

Methyl Phosphate Dianion Hydrolysis in Solution Characterized by Path Collective Variables Coupled with DFT-Based Enhanced Sampling Simulations

Davide Branduardi,[†] Marco De Vivo,[†] Nadia Rega,[‡] Vincenzo Barone,[§] and Andrea Cavalli^{*,†,||}

[†]Department of Drug Discovery and Development, Italian Institute of Technology, Via Morego 30, I-16163 Genoa, Italy

[‡]Dipartimento di Chimica, Università "Federico II", Complesso Universitario di Monte S. Angelo, via Cintia, I-80126 Napoli, Italy

[§]Scuola Normale Superiore, piazza dei Cavalieri 7, I-56126 Pisa, Italy; INFN, Sezione di Pisa, Italy

^{||}Department of Pharmaceutical Sciences, University of Bologna, Via Belmeloro 6, I-40126 Bologna, Italy

S Supporting Information

ABSTRACT: Herein, we propose a conceptually innovative approach to investigating reaction mechanisms. This study demonstrates the importance of considering explicitly the effects of large amplitude motions, aside from the intrinsic reaction coordinate, when tuning the free energy landscape of reaction pathways. We couple the path collective variables method with DFT-based enhanced sampling simulations to characterize the associative mechanism of the hydrolysis of the methyl phosphate dianion in solution. Importantly, energetics and mechanistic differences are observed when passing from the potential to the free energy surface.

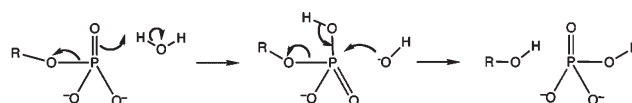
Traditionally, chemical reaction mechanisms are unraveled in terms of stationary points of potential energy surfaces (PESs) and their harmonic environment.¹ It is then straightforward to recover the free energy differences using standard statistical thermodynamics equations in the rigid rotor/harmonic oscillator approximation. Such calculations provide an order-parameter-free description and allow researchers to study concerted mechanisms, providing reliable reaction rates. However, even simple chemical reactions can experience anharmonic effects that influence the range of applicability of such approaches. While small anharmonic effects experienced by semirigid systems can be effectively considered using perturbative approaches,² proper consideration of large-amplitude motions for large systems is much more involved. Conversely, molecular dynamics (MD)-based free energy calculations can deal with arbitrarily large amplitude motions. However, such calculations are usually carried out using simple geometrical descriptors. Subsequently, complex free energy landscapes must be split into separated steps, preventing accurate investigation of multievent mechanisms. Here, we apply an MD-based approach that can identify the free energy landscape using a single collective descriptor. This approach continuously captures the complexity of concerted multievent reaction mechanisms. We use this novel approach to identify the free energy landscape of the associative mechanism of a phosphoryl transfer reaction, using methyl phosphate dianion hydrolysis as a prototypical case study. Our approach reveals energetics and mechanistic features of this important reaction with unprecedented detail.

Phosphoryl transfers are ubiquitous in biology since they represent the crucial chemical process of many metabolic pathways. The hydrolysis and subsequent transfer of phosphates is efficiently catalyzed by a huge number of enzymes involved in

energy production, replication of genetic material, biosynthesis, and protein control mechanisms.^{3,4} One possible phosphoryl transfer mechanism displays an associative transition state (TS). Here, the nucleophilic attack precedes the departure of the leaving group, which leads to a phosphorane-like TS geometry. This mechanism can be concerted (A_ND_N , see Scheme 1) when stable intermediates are not identified along the reaction pathway. In the past few decades, there has been significant computational and experimental effort to elucidate the different mechanisms of phosphoryl transfer in vacuo, in solution, and in enzymes.^{5–11} However, the energetics and mechanistic details are still under active investigation.^{5–11} As a paradigmatic case study, here we focus on the A_ND_N mechanism of the methyl phosphate dianion hydrolysis, shedding further light on the mechanistic details related to the reaction's free energy profile.

