second moment of the probability distribution $P(r)$ [$R_g^2 \equiv \int dr\, r^2\, P(r)$]. The comparison of the two methods provides an estimation of the actual error bar on R_g. The difference between the values of the R_g evaluated by the two methods from the same set of experimental data is about 3.6 Å for ADP-bHSc70 (closed) and 1.9 Å for ATP-bsHc70 (open) (Table 2). On the other hand, for ADP-bHsc70, two identical experiments gave different answers: $R_g = 34.9$ Å (Guinier) and $R_g = 37.5$ Å (Guinier), which points to a possible dependence on the protein sample preparation. The largest dispersion of the values of the gyration radius is observed

Table 2. Experimental Radii of Gyration R_g (Å)

	ADP			ATP		
protein	Guinier	$P(r)$	XRD	Guinier	$P(r)$	XRD
bHsc70[36]	34.9 ± 0.2	38.5 ± 0.5		32.8 ± 0.3	34.7 ± 0.8	
	37.5 ± 1.0	41.7 ± 0.5		33.4 ± 0.2	34.5 ± 0.6	
DnaK[32,37]	45.4 ± 1.5		37.6	39.8 ± 1.0	40.5 ± 1.0	
Sse1 (Hsp110)[18]						31.4
bHsc70 NBD (1−386)[24,36]	22.8 ± 0.2	22.6 ± 0.2	21.4	22.6 ± 0.1	22.2 ± 0.2	21.2
	23.1 ± 0.2	22.9 ± 0.2				
NBD *in* DnaKΔC[32]			21.1			
NBD *in* Sse1[18]						20.5
hHsp70 NBD[26]			21.3			
SBD *in* DnaK[32]			20.6			
SBD *in* Sse1[18]						38.9

for the full structures. For instance, the gyration radius of the full-length DnaK evaluated from the Guinier approximation from SAXS data is 45.4 Å, whereas the gyration radius of the template of the closed model (DnaKΔC) is 37.6 Å from the XRD data (the difference cannot be accounted for by the 30 missing residues in DnaKΔC, since the difference

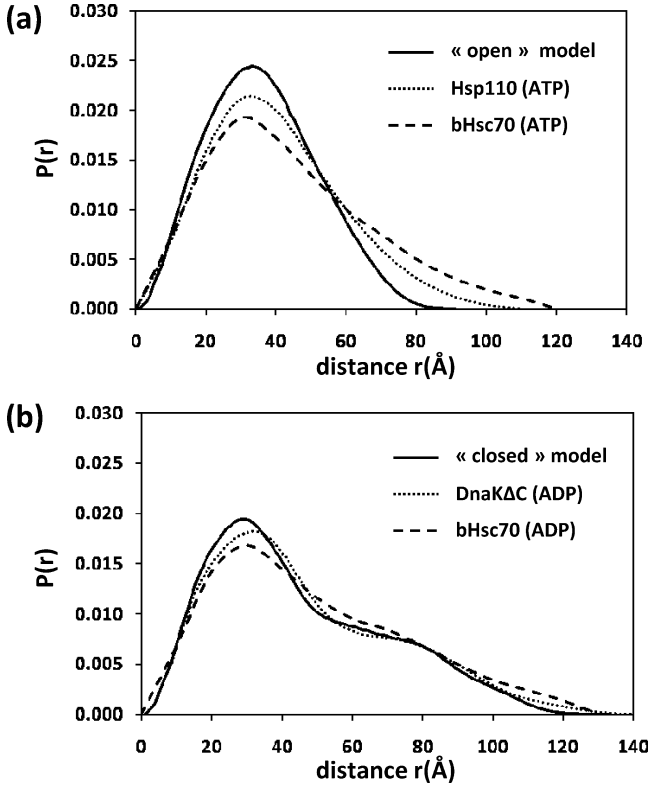


Figure 6. Comparison of the distance distribution functions $P(r)$ computed for the different models of hHsp70 relaxed by MD with those derived from SAXS data for bHsc70 and those computed for the homologue proteins Hsp110 and DnaK. The function $P(r)$ of bHsc70 in the ATP and ADP states was digitized from Figure 4 of ref 36. The function $P(r)$ for the homologue proteins Hsp110 (PDB ID: 3C7N A) in the ATP state and DnaKΔC in the ADP state (PDB ID: 2KHO) were computed from the structures deposited in PDB. (a) Distance distribution function $P(r)$ of the open relaxed model (apo; solid line), Hsp110 (locked in the ATP state; dotted line), and bHsc70 (locked in the ATP state; dashed line). (b) Distance distribution function $P(r)$ of the closed relaxed model (apo; solid line), DnaKΔC (locked in the ADP state; dotted line), and the bHsc70 (locked in the ADP state; dashed line).

Table 3. Computed Radii of Gyration R_g (Å)

MD simulations		hHsp70	NBD in hHsp70	SBD in hHsp70
open model	APO-1	26.4	21.2	28.9
	APO-2	27.5	20.4	26.7
	ADP-1	27.4	20.7	30.3
	ADP-2	27.0	20.5	29.9
closed model	APO	34.1	20.2	19.4

between the radius of gyration of the full length structure of hHsp70 simulated by MD in a closed state and the same structure of hHsp70 with the last 30 residues removed is only 1.1 Å). Taking all of this into consideration, we estimate an error bar on the radius of gyration for the full structure of Hsp70 of at least 3−4 Å. Contrary to our findings for the full structure, the radii of gyration for the isolated NBD of Hsp70 measured by SAXS in solution or by XRD in crystals differ only by 1−2 Å (Table 2).

In all of the SAXS experimental data,[36,37] ATP-Hsp70 is found to be systematically more compact than the ADP-Hsp70 state (Table 2). Other experimental data, such as the measurement of tryptophan fluorescence[17,59,60] and proteolytic susceptibility,[17,61] support the hypothesis of a change in protein shape between the ADP and ATP states of Hsp70. The difference in size is due to a change of the radii of gyration of the SBD between the two states. Indeed, the R_g's of the isolated NBD of bHc70 with ATP and ADP are very similar in both SAXS and XRD data (Table 2). The sizes of the NBD isolated and within the crystallized structure of DnaKΔC and that of the full length Hsp110 are also similar. On the other hand, the opening of the SBD increases its radius of gyration by about 20 Å, as we roughly estimated by comparing the size of the SBD within ADP-DnaKΔC and within ATP-Hsp110 (Table 2). The magnitude of the change is probably overestimated because in Hsp110 the "A+B" helix is not broken (Figure 1b).

In agreement with the SAXS data,[36,37] the average stable structures of hHsp70 relaxed by MD in an open state (Figure 3) are more compact than the structure in a closed state (Figure 5). The values of the R_g computed from MD (Table 3) for the open model varied between 26.5 and 27.5 Å between the four runs (for comparison, the closest value found in SAXS for ATP-bHsc70 is 32.8 Å, Table 2), and the R_g of the closed model was 34.1 Å (for comparison, the closest value found in SAXS for ADP-bHsc70 is 34.9 Å). The radius of gyration of the NBD domain within the full length hHsp70 model did not vary between the open and

closed states, in agreement with the values of the radii of gyration of ATP-bHsc70 and ADP-bHsc70 (Table 2), and its value is close to the experimental value. The size of the SBD within the closed model of hHsp70 is similar to that of DnaKΔC (Table 3). The opening of the SBD increases the radius of gyration of hHsp70 by about 9−10 Å (Table 3).

Distance Distribution Function P(r). The function $P(r)$ of ATP-bHsc70[36] measured by SAXS and computed for the structure of the yeast ATP-Hsp110[18] (PDB ID: 3C7N A), both in an open state, are compared to the $P(r)$ computed from the open model of hHsp70 relaxed by MD in Figure 6a [$P(r)$ from run APO-1 is shown; similar results were obtained from the four MD runs of the open model (data not shown)]. The $P(r)$ functions extracted from SAXS data and from the experimental structure of Hsp110 have one peak at about 35 Å (Figure 6a) which might reflect the binding of NBD on SBD. Indeed, the $P(r)$ computed for the open model of hHsp70 has a peak at the same position, and the overall shape of $P(r)$ agrees quite well with the one measured for bHsc70 by SAXS. The ATP-Hsp structures (Figure 6a) are substantially contracted structures by comparison with ADP-Hsp structures (Figure 6b). Indeed, the function $P(r)$ measured by SAXS for ADP-bHsc70[37] and the function $P(r)$ calculated from the experimental structure of ADP-DnaKΔC[32] (PDB ID: 2KHO; truncated by 30 residues in the C-terminal part) have two peaks: one at 30 Å and a distinct shoulder at 80 Å (Figure 6b), suggesting a structure with a bilobal shape. The function $P(r)$ for the closed model of hHsp70 has also a second peak around 80 Å as in the ADP molecular chaperones.

The agreement between theory and experiment is satisfactory taking into account that we compared simulations of human Hsp70 with experimental data of homologue proteins (bovine, bacteria, and yeast). In addition, one must take into account the necessary inaccuracy of any MD force field, the finite length of the MD simulations, and the difference between the idealized solvent simulated (pure water and ions) and the experimental one. Indeed, the experimental conditions can influence the size (R_g) and shape [$P(r)$] of the protein due to the fact that the solution contains both ADP and ATP in addition to buffers such as Tris·HCl,[36,37] whereas in our simulation, there are only water molecules and counterions. In addition, as stated by the authors in ref 37, the steady-state solution of ATP-bHsc70 is in fact a mixture of the protein in the open and closed conformations.

Comparison of the relaxed open (Figure 3) and the closed models (Figure 5) with all experimental data[36,37,59−62] confirms that the former is a good representative model of ATP-Hsp70, whereas the latter is a good representative model of ADP-Hsp70.

3.4. Conformational Changes between the Open and the Closed Conformations of Hsp70. The relaxed models of Hsp70 represent initial and final structures of a conformational transition of the functional cycle (see next section) of this chaperone. The conformational change is induced *in vivo* by several factors: nucleotide exchange (ATP↔ADP),[16,17] hydrolysis, and all steps being assisted by cochaperones.[18−20]

In an attempt to identify the functionally important motions for the transition between the open and closed state, we computed the collective modes of hHsp70 by applying the anisotropic network model (ANM)[41−46] to the open model and to the closed model obtained after relaxation by MD (see Methods). As shown for a few proteins, the collective modes of the structures in the initial and final states of a conformational change contain information on the dynamic of the transition.[56,63−68] Because the collective modes have a low restoring force constant and correspond to displacements of large structural subunits of a protein, a simplified elastic model (ANM) is sufficient to reproduce these modes.[41−46]

Transition from the Open State to the Closed State. First, we computed the collective modes of the open model and their relevance (involvement coefficient, see Methods) for the transition from the open state to the closed state of hHsp70. The theory is discussed in detail in the Methods section. In brief, a linear pathway interpolating between the two conformations (open and closed) was built. For each collective mode k, the projection of the atomic displacements within the mode k on the interpolated pathway defined the involvement coefficient I_k of the mode (the maximum value is 1, corresponding to a perfect match between the displacements of the atoms within the mode and the interpolating pathway). The sum of the square involvement coefficient of each mode up to an index k is the cumulative involvement coefficient CI_k (see Methods). The cumulative involvement coefficient indicates that the first 10 and the first 100 slow modes of a total of the 1917 modes of nonzero frequency of hHsp70 account for 45% ($CI_{10} = 0.45$) and 69% ($CI_{100} = 0.69$) of the displacement from the open to the closed state, respectively (Figure 7a).

The mode contributing the most to the transition between the open state and the closed state is the mode having the lowest nonzero "frequency" λ_1 (mode 1), which has an involvement coefficient of 0.62 for the transition from the open state to the closed one (Figure 7b). In this mode, the NBD essentially moves as a rigid unit, whereas the SBD is the most mobile part. The amplitudes of the C^α B factors (see eq 5 in the Methods) for this mode (Figure 8a) clearly show that only very large fluctuations are present in the SBD-β and in the SBD-α, which correspond to the substrate pocket and the lid, respectively. In SBD-β, C^α B factors show four peaks for residues 405, 433, 467, and 496 which correspond to important fluctuations in β strands. In SBD-α, the helix A and the C-terminal part are completely rigid (residues 508−525 for helix A and residues 616 to 641 for the C-terminal part in Figure 8a), and helix A forms a hinge region. The global motion described by the lowest frequency mode shows that the helix A of SBD-α serves as a hinge region around which SBD-β and the rest of SBD-α (helix B + C + D) move toward (Figure 9). SBD-β and SBD-α move in opposite directions from each other (Figure 9), and R_g is nearly constant (Figure 10a). For all of the structures of the open model relaxed by all-atom MD (runs APO-1, APO-2, ADP-1, and ADP-2), the transition from the open state to the closed state was typically described by the same collective mode (Figure 9): we observed the rotation of the

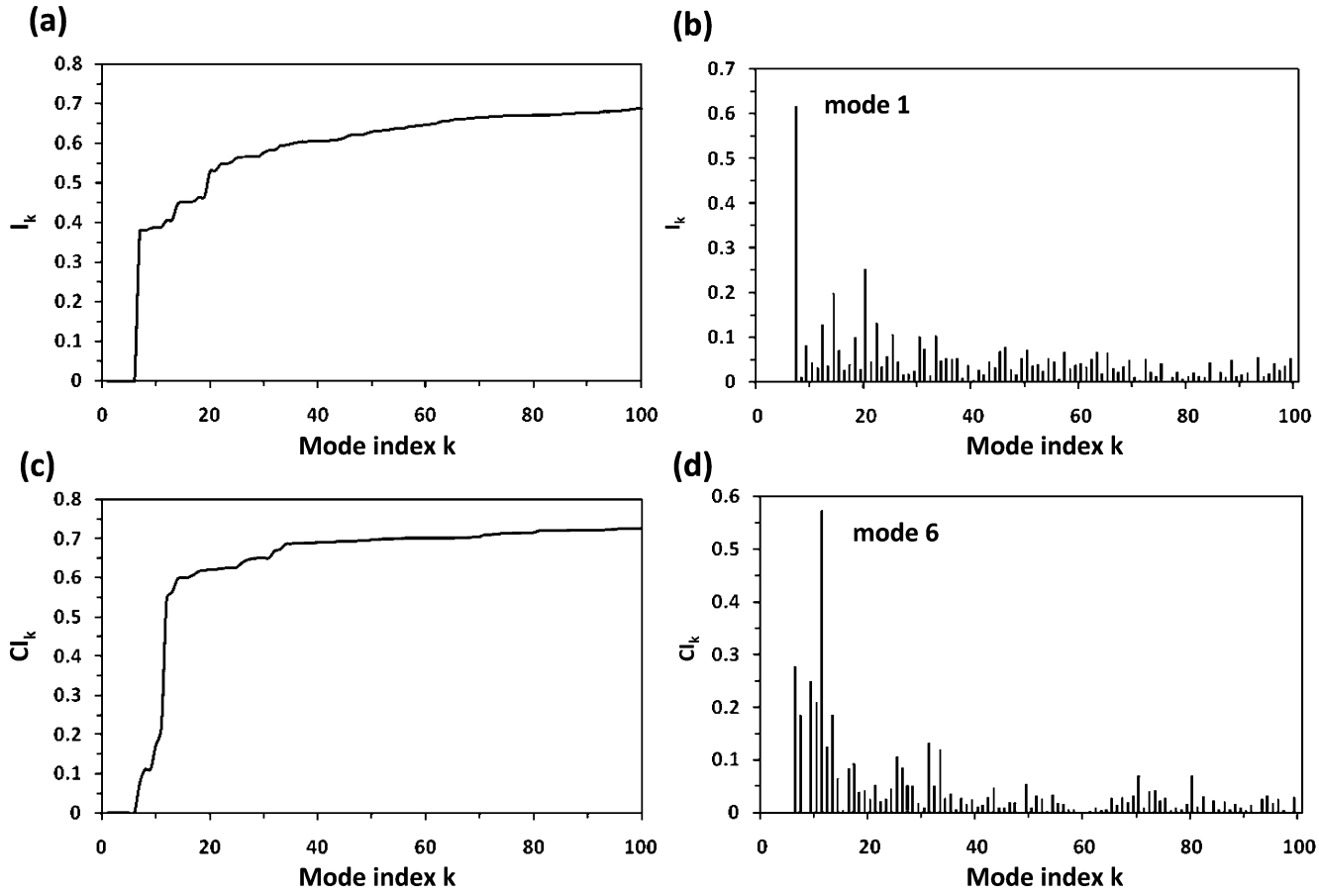


Figure 7. Individual and cumulative involvement coefficients (eqs 9 and 10 in Methods). The individual (a) and cumulative involvement coefficients (c) of the first 100 modes of nonzero frequency for the transition from the open state to the closed state of human Hsp70. The modes are computed for the average structure of the open model calculated between 100 and 400 ns of the MD run APO-1. The individual (b) and cumulative involvement coefficients (d) of the first 100 modes for the transition from the closed state to the open state of human Hsp70. The modes are computed for the average structure of the closed model calculated between 30 and 130 ns of the MD run.

SBD-α in the direction of the substrate binding pocket, which simulates the closing of the pocket by the lid.

Transition from the Closed State to the Open State. Second, we computed the modes for the relaxed closed model in order to study the transition from the closed to the open state. The first 10 and the first 100 collective modes of a total 1917 nonzero frequency modes of hHsp70 account for 60% ($CI_{10} = 0.6$) and 73% ($CI_{100} = 0.73$) of the displacement from the closed to the open state, respectively (Figure 7c). In the closed state, the mode contributing the most to the transition between the closed state and the open state is the mode having the lowest "frequency" λ_6 (mode 6), which has an involvement coefficient of 0.57 (Figure 7d). This motion tends to decrease ($p > 0$) or to increase ($p < 0$) the radius of gyration of the protein (Figure 10a). Indeed, this motion corresponds to a compression of the two domains ($p > 0$, Figure 11a), which restricts the mobility of the linker, or to an elongation of the two domains [p < 0, not shown]. This mode does not correspond to a direct opening of the lid although there are important fluctuations in the SBD as shown by the *B* factors computed for this mode (Figure 8b).

The analysis of the *B* factors shows clearly that mode 6 involves displacements of residues both in the NBD and in the SBD (Figure 8b). In particular, residues 190, 214, 247−249, and 289−291 in the NBD are very mobile. These residues belong to subdomain II: the most mobile residues (residues 247−248, 289−291) are in subdomain IIB, and the others (residues 190, 214) in subdomain IIA. The ribbon diagrams displayed in Figure 11b show the displacement of subdomain IIB in the closed to open transition. The large *B* factors of residues 190, 214, 247−249, and 289−291 are due to the motion of subdomain IIB of the NBD, which rotates and opens slightly the nucleotide pocket in NBD (Figure 11b). We quantified the opening of the NBD by computing the angle δ formed by the centers of mass of subdomains IB, IA/IIA, and IIB.[69] The angle δ increased (for $p > 0$) for mode 6, confirming the opening of NBD (Figure 10b). Contrarily, in the transition from the open state to the closed state of hHsp70, the angle δ did not vary for mode 1 of the open structure of hHsp70, as shown in Figure 10b. On the other hand, the average value of angle δ is larger when a nucleotide is bound to the NBD: the average value of δ of APO-hHsp70 was a few degrees larger than the average value of δ of ADP-hHsp70 in our MD simulations. In recent MD simulations of isolated NBD of hHsp70, angle δ of a nucleotide-free NBD was also found to be 10° larger than that of an ATP-NBD.[69]

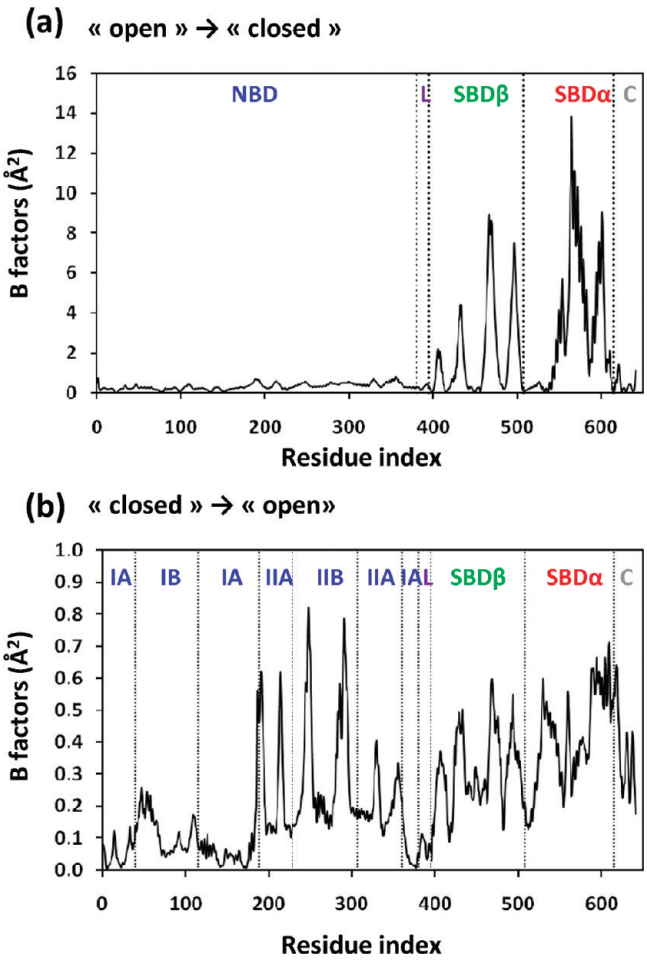


Figure 8. *B* factors (Å^2) computed at 300 K [eq 5 in the Methods]. (a) Computed *B* factors for the collective mode 1 of the open model of the hHsp70 chaperone having the largest involvement coefficient in the transition from the open state to the closed state (Figure 7b). (b) Computed *B* factors for the sixth slowest mode, 6, of the closed model of hHsp70 having the largest involvement coefficient in the transition from the closed state to the open state (Figure 7d). The dotted lines show the different subdomains of hHsp70 in agreement with Figure 1.

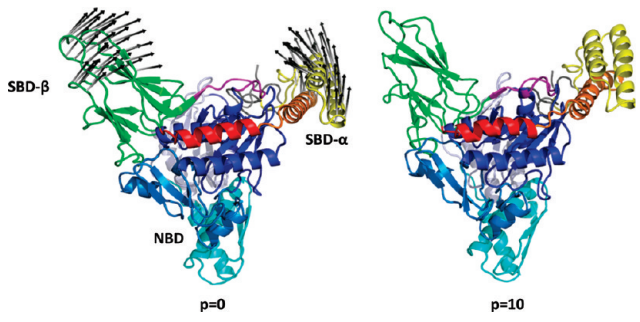


Figure 9. Graphical representation of the collective mode λ_1 of the open model of hHsp70 chaperone. The color code is the same as in Figure 1a. Eigenvectors are represented by black arrows. Vector norms have been uniformly scaled by an arbitrary factor, and only vectors with the higher norms are represented for clarity. The quantity "p" represents the amplitude of the mode in arbitrary units. This figure was prepared with PyMOL [http://www.pymol.org].

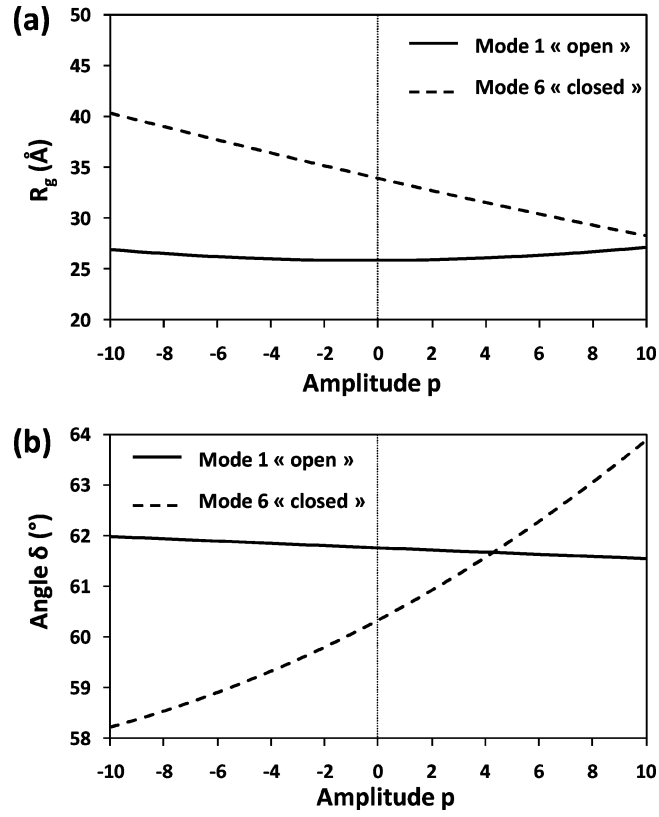


Figure 10. (a) Evolution of the radius of gyration (R_g) of the structure in the function of the amplitude of mode 1 (open → closed; solid line) and of mode 6 (closed → open; dashed line). (b) Evolution of the opening angle δ of the NBD (see text) as a function of the amplitude of modes 1 and 6. The representation is the same as in panel a.

The large *B* factors of residues 530 and 608−609 in mode 6 (Figure 8b) are due to motion that involves the displacement of the helix bundle (C + D) and the C-terminal part moving in the opposite direction of SBD-β (Figure 11c).

Three other modes contribute significantly to the transition from the closed to the open state (Figure 7d). The first (λ_1), the fourth (λ_4), and the fifth (λ_5) lowest frequency modes have involvement coefficients of 0.28, 0.25, and 0.21, respectively (Figure 7d). In the first mode (λ_1), the NBD and the SBD move toward each other with a "scissor" movement, in which the linker is a rigid unit (data not shown). The most mobile part in the NBD is subdomain IIA, whereas in the SBD, the SBD-α and the C-terminal part are the most mobile. The fourth mode (λ_4) corresponds to an opening of the SBD of the protein (data not shown). The most mobile part in the SBD is the SBD-α and the C-terminal part. Finally, in the fifth mode (λ_5), the NBD and the SBD move toward each other with an important motion in the two domains (data not shown): in the NBD, lobe II is the most mobile part, and the motion in the SBD corresponds to an opening of the lid with huge displacements in the SDB-α. These three modes have in common a motion within the SBD and a coupling between the motion of the SBD and the motion of the NBD, principally via the rotation of lobe II of the NBD.

The study of the transition from the closed state to each of the different other structures of the open model (runs APO-

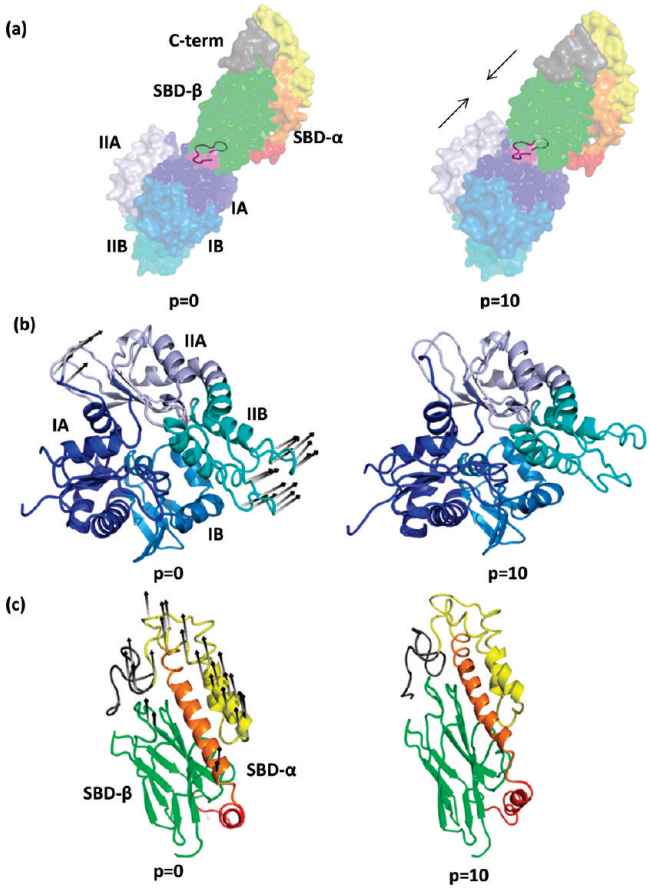


Figure 11. Graphical representation of mode 6 of the closed model of the hHsp70 chaperone. The color code is similar to the one in Figure 1a. (a) Surface diagram of the hHsp70 chaperone. The linker is also represented with a ribbon. (b) Ribbon diagram of the NBD. Eigenvectors are represented by black arrows. Vector norms have been uniformly scaled by an arbitrary factor, and only vectors with higher norms are represented for clarity. (c) Ribbon diagram of the SBD. The vectors representation is the same as in panel b. The quantity "p" represents the amplitude of the mode in arbitrary units. This figure was prepared with PyMOL [http://www.pymol.org]

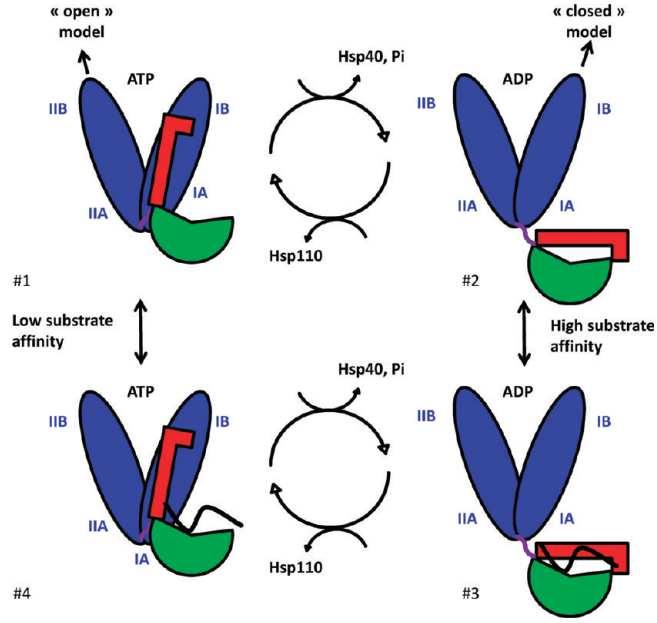


Figure 12. Schematic representation of the Hsp70 cycle. The color code is as follows: NBD = blue, SBD-β = green, SBD-α = red, the so-called linker = purple, and the substrate peptide in black. The two models of human Hsp70 were built by homology from the experimental data of the proteins (PDB ID: 3C7N, open model; 2KHO, closed model) and relaxed by MD, corresponding to steps 1 and 2 in the cycle, as indicated by arrows. Adapted with permission from ref 22, Figure 6a. Copyright 2007 Cell Press.

1, APO-2, ADP-1, and ADP-2) gave similar results (data not shown). For each run, the sixth lowest frequency mode contributes the most to the transition with the same motion shown in Figure 11. One emphasizes that the collective modes involved in the transition from the open state to the closed state and from the closed state to the open state are very different (Figures 8−11) as well as the dependence of the cumulative involvement coefficient with the mode index (Figure 7).

3.5. Hsp70 Chaperone Cycle and Interdomain Communication. The two nucleotides have similar binding energy to Hsp70; the affinity of ATP is slightly higher [the (apparent) Gibbs free energies at 25 °C of the binding of ATP and ADP to a mutant of Dnak lacking hydrolysis activity are $\Delta G = -39$ kJ/mol and $\Delta G = -37.7$ kJ/mol, respectively].[70] The binding of ATP to a nucleotide free Hsp70 induces a conformational change.[17,37,60] The hydrolysis of the bound ATP converts slowly ATP-Hsp70 into ADP-Hsp70.[17,70] An exchange reaction between the products of hydrolysis (ADP and inorganic P) and ATP in solution reinitializes the reaction. The rate of hydrolysis and the rate of the nucleotide exchange are increased by several orders of magnitude by cochaperones,[16] and the binding of the polypeptide to ATP-Hsp70 stimulates the ATP hydrolysis.[17]

A detailed mechanistic model of these different steps of the ATPase activity and of the chaperoning cycle (Figure 12]) is unknown despite important insights obtained recently from the structures of Hsp70 fragments and complexes of different species. Due to the large number of Hsp70 species, states, and complexes, studied by different experimental techniques, a qualitative scenario of the allosteric mechanism has been proposed.[17,18,21,22,71] However, the absence of structures of Hsp70+ATP leaves unanswered the mechanism by which the protein substrate triggers the conformational change observed by SAXS[36,37] and depicted by the elongation of a linker between the NBD and SBD in moving from steps 1 to 2 or 4 to 3 (Figure 12). Therefore, our two models and simulations of the conformational dynamic of human Hsp70 may provide important insights about transition pathways that experimental data could not reveal so far.

In the present paper, we consider only the chaperoning cycle without a substrate and without cochaperones corresponding to steps 1 and 2 in Figure 12. In ADP-Hsp70 without a substrate (step 2 in Figure 12), represented by hHsp70 in a closed state (Figure 5), the linker is bound to SBD-β. Consequently, the SBD and the NBD of hHsp70 are not highly bound and move rather freely: in our MD simulations, the lowest frequency collective modes of the closed state correspond to different global motions of the

NBD relative to the SBD. The mobility of the NBD relative to the SBD is in agreement with NMR data for ADP-DnaK.[22] When ADP is exchanged by ATP, from step 2 (ADP-Hsp70) to step 1 (ATP-Hsp70), the protein undergoes large conformational changes in both domains,[16,17,22,36,37] and the linker is sequestered from the solvent and binds in the IA/IIA surface cleft as we found for hHsp70 in the open state (Figure 3b). In the open state, we found that the SBD and the NBD are highly bound. The lowest frequency mode of the open state corresponds to oscillation of the substrate binding cleft with little variation of the structure size (Figure 10a).

The collective modes help us to understand the transition from step 2 to step 1 and the interdomain communication. Indeed, we observe a rotation of subdomains IIA and IIB in the NBD (Figure 11b) as well as a rotation of the SBD-α (Figure 11c) in mode λ_6. These two motions suggest that the transition from step 2 to 1 (Figure 12) is assisted by a collective motion which couples a conformational change in the ATPase domain of Hsp70 to the polypeptide binding site.

The rotation of subdomains IIA and IIB in the collective mode λ_6 driving the transition from step 2 to step 1 agrees with various experimental data.[22,71] Rotations in the subdomain IIB have been suggested to be the allosteric mechanism in Hsp70 as NMR measurement of truncated structures of bacteria Hsp70[60,71] reveal a rotation of this subdomain intrinsically or after nucleotide exchange. Several nucleotide exchange factors (cochaperones Bag-1 and Hsp110) induce a rotation of IIB upon binding to ADP-Hsp70.[18−20] In particular, subdomain IIB is a key domain of the interaction between Hsp70 and the nucleotide exchange factor (NEF) and cochaperone Hsp110.[18−20] The subdomains IIB, IIA, and IB of the NBD of Hsp110 make contact with subdomains IB, IA, and IIB of Hsp70, respectively. Moreover, the helix "A+B" of the SBD-α of Hsp110 (Figure 1b) contacts subdomain IIB of Hsp70. Therefore, subdomain IIB of Hsp70 is clamped between helix "A+B" of the SBD-α and subdomain IB of Hsp110, and a rotation by 20° allows the exchange of ADP by ATP.[18] The displacement of subdomain IIA in the collective mode from closed to open states (Figure 11b) is also of great interest,[71] since it affects a hydrophobic surface cleft between subdomains IA and IIA, which is likely involved in the binding of the linker of Hsp70 to its NBD.[21,22,72] The present study indicates that the rotation of subdomain IIB is excited thermally even in the absence of a cochaperone. Consequently, the slow modes which described the transition from the ADP (step 2) to the ATP state (step 1) in the chaperone cycle provide a detailed mechanistic description of the allosteric mechanism, in agreement with NMR data for isolated NBD domains of DnaK.[71] In the opposite transition from step 1 to step 2, there is a conformational change in the SBD, which is the closing of the substrate binding pocket by the lid. In the present study, the thermal excitation of mode λ_1 and the binding of ADP to the open model of hHsp70 do not generate a force large enough to displace the linker from the buried position adopted during the open state. This suggests that binding of the peptide substrate to the open state and hydrolysis might be necessary to close the SBD. On the other hand, the binding of the polypeptide substrate could be coupled to the collective mode λ_1 which closes the SBD pocket (Figure 9).

4. Conclusion

The purpose of the present paper was to obtain stable atomic models of hHsp70 in different states in the absence of cochaperones. Two typical structures corresponding to two steps (1 and 2) of the chaperoning cycle (Figure 12) were reported here. They correspond to human Hsp70, in an open state and a closed state, respectively. The full-length closed and open states models were built by homology modeling[47−50] using as templates the experimental structures of DnaK[32] (closed model; PDB ID: 2KHO) and the homologous protein Hsp110[18] (open model; PDB ID: 3C7N, chain A). The initial models were relaxed by using all-atom MD simulations at 300 K in explicit solvent (see Methods). We obtained stable structures (with and without ADP) by using four different initial conditions for the open state of hHsp70. In the four MD trajectories, the open model adopts a stable conformation in which the lid is open (Figure 3). Different stable binding sites were found for the SBD-α on the NBD, suggesting that hHsp70 in an open state might exist in different conformations. A comparison with experimental data, particularly SAXS[36,37] (Figure 6), suggests that the relaxed open model corresponds to an ATP state of the chaperone cycle. The closed model (2KHO template)[32] was relaxed by all-atom MD simulations at 300 K in explicit solvent (see Methods). The stable structure found by MD has the SBD-α closing the substrate binding pocket located in the SBD-β; this stable structure represents an ADP state of the hHsp70 chaperone.[32]

By assuming that the open state of hHsp70 represents a reasonable model for the ATP-hHsp70 state, we can suggest an explanation of the interdomain communication. The present study of collective modes involved in the transition from step 1 to step 2 of the chaperone cycle (Figure 12) and vice versa reveals one dynamical domain of Hsp70: the SBD. The first lowest frequency mode of the structure in an open state conformation contributing the most to the transition from 1 to 2 of the cycle (Figure 12) corresponds to a closing of the SBD-β by the SBD-α (Figure 9). The study of the transition from step 2 to step 1 of the Hsp70 cycle (Figure 12) is dominated by one mode of the structure in a closed state conformation which involved both the NBD and the SBD (Figures 8 and 11). The opening of the substrate binding pocket in this mode causes simultaneously a rotation of subdomains IIA and IIB in the NBD; therefore, it means that this mode provide a means for distant residues in the NBD and SBD to communicate at a low energy cost (weak restoring forces) in the transition from the closed to the open state of human Hsp70. The present calculations demonstrated, for the first time, that the rotation of subdomains IIA and IIB is thus related to the opening of the SBD-α and to an oscillation of the length structure (R_g).

Acknowledgment. P.S. thanks the Centre National de la Recherche Scientifique (CNRS) for support (délégation CNRS). This research was conducted using the computational resources of the "Centre de Ressources Informatiques" (CRI), Université de Bourgogne. This work was granted to the HPC

ressources of CINES under the allocation 2009-c2009076161 made by GENCI (Grand Equipement National de Calcul Intensif). A.N. thanks the French Ministry of Research and Education for a Ph.D. fellowship. The authors thank Thomas Gaillard for providing a PYMOL script to represent the normal modes.

References

(1) Hartl, F. U.; Hayer-Hartl, M. Molecular chaperones in the cytosol, from nascent chain to folded protein. *Science* **2002**, *295*, 1852–1858.

(2) Young, J. C.; Agashe, V. R.; Siegers, K.; Hartl, F. U. Pathways of chaperone mediated protein folding in the cytosol. *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 781–791.

(3) Bukau, B.; Deuerling, E.; Pfund, C.; Craig, E. A. Getting newly synthesized proteins into shape. *Cell* **2000**, *101*, 119–122.

(4) Saibil, H. R. Chaperone machines in action. *Curr. Opin. Struct. Biol.* **2008**, *18*, 35–42.

(5) Mayer, M. P.; Bukau, B. Hsp70 chaperones, cellular functions and molecular mechanism. *Cell. Mol. Life Sci.* **2005**, *62*, 670–684.

(6) Parsell, D. A.; Lindquist, S. The function of heat-schock proteins in stress tolerance, degradation and reactivation of damaged proteins. *Annu. Rev. Genet.* **1993**, *27*, 437–496.

(7) Qian, S.; McDonough, B.; Boellmann, H. F.; Cyr, D. M.; Patterson, C. CHIP-mediated stress recovery by sequential ubiquitination of substrates and Hsp70. *Nature* **2006**, *440*, 551–555.

(8) Young, J. C.; Barral, J. M.; Hartl, F. U. More than folding localized functions of cytosolic chaperones. *Trends Biochem. Sci.* **2003**, *28*, 541–547.

(9) Broadley, S. A.; Hartl, F. U. The role of molecular chaperones in human misfolding diseases. *FEBS Lett.* **2009**, *583*, 2647–2653.

(10) Warrick, J. M.; Chan, H. Y. E.; Gray-Board, G. L.; Chai, Y.; Paulson, H. L.; Bonini, N. M. Suppression of polyglutamine-mediated neurodegeneration in drosophila by the molecular chaperone Hsp70. *Nat. Genet.* **1999**, *23*, 425–428.

(11) Muchowski, P. J.; Schaffar, G.; Sittler, A.; Wanker, E. E.; Hayer-Hartl, M.; Hartl, F. U. Hsp70 and Hsp40 chaperones can inhibit self-assembly of polyglutamine proteins into amyloid-like fibrils. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 7841–7846.

(12) Jana, N. R.; Tanaka, M.; Wang, G.; Nukina, N. Polyglutamine length-dependent interaction of Hsp40 and Hsp70 family chaperones with truncated N-terminal huntingtin, their role in suppression of aggregation and cellular toxicity. *Hum. Mol. Genet.* **2000**, *9*, 2009–2018.

(13) Kim, S.; Nollen, E. A.; Kitagawa, K.; Bindokas, V. P.; Morimoto, R. I. Polyglutamine protein aggregates are dynamic. *Nat. Cell Biol.* **2002**, *4*, 826–831.

(14) Javid, B.; MacAry, P. A.; Lehner, P. J. Structure and function, Heat Shock Proteins and adaptative immunity. *J. Immunol.* **2007**, *179*, 2036–2040.

(15) Schmitt, E.; Gehrmann, M.; Brunet, M.; Multhoff, G.; Garrido, C. Intracellular and extracellular functions of heat shock proteins, repercussions in cancer therapy. *J. Leukocyte Biol.* **2007**, *81*, 15–27.

(16) Brehmer, D.; Rudiger, S.; Gassler, C. S.; Klostermeier, D.; Packschies, L.; Reinstein, J.; Mayer, M. P.; Bukau, B. Tuning of chaperone activity of Hsp70 proteins by modulation of nucleotide exchange. *Nature* **2001**, *8*, 427–432.

(17) Buchberger, A.; Theyssen, H.; Schröder, H.; McCarty, J. S.; Virgallita, G.; Milkereit, P.; Reinstein, J.; Bukau, B. Nucleotide-induced conformational changes in the ATPase and substrate binding domains of the DnaK chaperone provide evidence for interdomain communication. *J. Biol. Chem.* **1995**, *270*, 16903–16910.

(18) Schuermann, P. J.; Jiang, J. W.; Cuellar, J.; Llorca, O.; Wang, L. P.; Gimenez, L. E.; Jin, S. P.; Taylor, A. B.; Demeler, B.; Morano, K. A.; Hart, P. J.; Valpuesta, J. M.; Lafer, E. M.; Sousa, R. Structure of the Hsp110:Hsc70 nucleotide exchange machine. *Mol. Cell* **2008**, *31*, 232–243.

(19) Polier, S.; Dragovic, Z.; Hartl, F. U.; Bracher, A. Structural basis for the cooperation of Hsp70 and Hsp110 chaperones in protein folding. *Cell* **2008**, *133*, 1068–1079.

(20) Liu, Q.; Hendrickson, W. A. Insights into Hsp70 chaperone activity from a crystal structure of the yeast Hsp110 Sse1. *Cell* **2007**, *131*, 106–120.

(21) Vogel, M.; Mayer, M. P.; Bukau, B. Allosteric regulation of Hsp70 chaperones involves a conserved interdomain linker. *J. Biol. Chem.* **2006**, *281*, 38705–38711.

(22) Swain, J. F.; Dinler, G.; Sivendran, R.; Montgomery, D. L.; Stotz, M.; Gierasch, L. M. Hsp70 chaperone ligands control domain association via an allosteric mechanism mediated by the interdomain linker. *Mol. Cell* **2007**, *26*, 27–39.

(23) Flaherty, K. M.; DeLuca-Flaherty, C.; McKay, D. B. Three-dimensional structure of the ATPase fragment of a 70K heat-shock cognate protein. *Nature* **1990**, *346*, 623–628.

(24) Flaherty, K. M.; Wilbanks, S. M.; DeLuca-Flaherty, C.; McKay, D. B. Structural basis of the 70-kilodalton heat shock cognate protein ATP hydrolytic activity. *J. Biol. Chem.* **1994**, *269*, 12899–12907.

(25) Wisniewska, M.; Karlberg, T.; Lehtio, L.; Johansson, I.; Kotenyova, T.; Moche, M.; Schüler, H. Crystal structures of the ATPase domains of four human Hsp70 isoforms: HSPA1L/Hsp70-hom, HSPA2/Hsp70−2, HSPA6/Hsp70Bμ, and HSPA5/BiP/GRP78. *PLoS ONE* **2010**, *5* (1), e8625.

(26) Osipiuk, J.; Freeman, C.; Morimoto, R. I.; Joachimiak, A. Structure of a new crystal form of human Hsp70 ATPase domain. *Acta Crystallogr.* **1999**, *D55*, 1105–1107.

(27) Wang, H.; Kurochkin, V.; Pang, Y.; Hu, W.; Flynn, G. C.; Zuiderweg, E. R. P. NMR solution structure of the 21 kDa chaperone protein DnaK substrate binding domain, a preview of chaperone-protein interaction. *Biochem.* **1998**, *37*, 7929–7940.

(28) Chou, C. C.; Forouhar, F.; Yeh, Y. H.; Shr, H. L.; Wang, C.; Hsiao, C. D. Crystal structure of the C-terminal 10-kDa subdomain of Hsc70. *J. Biol. Chem.* **2003**, *278*, 30311–30317.

(29) Zhu, X.; Zhao, X.; Burkholder, W. F.; Gragerov, A.; Ogata, C. M.; Gottesman, M. E.; Hendrickson, W. A. Structural analysis of substrate binding by the molecular chaperone DnaK. *Science* **1996**, *272*, 1606–1614.

(30) Morshauer, R. C.; Hu, W.; Wang, H.; Pang, Y.; Flynn, G. C.; Zuiderweg, E. R. P. High-resolution solution structure of the 18 kDa substrate-binding domain of the mammalian chaperone protein Hsc70. *J. Mol. Biol.* **1999**, *289*, 1387–1403.

(31) Jiang, J.; Prasad, K.; Lafer, E. M.; Sousa, R. Structural basis of interdomain communication in the Hsc70 chaperone. *Mol. Cell* **2005**, *20*, 513–524.

(32) Bertelsen, E. B.; Chang, L.; Gestwicki, J. E.; Zuiderweg, E. R. P. Solution conformation of wild-type E. coli Hsp70 (DnaK) chaperone complexed with ADP and substrate. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 8471–8476.

(33) Schmitt, E.; Maingret, L.; Puig, P. E.; Ghiringhelli, F.; Hammann, A.; Solary, E.; Kroemer, G.; Garrido, C. Hsp70 neutralization exerts potent anti-tumor effects in animal models of colon cancer and melanoma. *Cancer Res.* **2006**, *66*, 4191–4197.

(34) Didelot, C.; Lanneau, D.; Brunet, M.; Joly, A. L.; De Thonel, A.; Ghiosis, G.; Garrido, C. Anti-cancer therapeutic approaches based on intracellular and extracellular heat shock proteins. *Curr. Med. Chem.* **2007**, *14*, 1–9.

(35) Leu, J. I. J.; Pimkina, J.; Frank, A.; Murphy, M. E.; George, D. L. A small molecular inhibitor of inducible Heat shock protein 70. *Mol. Cell* **2009**, *36*, 15–27.

(36) Wilbanks, S. M.; Chen, L.; Tsuruta, H.; Hodgson, K. O.; McKay, D. B. Solution small-angle x-ray scattering study of the molecular chaperone Hsc70 and its subfragment. *Biochem.* **1995**, *34*, 12095–12106.

(37) Shi, L.; Kataka, M.; Fink, A. L. Conformational characterization of DnaK and its complexes by small-angle x-ray scattering. *Biochem.* **1996**, *35*, 3297–3308.

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

(39) Kitao, A.; Go, N. Investigating protein dynamics in collective coordinate space. *Curr. Opin. Struct. Biol.* **2000**, *9*, 164–169.

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

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

(42) Yang, L.; Song, G.; Jernigan, R. L. How well can we understand large-scale protein motions using normal modes of elastic network models. *Biophys. J.* **2007**, *93*, 920–929.

(43) Kim, M. K.; Jernigan, R. L.; Chirikjian, G. S. Efficient generation of feasible pathways for protein conformational transitions. *Biophys. J.* **2002**, *83*, 1620–1630.

(44) Kim, M. K.; Jernigan, R. L.; Chirikjian, G. S. Rigid-cluster models of conformational transitions in macromolecular machines and assemblies. *Biophys. J.* **2005**, *89*, 43–55.

(45) Yang, L.; Song, G.; Jernigan, R. L. Protein elastic network models and the ranges of cooperativity. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 12347–12352.

(46) Navizet, I.; Lavery, R.; Jernigan, R. L. Myosin flexibility: structural domains and collective vibrations. *Proteins* **2004**, *54*, 384–393.

(47) Altschul, S. F.; Madden, T. L.; Schäffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. Gapped BLAST & PSI-BLAST, a new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402.

(48) Sali, A.; Blundell, T. L. Comparative protein modeling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234*, 779–815.

(49) Sali, A.; Potterton, L.; Yuan, F.; Vanvlijmen, H.; Karplus, M. Evaluation of comparative protein modeling by MODELLER. *Proteins* **1995**, *23*, 318–326.

(50) Sanchez, R.; Sali, A. Comparative protein structure modeling. Introduction and practical examples with MODELLER. *Methods Mol. Biol.* **2000**, *143*, 97–129.

(51) Shindyalov, I. N.; Bourne, P. E. Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Eng.* **1998**, *11*, 739–747.

(52) Lindahl, E.; Hess, B.; van der Spoel, D. Gromacs 3.0: A package for molecular simulation and trajectory analysis. *J. Mol. Mod.* **2001**, *7*, 306–317.

(53) Schuettelkopf, A. W.; Van Aalten, D. M. F. PRODRG - a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr.* **2004**, *D60*, 1355–1363.

(54) Ben-Avraham, D. Vibrational normal-mode spectrum of globular proteins. *Phys. Rev. B* **1993**, *47*, 14559–14560.

(55) Svanidze, A. V.; Sashin, I. L.; Lushnikov, S. G.; Gvasaliya, S. N.; Turoverov, K. K.; Kuznetsova, I. M.; Kojima, S. Inelastic incoherent neutron scattering in some proteins. *Ferroelectrics* **2007**, *348*, 556–562.

(56) Tama, F.; Sanejouand, Y. H. Conformational change of proteins arising from normal mode calculations. *Protein Eng.* **2001**, *14*, 1–6.

(57) Baba, A.; Komatsuzaki, T. Construction of effective free energy landscape from single-molecule time series. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 19297–19302.

(58) Maisuradze, G. G.; Senet, P.; Liwo, A.; Scheraga, H. A. Investigation of protein folding by coarse-grained molecular dynamics with the UNRES force field. *J. Phys. Chem. A* **2010**, *114*, 4471–4485.

(59) Palleros, D.; Reid, K. L.; McCarty, J. S.; Walker, G. C.; Fink, A. L. DnaK, Hsp73, and their molten globules. Two different ways heat shock proteins respond to heat. *J. Biol. Chem.* **1992**, *267*, 5279–5285.

(60) Ha, J.-H.; McKay, D. B. Kinetics of nucleotide-induced changes in the tryptophan fluorescence of the molecular chaperone Hsc70 and its subfragments suggest the ATP-induced conformational change follows initial ATP binding. *Biochem.* **1995**, *34*, 11635–44.

(61) Liberek, K.; Marszalek, J.; Ang, D.; Georgopoulos, C.; Zylicz, M. Escherichia coli DnaJ and GrpE heat shock proteins jointly stimulate ATPase activity of DnaK. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 2874–2878.

(62) Revington, M.; Zhang, Y.; Yip, G. N. B.; Kurochkin, A.; Zuiderweg, E. R. P. NMR investigations of allosteric processes in a two domain thermus thermophilius Hsp70 molecular chaperone. *J. Mol. Biol.* **2005**, *349*, 163–183.

(63) Ma, J.; Karplus, M. Ligand-induced conformational changes in ras p21, a normal mode and energy minimization analysis. *J. Mol. Biol.* **1997**, *274*, 14–131.

(64) Cui, Q.; Li, G.; Ma, J.; Karplus, M. A normal mode analysis of structural plasticity in the biomolecular motor F1-ATPase. *J. Mol. Biol.* **2004**, *340*, 345–372.

(65) Gaillard, T.; Martin, E.; San Sebastian, E.; Cossio, F. P.; Lopez, X.; Dejaegere, A.; Stote, R. H. Comparative normal mode analysis of LFA-1 integrin I-domains. *J. Mol. Biol.* **2007**, *374*, 231–249.

(66) Cecchini, M.; Houdusse, A.; Karplus, M. Allosteric communication in myosin V: from small conformational changes to large directed movements. *PLOS Comput. Biol* **2008**, *4* (8), e1000129.

(67) Kong, Y.; Ma, J.; Karplus, M.; Lipscomb, W. N. The allosteric mechanism of a yeast chorismate mutase, a dynamic analysis. *J. Mol. Biol.* **2006**, *356*, 237–247.

(68) Ma, J.; Karplus, M. The allosteric mechanism of the chaperone GroEL: a dynamic analysis. *Proc. Natl. Acad. Sci. U.S.A.* 199895, 8502–8507.

(69) Woo, H. J.; Jiang, J.; Lafer, E. M.; Sousa, R. ATP-induced conformational changes in Hsp70: molecular dynamics and experimental validation of a silico predicted conformation. *Biochem.* **2009**, *48*, 11470–11477.

(70) Taneva, S. G.; Moro, F.; Velazquez-Campoy, A.; Muga, A. Energetics of nucleotide-induced DnaK conformational states. *Biochemistry* **2010**, *49*, 1338–1345.

(71) Bhattacharya, A.; Kurochkin, A. V.; Yip, G. N. B.; Zhang, Y.; Bertelsen, E. B.; Zuiderweg, E. R. P. Allostery in Hsp70 chaperones is transduced by subdomain rotations. *J. Mol. Biol.* **2009**, *388*, 475–490.

(72) Zhang, Y.; Zuiderweg, E. R. P. The 70-kDa heat shock protein chaperone nucleotide-binding domain in solution unveiled as a molecular machine that can reorient its functional subdomains. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 10272–10277.

CT1002169

JCTC Journal of Chemical Theory and Computation

A Note on the Standard State's Binding Free Energy

Ignacio J. General*

Department of Computational and Systems Biology, School of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania 15260

Received May 14, 2010

Abstract: The relation between the equilibrium constant of a given chemical reaction and the associated free energy is an issue well studied in chemistry books, but when the reaction involves changes in the number of components in the system, as is the case in binding, things become a little more obscure since one needs to define the so-called standard state. This is reflected in the literature, especially in computational studies of binding, where contradicting approaches are followed when treating this problem. In this work, we present a detailed and unifying explanation of the concepts involved and derive the necessary relations to convert a binding free energy from an arbitrary state to some given standard state. This is done in three independent ways, from the point of view of (1) the dimensions of the quantities involved, (2) the energy and entropy of the molecules, and (3) their chemical potentials.

Introduction

Equilibrium constants are generally defined, in chemistry books, in terms of ratios of concentrations of reactants to products. This presents a problem, often ignored, when the number of reactants is not the same as the number of products, since the ratio will not be dimensionless. Many times, one needs to calculate the free energy associated with that equilibrium constant, K, and the usual way of doing so is through the well-known formula[1,2]

$$F = k_B T \ln(K) \tag{1}$$

where k_B stands for the Boltzmann constant, T is the temperature of the system under consideration, and F is the corresponding free energy. In this paper, we will be working in the canonical ensemble, so F represents the Helmholtz free energy, and we assume a constant volume in all calculations. This is not a big drawback since, in binding reactions where the system's volume is allowed to fluctuate, its change, ΔV, at normal values of concentration, pressure and temperature, is negligible.

Equation 1 shows that having units in the equilibrium constant forbids the use of the formula, since the argument of a logarithm must be dimensionless. In particular, when talking about the experimental binding of two molecules (e.g., protein A + ligand B), K^{exp}, defined by

$$K^{exp} = \frac{[A][B]}{[AB]} \tag{2}$$

is typically given in units of molar. The solution to this units problem is to first convert K^{exp} to a new K^0, without units. This can be achieved by doing the following:[3]

$$K^0 = \frac{\dfrac{[A]}{C^0}\dfrac{[B]}{C^0}}{\dfrac{[AB]}{C^0}} = \frac{K^{exp}}{C^0} \tag{3}$$

where C^0 is a constant with the same units as K^{exp}. We call this the *standard state equilibrium constant*. But note that C^0 is arbitrary; it is precisely here where the standard state is defined. In other words, in order to be able to calculate a free energy from the equilibrium constant, we must first define a standard state. Any free energies that we calculate later, via theory, simulation, or experiment, should also be referred to the same standard state, so that we can make a meaningful comparison between them.

This problem has already been correctly treated by other authors,[4,5] as related to computational studies of binding, but nevertheless, it is still easy for somebody who is confronted with this issue for the first time to get confused about it, since one can find some authors comparing results without taking care of the reference state at all,[6] some others considering corrections based on completely different grounds

* E-mail: ijgeneral@gmail.com.

10.1021/ct100255z © 2010 American Chemical Society
Published on Web 07/15/2010

and, thus, apparently contradicting each other (see, e.g., ref 7 for a correction based on the unbound sampled volume,[8] for a correction based on the binding site volume,[4,5] for one based on the chemical potential, and[9] for a model that contains its own correction), and yet some others explicitly denying the existence of such corrections.[10] It is the purpose of this work to review this issue in a comprehensive and rigorous way, trying at the same time to be as clear as possible, especially in those points where confusion is seen in the literature.

So the key question is: how do we convert a computational (or experimental or theoretical) free energy, F^{comp}, to a free energy in a given standard state, F^0 (usually, the standard state in which the experimental value is already known)? In the following sections, we will derive this transformation in three different ways: the first is a simple *dimensional* approach which, through the consideration of units, leads to the desired result; the second is an intuitive *physical* approach, based on the consideration of energy and entropy; and the last one is a more *chemical* approach, which considers the chemical potential of the molecules in the binding process. The first derivation is very illustrative, as it shows that the problem in comparing free energies is ultimately related to the matching of their units. The second derivation is, in a way, the best derivation since it makes very clear the physical meaning of the correction, from the point of view of physical magnitudes. The last derivation is presented for the sake of completeness, but different versions of it can be found in the literature.[4,5]

Conversion of a Free Energy to Some Standard State

Dimensional Approach. Let us define K as a dimensionless equilibrium constant, related to $\tilde{K}$, its equivalent with units, i.e.,

$$K = \frac{\tilde{K}}{C} \tag{4}$$

C is some concentration that we arbitrarily choose as a standard (typically 1 M), but there are some times when we are not able to choose it, but it is chosen automatically, instead. This is the case when we calculate F by simulation; there we get the free energy, which is related to some already dimensionless equilibrium constant, via the inverse of eq 1:

$$K = \exp\left(\frac{F}{k_B T}\right) \tag{5}$$

So now we could ask: In the experiment we obtain $\tilde{K}^{exp}$, and get K^{exp} by dividing by some standard concentration C^0, in order to get rid of the units. In the simulation, we are getting K^{comp}. Then, by what factor did the computer (or the theory) divide to get rid of the units? The only concentration available to the theory is that of the system we are simulating, so we must assume that we have

$$K^{comp} = \frac{\tilde{K}^{comp}}{C^{comp}} \tag{6}$$

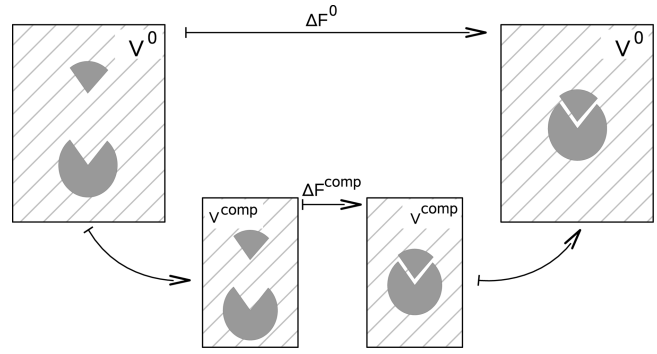


Figure 1. Relation between computational and standard state binding free energies. V^0 and V^{comp} represent the volumes of the corresponding systems.

where C^{comp} is the concentration we use in the simulation. But, of course, we want the experimental and computational constants to be the same (within some error) when they are expressed in the same units, so

$$\tilde{K}^{exp} = \tilde{K}^{comp} \Rightarrow \frac{\tilde{K}^{exp}}{C^0} = \frac{\tilde{K}^{comp}}{C^{comp}}\frac{C^{comp}}{C^0} \tag{7}$$

which, by taking $k_B T \ln(...)$, leads to

$$F^0 = F^{comp} + k_B T \ln\left(\frac{C^{comp}}{C^0}\right) \tag{8}$$

where we renamed F^{exp} as F^0, since it is in the standard state. This expression tells us, by way of dimensional arguments, that the computational free energy needs a correction term in order to be compared to a free energy calculated by other methods. This is necessary if the units of the different equilibrium constants are to match.

The magnitude of the correction is a function of the concentration of the molecules we have in the study. Typical ligand concentrations in computational binding studies are between 1 per 15 Å radius sphere and 1 per 100 Å radius sphere. Using a standard state of 1 M (1 ligand per 1661 $Å^3$), the corresponding corrections at 300 K fall between 1.3 and 4.7 kcal/mol. So, although these corrections will not, in general, drastically change the results, they are significant, especially if one is interested in getting precise values of free energy.

Equation 8 also suggests a way of checking the results of computational calculations. Note that F^0, being in a given standard state, must be independent of C^{comp} (it should not matter how it was calculated, as long as the standard state in which it is expressed is the same). This means that F^{comp} has to be a function of C^{comp}, so that, together with the last term in eq 8, this dependence on the computational concentration is canceled. Thus, the lack of dependence of F^{comp} on C^{comp}, sometimes observed in computational studies, indicates a problem, typically a convergence issue, such as the ligand not sampling the full available phase space.

Physical Approach. This approach is vaguely mentioned in the literature,[7,8,11] and to the author's knowledge, it has never been put on a solid foundation. Figure 1 shows the relation between the binding free energies of two molecules, as calculated in two different systems. We will call them

standard and computational, but the discussion below is correct for any two systems (standard−experimental, computational−experimental, even computational−computational, when each is calculated using different methods). The volumes there represented, V^{comp} and V^0, are not in general the same, so we should now find a way to convert F^{comp} to F^0. We can achieve this by considering the physical changes in the figure.

• ΔF^{0-comp}: change in F when the volume of the system varies from V^0 to V^{comp}. Assuming the molecules in the system to be electrically neutral, we can see that the energies of the two molecules, A and B, remain constant, while the entropies of the two do change; increasing (decreasing) the volume will produce an increase (decrease) in the entropy of these molecules, since they will have more (less) possibilities to move.

$$\Delta F^{0-comp} = (\Delta E - T\Delta S)_A + (\Delta E - T\Delta S)_B = (-T\Delta S)_A + (-T\Delta S)_B = -2Tk_B \ln \frac{V^{comp}}{V^0} \quad (9)$$

Above, A and B typically represent a protein−ligand pair.

• ΔF^{comp}: change in F as calculated computationally.

• ΔF^{comp-0}: change in F when the volume is changed back from V^{comp} to V^0. As in the first step, the only difference here is an entropic one. But this time, molecule B can only move inside the binding site. We can think of this AB complex as only one molecule with some extra freedom because the ligand can slightly move in the complex. But the amount of this extra freedom is the same in the simulation and the standard state boxes. So the ligand does not contribute to this change.

$$\Delta F^{comp-0} = (\Delta E - T\Delta S)_{AB} = -Tk_B \ln \frac{V^0}{V^{comp}} \quad (10)$$

So, finally, adding the three terms above we find the conversion term for going from the simulation to the standard state:

$$\Delta F^0 = \Delta F^{comp} - k_B T \ln \frac{V^{comp}}{V^0} \quad (11)$$

or, using $C_i = N_i/V_i$, we recover eq 8.

The above approach to the standard state correction, also provides a different kind of computational correction that we may sometimes need to consider, depending on the particular calculation we are doing. We are talking specifically about (alchemical) thermodynamic integrations. Over a decade ago, it was found[4,12,13] that it is computationally convenient to calculate binding free energies using thermodynamic cycles of the type shown in Figure 2, but performing the integration of the right-hand side leg of the cycle with a restraint applied on the ligand, so that it cannot leave the binding pocket, even when the interaction between the ligand and the rest of the system is decreased to zero (the so-called *double decoupling method*,[4] as opposed to the older *double annihilation method* of Jorgensen et al.[14]). This brings the problem that the change in free energy on the bottom of the cycle (where the ligand is altogether decoupled from the rest of the system) is not zero anymore (see Figure 2). But if we realize that the difference between those two systems is just the volume accessible to the ligand, then we can easily apply the correction derived above:

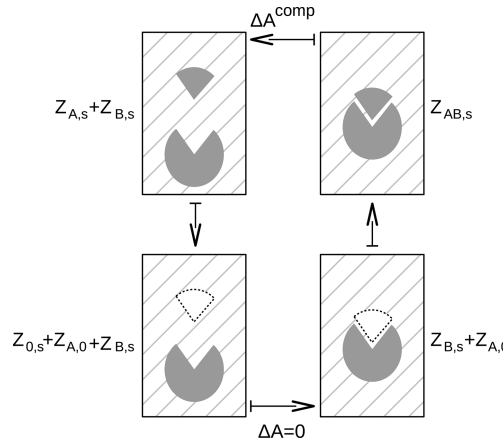


Figure 2. Thermodynamic cycle for protein−ligand binding. Each step in the cycle shows the partition functions associated with it. The white triangular shape, representing the ligand (A), is decoupled from the rest of the system, protein (B) and water, as it is usually done in computational simulations (e.g., thermodynamic integration). The hatched lines in the background represent a solvent.

$$k_B T \ln \frac{V^{comp}}{V^{pocket}} \quad (12)$$

Adding this term to the free energy corrects also for having a restraint applied on the ligand. Note that there could be an analogous correction, in angular space, if the restrained ligand is not able to freely rotate. But we should keep in mind that eq 12 is a different correction than the one studied above. Their forms are equal since both arise from the consideration of a change in volume, but they should not be confused as the same, something that often happens in the literature, because then there is going to be a missing contribution to the binding free energy. Sometimes, this second correction is put together with the standard state one:[8,15,16]

$$-k_B T \ln \frac{V^{comp}}{V^0} + k_B T \ln \frac{V^{comp}}{V^{pocket}} = -k_B T \ln \frac{V^{pocket}}{V^0} \quad (13)$$

but the authors[8] may fail to mention that there are two different contributions here, and they just call it the standard state correction. This has no numerical consequences, as the result is correct, but it may confuse the reader.

Chemical Approach. *Preliminaries.* We start by recalling the derivation of a basic result in thermochemistry. The Helmholtz free energy can be defined, in terms of temperature and volume, by $dF = -SdT - pdV$. Now, if we consider an ideal gas undergoing an isothermal volume change from V^0 to V, we have

$$dF = -pdV \Rightarrow F(V) = F(V^0) - \int_{V^0}^{V} \frac{Nk_B T}{V} dV = F(V^0) - Nk_B T \ln \frac{V}{V^0} \quad (14)$$

$$\Rightarrow \mu = \mu^0 - k_B T \ln \frac{V}{V^0} \tag{15}$$

In the last step, we used $\mu = F/N$. This expression, the relation between the chemical potential of a gas under two different conditions, can be generalized to nonideal gases and ideal and nonideal solutions. For example, for the case of some component i in a solution, the expression becomes

$$\mu_i = \mu_i^0 + k_B T \ln \frac{\gamma_i V_i^0}{V_i} \tag{16}$$

where γ_i is the activity coefficient of the substance we are considering (this coefficient contains the nonideality information) and V_i and V_i^0 are the volumes covered by component i of the solution, under the two external conditions.

The important point to note about this relation is that it expresses just a change in the conditions of the systems, as a change in volume or concentration. There is nothing special about the "0-state" or "standard state"; it is arbitrary, and thus we can choose any state that makes things easier for us.

With this relation, we can start the derivation of the conversion term.

Derivation. The following derivation of the conversion term is based on the ones given by Gilson et al.[4] and Boresch et al.[5] Consider a binding reaction, $A_s + B_s \leftrightarrow AB_s$, where the subscript s indicates that the molecules are immersed in a solvent. In the following, we are going to assume there is only one molecule of each component, i.e., $N_A = N_B = N_{AB} = 1$. We know that in equilibrium $\mu_A + \mu_B = \mu_{AB}$; thus, according to eq 16

$$\begin{aligned} &\mu_A^0 - k_B T \ln \frac{V_A}{V^0} + \mu_B^0 - k_B T \ln \frac{V_B}{V^0} = \mu_{AB}^0 - k_B T \ln \frac{V_{AB}}{V^0} \\ &\Rightarrow \Delta F^0 \equiv \mu_{AB}^0 - \mu_A^0 - \mu_B^0 = -k_B T \ln \frac{V_A V_B}{V^0 V_{AB}} \end{aligned} \tag{17}$$

where we have used $\gamma_i = 1$, the very low concentrations approximation. Since $C_i = N_i/V_i$, we can rewrite this as

$$\Delta F^0 = -k_B T \ln K;\ K \equiv \frac{C^0 C_{AB}}{C_A C_B} \tag{18}$$

recovering eq 1. But notice this relation is telling us that when we choose the value of C^0, so that we can calculate ΔF from K, we are determining the standard state. Consequently, in the equations above, when we write μ^0, V^0, or C^0, we are referring to the same standard state. In conclusion, the current derivation, started with eq 16, is just a conversion of some quantity from some standard state to a nonstandard state.

Let us now turn our attention to statistical mechanics. In the canonical ensemble

$$\mu_{i,s} = F_{i,s} - F_{0,s} = -k_B T \ln \frac{Z_{i,s}}{Z_{0,s}} \tag{19}$$

is the chemical potential of component i in the solution, s. $Z_{i,s}$ and $Z_{0,s}$ represent the partition functions of one and zero molecules of type i, respectively, in solution. Replacing this in eq 16

$$\begin{aligned} -k_B T \ln \frac{Z_{i,s}}{Z_{0,s}} &= \mu_i^0 + k_B T \ln \frac{\gamma_i C_i}{C_i^0} \\ \mu_i^0 &= -k_B T \ln\left(\frac{\gamma_i C_i}{C_i^0}\frac{Z_{i,s}}{Z_{0,s}}\right) \end{aligned} \tag{20}$$

And inserting this expression in eq 17:

$$\Delta F^0 = -k_B T \ln\left(\frac{\gamma_{AB}}{\gamma_A \gamma_B}\frac{C_{AB}}{C_A C_B}\frac{C_A^0 C_B^0}{C_{AB}^0}\frac{Z_{0,s} Z_{AB,s}}{Z_{A,s} Z_{B,s}}\right) \tag{21}$$

For the binding case, where we consider $C_A = C_B = C_{AB}$, we get

$$\Delta F^0 = -k_B T \ln\left(\frac{\gamma_{AB}}{\gamma_A \gamma_B}\frac{C_A^0}{C_A}\frac{Z_{0,s} Z_{AB,s}}{Z_{A,s} Z_{B,s}}\right) \tag{22}$$

To understand the meaning of this equation, let us assume infinite dilution ($\gamma_i = 1$), for simplicity, and rewrite it in the following way:

$$\begin{aligned} \Delta F^0 = -k_B T \ln\left(\frac{C_A^0}{C_A}\right) + [F_{AB,s} - F_{B,s} - F_{A,0}] + \\ [F_{0,s} + F_{A,0} + F_{B,s} - F_{A,s} - F_{B,s}] \end{aligned} \tag{23}$$

where F_i is the free energy associated with partition function Z_i, and we have added and subtracted a pair of terms, one involving $F_{A,0}$, which represents molecule A in no solution, meaning that it has been decoupled from the solution (it is in a gas phase), as it is done when using thermodynamic integration. Equation 23 can be easily interpreted in terms of the thermodynamic cycle in Figure 2. But that cycle is exactly the one used in computational simulations; consequently, eq 23 leads to

$$\Delta F^0 = -k_B T \ln\left(\frac{C_A^0}{C_A}\right) + \Delta F^{comp} \tag{24}$$

or

$$\Delta F^0 = -k_B T \ln\left(\frac{V_A}{V_A^0}\right) + \Delta F^{comp} \tag{25}$$

So we see that in order to convert the computational free energy to a given standard state, we need to add a correction term that takes into account the difference in volume between the two states (the simulation and the standard state volumes). And, of course, V^0 has to be the one used in eq 3, when the standardization of the experimental result is performed. Equation 25, with the subscripts A removed, coincides with eqs 8 and 11, derived using other approaches.

Summary and Conclusions

We have seen that binding free energies are defined in terms of some reference state, whose purpose is to make the equilibrium constant dimensionless. In this way, one can compare two constants calculated in different ways and obtain sensible results since they are in the same standard state.

This argument can also be put in terms of volumes and concentrations, instead of units, since the second derivation

showed that the correction appears as a consequence of considering different volumes, hence different entropies, when the free energy is calculated through different methods. The expression for the correction, given in eqs 8, 11, and 25, depends on the ratio of the volumes (or the concentrations) of the two states. In the binding case we use in the present work, they are the volume V^0 used to make the experimental equilibrium constant dimensionless (eq 3) and the volume V_A that was available to the type-A molecule, during the simulation, before binding (which was the same as that for type B: $V_A = V_B = V^{comp}$).

We also showed that the second formulation of the problem gives the correction that has to be added when using thermodynamic integration with restraints on the ligand. The reason for why the same type of correction can be used is that both problems, the standard state and the binding site problems, arise from comparing two systems with different volumes. Thus, the solution is the same, a factor that takes into account that difference.

Acknowledgment. The author would like to thank Prof. Jeffry Madura for helpful comments and Prof. Hagai Meirovitch for valuable discussions and for his support.

References

(1) Moore, W. J. *Physical Chemistry*, 5th ed.; Longmans: London; Prentice Hall: New York, 1972.

(2) Glasstone, S. Lewis, D. *Elements of Physical Chemistry*, 2nd ed.; Macmillan: London, 1964.

(3) Janin, J. *Proteins: Struct., Funct., Genet.* **1996**, *24*, i–ii.

(4) Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. *Biophys. J.* **1997**, *72*, 1047–1069.

(5) Boresch, S.; Tettinger, F.; Leitgeb, M.; Karplus, M. *J. Phys. Chem. B* **2003**, *107*, 9535–9551.

(6) Fujitani, H.; Tanida, Y.; Ito, M.; Jayachandran, G.; Snow, C. D.; Shirts, M. R.; Sorin, E. J.; Pande, V. S. *J. Chem. Phys.* **2005**, *123*, 084108.

(7) Doudou, S.; Burton, N. A.; Henchman, R. H. *J. Chem. Theory Comput.* **2009**, *5*, 909–918.

(8) Jayachandran, G.; Shirts, M. R.; Park, S.; Pande, V. S. *J. Chem. Phys.* **2006**, *125*, 084901.

(9) Fukunishi, Y.; Mitomo, D.; Nakamura, H. *J. Chem. Inf. Model.* **2009**, *49*, 1944–1951.

(10) Fujitani, H.; Tanida, Y.; Matsuura, A. *Phys. Rev. E* **2009**, *79*, 021914.

(11) Zhou, H.-X.; Gilson, M. K. *Chem. Rev.* **2009**, *109*, 4092–4107.

(12) Roux, B.; Nina, M.; Pomès, R.; Smith, J. C. *Biophys. J.* **1996**, *71*, 670–681.

(13) Hermans, J.; Wang, L. *J. Am. Chem. Soc.* **1997**, *119*, 2707–2714.

(14) Jorgensen, W. L.; Buckner, J. K.; Boudon, S.; Tirado-Rives, J. *J. Chem. Phys.* **1988**, *89*, 3742–3746.

(15) Rodinger, T.; Howell, P. L.; Pomès, R. *J. Chem. Phys.* **2008**, *129*, 155102–1.

(16) Mobley, D. L.; Graves, A. P.; Chodera, J. D.; McReynolds, A. C.; Shoichet, B. K.; Dill, K. A. *J. Mol. Biol.* **2007**, *371*, 1118–1134.

CT100255Z

JCTC Journal of Chemical Theory and Computation

Ab Initio Studies of Structural and Vibrational Properties of Protonated Water Cluster $H_7O_3^+$ and Its Deuterium Isotopologues: An Application of Driven Molecular Dynamics

Martina Kaledin* and Christopher A. Wood

Chemistry and Biochemistry, Kennesaw State University, 1000 Chastain Road, Box 1203, Kennesaw, Georgia 30144

Received March 3, 2010

Abstract: In this work, we present infrared (IR) spectra of $H_7O_3^+$ and its deuterium isotopomers calculated by direct molecular dynamics (MD) simulations at the B3LYP/6-31+G** computational level. The calculated spectra obtained at 100, 300, and 500 K were compared to available experimental observations, and spectral features were assigned using normal-mode analysis (NMA) and driven molecular dynamics (DMD). Spectral peaks at 2410 and 2540 cm^{-1} were assigned to asymmetric and symmetric stretches of the bridging hydrogen (BH) using NMA. The weak spectral features at 2166 and 2275 cm^{-1} were assigned to a combination band of BH asymmetric stretch, H_2O in phase wagging, OO stretch, and H_3O^+ rocking vibrations by DMD simulations. Our observation of BH stretch vibrations as low as 2166 cm^{-1} is in good agreement with the assignment of the low-resolution spectrum obtained by Schwarz at 2200−2300 cm^{-1} [Schwarz, H. A. *J. Chem. Phys.* **1977**, *67*, 5525−5534] and vibrational predissociation spectrum by Lee et al. ∼2300 cm^{-1} [Okumura, M.; Yeh, L. I.; Myers, J. D.; Lee, Y. T. *J. Chem. Phys.* **1990**, *94*, 3416−3427].

1. Introduction

The study of ions in water raises many interesting questions about charge transfer mechanisms and dynamics in liquid water.[1−4] The nature and transport mechanism of the proton in aqueous solutions and its spectroscopic signatures have been the subject of extensive discussions for many years.[5−7] Protons (H^+) and hydroxide ions (OH^-) exhibit anomalously high mobility in aqueous media compared to other ions, such as sodium (Na^+) and chloride (Cl^-) ions,[8] indicating that transport occurs via mechanism other than normal ionic diffusion. This phenomenon immediately introduces the crucial role of H_3O^+ and $H_5O_2^+$ that are ion cores of so-called Eigen[9] and Zundel[10] forms of the cations, respectively. Fluctuations between these species are thought to mediate the Grotthuss mechanism[11] for proton transport.

A powerful way to test the validity of various theoretical models is through the use of the cluster ions, e.g., $H^+(H_2O)_n$ and $OH^-(H_2O)_n$ which can be prepared and isolated in the laboratory. It is of significant interest to study clusters of larger size, because the structure and behavior of small hydrated proton clusters are quite different from those of larger size. The most obvious difference is that smaller clusters do not favor the ring structure which gives the maximum number of hydrogen bonds.[12] A recent study shows that proton transfer along hydrogen-bonded water chains can be facilitated by $H_7O_3^+$.[13]

The motivation for this study is to explain the proton transfer mechanics in $H_7O_3^+$ for which theoretical and experimental studies are limited and global potential energy surface is not available. DFT direct molecular dynamics simulations of $H_7O_3^+$ in the gas phase were first performed using the B3LYP functional for a short simulation time (2 ps)[14] and the spectral peaks assigned using NMA. Simulated IR spectra reproduced the observed IR pattern,[15] and the broad region between 2200−2300 cm^{-1} was assigned to a stretch of bridging protons. The $H_7O_3^+$ spectrum in the OH stretch region was studied experimentally by Lee et al.[16]

* Corresponding author phone: +1-770-423-6281; fax: +1-770-423-6744; e-mail: mkaledin@kennesaw.edu.

10.1021/ct100122s © 2010 American Chemical Society
Published on Web 07/20/2010

They observed triplet signals at 3637, 3667, and 3722 cm^{-1} corresponding to a free OH symmetric H_2O stretch, H_3O^+ OH stretch, and free OH asymmetric H_2O stretch, respectively, and compared them to the spectra of $H_7O_3^+ \cdot H_2$ at 3642, 3587, and 3726 cm^{-1} measured by vibrational predissociation spectroscopy.[17] When authors searched for spectra of $H_7O_3^+$ down to 2200 cm^{-1}, a broad absorption was found below 2500 cm^{-1}, but due to low laser energy, their results were inconclusive. In this experimental work,[17] the hydrogen bonded OH stretch in $H_9O_4^+$ at 2670 cm^{-1} was observed that largely agreed with Schwarz's results.[15] Isotope dependence of proton exchange and reactions of protonated water clusters $H^+(H_2O)_n$ ($n = 1-4$) with D_2O and $D^+(D_2O)_n$ ($n = 1-4$) with H_2O were studied using guided ion beam mass spectroscopy.[18] At low collision energies, H−D exchange product ions were observed as well as solvent exchange reaction products.[18]

Recently, Duncan et al.[19] reported photodissociation infrared spectra of water clusters $H^+(H_2O)_n$ ($n = 2-5$) with and without Ar tagging. Their experimental studies were complemented by computational work at the MP2/aug-cc-pVTZ level of theory. The $H_7O_3^+$ without Ar yields broad-defined resonances due to the fact that the photodissociation signal was weak. The spectrum measured with Ar contains three main bands at 3577, 3638, and 3722 cm^{-1}; those were compared to the triplet observed by Lee et al.[16] They also found a small, but reproducible, feature at 3098 cm^{-1} and assigned it to the overtone of the water scissors motion predicted at 1599/1623 cm^{-1}.

Spectral signatures of hydrated proton vibrations in water clusters $H^+(H_2O)_n$ ($n = 2-11$) were characterized by vibrational predissociation spectroscopy[20] and assigned by anharmonic VSCF calculations.[21] The vibrational spectrum of the Eigen cation $H^+(H_2O)_4$ shows a strong band at 2665 cm^{-1} that correspond to symmetric and asymmetric H_3O^+ stretches. After removal of one water from $H^+(H_2O)_4$, the spectrum of $H^+(H_2O)_3$ shows a dramatic splitting of the $H^+(H_2O)_4$ strong feature at 1880 and 3580 cm^{-1}. Therefore, anharmonic corrections are required to describe the large redshifts in $H^+(H_2O)_3$ stretching transitions. The peaks at 1880 and 3580 cm^{-1} were assigned to asymmetric H_3O^+ stretch and H_3O^+ free OH stretch vibrations using anharmonic VSCF calculations.

Experimental spectra are often analyzed with the aid of theoretical calculations. Ab initio molecular dynamics simulations contribute to the gaining of insight into the process of solvation of OH^- and H^+ ions. From these simulations, structural as well as dynamical properties of these systems can be investigated. We present IR spectra obtained from the direct MD approach. The accuracy and feasibility of this approach were tested and verified on the smaller well studied molecular system, $H_3O_2^-$.[22] Recently, a classical method of assignment of the spectral features using driven molecular dynamics (DMD) method[23] has been proposed. This method was originally developed to calculate the normal mode frequencies and normal mode vectors of large molecules.[24,25] We used the DMD method to analyze unassigned, anharmonic spectral features in the $H_5O_2^+$ infrared spectrum.[26] Simulated IR spectra reproduced the patterns in the experimental spectra and the assignment of IR bands. In this work, we use DMD to identify the shared proton vibrations in the IR spectrum of $H_7O_3^+$ beyond the harmonic normal-mode analysis.

2. Computational Details

A. Potential Energy Surface. We have studied the structural isomers, energetics, and IR spectra of $H_7O_3^+$ protonated water cluster. All electronic structure calculations were performed by using the Gaussian 03 program package.[27] The geometries of $H_7O_3^+$ stationary points were fully optimized using the density functional theory (DFT) with B3LYP[28−30] functional and 6-31+G** basis set. The second order Møllet−Plesset (MP2) perturbation theory[31−33] optimizations with the cc-pVTZ and aug-cc-pVTZ basis sets were performed as well, in order to validate predictions of DFT-B3LYP. The CCSD(T)/cc-pVTZ energies were obtained using the single-point energy calculations on the MP2/cc-pVTZ geometries. The vibrational frequencies for $H_7O_3^+$ and its D-substituted isotopologues were obtained using the standard NMA.[34] The dissociation energies of $H_7O_3^+$ to $H_2O + H_5O_2^+$ and $H_5O_2^+$ to $H_2O + H_3O^+$ were estimated at various levels of theory.

B. Direct Molecular Dynamics. IR spectra of $H_7O_3^+$ and D-substituted isotopologues were obtained by propagating trajectories on the fly at the B3LYP/6-31+G** level of theory. The interaction potential, forces, and dipole were computed at every time step. The simulations were carried out on our Linux computer cluster using a shell script that interfaces the Gaussian 03 program and our own suite of MD codes.

Two different kinds of direct MD simulations were performed to study the system. The starting point for all trajectories was the equilibrium structure with randomly sampled velocities from the Boltzmann distribution. Initially the system was equilibrated for the first 500 steps (250 fs). The first set of simulations was performed under constant temperature (NVT) conditions, and results from these simulations were used to analyze the structural features of the $H_7O_3^+$ water cluster. Ten trajectories were propagated at a constant temperature via Berendsen's thermostat.[35] The constant temperature was maintained with a 10 fs time response parameter.[35] A second set of simulations, also from 10 trajectories, was performed at constant energy (NVE) conditions. The results from these simulations were used to analyze the dynamical properties calculated using time correlation function. Spectra were collected for $J = 0$ using the expression for the intensity I

$$I(\omega) = \frac{\mathrm{Re}}{\pi} \int_0^\infty \mathrm{d}t e^{i\omega t} \langle \vec{\mu}(0) \vec{\mu}(t) \rangle_T \qquad (1)$$

where ω is the frequency, $\langle \vec{\mu}(0) \cdot \vec{\mu}(t) \rangle_T$ is the dipole−dipole correlation function at temperature T.

The initial coordinates and momenta for the NVE ensembles were sampled from the previously run equilibrated NVT trajectories, and then the spectra were calculated by the Fourier transform of the dipole−dipole correlation function multiplied by a quantum mechanical frequency-

dependent factor[36] equal to $\omega/[1 - \exp(-\omega/kT)]$. The correlation function in eq 1 was obtained as a time average along the trajectory[37] to yield better converged spectra. We checked conservation of the total angular momentum and the energy to ensure the accuracy of the propagator.

In order to facilitate the spectral assignment of the peaks, we calculate position−position spectra for certain atomic groups and dipole spectra from the dipole−dipole correlation function decomposed into parallel and perpendicular components $\langle \mu_y(0)\mu_y(t) + \mu_z(0)\mu_z(t)\rangle_T$ and $\langle \mu_x(0)\mu_x(t)\rangle_T$, respectively. Details of these two methods can be found in our recent paper on $H_3O_2^-$.[22]

C. Driven Molecular Dynamics. The most dominant spectral features were assigned using the driven molecular dynamics (DMD) method.[24] The DMD method was originally used to calculate the normal mode frequencies and normal mode vectors of large molecules for which the standard Hessian-based methods are not feasible due to the high memory and disk requirements.[25] The important advantage of DMD over the Hessian-based normal-mode analysis is the ability to study anharmonic motion and mode coupling.

The DMD method uses an external, sinusoidal driving source to identify resonant absorptions. To identify a resonant vibrational mode in the spectrum, the vibrational motion associated with that mode must produce a change in the dipole moment of the molecule. At resonant frequencies, the molecular motions induced by a weak driving force (small perturbation) correspond to infrared active frequencies, while harder driving (stronger perturbation) induces anharmonic motion.

In DMD simulations of vibrational IR spectra, the Hamiltonian of a molecular system, H

$$H(\mathbf{p},\mathbf{q},t;\omega) = H_0(\mathbf{p},\mathbf{q}) + U(\mathbf{q},t;\omega) \tag{2}$$

consists of the free molecular Hamiltonian, $H_0 = V(\mathbf{q}) + \mathbf{p}^2/2m$, where V is the molecular potential and a dipole-driving term $U(\mathbf{q}, t; \omega)$

$$U(\mathbf{q},t;\omega) = \vec{\varepsilon}_0\cos(\omega t)\cdot\vec{\mu}(\mathbf{q}) \tag{3}$$

$\mathbf{q}$ and $\mathbf{p}$ are the $3N$ atomic Cartesian coordinates and momenta, respectively, ω is frequency, t is time, $\vec{\varepsilon}_0$ determines the direction and strength of the electric field, and $\vec{\mu}(\mathbf{q})$ is the electric dipole of the molecule. Hamilton's equations of motion for Cartesian coordinate α of atom i are

$$\dot{q}_{\alpha,i} = \frac{\partial H}{\partial p_{\alpha,i}} = \frac{p_{\alpha,i}}{m_i},$$
$$\dot{p}_{\alpha,i} = -\frac{\partial H}{\partial q_{\alpha,i}} = -\frac{\partial V}{\partial q_{\alpha,i}} - \varepsilon_{0,\alpha}\cos(\omega t)\frac{\partial \mu_\alpha}{\partial q_{\alpha,i}};$$
$$i = 1,...N, \alpha = x,y,z \tag{4}$$

The dipole moment and dipole derivatives of the system are defined as follows:

$$\vec{\mu} = \sum_{i=1}^{N} \vec{q}_i Q_i$$
$$\frac{\partial \vec{\mu}}{\partial \vec{q}_i} = \delta_{\alpha,\beta} Q_i; \quad \alpha,\beta = x,y,z \tag{5}$$

Dipole derivative is not generally computed in single energy and gradient calculations. We approximated the dipole derivative using atomic charges from Mulliken population analysis[38] assuming that atomic charges vary slowly along the trajectory.

To identify resonant frequencies, the average internal energy of the molecule is obtained after a finite time of driving, which is given by

$$\langle E\rangle = \frac{1}{t - t_0}\int_{t_0}^{t} H(\tau)\mathrm{d}\tau \tag{6}$$

where t_0 is some initially chosen time elapsed after the start of propagation. At nonresonant frequencies, the absorbed energy is small and oscillatory with time, while on resonance, whether fundamental or nonfundamental frequency, it increases rapidly with time. DMD simulations were carried out for several frequencies of interest.

At a resonant frequency a DMD trajectory shows oscillations of atoms around their equilibrium positions with the amplitudes proportional to the absorbed energy. A procedure of assignment of particular frequency consists of analysis of the atomic motion from DMD simulations in terms of the standard normal modes.[26] We evaluate normalized mass-scaled atomic displacements from DMD simulations after a significant amount of energy is absorbed by the molecule, after time t_0 (see eq 6)

$$\Delta x_i(t) = \frac{m_i^{1/2}(q_{ti} - q_{ei})}{[\sum_{i=1}^{3N} m_i(q_{ti} - q_{ei})^2]^{1/2}} \tag{7}$$

Then, a comparison of two modes is done by calculation the inner product of the DMD displacement $\Delta x(t)$ and standard normal mode

$$O_j(t) = \Delta x(t)\cdot u_j = \sum_{i=1}^{3N} \Delta x_i(t)\cdot u_{ij} \tag{8}$$

where u_j is the jth normal mode vector. The overlap is expected to be 1.0 if the two sets of vectors are identical and 0.0 if they are uncorrelated. This procedure was tested extensively on $H_5O_2^+$ [26] and is capable of identifying combination bands which are absent in the NMA model.

3. Potential Energy Surface

A. Stationary Points. The orientation of $H_7O_3^+$ in the coordinate system and the atom labeling are shown in Figure 1. The global minimum of $H_7O_3^+$ is of C_s symmetry. The structural parameters (Table 1) obtained at the B3LYP/6-31+G**, MP2/cc-pVTZ, and MP2/aug-cc-pVTZ levels of theory are very similar. The bond lengths are within 0.01 Å except the proton stretch coordinate, $R(O_2-H_5)$. The proton transfer coordinate is particularly sensitive to the basis set and the method. According to the previous B3LYP calculations[14] using the relatively small basis set DZP basis set for hydrogen and TZP for oxygen atoms, this stretching coordinate is 1.415 Å, while at DFT-B3LYP/6-31+G** and

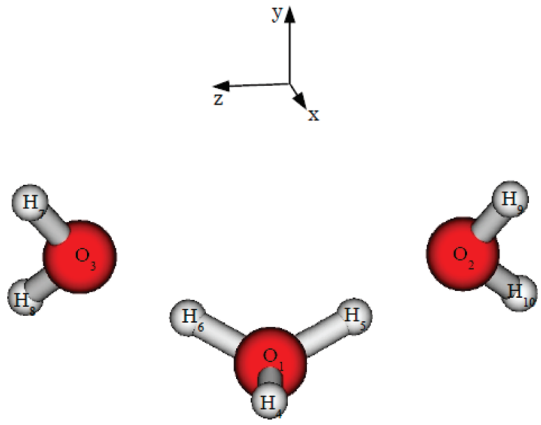


Figure 1. MP2 minimum structure (C_s) and the atom labeling.

MP2/aug-cc-pVTZ levels of theory $R(O_2-H_5)$ is 1.435 Å, 1.453 Å, respectively.

Vibrational frequencies for $H_7O_3^+$ are listed in Table 2. The vibrational frequencies and mode assignment are very similar for both B3LYP and MP2 methods, except of a few low frequency modes. The important qualitative features in the spectrum are the bridging hydrogen bending modes (agree within 40 cm^{-1}), bridging hydrogen stretching mode (within 80 cm^{-1}), and OH stretching modes (within 20 cm^{-1}). In order to assign the spectral peaks in our classical MD spectra, we report spectra for various deuterium isotopologues, $D_7O_3^+$, $DH_6O_3^+$, and $D_3H_4O_3^+$.

There are several stationary points on the potential energy surface of the protonated water cluster, $H_7O_3^+$ (Figure 2). The low-energy transition structure (TS1) of C_{2v} symmetry was found in the previous theoretical work.[14] This planar transition state connects two pyramidal H_3O^+ configurations. The reaction mode represents the umbrella motion of the H_3O^+ ion. The energy differences between the global minimum of C_s symmetry

Table 1. $H_7O_3^+$ Minimum Structure Using the DFT-B3LYP and MP2 Methods

internal coordinates[a]	B3LYP/6-31+G**	MP2/cc-pVTZ	MP2/aug-cc-pVTZ
$R_1(O_1-H_4)$ Å	0.969	0.965	0.967
$R_2(O_1-H_5)$ Å	1.047	1.038	1.038
$R_3(O_2-H_5)$ Å	1.435	1.443	1.453
$R_4(O_2-H_9)$ Å	0.967	0.962	0.964
$R_5(O_2-H_{10})$ Å	0.968	0.963	0.965
$\alpha_1(H_4-O_1-H_5)$ deg	114.1	112.1	112.1
$\alpha_2(H_5-O_1-H_6)$ deg	118.4	114.6	114.6
$\alpha_3(H_5-O_2-H_9)$ deg	127.3	125.4	125.5
$\alpha_4(H_5-O_2-H_{10})$ deg	123.4	121.0	121.3
$\alpha_5(H_9-O_2-H_{10})$ deg	108.8	106.8	106.9
$\alpha_6(O_1-H_5-O_2)$ deg	176.3	174.5	174.3
$\delta_1(H_4-O_1-H_5-O_2)$ deg	114.2	156.7	154.8
$\delta_2(H_5-O_1-H_6-O_3)$ deg	107.3	73.9	76.0
$\delta_3(H_7-O_3-H_6-O_1)$ deg	179.4	160.9	161.8
$\delta_4(H_8-O_3-H_6-O_1)$ deg	7.9	52.0	49.8

[a] The labeling of the atoms is shown on the Figure 1.

Table 2. $H_7O_3^+$ and $D_7O_3^+$ Vibrational Frequencies in cm^{-1}

	B3LYP/6-31+G**		MP2/aug-cc-pVTZ	
tentative assignment	$H_7O_3^+$	$D_7O_3^+$	$H_7O_3^+$	$D_7O_3^+$
MP_2, sym OO stretch; B3LYP, H_2O torsion	38 (A″)	28 (A″)	76 (A′)	71 (A′)
MP_2, H_2O torsion; B3LYP, sym OO stretch	68 (A′)	64 (A′)	86 (A″)	61 (A″)
H_2O twisting, in phase	124 (A′)	93 (A′)	130 (A′)	94 (A′)
H_2O wagging, out of phase	142 (A″)	104 (A″)	242 (A″)	184 (A″)
H_2O wagging, in phase	182 (A′)	133 (A′)	256 (A′)	192 (A′)
H_3O^+ wagging	341 (A′)	257 (A′)	352 (A′)	270 (A′)
MP2, asym OO stretch; B3LYP, H_3O^+ wagging	386 (A′)	332 (A′)	369 (A″)	293 (A″)
MP2, H_2O rocking; B3LYP, asym OO stretch	388 (A″)	297 (A″)	405 (A″)	343 (A″)
H_2O rocking	409 (A″)	358 (A″)	403 (A′)	354 (A′)
H_2O rocking, H_3O^+ umbrella	437 (A′)	350 (A′)	486 (A′)	375 (A′)
H_3O^+ rocking	613 (A″)	436 (A″)	615 (A″)	438 (A″)
H_3O^+ twisting	1142 (A″)	821 (A″)	1077 (A″)	775 (A″)
H_3O^+ wagging	1234 (A′)	896 (A′)	1266 (A′)	921 (A′)
H_2O bend, out of phase	1591 (A″)	1149 (A″)	1605 (A″)	1166 (A″)
H_2O bend, in phase	1629 (A′)	1189 (A′)	1633 (A′)	1201 (A′)
BH sym bend	1660 (A′)	1208 (A″)	1699 (A′)	1226 (A′)
BH asym bend	1668 (A″)	1218 (A′)	1702 (A″)	1235 (A″)
BH asym stretch	2387 (A″)	1802 (A″)	2464 (A″)	1852 (A″)
BH sym stretch	2554 (A′)	1839 (A′)	2620 (A′)	1888 (A′)
H_2O OH sym stretch, out of phase	3798 (A″)	2736 (A″)	3788 (A″)	2730 (A″)
H_2O OH sym stretch, in phase	3799 (A′)	2737 (A′)	3789 (A′)	2731 (A′)
H_3O^+ OH stretch	3823 (A′)	2786 (A′)	3810 (A′)	2775 (A′)
H_2O OH asym stretch, out of phase	3903 (A″)	2865 (A″)	3896 (A″)	2857 (A″)
H_2O OH asym stretch, in phase	3904 (A′)	2866 (A′)	3896 (A′)	2858 (A′)

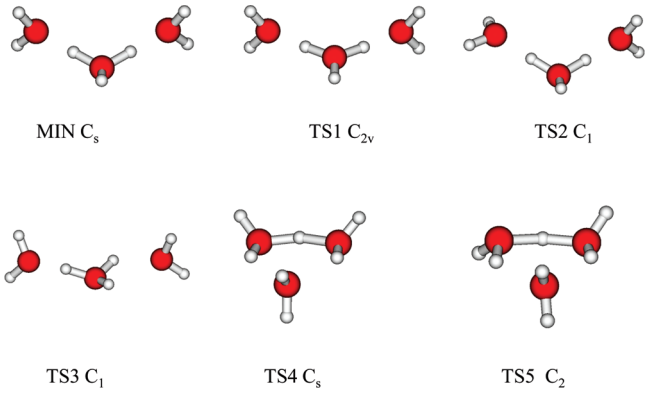


Figure 2. MP2 $H_7O_3^+$ stationary points.

Table 3. Energy Difference between the Stationary Points ($E_{C_s} - E_{C_{2v}}$) on the Potential Energy Surface of $H_7O_3^+$

method/basis set	$E_{C_s} - E_{C_{2v}}$ (kcal/mol)
B3LYP/6-31G**	0.61
B3LYP/6-31+G**	0.39
B3LYP/6-311G**	0.51
B3LYP/6-311+G**	0.43
MP2/cc-pVTZ	1.07
MP2/aug-cc-pVTZ	1.08
CCSD(T)/cc-pVTZ//MP2/cc-pVTZ	1.18
B3LYP/DZP(H)TZP(O)[14]	0.69

and the transition state of C_{2v} symmetry obtained by various methods are listed in Table 3. The barrier height is sensitive to the choice of the basis set and correlation method. Our best estimate for the barrier height is 1.18 kcal/mol obtained from the CCSD(T)/cc-pVTZ single point energy calculations at the MP2/cc-pVTZ geometries. We found a transition state of C_1 symmetry (TS2) representing rotation of one of two waters. The rotational barrier is 0.65 kcal/mol at the B3LYP/6-31+G** level and 0.36 kcal/mol at the CCSD(T)/cc-pVTZ//MP2/cc-pVTZ computational level. Several high-energy transition states (with barrier heights about 5−7 kcal/mol) are shown in Figure 2 (TS3, TS4, TS5). The high-energy transition state (TS3 C_1) is assigned to the rotation of the hydronium ion. The barrier for rotation of H_3O^+ is much higher (6.8 kcal/mol), due to the presence of two hydrogen bonds between H_3O^+ and two H_2O compared to the rotational barrier of 0.36 kcal/mol for TS2 C_1 stationary point representing rotation of one of two waters within $H_7O_3^+$. Pseudo-IRC calculations for the TS4 C_s and TS5 C_2 high-energy transition states show an interconversion of $H_7O_3^+$ from $H_5O_2^+$ to H_3O^+ ion cores.

B. Dissociation Energies. An important feature of the potential energy surface is the asymptotic region leading to the dissociation of $H_7O_3^+$ into the H_2O and $H_5O_2^+$ fragments. To estimate the influence of electron correlation on the dissociation energy, we carried out calculations of dissociation energy at various levels of theory (Table 4). For comparison, we report also $H_5O_2^+ \rightarrow H_2O + H_3O^+$ dissociation energy. With a comparison of our predictions of dissociation energy for $H_5O_2^+$ to the previous theoretical calculation[39] and experimental value,[40] our best estimate for the $H_5O_2^+$ dissociation energy is $D_e = 34.6$ kcal/mol at the MP2/aug-cc-pVTZ level of theory. Using the zero point energy (ZPE) of monomers and the ZPE of $H_5O_2^+$ calculated using harmonic approximation, we obtained $D_0 = 33.7$ kcal/mol that agrees well with the experimental value reported by Dalleska et al., 32.4 kcal/mol.[40] Similarly, the best estimate for the $H_7O_3^+$ dissociation energy is $D_e = 23.6$ kcal/mol at the MP2/aug-cc-pVTZ level of theory. The ZPE corrected value $D_0 = 20.7$ kcal/mol is in a very good agreement with the experimental value of 20.5 kcal/mol.

Table 4. $H_5O_2^+$ and $H_7O_3^+$ D_e Dissociation Energies and Zero Point Energy Corrected Values D_0 in kcal/mol

	$H_5O_2^+$		$H_7O_3^+$	
method/basis set	D_e	D_0	D_e	D_0
B3LYP/6-31G**	42.1	40.9	28.0	25.4
B3LYP/6-31+G**	37.6	36.6	25.0	22.5
B3LYP/6-311G**	40.8	39.7	27.8	25.3
B3LYP/6-311+G**	36.7	35.9	24.8	21.9
MP2/cc-pVTZ	36.9	35.9	25.2	22.4
MP2/aug-cc-pVTZ	34.6	33.7	23.6	20.7
CCSD(T)/cc-pVTZ//MP2/cc-pVTZ	36.1	35.2	24.9	22.1
CCSD(T)/aug-cc-pVTZ[39]	34.1			
exptl value of D_0[40]		32.4		20.5

4. Results from MD Simulations

A. Dipole Spectra. Dipole−dipole correlation functions were collected from 10 NVE trajectories propagated for 10 ps with a 0.5 fs time step at the energies roughly corresponding to each of the three temperatures, 100, 300, and 500 K. (See section 2B for details.) The error in the total energy was confined to less than 0.04% of the total energy for all trajectories.

It is usually sufficient to propagate a few such trajectories to achieve statistical convergence.[22] The convergence of the spectrum at the B3LYP level was tested by comparing the shape of the dipole−dipole spectra from two sets of trajectories. The overlap of the spectra from 5 and 10 trajectories was 97%. The convergence of the position−position spectra was also satisfactory, within 96%.

Figure 3 presents the temperature dependence of the dipole spectrum of $H_7O_3^+$ calculated at the MD-B3LYP/6-31+G** level of theory from 500 to 4000 cm^{-1}, a spectral region

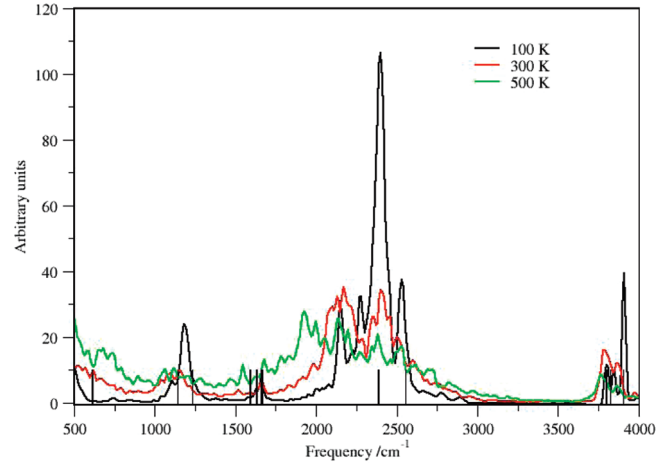


Figure 3. Temperature dependence of the MD B3LYP dipole−dipole spectra for $H_7O_3^+$ from 100 to 500 K. The IR spectra were obtained from 10 NVE trajectories propagated for 10 ps. The harmonic frequencies are shown as sticks in the spectrum.

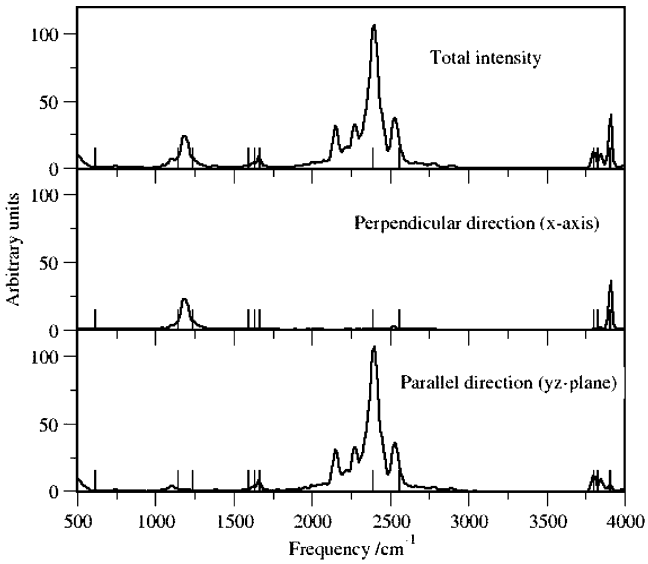


Figure 4. MD B3LYP spectrum for $H_7O_3^+$ at 100 K obtained from dipole−dipole correlation function $\langle\vec{\mu}(0)\cdot\vec{\mu}(t)\rangle_T$. The top panel shows the total intensity. The dipole−dipole correlation function was decomposed into perpendicular $\langle\mu_x(0)\cdot\mu_x(t)\rangle_T$ (along *x*-axis, middle panel) and parallel components $\langle\mu_y(0)\cdot\mu_y(t) + \mu_z(0)\cdot\mu_z(t)\rangle_T$ (along *yz* plane, bottom panel). The harmonic frequencies are shown as sticks in the spectrum.

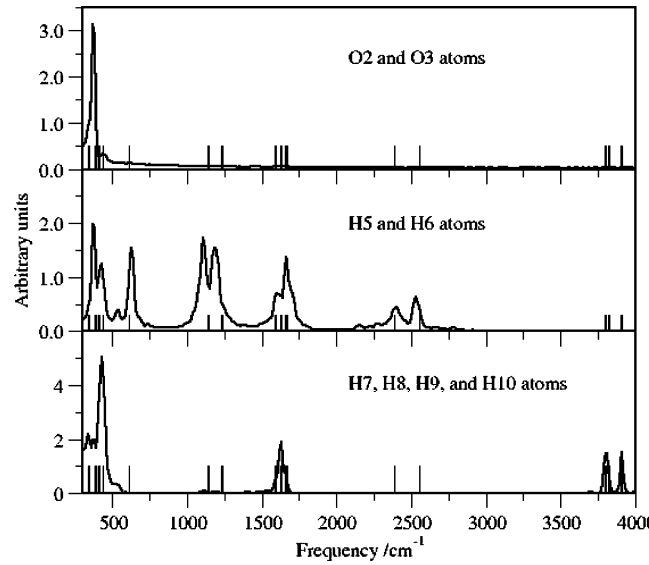


Figure 5. MD B3LYP position−position spectra decomposed to atomic groups at 100 K. The top panel shows the averaged spectrum for O2 and O3 oxygen atoms, the middle panel shows H5 and H6 bridging hydrogens, and the bottom panel shows H7, H8, H9, and H10 hydrogens. The labeling of the atoms is shown in Figure 1.

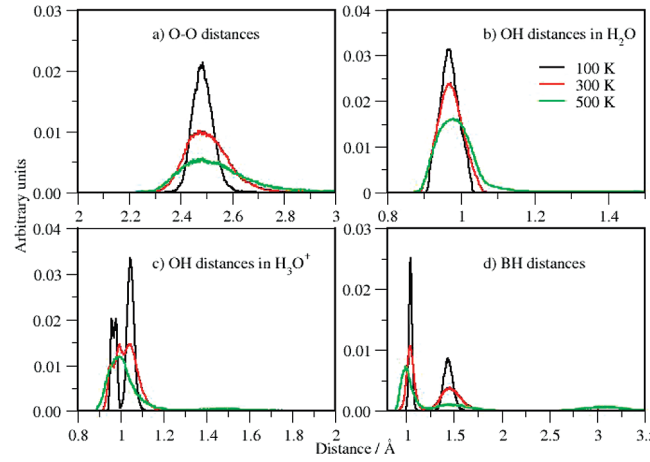


Figure 6. Temperature dependence of bond length distributions for $H_7O_3^+$ calculated from NVT MD-B3LYP trajectories at 100, 300, and 500 K. Each panel shows an average distribution over (a) O−O distances, (b) OH distances in H_2O, (c) OH distances in H_3O^+, and (d) bridging hydrogen (BH) distances from oxygen in H_2O.

relevant to the available experimental measurements. The harmonic frequencies are shown as sticks in the spectrum. At low temperature, 100 K, the MD simulations reproduce the harmonic spectrum. Significant deviation from the harmonic spectrum is seen between 2200 and 2800 cm^{-1} showing four peaks in the shared proton (bridging hydrogen, BH) stretch region at 100 K. At 500 K, this spectral feature significantly broadens.

In order to better understand the spectrum, the dipole spectrum at 100 K was decomposed into the parallel (*yz* plane) and perpendicular components (*x* axis) (Figure 4). All three oxygen atoms and two shared protons are positioned nearly in the *yz* plane. Therefore, the *yz* plane is coincident with the vibration of the shared proton between two oxygens (Figure 1). The most intense spectral peaks in the perpendicular dipole spectrum belong to the OH stretch vibrations. The broad spectral region between 2200 and 2800 cm^{-1} is due to the vibrations of those normal modes in the *yz* plane. The two peaks in this region at 2410 and 2540 cm^{-1} are assigned to the asymmetric and symmetric BH stretch vibration using NMA at the B3LYP/6-31+G** level of theory.

The identity of two weak spectral features at 2166 and 2275 cm^{-1} is scrutinized by means of the position−position spectra (Figure 5). This analysis reveals which atomic groups are involved in a particular vibration in a given spectral range. We analyzed three atomic groups, two oxygen atoms O2 and O3, two BH, H5 and H6 atoms, and four hydrogen atoms in two H_2O moieties, H7−H10. It is obvious that the most intense peaks around 390 cm^{-1} in the O2−O3 position−position spectrum and in the region 3800−3900 cm^{-1} in the H7−H10 position−position spectrum correspond to the OO stretch and OH stretch vibrations, respectively. There are several peaks present in the position−position spectrum for two shared protons H5 and H6: those correlating with the normal mode frequencies (shown as sticks in the spectrum) umbrella vibration, rocking, twisting, and wagging H_3O^+ vibrations, BH bend vibrations, and BH stretch vibrations. See mode assignment in Table 2. This figure shows some activity between 2150 and 2300 cm^{-1}, which means that these vibrations most likely engage a shared proton.

B. Bond Length Distributions. We examined the bond length distributions for $H_7O_3^+$ calculated from the coordinates saved along the NVT MD-B3LYP/6-31+G** trajectories to determine favorable atomic configurations at temperatures between 100 and 500 K (Figure 6). At low temperature, 100 K, the distributions peaked at distances that are close to the equilibrium distances listed in Table 1, while

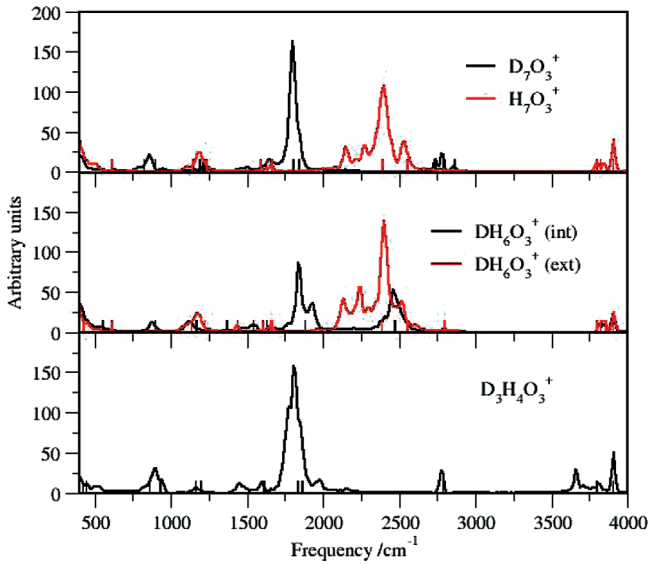


Figure 7. MD B3LYP dipole−dipole spectra at 100 K for $H_7O_3^+$ and its D-substituted isotopologues.

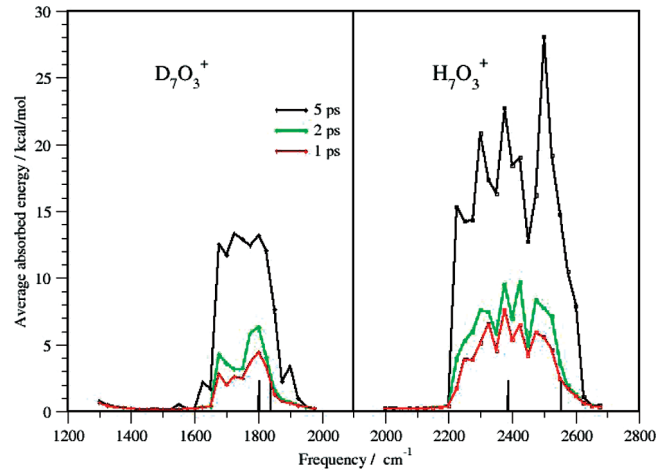


Figure 8. DMD B3LYP simulations: scanned spectrum with a 25 cm^{-1} step for $H_7O_3^+$ and $D_7O_3^+$. For each frequency, a trajectory was propagated for 5 ps using $\vec{\varepsilon}_0 = 50$ mV/bohr. The harmonic frequencies are shown as sticks in the spectrum.

at higher temperatures, the bond length distributions broaden significantly and reveal various bonding topology changes. H_3O^+ in $H_7O_3^+$ is slightly asymmetric at equilibrium showing maxima at 0.97 and 1.05 Å (Figure 6c, 100 K). At 500 K all OH distances become equal to 1.0 Å corresponding to pure H_3O^+ (hydronium cation). The average distribution of bond lengths for two bridging hydrogens hopping between oxygens is showing two peaks around 1.05 and 1.44 Å at 100 K. With increasing temperature, the first peak shifts to the lower distance 1.0 Å, indicating a presence of hydronium cation, while the intensity of the second peak decreases significantly, indicating the presence of the $H_5O_2^+$ Zundel form of cation. Thus, we assume that proton hopping primarily occurs through an Eigen → Zundel → Eigen transition. During our simulations we did not observe formation of H_3O^+ and water dimer configurations as suggested in the previous work.[14] However, at high temperature, 500 K, we observed frequent hydrogen exchange between H_3O^+ and H_2O as evidenced by a small peak around 3.25 Å (Figure 6d).

C. Effect of Isotopic Substitution. In the present study, we report vibrational spectra of D-substituted isotopologues $D_7O_3^+$, $DH_6O_3^+$, and $D_3H_4O_2^+$, calculated from MD-B3LYP trajectories at 100 K (Figure 7). In $DH_6O_3^+$, multiple isomers are possible as deuterium can be located in the shared position between oxygen atoms (interior position designated as int in Figure 7) or in one of two H_2O (exterior position, ext). In the $D_3H_4O_3^+$ isomer, three deuterium atoms are located on hydronium ion core. OH (OD) stretch region of the spectra can be easily identified using the NMA. The $D_7O_3^+$ spectrum displays the dominant peak at 1800 cm^{-1} that can be assigned using the NMA to asymmetric and symmetric shared proton vibrations separated by 37 cm^{-1} (see Table 2). In the $H_7O_3^+$ spectrum, these frequencies are well resolved about 167 cm^{-1} apart. The unassigned weak spectral features in the $H_7O_3^+$ at 2166 and 2275 cm^{-1} are seen in the $D_7O_3^+$ spectrum as weak spectral features at 1500 and 1650 cm^{-1}.

$DH_6O_3^+$ spectra for interior (int) and exterior (ext) positions of deuterium are shown in the middle panel of Figure 7. Because zero-point energies for interior and exterior isomers are very similar, 17 432 and 17 424 cm^{-1}, respectively, we assume that experimental measurements would show the signature of both isomers. Interestingly, deuterium in the exterior position increases the spectral intensity of shared proton vibrations by about 20%. Two peaks in the $DH_6O_3^+$ (int) spectrum positioned at 1800 and 2470 cm^{-1} correspond to the shared D (H) vibrations, respectively. The nonfundamental spectral features present in the $H_7O_3^+$ spectrum are absent in the $DH_6O_3^+$ (int) spectrum. As will be discussed below, these bands arise probably from mixing of low frequency modes and BH stretch vibrations.

5. Results from DMD Simulations

DMD spectra were obtained by propagating trajectories for 5 ps, with the time step 0.5 fs for each frequency. Each trajectory was started at rest at the equilibrium structure. We tested the electric field, $\vec{\varepsilon}_0$, for a few strongly absorbing frequencies to select an appropriate field strength. The optimal field for most frequencies was found to be 50 mV/bohr, which corresponds to about 5% of the $H_7O_3^+$ dissociation energy and 0.5% of the first electronic excitation energy and leads to mild driving. We note that the first excitation energy for $H_7O_3^+$ is singlet A″ at 8.94 eV, calculated using TD-DFT with the B3LYP functional and 6-31+G** basis set.

Previous experimental studies[15] concluded that bridging proton vibrations in $H_7O_3^+$ occur near 2200−2300 cm^{-1}. To capture IR activity in this region, we scanned the spectrum between 2000 and 2700 cm^{-1} and 1300 and 2000 cm^{-1} for $H_7O_3^+$ and $D_7O_3^+$, respectively, with a 25 cm^{-1} frequency step. Figure 8 shows DMD spectra for $H_7O_3^+$ and $D_7O_3^+$ (averaged absorbed energy as a function of frequency). The longer the propagation time of the trajectory, the more energy is absorbed by the molecule. We analyzed DMD spectra obtained from trajectories propagated up to at 1, 2, and 5 ps

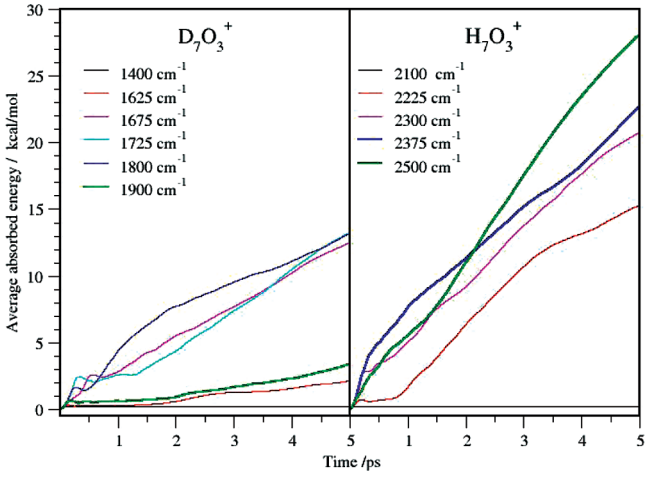


Figure 9. DMD B3LYP simulations: average absorbed energy as a function of time for $H_7O_3^+$ and $D_7O_3^+$ using $\vec{\varepsilon}_0 = 50$ mV/bohr.

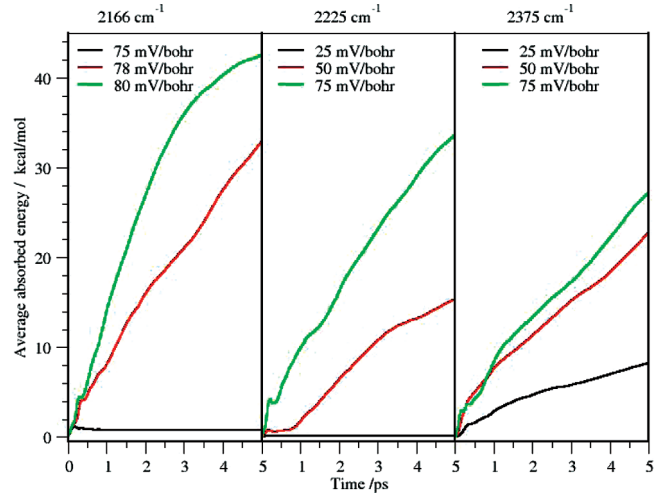


Figure 10. DMD B3LYP simulations: average absorbed energy in cm^{-1} as a function of time and an electric field strength for three driven frequencies ω= 2166, 2225, and 2375 cm^{-1}.

to identify anharmonic features. The short time simulations at or near fundamental frequencies reveal the harmonic vibrational spectra, while longer time driving at nonfundamental peaks causes the molecule to absorb enough energy to leave the harmonic region and give rise to absorptions due to mode coupling (see below).

The scanned DMD spectrum (Figure 8) for $D_7O_3^+$ shows peaks around 1700−1800 cm^{-1}. The normal mode fundamental frequencies (displayed as sticks in the spectrum) for asymmetric and symmetric bridging hydrogen stretch for $D_7O_3^+$ are equal to 1802 and 1839 cm^{-1}, respectively. At longer driving simulation times, these peaks broaden, and weak features are seen at 1550 and 1900 cm^{-1}, similar to the IR spectrum shown in Figure 7. The maximum peaks of the average absorbed energies (eq 6) as a function of time are shown in Figure 9. At nonfundamental and absorbing frequencies (1400 cm^{-1} for $D_7O_3^+$, 2100 cm^{-1} for $H_7O_3^+$), the absorbed energy is small and oscillatory with time, while at fundamental frequencies (maximum peaks at 1800 cm^{-1} for $D_7O_3^+$, 2375 and 2500 cm^{-1} for $H_7O_3^+$) the absorbed energy increases rapidly with time. The absorption profiles for several nonfundamental frequencies (1625, 1675, 1725, 1900 cm^{-1} for $D_7O_3^+$ and 2225, 2300 cm^{-1} for $H_7O_3^+$) are more complex. The absorbed energy is small and oscillating until the system absorbs enough energy and escapes the harmonic region of the potential energy surface. The moderate or rapid absorption is observed then. These frequencies are expected to be combination bands due to mode coupling.

Analysis of absorbing modes depends on the driving time as well as on the strength of the field. To study the effect of the strength of the electric field on the DMD simulations, we carried out a set of driven simulations using electric field strengths $\vec{\varepsilon}_0$= 25, 50, and 75 mV/bohr for maximum peaks at 2225 and 2375 cm^{-1} in the DMD spectrum (Figure 8). Figure 10 shows the energy profiles as a function of driven time. The DMD peak at 2375 cm^{-1} represents a resonance with the normal mode at 2387 cm^{-1}, and it shows very little effect of the electric field strength on the absorbed energy profile, while driving at nonfundamental frequency 2225 cm^{-1}, the stronger the electric field, more energy is absorbed by the molecule at much shorter time. To resolve the weak spectral feature at 2166 cm^{-1} found in the MD spectrum (Figure 4), we gradually increased the electric field, until the absorption of energy was identified. Dipole driving does not show much response for the electric field up to $\vec{\varepsilon}_0 = 75$ mV/bohr, while the system started to absorb rapidly when the electric field $\vec{\varepsilon}_0 = 78$ mV/bohr was applied.

To quantify similarities between the DMD displacements and normal mode vectors, the average overlaps (eq 8) and average weights[26] were calculated for each driven frequency. We analyzed the overlap matrix for two fundamental frequencies 2387 and 2554 cm^{-1} ($\vec{\varepsilon}_0 = 50$ mV/bohr) corresponding to asymmetric and symmetric stretches of bridging hydrogen and for two nonfundamental frequencies 2166 ($\vec{\varepsilon}_0 = 78$ mV/bohr) and 2225 cm^{-1} ($\vec{\varepsilon}_0 = 50$ mV/bohr). At short simulation times, up to 100 fs, the average weights (diagonal elements of the overlap matrix) for fundamental frequencies 2387 and 2554 cm^{-1} are 0.435 and 0.204, respectively. Off-diagonal elements of the overlap matrix are relatively small (less than 0.1). The average absorbed energy was only 0.7 and 0.5 kcal/mol, respectively. At longer driving times, 250 fs, the average absorbed energies have reached values 2.9 and 1.3 kcal/mol for frequencies 2387 and 2554 cm^{-1}, respectively, and the average weights for these two normal modes are only 0.299 and 0.247. Also, we observed increasing torsion motion at all resonant frequencies. The analysis of the overlap matrix is meaningful shortly after the system hits the resonance and starts to absorb energy. We now analyze dipole driving for nonfundamental peaks at 2166 cm^{-1} using the electric field $\vec{\varepsilon}_0 = 78$ mV/bohr, 2225 and 2300 cm^{-1} using $\vec{\varepsilon}_0 = 50$ mV/bohr. Evident mixing among the modes is reflected in overlaps. The overlap matrix shows involvement of the H_2O in phase wagging, OO stretch, H_3O^+ rocking, and BH asymmetric stretch vibrations for nonfundamental frequencies. The overlaps with the normal mode vectors are relatively small, about 0.1 for H_2O in phase wagging, OO stretch, and H_3O^+ rocking modes, and 0.3 for BH asymmetric stretch vibrations.

To clarify the spectral assignment of the nonfundamental frequencies, we also monitored structural parameters along

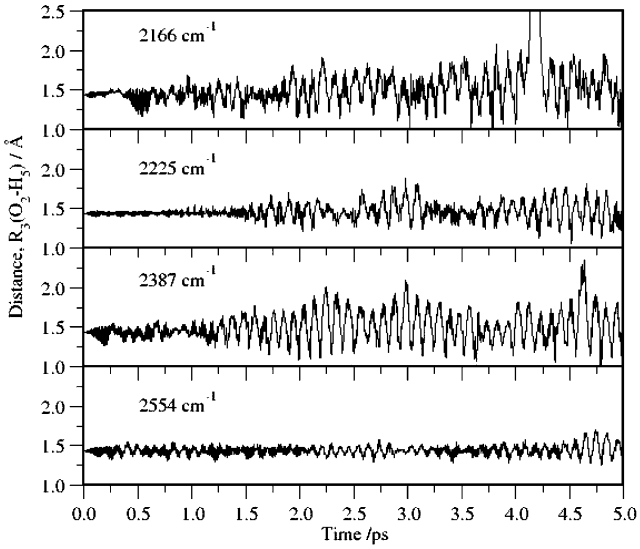


Figure 11. Results of DMD simulations: time dependence of the hydrogen bond $R_3(O_2-H_5)$. Atom numbering is shown in Figure 1. For driven frequencies $\omega = 2225$, 2387, and 2554 cm^{-1}, the electric field was $\vec{\varepsilon}_0 = 50$ mV/bohr, while for driven frequency $\omega = 2166$ cm^{-1}, $\vec{\varepsilon}_0 = 78$ mV/bohr.

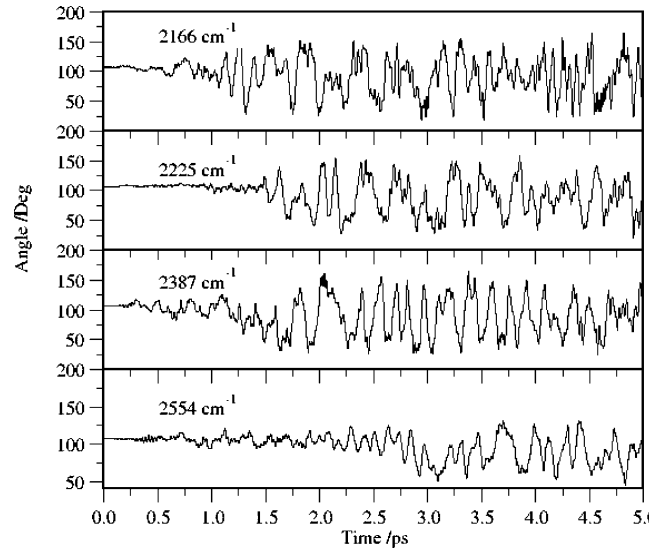


Figure 12. Time dependence of the torsion angle between two planes H5−O1−H4 and H9−O2−H10. For driven frequencies $\omega = 2225$, 2387, and 2554 cm^{-1}, the electric field was $\vec{\varepsilon}_0 = 50$ mV/bohr, while for driven frequency $\omega = 2166$ cm^{-1}, $\vec{\varepsilon}_0 = 78$ mV/bohr.

the DMD trajectory. Figure 11 shows time dependence of the hydrogen bond between O2 and H5 atoms. The atomic fluctuations increase with time until the system absorbs sufficient energy, and then O2−H5 distance fluctuates around the equilibrium value 1.435 Å. In the case of 2387 cm^{-1} the atomic fluctuations are large comparing to the 2554 cm^{-1} frequency. Driving at 2166 cm^{-1} nonfundamental frequency shows a similar pattern of atomic fluctuations to ones at 2387 cm^{-1}, supporting our spectral assignment of 2166 cm^{-1} frequency to BH stretch. In the DMD trajectory propagated with $\omega = 2166$ cm^{-1}, we observed proton exchange in the system at 4.2 ps. In this simulation, the maximum absorbed energy reaches 28 kcal/mol. In the case of nonfundamental frequency 2225 cm^{-1}, the bridging hydrogen fluctuations are relatively small up to 1.5 ps. The maximum absorbed energy at this point was about 2.8 kcal/mol. This is probably a threshold energy the system must absorb in order to activate this vibration. Visualization of the DMD trajectory demonstrates that the atomic motion acquires significant torsion character almost at all resonant frequencies. We plot the angle between the two planes H5−O1−H4 and H9−O2−O10 (Figure 12) to describe the torsion motion. Fluctuations of bridging hydrogen and torsion angle show similar patterns along the DMD trajectory that imply possible mixing between these modes.

6. Conclusions

(a) In this work, we present new results on analysis of the potential energy surface. Several stationary points on the PES of $H_7O_3^+$ were identified. They are the high energy transition states with barrier heights about 5−7 kcal/mol and the low energy transition structures that represent the umbrella motion of the H_3O^+ ion and rotation of one of two waters in $H_7O_3^+$ with 1.18 and 0.36 kcal/mol barrier heights at CCSD(T)/cc-pVTZ //MP2/cc-pVTZ level of theory, respectively.

The dissociation energies of $H_3O^+ \cdot (H_2O)_n$ for loss of one H_2O are 33.7 and 20.7 kcal/mol for $n = 1$ and $n = 2$ at the MP2/aug-cc-pVTZ level compared to 32.4 and 20.5 kcal/mol experimental values.[40]

(b) The dipole spectra were obtained for 100, 300, and 500 K at the B3LYP/6-31+G** level of theory and compared to previous theoretical calculations and experimental measurements. In this work, our focal point was analysis of shared proton vibrations. The spectra are decomposed to the parallel and perpendicular components to identify symmetry related vibrations. We also evaluated position−position spectra to analyze vibrations in terms of atomic groups. On the basis of NMA, the spectral peaks at 2410 and 2540 cm^{-1} were assigned to the asymmetric and symmetric bridging hydrogen stretch vibrations similar to the assignment in the previous theoretical work.[14] However, the previous theoretical work did not analyze the weak spectral feature around 2000 cm^{-1}. Using polarized dipole−dipole spectra and position−position spectra, we assigned the weak spectral features at 2166 and 2275 cm^{-1} to bridging hydrogen vibrations.

The effect of isotopic substitution was analyzed. We reported vibrational spectra of several D-substituted isotopologues $D_7O_3^+$, $DH_6O_3^+$, and $D_3H_4O_2^+$. Our simulations can guide future experimental measurements similar to experimental H/D isotopic studies done for $H_5O_2^+ \cdot Ar$.[41,42]

(c) Bond length distributions for $H_7O_3^+$ for temperatures 100, 300, and 500 K indicate the presence of both Eigen-based and Zundel ion cores in the MD simulations of $H_7O_3^+$. The simulations do not reveal formation of H_3O^+ and water dimer configurations as suggested in the previous work;[14] however, at high temperature 500 K, we observed frequent hydrogen exchange between H_3O^+ and H_2O.

(d) The emphasis of this paper was the development and implementation of the driven molecular dynamics method to assign the spectral features. This is a first attempt to run

direct DMD simulations. The challenging part is to select the electric field strength and to determine the simulation time to obtain meaningful results. DMD simulations with weak electric field and short simulation times reproduce the harmonic features, while using higher intensity electric field and longer simulation times reveals anharmonic features. The scanned DMD spectrum can be correlated directly to the spectrum from the classical MD simulations. The average absorbed total energy was monitored as a function of driven frequency. DMD simulations confirm the assignment of the dominant peaks at 2410 and 2540 cm^{-1} to asymmetric and symmetric bridging hydrogen vibrations. The nonfundamental frequencies at 2166, 2225, and 2300 cm^{-1} were assigned to a combination band of BH asymmetric stretch, H_2O in phase wagging, OO stretch, and H_3O^+ rocking vibrations by DMD simulations. We evaluated overlap matrix to quantify the involvement of a certain vibration in the spectral feature. Analysis of the molecular motion along the DMD trajectories also guided the assignment of the spectral features. Visualization of the DMD trajectory demonstrates that the atomic motion acquires significant torsion character almost at all resonant frequencies.

(e) Our observation of BH stretch vibrations as low as 2166 cm^{-1} is in good agreement with the assignment of the low-resolution spectrum obtained by Schwarz at 2200−2300 cm^{-1} and the assignment of the $H_7O_3^+ \cdot H_2$ spectrum by Lee at ~2300 cm^{-1}. However, we did not observe asymmetric and symmetric bridging hydrogen vibrations at 1880 and 3580 cm^{-1} as reported by recent Ar-tagged predissociation spectra.[20] The authors concluded that anharmonic corrections are required to qualitatively recover the large observed red shifts to describe $H^+(H_2O)_3$ stretching transitions.

Quantum dynamics study would help to resolve uncertainties in the experimental observations of the shared proton vibrations; however, this would be computationally challenging. Also, the perturbations of the $H_7O_3^+$ spectrum by attached argon atom have to be evaluated. Previous calculations indicate a strong effect of an Ar messenger atom on $H_5O_2^+$ spectra.[43] The $O-H^+-O$ asymmetric stretch mode is significantly blue-shifted in $H_5O_2^+ \cdot Ar$. We believe that the present study provides new insight into the shared proton vibrations of $H_7O_3^+$ system and its D-substituted isotopologues and describes the application of the DMD method to assign spectral features.

Acknowledgment. The authors thank the Research Corporation (Cottrell College Science Award 7725) for support of this work.

References

(1) Marx, D.; Tuckerman, M. E.; Hutter, J.; Parrinello, M. *Nature* **1999**, *397*, 601–604.

(2) Marx, D.; Tuckerman, M. E.; Parrinello, M. *J. Phys.: Condens. Matter* **2000**, *12*, A153–A159.

(3) Paesani, F.; Zhang, W.; Case, D. A.; Cheatham, T. E., III; Voth, G. A. *J. Chem. Phys.* **2006**, *125*, 184507.

(4) Iyengar, S. S.; Petersen, M. K.; Day, T. J. F.; Burnham, C. J.; Teige, V. E.; Voth, G. A. *J. Chem. Phys.* **2005**, *123*, 084309.

(5) Zundel, G. In *The Hydrogen Bond: Recent Developments in Theory and Experiment*; Schuster, P., Zundel, G., Sandorfy, C., Eds.; North Holland: Amsterdam, 1976; Vol. 2, p 683.

(6) Giguere, P. A.; Turrell, S. *Can. J. Chem.* **1976**, *54*, 3477–3482.

(7) Librovich, N. B.; Sakun, V. P.; Sokolov, N. D. *Chem. Phys.* **1979**, *39*, 351–366.

(8) Atkins, P. W.; de Paula, J. *Physical Chemistry*, 8th ed., Freeman: New York, 2002; p 763.

(9) Eigen, M.; Wicke, E. *J. Phys. Chem.* **1954**, *58*, 702–714.

(10) Zundel, G.; Metzger, H. *Z. Phys. Chem. (Munich)* **1968**, *58*, 225–241.

(11) Agmon, N. *Chem. Phys. Lett.* **1995**, *244*, 456–462.

(12) Wei, D.; Salahub, D. R. *J. Chem. Phys.* **1994**, *101*, 7633–7642. Wei, D.; Salahub, D.R. *J. Chem. Phys.* **1997**, *106*, 6086–6094.

(13) Yan, S.; Zhang, L.; Cukier, R. I.; Bu, Y. *Chem. Phys. Phys. Chem.* **2007**, *8*, 944–954.

(14) Termath, V.; Sauer, J. *Mol. Phys.* **1997**, *91*, 963–975.

(15) Schwarz, H. A. *J. Chem. Phys.* **1977**, *67*, 5525–5534.

(16) Yeh, L. I.; Okumura, M.; Myers, J. D.; Price, J. M.; Lee, Y. T. *J. Chem. Phys.* **1989**, *91*, 7319–7390.

(17) Okumura, M.; Yeh, L. I.; Myers, J. D.; Lee, Y. T. *J. Chem. Phys.* **1986**, *85*, 2328–2329. Okumura, M.; Yeh, L. I.; Myers, J. D.; Lee, Y. T. *J. Chem. Phys.* **1990**, *94*, 3416–3427.

(18) Honma, K.; Armentrout, P. B. *J. Chem. Phys.* **2004**, *121*, 8307–8320.

(19) Douberly, G. E.; Walters, R. S.; Cui, J.; Jordan, K. D.; Duncan, M. A. *J. Phys. Chem. A* **2010**, *114*, 4570–4579.

(20) Headrick, J. M.; Diken, E. G.; Walters, R. S.; Hammer, N. I.; Christie, R. A.; Cui, J.; Myshakin, E. M.; Duncan, M. A.; Johnson, M. A.; Jordan, K. D. *Science* **2005**, *308*, 1765–1769.

(21) Chaban, G. M.; Jung, J. O.; Gerber, R. B. *J. Chem. Phys.* **1999**, *111*, 1823–1829.

(22) Kaledin, M.; Moffitt, J. M.; Clark, C. R.; Rizvi, F. *J. Chem. Theory Comput.* **2009**, *9*, 1328–1336.

(23) Bowman, J. M.; Zhang, X.; Brown, A. *J. Chem. Phys.* **2003**, *119*, 646–650.

(24) Kaledin, M.; Brown, A.; Kaledin, A. L.; Bowman, J. M. *J. Chem. Phys.* **2004**, *121*, 5646–5653.

(25) Kaledin, M.; Kaledin, A. L.; Brown, A.; Bowman, J. M. In *Normal Mode Analysis: Theory and Applications to Biological and Chemical Systems*; Cui, Q., Bahar, I., Eds.; CRC Press: Boca Raton, FL, 2006.

(26) Kaledin, M.; Kaledin, A. L.; Bowman, J. M. *J. Phys. Chem. A* **2006**, *110*, 2933–2939.

(27) 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.; and Pople, J. A. *Gaussian 03, Revision C.02*; Gaussian, Inc.: Wallingford, CT, 2004.

(28) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.

(29) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372–1377.

(30) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(31) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. *Chem. Phys. Lett.* **1988**, *153*, 503–506.

(32) Saebo, M. J. S.; Almlöf, J. *Chem. Phys. Lett.* **1989**, *154*, 83–89.

(33) Frisch, M. J.; Head-Gordon, M.; Pople, J. A. *Chem. Phys. Lett.* **1990**, *166*, 275–280.

(34) Wilson, E. B.; Decius, J. C.; Cross, P. C. *Molecular Vibrations*; McGraw-Hill: New York, 1955.

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

(36) Berens, P. H.; Wilson, K. R. *J. Chem. Phys.* **1981**, *74*, 4872–4882.

(37) Haile, J. M. *Molecular Dynamics Simulation: Elementary Methods*; Wiley & Sons: New York, 1992; p 277.

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

(39) Huang, X.; Braams, B. J.; Bowman, J. M. *J. Chem. Phys.* **2005**, *122*, 044308.

(40) Dalleska, N. F.; Honma, K.; Armentrout, P. B. *J. Am. Chem. Soc.* **2003**, *115*, 12125–12131.

(41) McCunn, L. R.; Roscioli, J. R.; Johnson, M. A.; McCoy, A. B. *J. Phys. Chem. B* **2008**, *112*, 321–327.

(42) McCunn, L. R.; Roscioli, J. R.; Elliott, B. M.; Johnson, M. A.; McCoy, A. B. *J. Phys. Chem. A* **2008**, *112*, 6074–6078.

(43) Park, M.; Shin, I.; Singh, N. J.; Kim, K. S. *J. Phys. Chem. A* **2007**, *111*, 10692–10702.

CT100122S

JCTC Journal of Chemical Theory and Computation

Geometry Optimization of Large and Flexible van der Waals Dimers: A Fragmentation−Reconstruction Approach

Ivo Cacelli,[†] Antonella Cimoli,[†] and Giacomo Prampolini*[,†,‡]

Dipartimento di Chimica e Chimica Industriale, Università degli Studi di Pisa, via Risorgimento 35, I-56126 Pisa, Italy, and Scuola Normale Superiore, piazza dei Cavalieri 7, I-56126 Pisa, Italy

Received March 30, 2010

Abstract: A novel approach for exploring the energy minima of the potential energy surface of large and flexible van der Waals dimers is proposed and tested. The total dimer energy is divided into intra- and intermolecular contributions, which can be computed at different levels of theory. The intermolecular energy, which is the time-consuming part of the calculation, is computed by means of the fragmentation reconstruction method (FRM), making possible the calculation of the interaction energy of large molecules. The method is validated by performing geometry optimizations through a quasi-Newton technique on two benchmark medium-sized systems, where the comparison with a direct *ab initio* calculation is still computationally feasible. In both cases, good agreement is achieved between geometries and energies of the resulting energy minima.

1. Introduction

van der Waals (vdW) interactions[1−5] are known to play a relevant role in many different fields of science, ranging from soft matter[6−9] to biochemistry,[10−12] and from molecular recognition[13] to nanotechnologies[14−16] and astrobiology.[17] Nevertheless, for molecular dimers of medium to large dimensions, the accurate calculation of such interactions with standard quantum mechanical (QM) or *ab initio* methods still remains a grand challenge. The main reason for these difficulties arises from the purely quantum mechanical nature of the vdW interactions, which strongly relies on a correct representation of the dynamical electron correlation. For this reason, within the *ab initio* methods, one has to employ post-Hartree−Fock (HF) techniques, whose computational costs dramatically increase along with molecular dimensions. This picture is made even worse when dealing with flexible molecules, whose isolated equilibrium conformation can be altered by the intermolecular interactions. In these cases, dimer energy minimizations should not be performed in a rigid monomer geometry approximation. Moreover, soft matter is characterized[6−8] by molecular and collective conformational changes driven by thermal fluctuations, and an accurate sampling of several local energy minima is therefore required to spot the most probable dimer conformations. A striking example of such features can be found even in a relatively small molecule such as biphenyl, where the torsional angle between the two phenyl rings assumes different values according to the phase (crystal, liquid, or gas) in which it is measured (see ref 18 and references therein). In more complex systems, for instance polymers, liquid crystals, or protein−ligand complexes, the determination of the most stable dimer structures is of fundamental importance in the understanding of the condensed phase properties. A straightforward route to accessing this information could be the use of molecular mechanics (MM), which is often adopted in the case of biomolecular dimers and allows for the computation of very large molecules. Unfortunately, it has recently been shown[19−21] that literature force fields are not able to provide dimer structures in agreement with reference QM data for a large variety of systems.

To overcome the computational problems of QM calculations on large vdW molecules, several fragmentation-based

* To whom correspondence should be addressed. E-mail: giacomo@dcci.unipi.it.

[†] Universita degli Studi di Pisa.

[‡] Scuola Normale Superiore.

10.1021/ct100172w © 2010 American Chemical Society
Published on Web 07/12/2010

strategies have been proposed[22−32] that essentially rely on the possibility of dividing the whole system into small subsystems and performing the calculations for each subsystem. Among these, the fragment molecular orbital (FMO) method[29] and the fragmentation reconstruction method (FRM)[22−24] were devised to account for dispersion energy, whereas other approaches are either coupled with calculations at the HF or DFT level[25,26] or implemented for monomer optimizations.[27,28] To the best of our knowledge, large dimer optimizations were performed only in the FMO approach,[30,33,34] but unfortunately gradients are not yet implemented in FMO coupled with post-HF methods. Furthermore, in all cases, dimer geometry optimizations were performed in a rigid monomer approach.

In this paper, we report a novel approach, based on the FRM[22−24] route to intermolecular energy previously developed in our group, that can be used to study vdW complexes by optimizing the relative position as well as the intramolecular geometry of each monomer, at least for the most flexible internal coordinates. This approach has been implemented in original software code written by us and named POLDO and is able to perform geometry optimizations of large and flexible vdW dimers otherwise impracticable with standard methods. On the other hand, even if FRM reduces the dimensions of the problem (with respect to the whole dimer), a reasonable compromise between accuracy and computational cost is still a major requirement, in view of the large number of calculations possibly involved in a dimer geometry optimization.

In the past decade, along with the impressive development of the computational resources, several groups have reported highly accurate calculations for model $\pi\cdots\pi$ interacting systems. Among others, the groups of Tsuzuki,[35,36] Hobza,[12,20,37,38] and Sherrill[21,39−44] have reported interaction energies at the CCSD(T) level, extrapolated at the complete basis set limit, which is now often referred to as the "gold standard" of quantum chemistry. The major drawback of this approach is the extremely high computational cost of the CCSD(T) method, which scales approximatively as N^7, making calculations rapidly unfeasible with the increase of molecular dimensions. Furthermore, to the best of our knowledge, no geometry optimization at this level of theory has yet been reported in the literature. Computational convenience would rather suggest resorting to a cheaper method, either density- or wavefunction-based.

The methods based on density functional theory (DFT) are generally less expensive. Unfortunately, none of the standard density functionals was shown to be able to reproduce even benzene dimer interaction curves at a qualitative level[45,46] because of an incorrect evaluation of dispersion contribution. Only recently did Truhlar and Zhao[47−49] succeed in reparameterizing the DFT functional (M06-2X) in order to take dispersion into account, achieving a reasonable agreement[49,50] with reference data for benchmark systems. However, very recent results[38] have put into evidence some defects in the overall representation of the computed interaction energy curves. In the same paper,[38] somewhat better performance was reported for the dispersion corrected DFT-D approach, where the standard DFT energy is corrected with an empirical term[43,44,51,52] to reproduce most of the dispersion interaction energy of aromatic dimers.

Finally, among mixed DFT-wavefunction techniques, it is worth mentioning that the SAPT-DFT approach has been shown[38,53,54] to yield results in good agreement with the most accurate *ab initio* values. However, despite its scaling properties with both the molecular dimensions and basis sets being more favorable than CCSD(T) methods,[53] this approach is still computationally too expensive for its straightforward application to the calculation of large vdW dimers.

As far as the wavefunction-based methods are concerned, a computationally convenient post-HF method is the Møller−Plesset second order perturbation (MP2) theory. Besides its affordable costs, another advantage relies in the availability of the energy gradients, which can be straightforwardly used in the POLDO code. On the other hand, a major drawback of the MP2 method is the remarkable overbinding found for aromatic interactions,[35,36,43,55−57] when large basis sets are employed. To overcome this lack, a few MP2-based methods have been recently proposed. Among these, the spin-component-scaled MP2 (SCS-MP2)[58] has been recently compared to CCSD(T) reference values for model systems,[43,57] resulting in a good agreement for $\pi\cdots\pi$ interactions but not for H-bonded systems. The scaled MP3 method (MP2.5)[57] appears to be more accurate, but the computational cost of this method is about 1 order of magnitude higher than standard MP2. A promising method,[59] which combines MP2 and time-dependent TDDFT response theory, was proposed to improve the agreement with the reference data on a set of benchmark aromatic dimers, at a computational cost slightly higher than standard MP2. However, it is worth noticing that the best agreement was obtained making use of rather large basis sets. An alternative choice is the adoption of the standard MP2 method, coupled with a small basis set (6-31G*), where the polarization exponent has been modified to a smaller value (0.25 instead of the standard 0.80) and therefore named 6-31G*(0.25).[60] The comparison with more expensive computational methods, recently reported for molecular complexes with stacking interactions,[19,38,61,62] is positive, since the intermolecular energy overestimation found with larger basis sets is avoided. The MP2/6-31G*(0.25) method was successfully employed by our group in sampling the QM PES of several dimers, among which are benzene[19] and 4-*n*-pentyl,4′cyano-biphenyl (5CB),[22,63] a common mesogenic molecule. In the last case, MP2/6-31G*(0.25) calculations were performed in the FRM scheme.

In this paper, the dimer geometry optimizations are performed with the POLDO code, which uses the FRM approach at the MP2/6-31G*(0.25) level. To test and validate the proposed method, two benchmark medium-sized molecules are chosen, namely, biphenyl and 5CB. The former can be fragmented into benzene moieties for which MP2/6-31G*(0.25) has already been validated. Moreover, its molecular dimensions allow us to perform direct (i.e., without fragmenting) optimizations at the same level of theory that are used to validate the proposed procedure. 5CB is one of the smaller prototypes of vdW dimers containing three

different moieties, namely, aromatic (phenyl), substituted aromatic (cyano-phenyl), and aliphatic (pentyl) bricks. The paper is organized as follows: Section 2 contains the method, theory, and main computational details. The results of POLDO optimizations are discussed in section 3, whereas main conclusions are collected in the last section.

2. Methods and Computational Details

2.1. Dimer Optimizations. The absolute energy E_{tot} of a vdW A···B dimer can be expressed as

$$E_{tot}(AB) = \Delta E(AB) + E(A_0) + E(B_0) \quad (1)$$

where the labels A_0 and B_0 indicate the monomers in their isolated minimum energy conformations, whereas AB stands for the dimer geometry. As the last two terms of eq 1 are constant, the geometry is optimized considering the ΔE(AB) term, which can be seen as the sum of two separate parts, namely an intermolecular (ΔE_{inter}(AB)) and an intramolecular (ΔE_{intra}(AB)) contribution:

$$\Delta E(AB) = \Delta E_{intra}(AB) + \Delta E_{inter}(AB) \quad (2)$$

where

$$\Delta E_{intra}(AB) = [E(A) - E(A_0)] + [E(B) - E(B_0)] \quad (3)$$

and

$$\Delta E_{inter}(AB) = E(AB) - E(A) - E(B) \quad (4)$$

The main reason for considering two distinct contributions in eq 2 is that accurate calculation of ΔE_{intra} and ΔE_{inter} presents a different level of difficulty, and in the proposed approach they are therefore computed in different ways. The former term may be evaluated employing standard QM techniques or by a suitable and accurate MM intramolecular force field. In this last case, only the internal geometrical degrees of freedom to be optimized need to be included in the set of intramolecular coordinates. Conversely, the intermolecular term ΔE_{inter} is computed using the FRM approach, which includes the counter poise (CP) correction[64] to the basis set superposition error (BSSE).

The minimization of E_{tot}(AB) vs the geometrical degrees of freedom has been performed with the POLDO program. POLDO is a FORTRAN code written by the authors and is freely available upon request. The geometrical quantities taken into account are the following:

(1) The translational vector $\vec{R} \equiv (X, Y, Z)$, connecting a defined point of monomer A with a defined point of monomer B. These points may be the monomer barycenter (default choice), a defined atom, or the midpoint of an assigned bond.

(2) Vectors $\vec{\Omega}_i \equiv (\alpha, \beta, \gamma)$ of each monomer i. $\vec{\Omega}_i$ (i = A, B) defines the absolute orientation of one monomer in a fixed reference frame through the three Euler angles α, β, and γ. The rotation center is a point defined within each monomer. As the internal geometry of both monomers can change during the optimization path, the values of the Euler angles are not sufficient to describe the effective rotation of each monomer, and should be merely considered as degrees of freedom. Moreover, since the interaction energy only depends on the reciprocal orientation, six orientational parameters are a redundant set. However, the inclusion in the optimization of Euler angles for both monomers, though not strictly necessary, was found to speed up the convergence toward the absolute minimum.

(3) Vector $\vec{Q}_i$ (i = A, B), containing the internal coordinates of both monomers. Although POLDO can handle in principle any type and number of coordinates, it can sometimes be preferable[65] to optimize only "soft" degrees of freedom, i.e., those internal coordinates which exhibit rather flat energy profiles, for instance, torsional angles along σ bonds. Indeed, considering the low values of the interaction energies as compared with the intramolecular contribution, it is expected that in common vdW complexes only the "soft" degrees of freedom can be significantly affected by the presence of the second molecule.

The optimization is performed using a gradient-based technique in which at each step the new geometry is determined by a quasi Newton method. The QM Cartesian gradients are converted to energy derivatives with respect to the chosen degrees of freedom. The needed transformation matrices are computed by numerical techniques using a five point interpolation formula, which assures a sufficient accuracy and avoids singularities connected with the use of analytical derivatives involving quaternion, in the case of rotational motion. In particular, for the derivative with respect to the translational degrees of freedom, in which the A monomer is fixed in space and B is moved, one has

$$\frac{\partial E_{tot}}{\partial I} = \sum_{i \in B} \left(\frac{\partial(\Delta E_{inter}(AB))}{\partial r_i} \right) \left(\frac{\partial r_i}{\partial I} \right); \quad I = X, Y, Z \quad (5)$$

where r_i ($i = 1, ..., 3N_B$) are the Cartesian coordinates of the atoms belonging to the molecule B formed by N_B atoms. Similarly, for the rotational degrees of freedom

$$\frac{\partial E_{tot}}{\partial \alpha_{B(A)}} = \sum_{i \in B(A)} \left(\frac{\partial(\Delta E_{inter}(AB))}{\partial r_i} \right) \left(\frac{\partial r_i}{\partial \alpha_{B(A)}} \right) \quad (6)$$

and analogous equations hold for β and γ. Finally, the gradients for an internal coordinate $Q_{A(B)}$ of the molecule A(B) is computed by

$$\frac{\partial E_{tot}}{\partial Q_{A(B)}} = \sum_{i \in A(B)} \left(\frac{\partial E(A(B))}{\partial r_i} \right) \left(\frac{\partial r_i}{\partial Q_{A(B)}} \right) + \sum_{i \in A(B)} \left(\frac{\partial(\Delta E_{inter}(AB))}{\partial r_i} \right) \left(\frac{\partial r_i}{\partial Q_{A(B)}} \right) \quad (7)$$

As many of the internal coordinates are kept frozen, the dummy index i in eq 7 runs over all atoms involved in a change of the $Q_{A(B)}$ coordinate and not only over those involved in its definition. For instance, for biphenyl (see next section), the only internal coordinate $Q_{A(B)}$ chosen for optimization is the inter-ring dihedral angle $\phi_{A(B)}$. In this case, the transformation matrix $(\partial r_i/\partial \phi_{A(B)})$ is computed numeri-

cally by small rigid rotation $\delta\phi_{A(B)}$ of all atoms of the rings and detecting the corresponding changes δr_i.

The POLDO code is interfaced with the Gaussian 03[66] software and automatically starts a Gaussian run for each QM calculation. The full minimization proceeds through the following steps:

(1) The energy of each isolated monomer is optimized with respect to all its internal coordinates using the method chosen for the intramolecular contribution and reference energies, $E(A_0)$ and $E(B_0)$, are recovered.
(2) The energy of each monomer at the starting geometry, $E(A)$ and $E(B)$, is computed with the same method chosen for step 1. POLDO recovers energies and relative Cartesian gradients and computes the $\Delta E_{intra}(AB)$ contribution according to eq 3.
(3) POLDO builds all the Gaussian input files needed by the FRM calculation (see the next section for details). Thereafter, all QM runs are automatically performed. Finally, all of the resulting energies and Cartesian gradients are recovered from Gaussian output files and combined in order to reconstruct the FRM interaction energy term $\Delta E_{inter}(AB)$.
(4) The full energy E_{tot} is computed as a sum of internal and intermolecular contributions according to eqs 1 and 2.
(5) All Cartesian gradients are converted into derivatives with respect to the chosen degrees of freedom, and the Hessian matrix is updated for the current geometrical arrangement.
(6) A new set of coordinates is computed according to the Newton−Raphson algorithm. If the total energy gradient and the computed displacements are small enough, the program ends.
(7) Steps 2−4 are repeated with the new geometry. If the total energy decreases, POLDO proceeds through steps 5 and 6; otherwise POLDO performs a line search in the direction of the predicted displacement and goes back to step 2.

2.2. FRM Calculations. The FRM[22−24] was originally developed in our group to calculate the interaction energy of large molecules. It basically relies on the hypothesis that the interaction energy of a vdW dimer can be approximated to a good level of accuracy as a sum of energy contributions between each pair of fragments into which the two monomers can be decomposed. The basic criterion behind this fragmentation scheme is that the ground state electronic density around the atoms of each fragment has to be as close as possible to that around the same atoms in the whole molecule. The main advantage of this approach lies in the possibility of performing calculations between moieties much smaller than the whole molecules under study. In previous applications,[23,24,63,67] this has allowed us to include electronic correlation effects and to obtain a good estimate of the dispersion energy which accounted for a large fraction of the attractive intermolecular energy of the considered dimers. The information gained about the interaction energy between the fragments pairs is then used to set up the full intermolecular energy.

The first step of the FRM is a decomposition of the whole molecule into moieties by fragmenting it along properly chosen bonds. The valence of the resulting fragments is then saturated by suitable atoms or small groups named "intruders." The intermolecular energy is thereafter expressed as a sum of the energy contributions arising from all of the fragment pairs. It is worth mentioning that, in order to recover the total dimer interaction energy, the intruder groups have to be subsequently canceled from the whole dimer and their energy contributions properly subtracted. More details about FRM can be found in the original papers.[22,24]

2.3. Computational Details. All isolated monomer calculations (eq 3) were performed at the DFT level of theory, using the B3LYP functional with a correlation consistent cc-pVDZ basis set. In the case of biphenyl, the intramolecular contribution $E(k)$ (k = A, B) was computed at the MM level, using the following expression for the inter-ring dihedral:

$$E(k) = \sum_n C_n(1 + \cos(n\phi_k)); \quad k = A, B \tag{8}$$

where all the remaining internal coordinates are fixed at their equilibrium values. The set of force-field parameters C_n was taken from ref 18, where they were derived from accurate QM calculations.

All FRM fragment−fragment calculations were performed at the MP2/6-31G*(0.25) level, in the supermolecule approach and corrected for the BSSE with the CP correction. Consequently, all MP2 direct optimizations and single point calculations, performed on the whole dimers, were correspondingly computed with the same basis set and CP corrected.

3. Results and Discussion

3.1. MP2/6-31G*(0.25) Validation. A mandatory test for vdW dimer optimizations is the validation of the level of theory to be used for the intermolecular energy calculations. It is worth pointing out that large and flexible vdW dimers may exhibit several energy minima, in which their reciprocal arrangements can be very different from simple stacked geometries. For this reason, all methods should be evaluated in several geometries, where, for instance, both $\pi\cdots\pi$ and $H\cdots\pi$ interaction types play different roles. The benzene dimer is a good candidate for this validation, as it can be considered as a prototype of vdW interactions, and at the same time, its dimensions allow high-level reference calculations.

To this aim, MP2/6-31G*(0.25) data are computed and results validated vs those of refs 43 and 44, where a detailed comparison among several methods has been reported for different geometries of the benzene dimer. Indeed, in these papers, Sherrill and co-workers report the interaction energy curves for sandwich (SW), T-shaped (TS), and parallel displaced (PD) benzene dimers at the CCSD(T) level, extrapolated at the complete basis set limit (CCSD(T)/CBS), which can be considered the reference values. In Figure 1, some of those curves are shown and compared with those obtained in this work using the MP2 method with the modified 6-31G*(0.25) basis set.

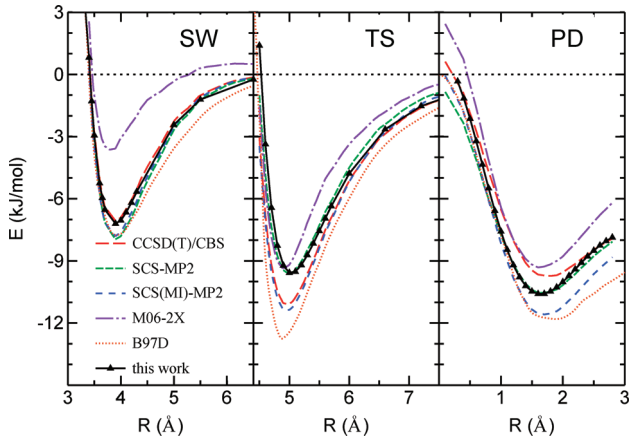


Figure 1. Interaction energies for sandwich (SW), T-shaped (TS), and parallel displaced (PD) arrangements of the benzene dimer at different levels of theory. All energies were taken from refs 43 and 44 where dimer geometries are described in detail (PD displacement is 3.4 Å). In this work, the same curves have been calculated at the MP2/6-31G*(0.25) level and reported with triangles in the above figure.

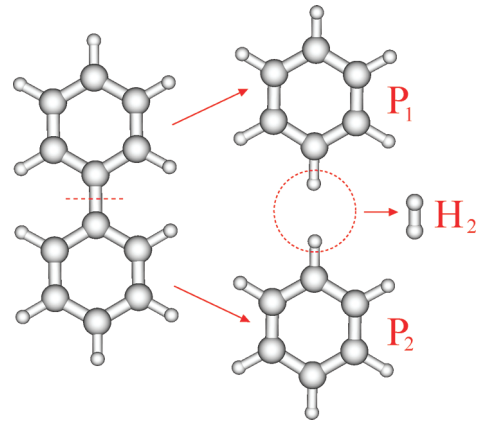


Figure 2. Fragmentation scheme for biphenyl. P_1 and P_2 are two benzenes, and H_2 is the intruder molecule. In the above picture, both moieties and intruder molecule have been shifted (red arrows) only for the sake of clarity (see text). For instance, both H atoms of the intruder molecole H_2 lie on the C−C bond connecting the two aromatic moieties. These moieties will be labeled with a further subscript A or B (that is $P_{1A(B)}$, $P_{2A(B)}$, $H_{2A(B)}$) depending on which of the two biphenyl monomers they belong.

Besides maintaining the correct relative order (i.e., $E_{TS} \simeq E_{PD} < E_{SW}$), MP2/6-31G*(0.25) shows an overall good agreement with the reference data: the accord is quantitative for the SW arrangement, whereas it slightly overestimates and underestimates CCSD(T)/CBS binding energies for PD and TS geometries by ∼0.9 and ∼1.5 kJ/mol, respectively. As far as the intermolecular separation at the minimum is concerned, a good agreement results for all geometries. Most important, the comparisons reported in refs 43 and 44 show that all other methods taken into account perform either worse or comparably. Therefore, at least for the benzene dimer, the MP2/6-31G*(0.25) method seems a good compromise between accuracy and computational cost (also considering the large number of calculations requested in an optimization procedure).

3.2. Biphenyl. After some preliminary tests (whose results are reported as Supporting Information), the first validation of the proposed procedure was attemped on the biphenyl molecule, one of the smaller prototypes of a fragmentable vdW dimer. Biphenyl is composed of two phenyl moieties (P_1 and P_2, see Figure 2) connected by a quasi-σ bond, seeing that the dihedral angle ϕ between the two rings is ∼42° in the gas phase.[18]

The intramolecular contribution, ΔE_{intra}, was computed at the MM level according to eq 8. As far as the intermolecular energy is concerned, the FRM scheme was applied to each biphenyl monomer by a cut along the inter-ring bond (see Figure 2). Indeed, this is the only possible fragmentation choice that respects the electronic density criteria exposed in the previous section. H atoms were then added to saturate the valence of the resulting C_6H_5 moieties, as shown in Figure 2. It is worth noticing that the spatial positions of the fragments are unchanged with respect to the whole molecule, and the location of the intruders H_{2A} and H_{2B} is unambiguously determined by the internal geometry of the saturated fragments. In other words, the Cartesian coordinates of all atoms in the fragments are unchanged with respect to the ones in the whole molecule, resulting in a slightly altered H_2 bond distance (0.68 Å, instead of the equilibrium value of 0.74 Å). The intermolecular energy of the whole biphenyl dimer is thus made up of the following contributions (see Figure 2):

$$\Delta E_{inter}(AB) = \underbrace{\Delta E(P_{1A}\cdots P_{1B}) + \Delta E(P_{1A}\cdots P_{2B}) + \Delta E(P_{2A}\cdots P_{1B}) + \Delta E(P_{2A}\cdots P_{2B})}_{PhH\cdots PhH} - \underbrace{(\Delta E(H_{2A}\cdots P_{1B}) + \Delta E(H_{2A}\cdots P_{2B}) + \Delta E(P_{1A}\cdots H_{2B}) + \Delta E(P_{2A}\cdots H_{2B}))}_{H_2\cdots PhH} + \Delta E(H_{2A}\cdots H_{2B}) \quad (9)$$

where $\Delta E(i_A\cdots j_B)$ is the interaction energy between fragment i, belonging to A, and fragment j, belonging to B.

The starting geometry was created by randomly displacing the B monomer and setting both ϕ's at 28°, as shown in Figure 3; the energy of the obtained arrangement is −10.8 kJ/mol, resulting from a sum of $\Delta E_{inter} = -15.4$ kJ/mol and $\Delta E_{intra} = 4.0$ kJ/mol. Given the reduced flexibility of the ring internal coordinates, this dimer conformation was optimized with the POLDO software by varying the translational ($\vec{R}$) and rotational ($\vec{\Omega}_B$) sets of coordinates together with both internal dihedrals ϕ_A and ϕ_B. The most striking feature during optimization is the rapid decrease to 0° of the β_B angle, which indicates a tendency of the two biphenyl monomers to realign their *para* axes, thus maximizing the $\pi\cdots\pi$ interactions. The inclusion of translational degrees of freedom allows the molecules to move along Y, reducing their distance to 3.6 Å, with a small displacement (−1.7 Å) along the X direction. Finally, the torsional dihedrals both reach the equilibrium value, found for the isolated monomer. The result is a local dimer minimum, almost isoenergetic (−31.5 kJ/mol vs −32.9 kJ/mol) with that found in preliminary optimizations (see test 2 in the Supporting Information), where the biphenyl long axes were imposed to be parallel. Some relevant geometrical data of the final optimized dimer are reported in Table 1.

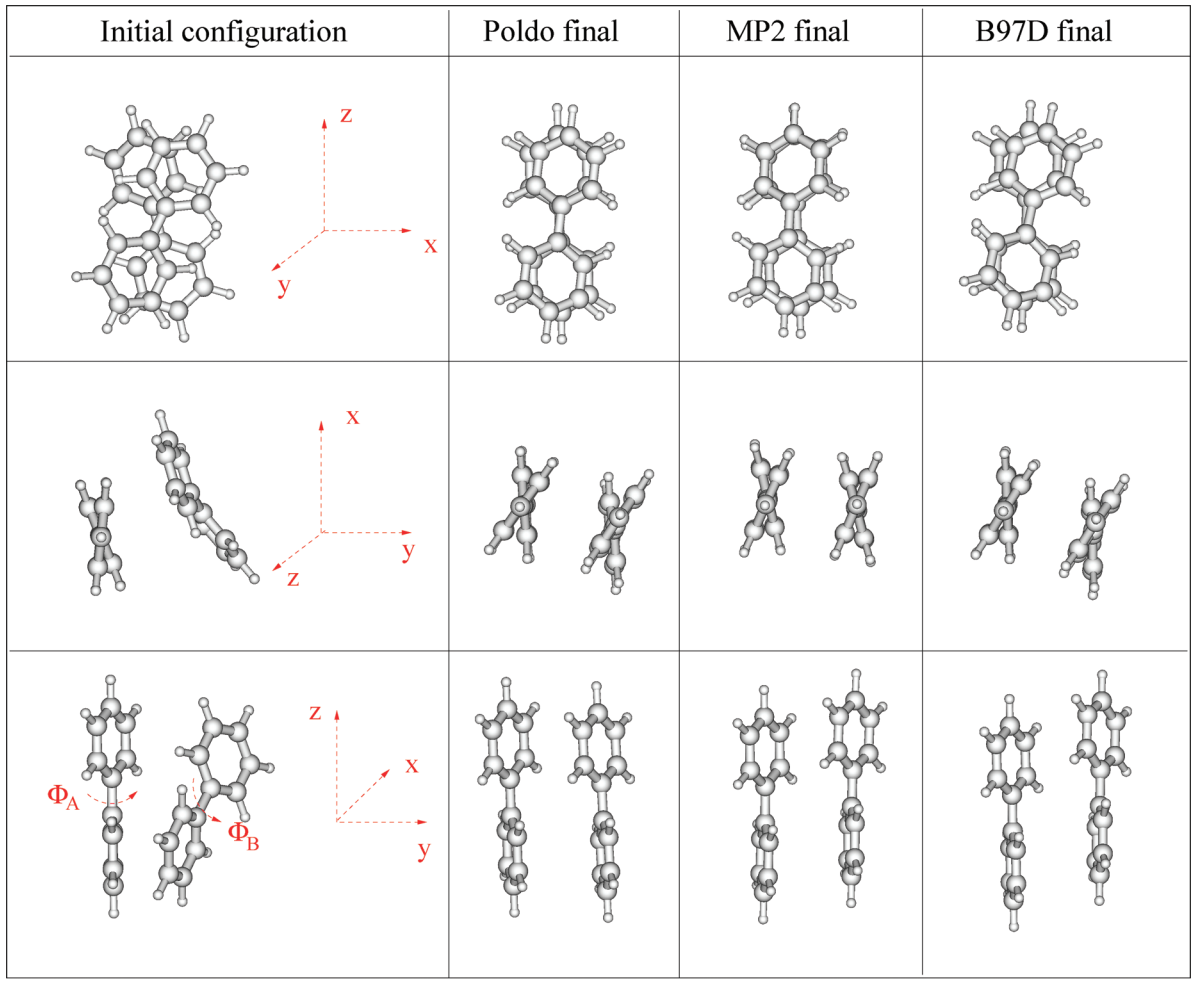


Figure 3. Complete optimization of a randomly placed biphenyl dimer. Starting (left column) and POLDO optimized (second column) geometries are shown from different points of view. The third and fourth columns refer to MP2 and B97D direct optimizations of the whole dimer, respectively.

Table 1. Optimized Biphenyl Dimer: Total Energy and Geometrical Values Obtained with Direct MP2 (second column), POLDO (third column), and B97D Optimizations

	Direct MP2	POLDO	B97D
ϕ_A	37.0°	40.0°	34.4°
ϕ_B	37.0°	40.0°	34.6°
$H_4^A\widehat{C_1^AC_1^B}H_4^B$	4.6°	8.0°	9.6°
$H_4^AH_4^B$	3.85 Å	4.04 Å	3.91 Å
$H_{4'}^AH_{4'}^B$	3.84 Å	3.62 Å	3.92 Å
$C_4^AC_4^B$	3.81 Å	3.97 Å	3.87 Å
$C_{4'}^AC_{4'}^B$	3.80 Å	3.64 Å	3.88 Å
$C_1^AC_1^B$	3.70 Å	3.81 Å	3.79 Å
$C_{1'}^AC_{1'}^B$	3.70 Å	3.74 Å	3.79 Å
$E(AB)$	-34.2 kJ/mol	-31.5 kJ/mol	-31.2 kJ/mol

As biphenyl molecular dimensions still allow for an MP2 direct optimization on the whole dimer, this was performed for comparison purposes and for the validation of our method. The direct MP2 optimization was performed starting from the final POLDO geometry, and each optimization step was CP corrected for BSSE. The final energy found with the direct MP2 route is −34.2 kJ/mol: 2.7 kJ/mol more attractive with respect to the minimum computed by our algorithm. This small difference can be ascribed to the different level of theory used for ΔE_{intra} and to small rearrangements of the internal geometry which are allowed in direct MP2 but not in POLDO. It is worth stressing that this energy difference cannot be attributed to FRM inaccuracies, as the direct MP2 energy of the final geometry obtained by POLDO is −31.9 vs −31.5 kJ/mol. A further test was performed by optimizing the whole biphenyl dimer using a DFT-D technique, namely, the B97D functional with the TZV2P basis set as first proposed by Grimme.[52] Although, up to our knowledge, the B97D functional was never employed for the biphenyl dimer, the performances of this dispersion corrected approach were recently tested[44] on the benzene dimer, giving results comparable with the MP2/6-31G*(0.25) method, as reported in Figure 1. In fact, if a B97D calculation is performed on the final geometry optimized by POLDO, an intermolecular energy of −29.7 kJ/mol is found. Conversely, the final optimized energy found with this method is −31.2 kJ/mol, not far from the direct MP2 value, and very close to the one obtained with the POLDO procedure.

In Table 1, geometrical values obtained with the direct MP2 and B97D optimizations are reported for comparison, while in Figure 3, POLDO, direct MP2, and B97D final geometries are sketched in the second, third, and fourth columns, respectively. Notwithstanding the good agreement

Table 2. Fragment−Fragment Energy Contributions (kJ/mol) in the POLDO Optimization of the Biphenyl Dimer[a]

Fragments $i_A \cdots j_B$		ΔE_{inter}	ΔE^{HF}_{inter}
$P_{1A} \cdots P_{1B}$	(PD)	−9.70	26.72
$P_{2A} \cdots P_{2B}$		−10.67	18.06
$P_{1A} \cdots P_{2B}$	(CD)	−4.59	−0.07
$P_{2A} \cdots P_{1B}$		−5.71	2.02
$H_{2A} \cdots H_{2B}$		0.05	0.09
$P_{1A} \cdots H_{2B}$		0.01	0.36
$P_{2A} \cdots H_{2B}$		0.41	0.88
$H_{2A} \cdots P_{1B}$		0.42	0.94
$H_{2A} \cdots P_{2B}$		0.03	0.35
Biphenyl - Biphenyl		-31.51	49.35

[a] The second and third column report intermolecular MP2 and HF energies, respectively.

among the three energy values, the final optimized geometries show some minor differences, suggesting that the PES around the minimum is rather flat. All algorithms predict a nonplanar value for ϕ_A and ϕ_B and an intermolecular distance between 3.7 and 3.9 Å, seeing that the biphenyl *para* axes are slightly more aligned in the dimer optimized with the direct MP2 algorithm. Indeed, an estimate of how much the A and B long axes are collinear can be gained by measuring the dihedral angle defined by the quadruplet of atoms $H_4^A-C_1^A-C_1^B-H_4^B$. The latter goes from an initial value of 38° to a final one of 5°, 8°, and 10° in the direct MP2, POLDO, and B97D optimized geometries, respectively.

Within the FRM, the intermolecular energies of all fragment pairs are directly available and can be used to get a deeper insight into the forces driving the optimization. Referring to the fragment labels defined in Figure 2, it can be seen that in the final geometry the benzenes of A and B are in parallel (PD) and cross-displaced (CD) arrangements. In Table 2, energy contributions for all fragment pairs are reported.

From these values it results that, notwithstanding that the major contribution comes from the PD arrangements, the moieties placed in CD geometries do not give a negligible contribution to the intermolecular energy.

4. 5CB

As a second benchmark calculation, the 5CB dimer was chosen. In this case, both intra- and intermolecular energies were computed with QM methods. The intramolecular energy was evaluated at the DFT level (B3LYP/cc-pVDZ). The intermolecular calculations were performed according to the fragmentation pattern described in the following and reported in Figure 4, and using the MP2/6-31G*(0.25) level of theory for the fragment−fragment interaction energies. Each 5CB monomer is fragmented by a cut along the inter-ring bond and the aromatic−aliphatic linkage. The valence of all fragments is saturated with H atoms, obtaining for each monomer three moieties (cyano-benzene, benzene, and *n*-pentane) and two H_2 intruder molecules.[63]

Energy minima were searched starting from different configurations. According to the mutual orientation of the two molecular long axes (which can be considered approximatively as lying along the cyano−biphenyl *para* axis), the

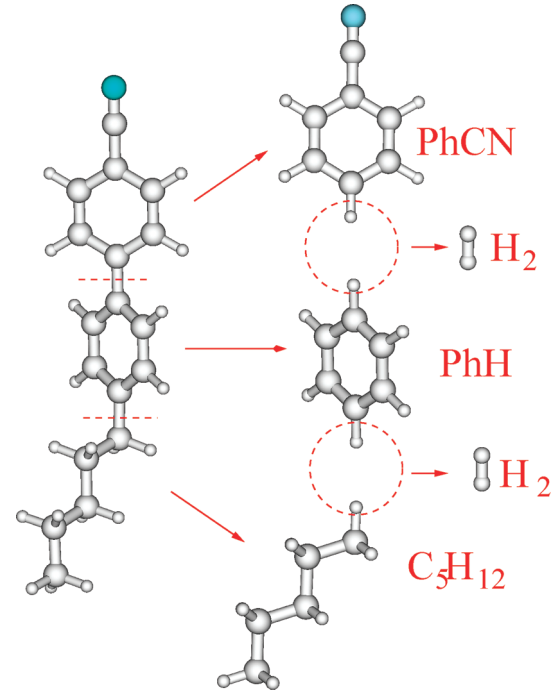


Figure 4. 5CB fragmentation scheme. All moieties will be labeled with a subscript A or B (i.e., $PhH_{A(B)}$) depending on which of the two 5CB monomers they belong to.

Table 3. 5CB, Summary of Optimized Energies for the Considered Arrangements[a]

geometry	ΔE(AB)	ΔE_{intra}(A)	ΔE_{intra}(B)	ΔE_{inter}(AB)
PFF1	−44.40	0.77	1.30	−46.47
PFF2	−47.23	0.79	2.85	−50.87
PSF	−31.05	0.37	0.04	−31.46
PD	−39.86	0.00	0.00	−39.86
An	−51.16	0.33	0.02	−51.51

[a] All energies are in kJ/mol.

investigated geometries can be referred to as parallel (P: long axes pointing in the same direction) or antiparallel (An: they point in opposite directions) arrangements. Among all possible P dimers, face-to-face (PFF) and side-to-face (PSF) and displaced (PD) geometries were investigated. In the PFF case, the aromatic rings of the two monomers are in a sandwich spatial disposition, whereas in the PSF they are found in a T-shaped geometry. Parallel displaced geometries were instead obtained from PFF by displacing one monomer along its long axis. Final energy contributions and selected geometrical quantities for all investigated arrangements are summarized in Tables 3 and 4, respectively.

PFF1 and PFF2 arrangements were obtained as shown in Figure 5 and are described in the following. Two 5CB monomers, in their isolated optimized conformation, were superimposed, and monomer B was then shifted by 5.5 Å along the vector normal to the central ring plane. Thereafter an, $\alpha_B = 180°$ rotation was performed and the obtained dimer geometry labeled PFF1 (geometry a in Figure 5). During optimization, PFF1 turns into a sort of PD arrangement: the aromatic rings come nearer, but a displacement takes place along the long molecular axis and the *X* direction (see Figure 5). In the final (geometry b in Figure 5) configuration, 5CB long molecular axes are not parallel anymore, and the

Table 4. Characterization of the Resulting 5CB Dimer Minima by Selected Geometrical Quantities[a]

	PFF1	PFF2	PSF	PD	An
$\phi_A - \phi_B$	29.0 - 30.0	30.0 - 27.0	31.8 - 37.5	34.0 - 33.0	32.0 - 34.0
$C_1^A\widehat{C_{1'}^A C_{1'}^B} C_1^B$	-20.7	-31.4	-24.0	13.7	-14.7
N-N	4.2	4.8	6.4	5.95	13.6
C_4-C_4	3.8	4.1	5.7	5.7	3.7 (C_4-$C_{4'}$)
C_1-C_1	3.7	3.7	5.1	5.4	3.7 (C_1-$C_{1'}$)
$C_{1'}$-$C_{1'}$	3.7	3.8	4.9	5.2	3.7 ($C_{1'}$-C_1)
$C_{4'}$-$C_{4'}$	3.9	4.3	4.7	5.0	3.9 ($C_{4'}$-C_4)

[a] Angle are reported in degrees and distances in Å.

dihedral angle defined by the quadruplet $C_1^A-C_{1'}^A-C_{1'}^B-C_1^B$ changes from 0° to 20°. A possible cause of this rotation may be the CN dipole−dipole repulsion. As far as dihedral angles are concerned, it is interesting to note that inter-ring (ϕ) values are different from the ones computed in the gas phase (∼42°) but similar to those expected in the condensed phase[68] (∼30°). Conversely, no conformational changes are induced in side chains by dimer interactions. The total energy of the dimer, $E_{tot}(AB)$, decreases from −17.5 to −44.4 kJ/mol, seeing that A's and B's intramolecular energies are 0.8 and 1.3 kJ/mol, respectively, with $\Delta E_{inter}(AB) = -46.5$ kJ/mol. As reported in Table 5, FRM again correctly reconstructs the interaction energy since an MP2 direct calculation, performed on the final POLDO optimized whole dimer geometry, gives $\Delta E_{inter}(AB) = -47.5$ kJ/mol.

The most relevant fragment−fragment contributions to ΔE_{inter} are reported in Table 5. As expected, the main contributions to the total intermolecular energy come from the $\pi\cdots\pi$ pair interactions. The cyano−benzene pair is found in a PD-like configuration whose large stability is little affected by the repulsive dipole−dipole interaction, due to the aforementioned rotation of the CN groups. The second contribution comes from the benzene pair, which takes advantage of the SW to PD displacement.

In the PFF1 optimized geometry, chain−chain interactions do not contribute significantly, as they point in opposite directions. If one of them is rotated by 180° around the 5CB's *para* axis, PFF2 geometry is generated as displayed in Figure 5. After optimization, the chain−chain energy amounts to −5.35 kJ/mol (see Table 5), and since all other fragment pair interactions are little affected by this chain rotation, PFF2 total interaction energy decreases from −46.47 kJ/mol (PFF1-opt) to −50.87 (PFF2-opt).

The next optimization was performed selecting as the starting geometry a PSF arrangement (see Figure 6, left panel) which is characterized by a T-shaped initial disposition of the aromatic rings. Owing to the stability of the T-shaped benzene dimer, an interaction energy comparable with the previous ones might be expected. However, as reported in Table 5, this is not the case. The main reason resides in a reduced stability of the PhCN dimer in the T-shaped configuration with respect to a PD one. On the contrary, the T-shaped benzene pair accounts for the maximum contribution (−8.6 kJ/mol) to the total intermolecular energy. This global lower interaction is consistent with the ϕ_A and ϕ_B values being more close to those of the isolated molecule.

Table 5. 5CB, More Relevant Optimized Fragment−Fragment Contributions to the Dimer $\Delta E_{inter}(AB)$ in Different Geometries[a]

fragment $i_A\cdots j_B$	PFF1	PFF2	PSF	PD	An
PhCN···PhCN	−14.50	−12.97	−7.37	−9.86	−3.54
PhCN···PhH	−7.77	−8.62	−5.95	−0.53	−14.09
PhCN···C_5H_{12}	−0.58	−0.61	−0.44	−0.05	−5.25
PhH···PhCN	−4.86	−5.23	−1.75	−13.75	−14.81
PhH···PhH	−10.36	−9.74	−8.65	−6.56	−6.09
PhH···C_5H_{12}	−2.62	−2.44	−3.39	−0.38	−0.45
C_5H_{12}···PhCN	−0.20	−0.18	−0.13	−2.34	−4.72
C_5H_{12}···PhH	−1.63	−3.27	−1.27	−4.58	−0.42
C_5H_{12}···C_5H_{12}	−1.34	−5.35	−2.62	−1.15	−0.02
$\Delta E_{inter}(AB)$ total	−46.47	−50.87	−31.46	−39.86	−51.51
direct MP2	−47.50	−51.09	−31.63	−40.17	−52.57

[a] In the last row, the MP2 energy computed on the whole POLDO optimized geometry is reported. All energies are in kJ/mol.

A last test on another parallel geometry was devised by displacing the monomer B of PFF1 dimer by ∼4 Å along *Z*, obtaining a PD arrangement. During optimization, a β rotation of B takes place, together with minor translational shifts that bring the C_4 atom of B in front of the $C_{1'}$ of the A monomer. A better comprehension of spatial displacement of A and B can be gained looking at the intermolecular energies of fragment−fragment pairs, as reported in Table 5. Benzene−cyanobenzene pairs contribute −13.75 and −0.53 kJ/mol, since the displacement along *Z* has led the former pair in a stacked conformation and the latter to large distance. The optimized PD dimer is also stabilized by $\pi\cdots\pi$ homodimer interactions (−9.86 and −6.56 kJ/mol) as well as one pentane−cyanobenzene pair (−4.58 kJ/mol). The final interaction energy, $\Delta E_{inter}(AB)$, sums up to −39.86 kJ/mol, showing that the PD optimized geometry is a local minimum more stable than the PSF dimer, but less favorable than the PFF ones.

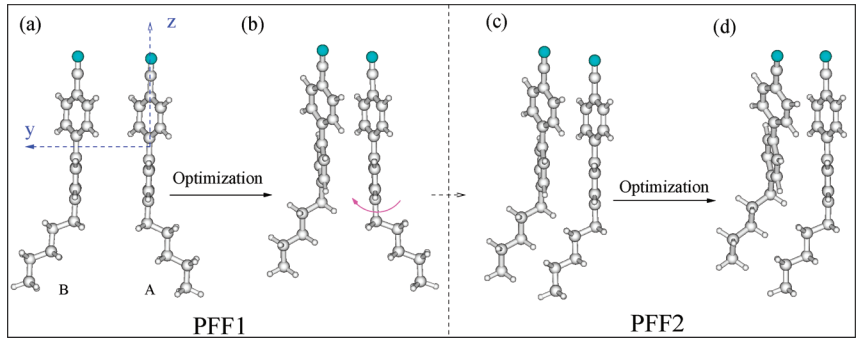


Figure 5. 5CB PFF1 (starting a and final b) and PFF2 (starting c and final d) geometries. Part c was obtained from part b through rotation of the first dihedral of the side chain.

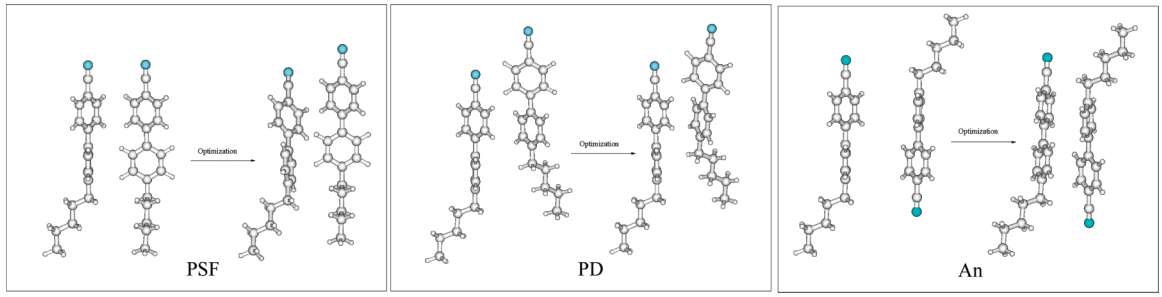


Figure 6. Parallel side-face (PSF, left panel), parallel displaced (PD, central panel), and antiparallel (An, right panel) starting and final geometries.

Finally, an antiparallel (An) 5CB dimer geometry was generated. This configuration is reported in Figure 6, and it was obtained rotating the B monomer of 180° around β and displacing it at 6 Å farther from A, along *Y*. In this way, the C_4 atom of one molecule faces the $C_{4'}$ of the other, and the C_1 atom of A(B) overlooks the $C_{1'}$ atom of B(A). During the optimization, the *Y* distance decreases to 3.7 Å, and B is subjected to a small β rotation, causing an imperfect stacking of the aromatic rings in the final configuration. From Table 5, it appears that, due to the antiparallel arrangement and difference from previous cases, the most important contributions are the cyanobenzene−benzene interactions, which each account for ∼−14 kJ/mol. The cyanobenzene and benzene homodimer contributions are much less, whereas the aliphatic−aromatic interactions now play an important role. Once again, side chain dihedrals are not affected by the presence of the second molecule, whereas the inter-ring dihedral is found around 33°. The final interaction energy ΔE_{inter}(AB) is −50.1 kJ/mol, suggesting that, at least for the phase space explored here, dimers of the An type are the most stable conformations. This conclusion seems to be consistent with the experimental observation of 5CB's nematic phases with zero permanent dipole (see ref 69 and references therein). This is an example of how information on dimers can help in a deeper understanding of condensed phase properties: in this case, the origin of the lack of permanent dipole in 5CB condensed ordered phases can, in fact, be tracked down to the greater stability of the antiparallel configuration.

5. Conclusions

In this paper, a new method for accurate geometry optimizations has been presented. The approach is explicitely devised for van der Waals complexes, even of large dimensions, which can be studied at affordable computational costs. This method, implemented in an original code named POLDO, can be used either to search for the most stable configuration or to investigate the relative stability of different local minima, in case one wants to explore some relevant portions of the dimer phase space. The optimization algorithm is guided by the energy gradients and an approximate Hessian, in the framework of the so-called quasi-Newton methods. The pivot of the whole procedure is the fragmentation reconstruction method (FRM), which involves a decomposition of the two molecules of the dimer into smaller fragments, the calculation of the interaction energies of all resulting fragment pairs and the final reconstruction of the whole intermolecular energy, thus avoiding the energy calculation on the entire system. This shortcut heavily cuts down computational costs, allowing calculations on molecules made up of a large number of atoms, that would be otherwise unaffordable.

The method was first tested on the biphenyl dimer, whose dimensions allowed us a comparison with the energy and geometries obtained by an optimization performed directly on the whole dimer. The intermolecular energies computed via FRM and on the whole dimer agree quantitatively (∼0.4 kJ/mol), and at the same time, the computational cost is lowered by almost 4 times. A good accord is also found between POLDO and direct MP2 optimization results, where the difference of 2.7 kJ/mol was ascribed to the different dimensionality of the optimized PES. As a second benchmark, a common nematogenic molecule, whose dimer is formed by 268 electrons, was chosen. The optimization was performed both on intermolecular coordinates and on a reduced set of internal degrees of freedom. As in the previous case, the comparison between FRM and direct (single point) calculations is favorable, seeing that the average error is ∼0.5 kJ/mol. Most important, the ratio of the CPU costs is 5; that is, it becomes more favorable with the increase of the molecular dimensions. It should also be stressed that 5CB dimensions make an MP2 direct optimization almost unfeasible on standard workstations.

In conclusion, the proposed optimization procedure, implemented in the POLDO code, grants some new advantages with respect to full and/or other fragmentation based minimization methods that can be summarized as follows:

(i) Inter- and intramolecular interactions can be treated at different levels of theory. Since the latter type is often less sensitive to the method/basis set employed, this may allow us to save computer time with a negligible loss of accuracy. Moreover, a FF description can also be employed, if accurate potentials are available.

(ii) The intermolecular energy is computed by means of the FRM approach, which does not rely on any specific method/basis set. Although in this and previous papers the FRM approach has been coupled with the MP2/6-31G*(0.25) level of theory, there are no obstacles to joining FRM with other promising methods, accurate while computationally affordable, such as those recently proposed in the literature and detailed in previous sections.

(iii) The fragmentation strategy on which the whole procedure is based allows for a deeper insight into experimental nonaccessible data, as it readily yields detailed information on local fragment−fragment interactions which concur with the formation of the whole vdW dimer.

In the near future, we plan to employ this method in two different contexts. The first one involves the parametrization of intermolecular force fields, suitable for classical computer simulations, from a database built using quantum mechanical calculations. In this context, the fragment−fragment interaction energies and their gradients, provided by FRM calculations, can be added to the database, yielding further information that can be exploited to reduce the redundancy of the FF parameters in the fitting procedure, therefore increasing the robustness of the set of parameters. The second project concerns the geometry optimization of large systems with particular attention to molecules of biological interest. Indeed, this approach allows for taking into account, during the optimization process, selected internal degrees of freedom while retaining a fixed geometry for more rigid parts of the monomer under study. This possibility seems, for instance, to fit the case of nucleobase quadruplets, where the flexibility around H bonds of two base pairs can be taken into account together with the intermolecular separation of the stacked pairs, whereas the planarity of the aromatic rings may be kept frozen during optimization.

Supporting Information Available: Additional text regarding preliminary tests. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Margenau, H.; Kestner, N. R. *Theory of Intermolecular Forces*; Pergamon Press: Braunschweig, Germany, 1969.

(2) Stone, A. J. *The Theory of Intermolecular Forces*; Oxford University Press: Oxford, U. K., 1996.

(3) Kaplan, I. G. The Interatomic Potential Concept and Classification of Interactions. In *Handbook of Molecular Physics and Quantum Chemistry*; Wilson, S., Bernath, P. F., McWeeny, R., Eds.; Wiley: Chichester, West Sussex, England, 2003.

(4) Israelachvili, J. *Intermolecular and Surface Forces*, 3rd ed.; Academic Press: London, 2010.

(5) Parsegian, V. A. *Van der Waals Forces: A Handbook for Biologists, Chemists, Engineers, and Physicists*; Cambridge University Press: Cambridge, U. K., 2005.

(6) Kleman, M.; Laverntovich, O. *Soft Matter Physics An Introduction*; Springer: Berlin, 2003.

(7) Hamley, I. *Introduction to Soft Matter: Synthetic and Biological Self-Assembling Materials*, Revised ed.; Wiley: Chichester, West Sussex, England, 2007.

(8) Collings, P. J.; Hird, M. *Introduction to Liquid Crystals*; Adam Hilger: Bristol, U. K., 1997.

(9) Demus, D.; Goodby, J.; Gray, G. W.; Spiess, H. W.; Vill, V. *Handbook of Liquid Crystals*; Wiley-VCH: Weinheim, Germany, 1998; Vol. 1: Fundamentals.

(10) McGaughey, G. B.; Gagné, M.; Rappé, A. *J. Biol. Chem.* **1998**, *273*, 15458.

(11) Hobza, P.; Sponer, J. *J. Am. Chem. Soc.* **2002**, *124*, 11802.

(12) Sponer, J.; Riley, K.; Hobza, P. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2595.

(13) Meyer, E. A.; Castellano, R. K.; Diederich, F. *Angew. Chem., Int. Ed.* **2003**, *42*, 1210.

(14) Bushby, R.; Lozman, O. *Curr. Opin. Colloid Interface Sci.* **2002**, *7*, 343.

(15) Laschat., S.; Baro, A.; Steinke, N.; Giesselmann, F.; Hagele, C.; Scalia, G.; Judele, R.; Kapatsina, E.; Sauer, S.; Schreivogel, A.; Tosoni, M. *Angew. Chem., Int. Ed.* **2007**, *46*, 4832.

(16) van de Craats, A. M.; Warman, J. M.; Müllen, K.; Geerts, Y.; Brand, J. D. *Adv. Mater.* **1998**, *10*, 36.

(17) Ehrenfreund, P.; Rasmussen, S.; Cleavesand, J.; Chen, L. *Astrobiology* **2006**, *6*, 490.

(18) Cacelli, I.; Prampolini, G. *J. Phys. Chem. A* **2003**, *107*, 8665.

(19) Cacelli, I.; Cinacchi, G.; Prampolini, G.; Tani, A. *J. Am. Chem. Soc.* **2004**, *126*, 14278.

(20) Valdes, H.; Pluháčková, K.; Pitoňák, M.; Řezáč, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2747.

(21) Sherrill, C. D.; Sumpter, B.; Sinnokrot, M. O.; Marshall, M.; Hohenstein, E.; Walker, R.; Gould, I. *J. Comput. Chem.* **2009**, *30*, 2187.

(22) Amovilli, C.; Cacelli, I.; Campanile, S.; Prampolini, G. *J. Chem. Phys.* **2002**, *117*, 3003.

(23) Amovilli, A.; Cacelli, I.; Cinacchi, G.; De Gaetani, L.; Prampolini, G.; Tani, A. *Theor. Chim. Acc.* **2007**, *117*, 885.

(24) Cacelli, I.; Cimoli, A.; DeGaetani, L.; Prampolini, G.; Tani, A. *J. Chem. Theory Comput.* **2009**, *5*, 1865.

(25) Zhang, D. W.; Zhang, J. Z. H. *J. Chem. Phys.* **2003**, *119*, 3599.

(26) Xiang, Y.; Zhang, D. W.; Zhang, J. Z. H. *J. Comput. Chem.* **2004**, *25*, 1431.

(27) Li, S.; Fang, T. *J. Am. Chem. Soc.* **2005**, *127*, 7215.

(28) Ganesh, V.; Dongare, R.; Balarayan, P.; Gadre, S. *J. Chem. Phys.* **2006**, *125*, 104109.

(29) Fedorov, D.; Kitaura, K. *The Fragment Molecular Orbital Method: Pratical Applications to Large Molecular Systems*; CRC press: Boca Raton, FL, 2009.

(30) Fedorov, D.; Ishida, T.; Kitaura, K. *J. Phys. Chem. A* **2007**, *111*, 2722.

(31) Dahlke, E.; Truhlar, D. *J. Chem. Theory Comput.* **2007**, *3*, 46.

(32) Řezáč, J.; Salahub, D. R. *J. Chem. Theory Comput.* **2010**, *6*, 91.

(33) Chiba, M.; Fedorov, D.; Nagata, T.; Kitaura, K. *Chem. Phys. Lett.* **2009**, *474*, 227.

(34) Li, H.; Fedorov, D.; Nagata, T.; Kitaura, K.; Jensen, J.; Gordon, M. *J. Comput. Chem.* **2010**, *31*, 778.

(35) Tsuzuki, S.; Uchimaru, T.; Matsamura, K.; Mikami, M.; Tanabe, K. *Chem. Phys. Lett.* **2000**, *319*, 547.

(36) Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. *J. Am. Chem. Soc.* **2002**, *124*, 104.

(37) Hobza, P.; Zahradník, R.; Müller-Dethlefs, K. *Collect. Czech. Chem. Commun.* **2006**, *71*, 443.

(38) Riley, K.; Pitoňák, M.; Černý, J.; Hobza, P. *J. Chem. Theory Comput.* **2010**, *6*, 66.

(39) Sinnokrot, M. O.; Valeev, E. F.; Sherrill, C. D. *J. Am. Chem. Soc*. **2002**, *124*, 10887.

(40) Sinnokrot, M. O.; Sherrill, C. D. *J. Phys. Chem. A* **2004**, *108*, 10200.

(41) Sinnokrot, M. O.; Sherrill, C. D. *J. Phys. Chem. A* **2006**, *110*, 10656.

(42) Arnstein, S.; Sherrill, C. D. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2646.

(43) Sherrill, C. D.; Takatani, T.; Hohenstein, E. *J. Phys. Chem. A* **2009**, *113*, 10146.

(44) Vasquez-Mayagoitia, A.; Sherrill, C. D.; Aprà, T.; Sumpter, B. G. *J. Chem. Theory Comput.* **2010**, *6*, 727.

(45) Meijer, E. J.; Sprik, M. *J. Chem. Phys.* **1996**, *105*, 8684.

(46) Tsuzuki, S.; Lüthi, H. P. *J. Chem. Phys.* **2001**, *114*, 3949.

(47) Zhao, Y.; Truhlar, D. *J. Chem. Theory Comput.* **2007**, *3*, 289.

(48) Zhao, Y.; Truhlar, D. *J. Chem. Phys.* **2006**, *125*, 194101.

(49) Zhao, Y.; Truhlar, D. *Theor. Chim. Accounts* **2008**, *120*, 215.

(50) Honenstein, E.; Chill, S.; Sherrill, C. *J. Chem. Theory Comput.* **2008**, *4*, 1996.

(51) Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463.

(52) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787.

(53) Hasselmann, A.; Jansen, G. *J. Chem. Phys.* **2005**, *122*, 014103.

(54) Podeszwa, R.; Bukowski, R.; Szalewicz, K. *J. Phys. Chem. A* **2006**, *110*, 10345.

(55) Jaffe, R. L.; Smith, G. D. *J. Chem. Phys.* **1996**, *105*, 2780.

(56) Hobza, P.; Selzle, H. L.; Schlag, E. W. *J. Phys. Chem.* **1996**, *100*, 18790.

(57) Pitoňák, M.; Neogrády, P.; Černý, J.; Grimme, S.; Hobza, P. *ChemPhysChem* **2009**, *10*, 282.

(58) Grimme, S. *J. Chem. Phys.* **2003**, *118*, 9095.

(59) Pitoňák, M.; Hesselmann, A. *J. Chem. Theory Comput.* **2010**, *6*, 168.

(60) Špóner, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem.* **1996**, *100*, 5590.

(61) Riley, K.; Hobza, P. *J. Phys. Chem. A* **2007**, *111*, 8257.

(62) Rutledge, L.; Durst, H.; Wetmore, S. *J. Chem. Theory Comput.* **2009**, *5*, 1400.

(63) Bizzarri, M.; Cacelli, I.; Prampolini, G.; Tani, A. *J. Phys. Chem. A* **2004**, *108*, 10336.

(64) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.

(65) Fuhrmann, J.; Rurainski, A.; Lenhof, H.-P.; Neumann, D. *J. Comput. Chem.* **2009**, *30*, 1371.

(66) 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.

(67) Cacelli, I.; Lami, C.; Prampolini, G. *J. Comput. Chem.* **2009**, *30*, 366.

(68) Cacelli, I.; Prampolini, G.; Tani, A. *J. Phys. Chem. B* **2005**, *109*, 3531.

(69) Cacelli, I.; De Gaetani, L.; Prampolini, G.; Tani, A. *J. Phys. Chem. B* **2007**, *111*, 2130.

CT100172W

JCTC Journal of Chemical Theory and Computation

Ensemble Docking from Homology Models

Eva Maria Novoa,[†] Lluis Ribas de Pouplana,[‡,§] Xavier Barril,[§,||] and Modesto Orozco*[,†,⊥]

Joint IRB-BSC Research Program in Computational Biology, Institute for Research in Biomedicine, Josep Samitier 1−5, Barcelona 08028, Spain, Cell and Developmental Biology, Institute for Research in Biomedicine, Josep Samitier 1−5, Barcelona 08028, Institució Catalana per la Recerca i Estudis Avançats, Passeig Lluis Companys 23, Barcelona 08010, Spain, Departament de Fisicoquímica, Facultat de Farmàcia, Avgda Diagonal sn, Barcelona 08028, Spain, and Structural Bioinformatics Node Instituto Nacional de Bioinformática, Institute of Research in Biomedicine, Josep Samitier 1−5, Barcelona 08028, Spain

Received May 10, 2010

Abstract: We present here a systematic exploration of the quality of protein structures derived from homology modeling when used as templates for high-throughput docking. It is found that structures derived from homology modeling are often similar in quality for docking purposes than real crystal structures, even in cases where the template used to create the structural model shows only a moderate sequence identity with the protein of interest. We designed an "ensemble docking" approach based on the use of multiple homology models. The method provides results which are usually of better quality than those expected from single experimental X-ray structures. The use of this approach allows us to increase around five times the universe of use of high-throughput docking approaches for human proteins, by covering over 75% of known human therapeutic targets.

Introduction

New algorithms and computers are making possible the use of atomistic docking approaches in a high-throughput (HTD) regime, being possible to screen in silico libraries containing 10^5-10^6 compounds against a limited number of protein targets.[1−3] However, we cannot ignore that the requirement of computational efficiency implies the introduction of severe simplifications in both the description of molecular interactions and the coverage of the conformational space of ligands and proteins.[4−6] As a result, docking methods have problems in representing ligand-induced conformational changes in the protein, and in general the quality of docking algorithms decreases as the docked drug differs from that bound in the crystal structure.[7,8] However, despite all these limitations, the power of current docking algorithms is beyond all doubt, and many authors have demonstrated that their use largely enriches the possibility to find a good binder from a large library of decoys and that the proposed optimal poses are good starting points for lead-optimization processes.[9−12] It is not surprising, then, that virtual screening based on docking algorithms is a routine task in medicinal chemistry laboratories.[11,13,14]

The inputs of docking algorithms are ligand and protein structures, and the outputs are a series of "poses", i.e., possible configurations of the protein−ligand complex, which are then scored using an empirically refined function yielding to a small subset of preferred binding modes with the associated binding affinity.[15−17] Given the number of approximations done in a docking algorithm, the practical purpose of HTD is not the accurate ranking of potential

* Corresponding author phone: 0034-93-4037155; e-mail: modesto@mmb.pcb.ub.es.

[†] Joint IRB-BSC Research Program in Computational Biology, Institute for Research in Biomedicine.

[‡] Cell and Developmental Biology, Institute for Research in Biomedicine.

[§] Institució Catalana per la Recerca i Estudis Avançats.

[||] Facultat de Farmàcia.

[⊥] Structural Bioinformatics Node Instituto Nacional de Bioinformática, Institute of Research in Biomedicine.

10.1021/ct100246y © 2010 American Chemical Society
Published on Web 07/16/2010

binders, but the enrichment of true binders among the top-ranked compounds and the recovery of good leads for refinement.

The need to have a three-dimensional structure of the target protein strongly limits the use of docking algorithms, and despite the impressive advance of structural genomics, the number of proteins for which experimental structure is known represents only a small fraction of the total proteome. Thus, the 2010 version of the Protein Data Bank (PDB) contains around 60000 entries, but only 42.5% (25560) of them correspond to unique proteins from which only 15% (3935) are human.[18,19] In comparison, sequence analysis suggests that the total number of human proteins ranges between 20332—Swissprot[20]—and 93110—RefSeq[21]—probably twice or more if spliced forms are considered,[22] which means that the PDB covers only between 2 and 19% of the human proteins. The gap between structure and sequence becomes even larger if we consider proteins from virus, bacteria, or other pathogens for which little structural information exists.

Protein structure can be predicted by a variety of computational methods,[23] homology modeling (also named comparative modeling) being the most accurate one in cases where there is a clear sequence identity between the target protein and at least one template with known three-dimensional structure.[24,25] The quality of the structure derived from homology modeling roughly correlates with the sequence identity between the target protein and template proteins.[26] Thus, it is accepted that for sequence identities below 30% less than half of the residues have their C_α correctly placed.[27,28] The percentage of correctly placed residues increases to 85% for identities ranging from 30 to 50%, and most of the C_αs are well-positioned for sequence identities above 50%. Inside the high-quality range no direct correlation exists between the accuracy of the model and the sequence identity with the template, and evaluation of the expected quality of a model is still an unsolved problem.[29] In fact, the concept of "goodness" is not unique, since it depends on its planned use.[30] For example a model with an accuracy around 3.5 Å in backbone positioning may be good enough for understanding protein function or designing mutations but is expected to be of small utility for prediction of ligand binding.[26−31]

Different authors have tried to evaluate the quality of homology models for docking experiments. Thus, McGovern and Shoichet performed high-throughput docking on 10 target enzymes for which apo, holo, and homology model structures were available, finding that they were useful for enriching the screening, but not as powerful as the holo-crystal structure.[32] Diller and Li reported good enrichments (in some cases similar to those obtained with the crystal structure) when model structures of six kinases obtained for identities in the range of 30−50% were used to screen a large library.[33] Similar results were obtained by Oshiro et al.[34] in the study of two targets (CDK2 and factor VIIa), by Gilson's group with a set of five targets,[35] and by Ferrara and Jacoby in the analysis of insulin growth factor I receptor.[36] In a very recent paper Fan et al.[37] found good results when ensembles of homology models of several proteins were used to screen for ligands in the DUD database[38] using the DOCK computer program.[39] All these studies illustrate the power of homology models to guide docking experiments but also underline their limitations related to the lack of "a priori" evaluations on the quality of the model for docking purposes and on the problems of selecting a priori a structural model from the battery of solutions given by homology modeling routines (for discussion see ref 36).

The introduction of protein flexibility is the next step in docking, and there is a significant amount of work focused in this direction.[40,41] Among the different approaches suggested, "ensemble docking" (also known as multiple docking) is one of the most popular ones. It assumes that the effect of target flexibility in docking can be represented by using a Boltzmann ensemble of conformations for the protein instead of just a single rigid structure. Different methods for generating ensembles have been proposed, including molecular dynamics[42,43] (from a known experimental structure of the target), crystallographic (X-ray),[44−47] and spectroscopic (NMR).[48,49] All these approaches require experimental knowledge of protein structure and are then able to cover just a small fraction of proteome. In this contribution, following the pioneering work by Fan et al.,[37] we explore the possibility of using ensembles derived from homology (comparative) modeling. This approach is simple and fast and, if successful, would allow us to dramatically expand the range of applicability of ensemble docking approaches. We explored, with a wide range of metrics and for a large number of proteins, not only the ability of the approach to enrich in active ligands drug libraries but also the structural quality of the docking predictions, a crucial element in lead optimization procedures. We designed and tested a procedure to perform ensemble docking based on the combination of Modeller[50] and Glide,[51] finding that the results are in general of better quality than those expected when a single-crystal structure is used as a template in docking experiments.

Methods

Protein Data Sets. We defined two sets of proteins of our study: one for training and another for testing. The training set was defined considering proteins for which at least 30 crystal ligand-bound structures were available in PDB (with the same sequence or at most one single mutation). PDBs with point mutations were only used to build the set of active ligands but were not included in the set of docked proteins. This set of proteins includes thrombin (2cn0, 1ay6, 1bmm, 1tom, 1xm1), renin (2g24, 1bil, 1hrn, 1rne, 2g1r), cyclin-dependent kinase 2—CDK2—(1aq1, 1e1v, 1gz8, 1jsv, 3ddq), and protein tyrosine phosphatase 1B—PTP-1B—(2f71, 1c83, 1g7g, 1ony, 2h4g). The test set was created using less restrictive conditions in terms of the number of crystallized structures available—at least eight—and contained α-momorcharin (1mrg, 1aha, 1mom, 1f8q, 1ahb), trypsin (1tng, 1tnl, 1f0t, 1lqe, 2by5), p38 kinase (3hp2, 1w7h, 2baj, 3c5u, 3cg2), HIV retrotranscriptase (3jyt, 1dtq, 1rt1, 1s6p, 3dol), factor Xa (2vvc, 1ezq, 1fax, 1lpk, 1nfu), and heat shock protein 90—HSP90—(1yet, 1osf, 1uy6, 1yc4, 2ccs).

Homology Modeling. The derivation of model structures was performed using scripts designed for HTD production

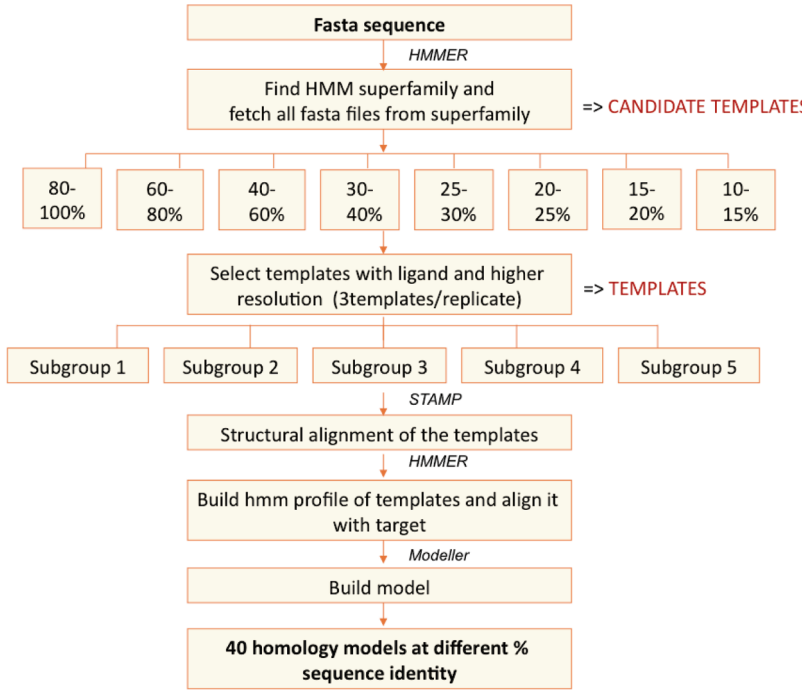


Figure 1. Comparative model building work flow. The process is automated such that a FASTA sequence is given as input, and a total of 40 homology models that range from 10 to 100% sequence identity are obtained. The software used at each step is detailed in Methods.

trying to mimic the standard expert procedure for homology modeling (see Figure 1). We are aware that by using automatic protocols homology modeling might be prone to errors, related mostly to misalignment problems, which can be easily corrected by manual refinement. However, to evaluate a pure HTD scenario, no human refinement was done here, which means that results presented here can be considered a lower limit of accuracy for the technique. Accordingly, the sequence of each target protein was extracted from the PDB, transformed to FASTA-format, and launched against the Pfam-A database[52] using HMMER[53] to assign the sequence to a superfamily. All the FASTA files for which there is a PDB corresponding to this same superfamily—all the candidate templates—were retrieved and aligned to the target sequence using ClustalW.[54] After this procedure each template was placed into different categories depending on its percentage sequence identity: 80−100, 60−80, 40−60, 30−40, 25−30, 20−25, 15−20, and 10−15%. For each sequence-identity category we selected 15 templates considering only proteins bound to ligand and solved at the highest resolution possible. Each set of 15 templates was divided into five subsets in order to build five different models per sequence-identity category. It is important to remark that the five models per sequence-identity category were built on the basis of different templates. Such templates were structurally aligned by STAMP,[55] creating then a profile using HMMER, which was introduced as a meta-template for alignment of the target sequence (see Figure 1). Finally, the 9v5 version of MODELLER[50] was used to create structural models using default options.

Ligand Selection. The active ligands to dock were downloaded from the PDB database (www.rcsb.org), by selecting all available X-ray ligands from PDB complexes for each of the proteins of the study. All the available ligands were subjected to similarity analysis using MOE[56] implementation of MACCS structural fingerprints[57] and distributed in 80% identity clusters. Only one compound per cluster was selected, which guarantees the diversity of the ligands, avoiding bias derived from the overrepresentation of the same scaffold. The set of known ligands was mixed with 1000 diverse "decoys" (molecules not described as binders for these proteins) which were selected from the most populated clusters obtained using Reynolds' algorithm at a similarity cutoff level of 60%[58] on a local database containing 1.7 million commercially available compounds—already filtered by drug-likeness criteria: Lipinski rules, Veber rules, and lack of reactive groups.[59−61] The percentage of active ligands ranged from 0.5 to 10%, depending on the protein.

Docking Procedure. Ligand screening and docking was performed using the Glide 5.0 program.[51] The extraprecision Glide docking (Glide XP) protocol was used for the training set, while the standard-precision (Glide SP) protocol was used for the test set, trying then to mimic a normal HTD procedure (in practice, we found very small differences between both scoring functions). Starting from the PDB structures, ligands were prepared using the LigPrep[62] facility in Schrödinger utility MAESTRO,[63] by generating low-energy ionization and tautomeric states within the range of pH 7.0 ± 2.0. All ligands were energy-minimized using the OPLS_2005 force field implemented in MAESTRO.[63] The setup of proteins was done with the Protein Preparation Wizard facility, which included hydrogen optimization, protonation, and geometry optimization using again the OPLS_2005 force field. The receptor grid defining the docking universe was built centered on the crystallographic ligand, which was then removed as any other nonprotein molecule.

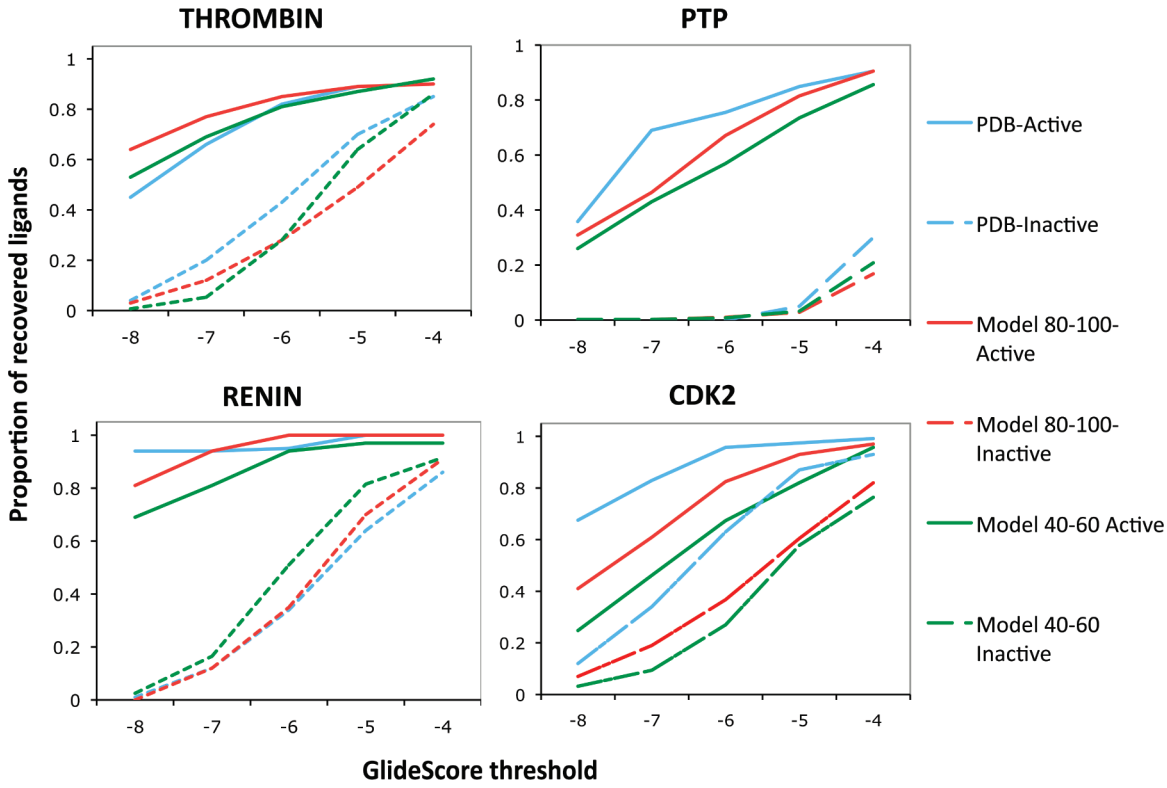


Figure 2. Recovery of active (nondashed lines) and inactive ligands (dashed lines) for each of the proteins of the training set. As can be seen in the plots, both active and inactive ligand recovery increases as the GS threshold decreases. The different colors correspond to different sequence identity ranges: blue (PDB), red (model 80−100), and green (model 40−60). The selected PDBs are all high-resolution holo conformations: 2cn0, 1aq1, 2g24, and 2f71, which correspond to thrombin, cdk2, renin, and PTP-1B, respectively.

Metrics for Preevaluation of Model Quality. The structural quality of the model was evaluated using both global and local parameters. The global quality indexes included global root-mean-square deviation (rmsd; model-reference PDB), global sequence identity, number of gaps in the alignment, and sequence coverage of the model. The local parameters were always referred to the binding site (defined as the set of residues with at least one atom at less than 5 Å from the crystal ligand) and included binding site rmsd, binding site sequence identity, and atom conservation in binding site structure. All rmsd measures were computed using the MMTSB tool set.[64]

Metrics for Evaluation of Success in Docking. The success of docking was measured by analyzing the following: (1) the ability of the models to predict the structure of the ligand−protein complex and (2) the applicability of the models for virtual screening purposes. The ability of models to predict the structure of the complex was assessed by (i) measuring the proportion of docked poses with rmsd below 2 Å from crystal structure using an SVL script in MOE, (ii) measuring the rmsd obtained when comparing the best-docked pose (rmsd-based selection) and the best-ranked pose (GlideScore-based selection) with the crystallographic ligand, and (iii) measuring the similarity between ligand−protein contact maps in models and crystal structures, which are determined by comparing the number of atoms that are conserved from those found at less than 5 Å from the docked ligand—compared to the original PDB where the docked ligand is found. Thus, for each docked ligand, a different ligand−protein map is built and compared to its corresponding PDB.

The utility of the models for virtual screening purposes was evaluated by assessing the performance of the homology models to discriminate between active compounds and decoys (inactive). A virtual screening run selects a list of molecules (n) from a given database of N entries, which includes both actives (true positive compounds, TP) and decoys (false positive compounds, FP). Actives (A) that have not been found by the screening method are false negatives (FN), and decoys that have not been selected are true negatives (TN). The optimum screening is that able to recover all true positives, without recovering any false positive.

Many different enrichment descriptors described in the literature have been considered in this work.[65,66] First we computed the *sensitivity* (true positive rate; TPR; see eq 1) and the *specificity* (true negative rate; TNR see eq 2) indexes. The first indicates the ability of the method to recover the real ligands, while the second informs on its ability to avoid decoys.

$$\text{sensitivity} = \text{TPR} = \frac{\text{TP}}{(\text{TP} + \text{FN})} \quad (1)$$

$$\text{specificity} = \text{TNR} = \frac{\text{TN}}{(\text{FP} + \text{TN})} = 1 - \text{FPR} \quad (2)$$

where FPR stands for false positive rate.

The *accuracy* (Acc; eq 3) index was used to describe the percentage of molecules which have been correctly classified by the screening protocol, while *precision* (positive predictive value; PPV) was used to describe the proportion of true positives among the list of selected compounds given by the docking (eq 4).

$$\mathrm{Acc} = \frac{\mathrm{TP} + \mathrm{TN}}{N} = \frac{A}{N}\mathrm{TPR} + \left(1 - \frac{A}{N}\right)\mathrm{TNR} \qquad (3)$$

$$\mathrm{PPV} = \frac{\mathrm{TP}}{(\mathrm{TP} + \mathrm{FP})} \qquad (4)$$

To assess the ability of the models to obtain true actives among the first ranked compounds (an extra requirement in HTD studies[67]), the *enrichment factor* (EF, eq 5) was used.

$$\mathrm{EF} = \frac{\mathrm{TP}/n}{A/N} \qquad (5)$$

ROC (receiver operating characteristic; true positive versus false positive rates) curves and the associated AUC curves (area under the ROC curve) have also been used to determine the discriminatory power of the virtual screening procedure. These metrics are especially powerful since they are not dependent on the ratio of actives to decoys of the database.[68−70]

Results and Discussion

Structural Quality of the Models. Modeller[50] provides good global models when using structural templates with sequence identities above 25% (Supporting Information Figures S1 and S2). The use of templates with sequence identities below such a threshold can yield wrong structures due mainly to alignment errors or to the presence of large unfolded regions. The atom conservation—i.e., the similarity between ligand−protein contact maps—at binding sites grows faster than global sequence identity, and for identities as small as 25−30% around 60−70% of the atoms at the experimental binding site are conserved in the model (Supporting Information Figure S3). The heavy-atoms rmsd between model and real binding sites are typically below 2 Å for sequence identities above 25% (Supporting Information Figure S4). Clearly, then, structural models created using homology modeling not only reproduce well global protein structure but also provide quite important details of the binding site. Whether or not the quality of these details is enough for drug docking studies will be the main subject of discussion in the remaining of our communication.

Docking Enrichment Using Single-Structure Homology Models. The second point to analyze was the quality of single homology models when used to recover specifically active ligands from a mixture of ligands and decoys. Within the Glide framework the number of hits recovered in a docking depends on the scoring (GS) threshold. For very restrictive GS values very few decoys (false positives) are recovered, but many real ligands might be lost. On the contrary, when very permissive GS values are used, all real ligands are recovered, but at the expense of increasing dramatically the number of incorrectly selected decoys. Results shown in Figure 2 demonstrate that using a single PDB structure as

Table 1. Training Set Enrichment Descriptors[a]

	PDB			model 80−100		model 60−80		model 40−60		model 30−40		model 25−30		model 20−25		model 15−20		model 10−15	
	GS	SG	ENS	SG	ENS	SG	ENS	SG	ENS	SG	ENS	SG	ENS	SG	ENS	SG	ENS	SG	ENS
sensitivity	−8	0.56	0.81	0.49	0.77	0.43	0.74	0.38	0.67	0.32	0.60	0.26	0.55	0.22	0.37	0.14	0.26	0.11	0.20
	−7	0.76	0.92	0.66	0.86	0.62	0.88	0.57	0.83	0.47	0.75	0.45	0.79	0.39	0.66	0.27	0.50	0.20	0.38
	−6	0.87	0.97	0.82	0.94	0.78	0.92	0.74	0.91	0.63	0.88	0.62	0.91	0.54	0.80	0.41	0.70	0.30	0.51
specificity	−8	0.95	0.85	0.94	0.88	0.96	0.88	0.96	0.88	0.97	0.91	0.98	0.93	0.98	0.96	0.99	0.97	0.99	0.97
	−7	0.88	0.71	0.88	0.76	0.90	0.76	0.90	0.75	0.94	0.80	0.94	0.77	0.94	0.81	0.91	0.86	0.94	0.90
	−6	0.69	0.44	0.73	0.51	0.74	0.48	0.71	0.45	0.80	0.51	0.81	0.49	0.82	0.56	0.80	0.66	0.85	0.75
EF (1%)	–	24.40	26.49	21.72	26.09	22.3	24.58	19.01	21.49	17.95	23.08	15.19	19.66	17.67	21.29	11.35	13.62	6.84	9.74
accuracy	−8	0.93	0.85	0.91	0.87	0.92	0.87	0.92	0.87	0.93	0.90	0.93	0.91	0.93	0.92	0.93	0.92	0.93	0.92
	−7	0.88	0.72	0.87	0.77	0.88	0.77	0.88	0.76	0.91	0.80	0.91	0.78	0.90	0.80	0.87	0.83	0.88	0.86
	−6	0.70	0.48	0.74	0.53	0.74	0.51	0.71	0.48	0.80	0.54	0.80	0.52	0.80	0.57	0.77	0.66	0.81	0.73
PPV	−8	0.52	0.37	0.33	0.29	0.40	0.31	0.39	0.32	0.40	0.34	0.47	0.50	0.49	0.45	0.48	0.41	0.48	0.41
	−7	0.42	0.31	0.36	0.27	0.39	0.30	0.43	0.33	0.53	0.35	0.46	0.39	0.49	0.39	0.42	0.33	0.41	0.33
	−6	0.26	0.18	0.26	0.17	0.25	0.16	0.25	0.16	0.26	0.16	0.28	0.22	0.28	0.20	0.23	0.18	0.23	0.18
AUC	–	0.85	0.95	0.86	0.92	0.84	0.90	0.83	0.89	0.82	0.88	0.78	0.88	0.73	0.83	0.70	0.78	0.56	0.59

[a] Six different enrichment descriptors (sensitivity, specificity, EF for the top 1% ranked compounds, accuracy, PPV and AUC) have been computed for models and PDBs using both single docking (SG) and ensemble docking (ENS) approaches. For each of the cases, enrichment descriptors have been quantified taking different scoring (GS) thresholds: −8, −7, and −6, allowing us to see the difference between models and PDBs not only depending on the docking approach used—single or ensemble—but also depending on the chosen GS threshold. In the case of the EF (1%) and AUC enrichment descr iptors, no GS threshold has been used, given that these descriptors are GS-threshold independent.

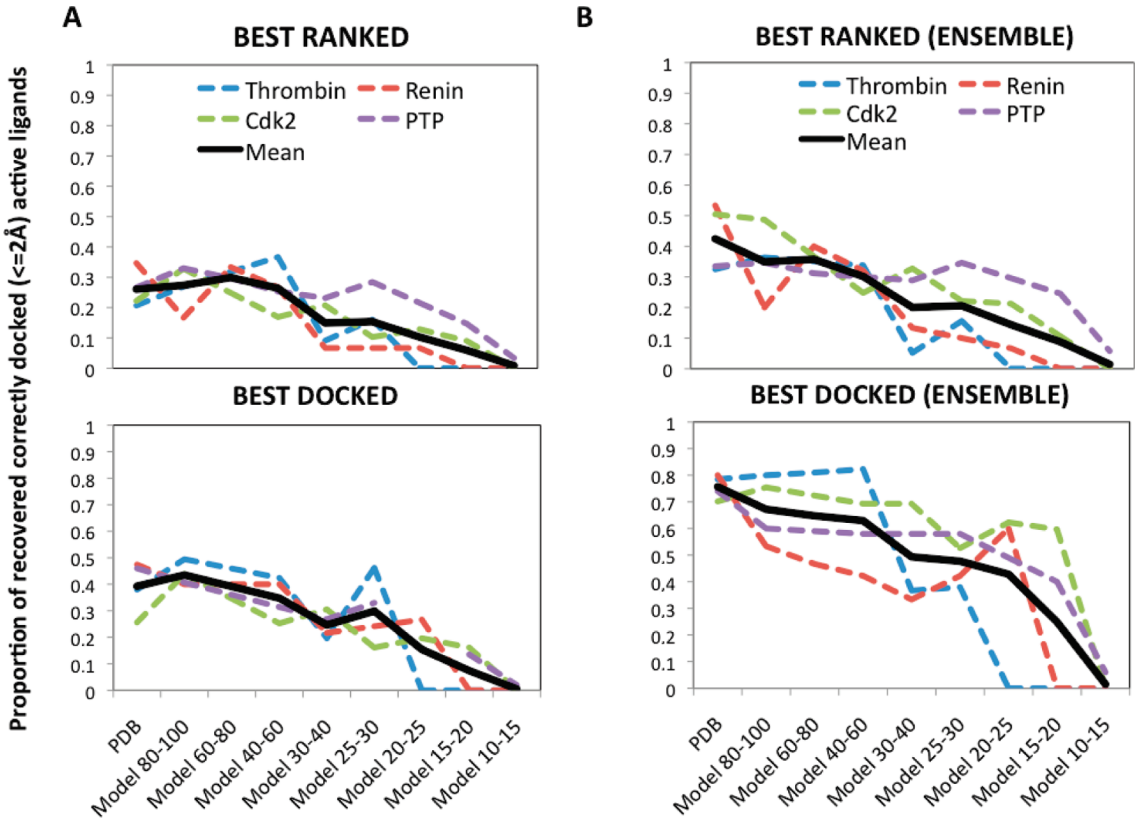


Figure 3. Recovery of correctly docked ligands versus sequence identity of the models. The recovery is defined as the fraction of correctly active docked ligands—less than 2 Å rmsd from the crystal structure—with respect to the total active docked and scored ligands. (A) In the upper plot, the *best-ranked* active ligand pose is chosen from all the proposed poses by using a score-based selection, whereas in the lower plot the *best-docked* ligand pose is chosen by using an rmsd-based selection. (B) Recovery of correctly docked ligands versus sequence identity when using an ensemble docking approach. Both score-based selection—i.e., best ranked—and rmsd-based selection—i.e., best docked—are shown. Each protein of the training set is labeled accordingly, and the mean value of the four training set proteins is shown in black.

template Glide is able to recover typically between 40 and 90% of the real ligands with a small number of false positives for a very strict scoring function threshold ($-GS = 8$). The ratio of true positive increases about 10 percentile points for $-GS = 7$ and 5−10 extra points for $-GS = 6$, keeping still an acceptable rate of recovery of false positives; for larger $-GS$ values the rate of false positives becomes unacceptable. In any case, the improvement with respect to random selection is very clear, demonstrating the performance of the Glide docking algorithm.

When homology models are used for docking, the performance of Glide is not lost (Figure 2 and Table 1), even in cases where the models are built using proteins with a modest level of homology as templates. It is especially encouraging that in some cases homology models outperform experimental structures for drug docking, a result already found by other authors[32,37] and which encourages the use of modeled protein structures for drug design experiments. The fact that homology models outperform X-ray structure for thrombin might appear surprising but is on the line of previous works with this protein[32] which demonstrated that probably the holo structure of thrombin is overspecialized for ligand binding, with problems arising in cross-docking experiments similar to those performed here. Homology models, less refined for a particular ligand binding mode, are then more successful.

Structural Quality of the Docking Poses Obtained Using Single-Structure Homology Models. The ability of the docking algorithm to capture specifically the maximum of active ligands is the major requirement for hit finding. However, to guide the optimization of the hit, there is an additional requirement: the drug needs to be correctly placed at the binding site. When using an experimental PDB structure as template, Glide is able to find poses that are very close (rmsd < 2 Å) to the bound conformation found in crystal in around 50% of cases, and in fact in more than 30% of cases the best scored poses (typically $-GS > 8$) match the experimental conformation (Figure 3A). Very interestingly, the global performance of the method does not change significantly when single homology models built from sequence identities above 40% are used, and even models built from templates with sequence identities around 25% can provide reasonable results. Again, it is remarkable that for some proteins homology models can provide more accurate binding mode predictions than the experimental structure—e.g., thrombin homology models recover on average 20% more correctly docked ligands compared to the crystallographic structures.

Ensemble Docking versus Single-Structure Docking. Proteins adapt their structure to the bound ligand, which explains the problems of docking methods to recognize active

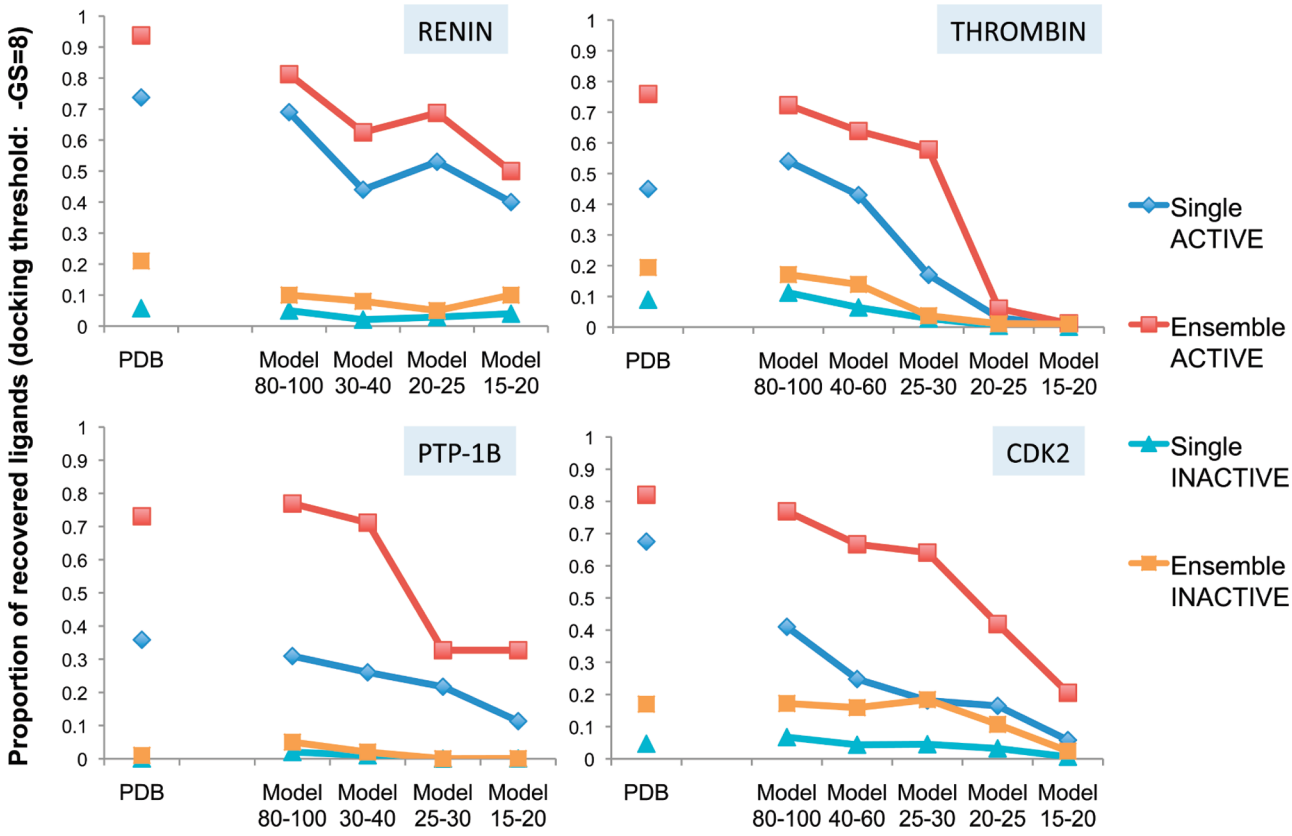


Figure 4. Ensemble docking versus single docking approach. The performance of both approaches is being compared in terms of recovered active ligands and decoys for the four proteins of the training set. The single docking approach performance is shown with blue and cyan lines, which correspond to the recovery of active and inactive ligands, respectively. Similarly, the red and orange lines correspond to the active and inactive ligand recovery, respectively, when using an ensemble docking approach. In all cases, the difference between active and inactive recovery is higher when using ensemble docking. Results shown correspond to $-GS = 8$.

ligands when the protein structure has been solved in the presence of a very different compound. This problem is graphically illustrated in Figures 3B and Supporting Information Figure S5, which show the dispersion of results that can be obtained for a given protein when different high-resolution X-ray structures are used for docking. We can alleviate this problem by docking the drug against all the protein structures, selecting then as optimal docking mode that with the best scoring. This strategy is known as multiple docking or ensemble docking, which has been used and described in previous papers.[40,71−76] An ensemble of receptor conformations provides a structural degree of freedom that cannot be achieved with other flexible-receptor docking methods, such as induced-fit docking (IFD).[77] In our ensemble docking procedure, we have used five different structures, which is in accordance with the number of receptor structures used in previous papers.[71,73] This ensemble docking procedure (using at this point only experimental structures) leads to a clear improvement with respect to the average situation found when docking was done for single structures if a restrictive GS threshold is used (see Table 1). In fact, for strict threshold values the ensemble docking approach yields in most cases better results than those obtained by using the best "dockable" experimental structure, while the performance can decay for permissive thresholds due to the retrieval of false positives. It is also worth noting that the ensemble docking approach improves also the chances to recover good structural models for lead optimization procedures (compare Figure 3A with Figure 3B, and see Supporting Information Figure S6).

Ensemble Docking from Homology Models. The preceding analysis suggests that in general better docking results are obtained if all the experimental structural information of a protein is used as input for an ensemble docking procedure. The question is now, whether or not this situation is maintained for the less accurate ensembles generated by comparative modeling. Results in Table 1 demonstrate that the use of ensembles increases very significantly sensitivity (70−100%) with respect to single models, decreasing only slightly the specificity (around 6% for $-GS \geq 8$), leading to an overall improvement in the docking results. Thus, improvement made by the use of ensemble docking is more important in cases where the initial structures have lower AUCs, such as those in homology models.

Homology-modeling based ensemble docking coupled with good structural models and strict scoring thresholds outperforms in most cases single-structure docking performed using experimental structures (Table 1 and Figure 4). In fact, the quality of the ensemble docking results for accurate homology models (sequence identity above 80%) is indistinguishable from those obtained using experimental ensembles, and on average more than 80% of active ligands are recovered with a small percentage of recovered decoys

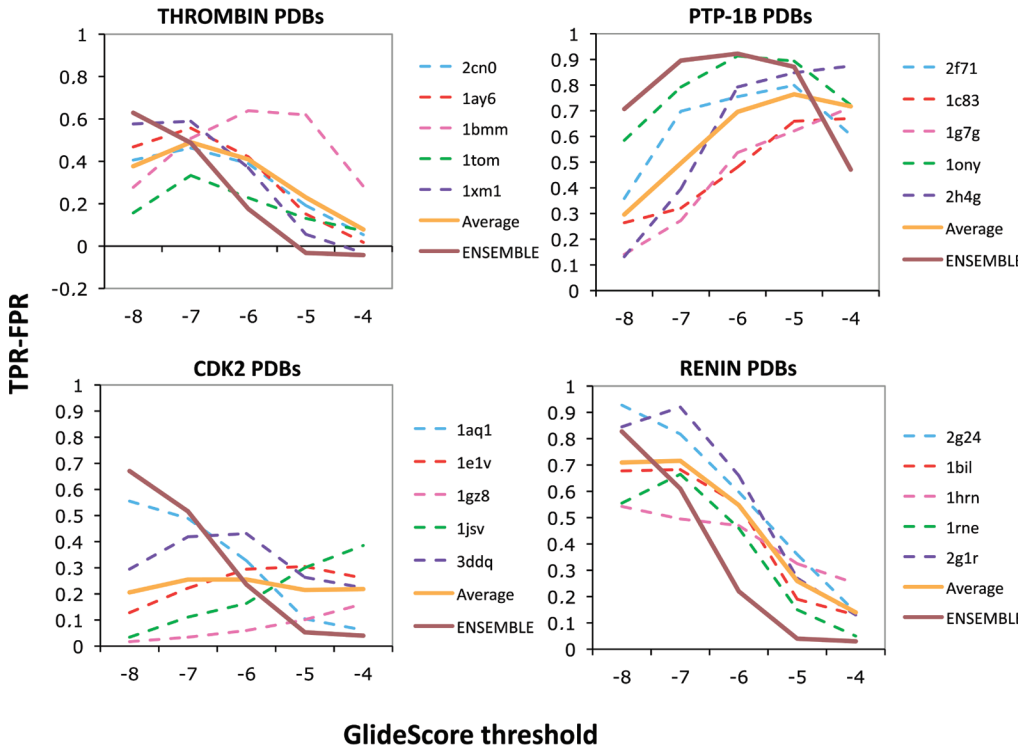


Figure 5. Performance of ensemble PDBs versus single PDBs. Performance (*y*-axis) is measured as the difference between the true positive rate (TPR) and the false positive rate (FPR). The dashed lines correspond to single PDBs, whereas the nondashed lines correspond to the average and ensemble of the PDBs, labeled with orange and brown, respectively. In all cases except for renin, the ensemble performs better than any single PDB at strict GS thresholds. However, in all proteins of the training set, the ensemble's performance decreases more rapidly—i.e., has higher slope—than any of the single PDBs.

(Figure 4) when homology ensembles are used. The ensemble docking protocol is very robust to the decrease in sequence identity, given that models with sequence identities in the range of 30−40% still provide good results. On the contrary, the protocol outlined here is very sensitive to the scoring threshold used, and less strict GS values increase excessively the recovery of false positives (Table 1 and Supporting Information Figures S7 and S8).

Finally, it is worth noting the large structural quality of the complexes obtained in homology-derived ensemble docking even when templates were not very homologous (Figure 5 and Supporting Information Figure S6). In fact, in most cases docking using ensembles of homology models outperform single experimental structure docking (Figure 4).

Validation of Results. Analysis on four proteins for which a large amount of structural data exist suggested (see above) that ensemble docking using homology models with sequence identity above 30−40% displayed a good ability to specifically recover active ligands when used as input for Glide calculations. Furthermore, the suggested complexes were in general reasonably close to the experimental binding modes, suggesting that the derived poses could be safely used in lead-optimization procedures. Analysis of the data suggests that the best balance between sensitivity and specificity is obtained when strict Glide scoring values were used to discriminate between active and decoy complexes. It is however unclear whether these results are general or specific for the proteins considered up to now. To analyze this point, we studied the ability of Glide on homology modeling ensembles of six unrelated proteins (see Methods).

Results summarized in Figure 6 demonstrate the good screening performance of the ensemble-docking approach performed with homology models also in the completely unrelated set of proteins used for validation. It is difficult to extend results of this small set of proteins to the entire proteome, but results suggest that docking performed using ensembles of homology models created using templates with sequence identity in the range of 30−40% leads to results which are of similar quality (according to most metrics) than those obtained using a single experimental structure. The screening performance of docking using ensembles of high-quality homology models is in general superior to that of docking using a single experimental structure and similar to docking procedures using an ensemble of experimental structures. Finally, Figure 7 confirms the geometrical quality of the complexes resulting from the homology model based docking procedure and accordingly its potential use in lead optimization processes. Our results indicate that the use of ensembles of homology models—built with Modeller—as input for Glide—using strict scoring thresholds—improves both the retrieval of active ligands from a chemical library and also the recovery of good structural complexes for lead optimization processes.

Gain in the Coverage of the Dockable Proteome. Results above suggest that an identity range of 30−40% is enough to build ensembles of homology models which can significantly enrich chemical libraries in active ligands. These results allow us to expand the applicability of structure-based drug design to a large universe of targets. Thus, while only 19% of (20332—Swissprot-annotated) human proteins can

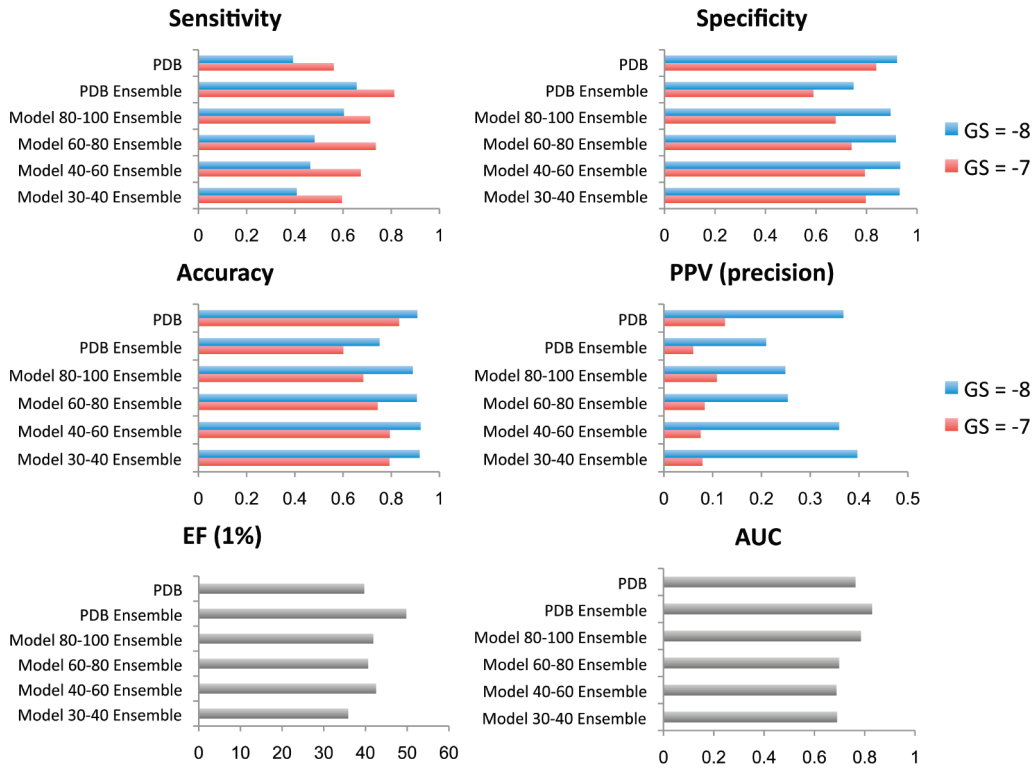


Figure 6. Enrichment descriptors for the test set. Only ensemble results for sequence identities >30% are shown for simplification. In the four top plots (sensitivity, specificity, accuracy, and PPV), enrichment descriptors are computed for −GS = 8 (blue) and −GS = 7 (red).

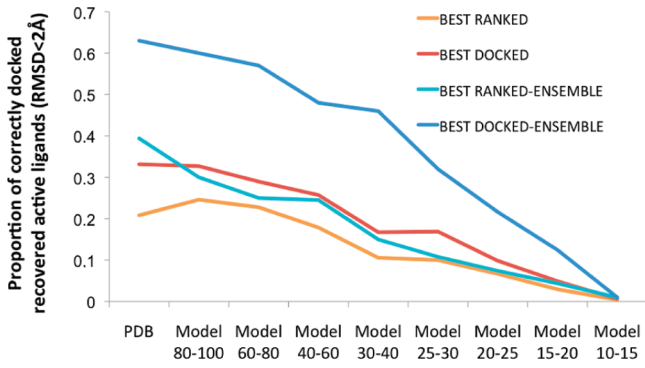


Figure 7. Recovery of correctly docked active ligands of the test set. A ligand is considered as correctly docked when its rmsd with the crystallographic ligand is below 2 Å. Both score-based selection—i.e., best ranked—and rmsd-based selection—i.e, best docked—are shown. Single docking averages are shown in red and orange, whereas ensemble docking averages are shown in blue and cyan.

be subjected to docking experiments using experimental structures, around 55% of (Swissprot) known human proteins can be studied by ensemble docking using homology models built from templates with 40% identity (Supporting Information Figure S9). Furthermore, less than 50% of human proteins of pharmacological interest have crystal structure available (DrugBank[78]). This coverage increases 41%—i.e., covering over 75% of the human drug targets—when using homology models up to 30% identity (see Supporting Information Figure S10).

With all the required cautions needed in the use of homology models for docking purposes (related mostly to the problems in finding good templates and in determining "a priori" the quality of the model), we suggest that the use of comparative models can enlarge dramatically the universe of applicability of small-molecule docking approaches, opening the possibility to analyze all potential cross-interactions of drug candidates, warning on potential adverse effects, opening new horizons both in the development of "dirty" drugs and in the determination of new indications for already annotated drugs.

ABBREVIATION. A, actives; Acc, accuracy; AUC, area under ROC curve; CDK2, cyclin-dependent kinase 2; EF, enrichment factor; ENS, ensemble; FN, false negatives; FP, false positives; FPR, false positive rate; GS, glide score; HIV, human immunodeficiency virus; HSP90, heat shock protein 90; HTD, high throughput docking; IFD, induced-fit docking; MACCS, molecular access system; PPV, positive predictive value; PTP-1B, protein tyrosine phosphatase 1B; rmsd, root-mean-square deviation; ROC, receiver operating characteristic; seq id, sequence identity; SVL, scientific vector language; TN, true negatives; TNR, true negative rate; TP, true positives; TPR, true positive rate.

Acknowledgment. This work has been supported by the Spanish Ministry of Science (Grants BIO2009-10964 and SAF2009-08811), the Instituto Nacional de Bioinformática, the Consolider E-science project, the ISCIII- COMBIOMED project, and the Fundación Marcelino Botin.

Supporting Information Available: Figures S1−S10 showing global rmsd between the homology models and the reference pdb, the correlation between the percentages of sequence identity and their sequence coverage, correlation

between the binding site sequence conservation and the percentage of sequence identity of the model, rmsd of the binding site, ROC curve plots for thrombin pdbs, similarity between ligand and protein contact maps, ensemble versus single docking approach, coverage of the human proteome, and structural coverage of human targets of pharmaceutical interest. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Schneider, G.; Bohm, H. J. *Drug Discovery Today* **2002**, *7*, 64–70.

(2) Alvarez, J. C. *Curr. Opin. Chem. Biol.* **2004**, *8*, 365–370.

(3) Lyne, P. D. *Drug Discovery Today* **2002**, *7*, 1047–1055.

(4) Mohan, V.; Gibbs, A. C.; Cummings, M. D.; Jaeger, E. P.; DesJarlais, R. L. *Curr. Pharm. Des.* **2005**, *11*, 323–333.

(5) Cozzini, P.; Kellogg, G. E.; Spyrakis, F.; Abraham, D. J.; Costantino, G.; Emerson, A.; Fanelli, F.; Gohlke, H.; Kuhn, L. A.; Morris, G. M.; Orozco, M.; Pertinhez, T. A.; Rizzi, M.; Sotriffer, C. A. *J. Med. Chem.* **2008**, *51*, 6237–6255.

(6) Jacobson, M. P.; Sali, A. *Annual Reports in Medicinal Chemistry*; Academic Press: London, 2004; pp 259−276.

(7) Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. *Proteins* **2006**, *65*, 15–26.

(8) Warren, G. L.; Andrews, C. W.; Capelli, A. M.; Clarke, B.; LaLonde, J.; Lambert, M. H.; Lindvall, M.; Nevins, N.; Semus, S. F.; Senger, S.; Tedesco, G.; Wall, I. D.; Woolven, J. M.; Peishoff, C. E.; Head, M. S. *J. Med. Chem.* **2006**, *49*, 5912–5931.

(9) Abagyan, R.; Totrov, M. *Curr. Opin. Chem. Biol.* **2001**, *5*, 375–382.

(10) Cavasotto, C. N.; Orry, A. J. *Curr. Top. Med. Chem.* **2007**, *7*, 1006–1014.

(11) Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J. *Nat. Rev. Drug Discovery* **2004**, *3*, 935–949.

(12) Shoichet, B. K. *Nature* **2004**, *432*, 862–865.

(13) Leach, A. R.; Shoichet, B. K.; Peishoff, C. E. *J. Med. Chem.* **2006**, *49*, 5851–5815.

(14) Jorgensen, W. L. *Science* **2004**, *303*, 1813–1818.

(15) Brooijmans, N.; Kuntz, I. D. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 335–373.

(16) Halperin, I.; Ma, B.; Wolfson, H.; Nussinov, R. *Proteins* **2002**, *47*, 409–443.

(17) Taylor, R. D.; Jewsbury, P. J.; Essex, J. W. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 151–166.

(18) 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.

(19) O'Donovan, C.; Apweiler, R.; Bairoch, A. *Trends Biotechnol.* **2001**, *19*, 178–181.

(20) Bairoch, A.; Apweiler, R. *Nucleic Acids Res.* **2000**, *28*, 45–48.

(21) Pruitt, K. D.; Tatusova, T.; Maglott, D. R. *Nucleic Acids Res.* **2005**, *33*, D501–D504.

(22) Clark, F.; Thanaraj, T. A. *Hum. Mol. Genet.* **2002**, *11*, 451–464.

(23) Zhang, Y. *Curr. Opin. Struct. Biol.* **2008**, *18*, 342–348.

(24) Marti-Renom, M. A.; Stuart, A. C.; Fiser, A.; Sanchez, R.; Melo, F.; Sali, A. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 291–325.

(25) Koehl, P.; Levitt, M. *Nat. Struct. Biol.* **1999**, *6*, 108–111.

(26) Marti-Renom, M. A.; Madhusudhan, M. S.; Fiser, A.; Rost, B.; Sali, A. *Structure* **2002**, *10*, 435–440.

(27) Eswar, N.; Sali, A. Comparative Modeling of Drug Target Proteins. In *Computer-Assisted Drug Design, Comprehensive Medicinal Chemistry II*; Taylor, J., Triggle, D., Mason, J. S., Eds.;Elsevier: Oxford, U.K., 2007; Vol. 4, pp 215−236.

(28) Sanchez, R.; Pieper, U.; Melo, F.; Eswar, N.; Marti-Renom, M. A.; Madhusudhan, M. S.; Mirkovic, N.; Sali, A. *Nat. Struct. Biol.* **2000**, *7*, 986–990.

(29) Eramian, D.; Eswar, N.; Shen, M. Y.; Sali, A. *Protein Sci.* **2008**, *17*, 1881–1893.

(30) Cavasotto, C. N.; Phatak, S. S. *Drug Discovery Today* **2009**, *14*, 676–683.

(31) Baker, D.; Sali, A. *Science* **2001**, *294*, 93–96.

(32) McGovern, S. L.; Shoichet, B. K. *J. Med. Chem.* **2003**, *46*, 2895–2907.

(33) Diller, D. J.; Li, R. *J. Med. Chem.* **2003**, *46*, 4638–4347.

(34) Oshiro, C.; Bradley, E. K.; Eksterowicz, J.; Evensen, E.; Lamb, M. L.; Lanctot, J. K.; Putta, S.; Stanton, R.; Grootenhuis, P. D. *J. Med. Chem.* **2004**, *47*, 764–767.

(35) Kairys, V.; Fernandes, M. X.; Gilson, M. K. *J. Chem. Inf. Model.* **2006**, *46*, 365–379.

(36) Ferrara, P.; Jacoby, E. *J. Mol. Model.* **2007**, *13*, 897–905.

(37) Fan, H.; Irwin, J. J.; Webb, B. M.; Klebe, G.; Shoichet, B. K.; Sali, A. *J. Chem. Inf. Model.* **2009**, *49*, 2512–2527.

(38) Huang, N.; Shoichet, B. K.; Irwin, J. J. *J. Med. Chem.* **2006**, *49*, 6789–6801.

(39) Shoichet, B. K.; Bodian, D. L.; Kuntz, I. D. *J. Comput. Chem.* **1992**, *13*, 380–397.

(40) Totrov, M.; Abagyan, R. *Curr. Opin. Struct. Biol.* **2008**, *18*, 178–184.

(41) B-Rao, C.; Subramanian, J.; Sharma, S. D. *Drug Discovery Today* **2009**, *14*, 394–400.

(42) Paulsen, L. P.; Anderson, A. C. *J. Chem. Inf. Model.* **2009**, *49*, 2813–2819.

(43) Armen, R. S.; Chen, J.; Brooks, C. L. *J. Chem. Theory Comput.* **2009**, *5*, 2909–2923.

(44) Rao, S.; Sanschagrin, P. C.; Greenwood, J. R.; Repasky, M. P.; Sherman, W.; Farid, R. *J. Comput.-Aided Mol. Des.* **2008**, *22*, 621–627.

(45) Huang, S. Y.; Zou, X. *Proteins* **2007**, *66*, 399–421.

(46) Rueda, M.; Bottegoni, G.; Abagyan, R. *J. Chem. Inf. Model.* **2009**, *50*, 186–193.

(47) Craig, I. R.; Essex, J. W.; Spiegel, K. *J. Chem. Inf. Model.* **2010**, *50*, 511–524.

(48) Damm, K. L.; Carlson, H. A. *J. Am. Chem. Soc.* **2007**, *129*, 8225–8235.

(49) Huang, S. Y.; Zou, X. *Protein Sci.* **2007**, *16*, 43–51.

(50) Sali, A.; Blundell, T. L. *J. Mol. Biol.* **1993**, *234*, 779–815.

(51) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. *J. Med. Chem.* **2004**, *47*, 1739–1749.

(52) Finn, R. D.; Mistry, J.; Schuster-Bockler, B.; Griffiths-Jones, S.; Hollich, V.; Lassmann, T.; Moxon, S.; Marshall, M.; Khanna, A.; Durbin, R.; Eddy, S. R.; Sonnhammer, E. L.; Bateman, A. *Nucleic Acids Res.* **2006**, *34*, D247–D251.

(53) Eddy, S. R. *Bioinformatics* **1998**, *14*, 755–763.

(54) Thompson, J. D.; Higgins, D. G.; Gibson, T. J. *Nucleic Acids Res.* **1994**, *22*, 4673−4680.

(55) Russell, R. B.; Barton, G. J. *Proteins* **1992**, *14*, 309–323.

(56) *Molecular Operating Environment (MOE)*, Version 2007 09; Chemical Computing Group: Montreal, Quebec, Canada, 2007.

(57) *MACCS Structural Keys*; Symyx Software: San Ramon, CA, 2002.

(58) Reynolds, C. H.; Druker, R.; Pfahler, L. B. *J. Chem. Comput. Sci.* **1998**, *38*, 305–312.

(59) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug. Delivery Rev.* **2001**, *46*, 3–26.

(60) Veber, D. F.; Johnson, S. R.; Cheng, H. Y.; Smith, B. R.; Ward, K. W.; Kopple, K. D. *J. Med. Chem.* **2002**, *45*, 2615–2623.

(61) Hann, M.; Hudson, B.; Lewell, X.; Lifely, R.; Miller, L.; Ramsden, N. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 897–902.

(62) *LigPrep*, Version 2.2; Schrödinger: New York, NY, 2008.

(63) *Maestro*, Version 8.5; Schrödinger: New York, NY, 2008.

(64) Feig, M.; Karanicolas, J.; Brooks, C. L. 3rd. *J. Mol. Graphics Modell.* **2004**, *22*, 377–395.

(65) Langer, T.; Hoffmann, R. D., *Pharmacophores and Pharmacophore Searches*; Wiley-VCH: Weinheim, Germany, 2006.

(66) Kirchmair, J.; Markt, P.; Distinto, S.; Wolber, G.; Langer, T. *J. Comput.-Aided Mol. Des.* **2008**, *22*, 213–228.

(67) Truchon, J. L.; Bayly, C. I. *J. Chem. Inf. Model.* **2007**, *47*, 488–508.

(68) Nicholls, A. *J. Comput.-Aided Mol. Des.* **2008**, *22*, 239–255.

(69) Jain, A. N.; Nicholls, A. *J. Comput.-Aided Mol. Des.* **2008**, *22*, 133–139.

(70) Witten, I. H.; Frank, E. Credibility: Evaluating what's been learned. In *Data mining—Practical machine learning tools and techniques*, 2nd ed.; Morgan Kaufmann: San Francisco, CA, 2005; pp 161−176.

(71) Knegtel, R. M.; Kuntz, I. D.; Oshiro, C. M. *J. Mol. Biol.* **1997**, *266*, 424–440.

(72) Yoon, S.; Welsh, W. J. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 88–96.

(73) Cavasotto, C. N.; Abagyan, R. A. *J. Mol. Biol.* **2004**, *12*, 209–225.

(74) Duca, J. S.; Madison, V. S.; Voigt, J. H. *J. Chem. Inf. Model.* **2008**, *48*, 659–668.

(75) Sperandio, O.; Mouawad, L.; Pinto, E.; Villoutreix, B. O.; Perahia, D.; Miteva, M. A. *Eur. Biophys. J.*, in press.

(76) Barril, X.; Morley, S. D. *J. Med. Chem.* **2005**, *48*, 4432–4443.

(77) Sherman, W.; Day, T.; Jacobson, M. P.; Friesner, R. A.; Farid, R. *J. Med. Chem.* **2006**, *49*, 534–553.

(78) Wishart, D. S.; Knox, C.; Guo, A. C.; Cheng, D.; Shrivastava, S.; Tzur, D.; Gautam, B.; Hassanali, M. *Nucleic Acids Res.* **2008**, *36*, D901–D906.

CT100246Y