We have performed a detailed description of the reaction's free energy landscape, coupling for the first time the path collective variables (PCVs) method¹² with ADMP (atom-centered density matrix propagation)¹³ available in G09¹⁴ for DFT-based MD simulations, which is similar to the Car–Parrinello scheme¹⁵ in a localized basis set. The integration time step adopted was 0.2 fs. A

Scheme 1. Concerted Associative Mechanism (A_ND_N) for Phosphoryl Transfer



Received: September 23, 2010

Published: January 05, 2011

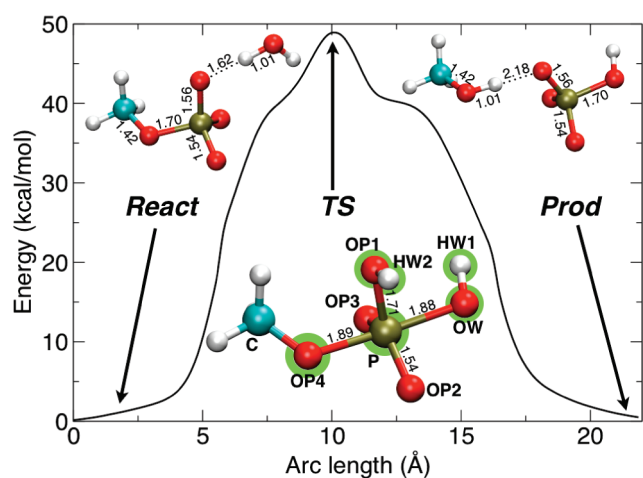


Figure 1. Reaction mechanism as extracted from TS optimization and IRC calculations. Bond distances are in Å. Green shadowed atoms are used for tracing the PCVs.

canonical ensemble was enforced by a stochastic rescaling thermostat¹⁶ with a period of 50 fs. Solvation effects were accounted for using a C-PCM model,¹⁷ one of the implicit solvation approaches which have been reported to reproduce experimental results well for these kinds of reactions.^{6,10,18} The PLUMED package¹⁹ was implemented in a special stand-alone version to add enhanced sampling (ES) capabilities to the G09 program. In this way, the ADMP module could perform the umbrella sampling (US) and steered-MD needed to depict the free energy landscape of the reaction in question. Importantly, the PCVs approach simultaneously considers proton transfer (PT) events, dipole moment reorientation, and conversion of the scissile phosphate along the reaction with a single collective variable.

The model system was formed by the methyl phosphate and the nucleophilic water molecule. The PCVs method aims to find low energy pathways that connect reactants and products. It requires an initial guess path in terms of Cartesian coordinates. Thus, we first identified a reactive path for the $A_N D_N$ TS that was obtained by B3LYP calculations in continuum solvent, followed by intrinsic reaction coordinate (IRC) calculations initiated from the TS imaginary frequency.²⁰ This provided the reactants-to-products pathway on the PES (Figure 1). The reaction mechanism was characterized by a very compact TS, in which the lengths of the forming (OW–P) and breaking (OP4–P) bonds were 1.88 Å and 1.89 Å, respectively. The water molecule was already fully deprotonated in the TS, while the HW2 proton was fully transferred to the oxygen OP1 of the scissile phosphate (OP1–HW2 = 0.97 Å and OW–HW2 = 2.79 Å). Overall, the phosphorane-like TS geometry resembled the typical S_N2 -like mechanism that characterizes the concerted $A_N D_N$ associative mechanism, with no intermediate along the reaction path (Scheme 1). The activation energy was about 49 kcal/mol.

The $A_N D_N$ pathway found on the PES was then used as a starting guess for the reaction mechanism for subsequent PCVs/ES calculations carried out on the free energy surface (FES). PCVs consist of two collective variables (S and Z) that describe the progression along (S), and the distance from (Z), the reaction pathway in terms of mean square deviations measured on a few key atoms involved in the reactivity.²¹ These atoms are the most

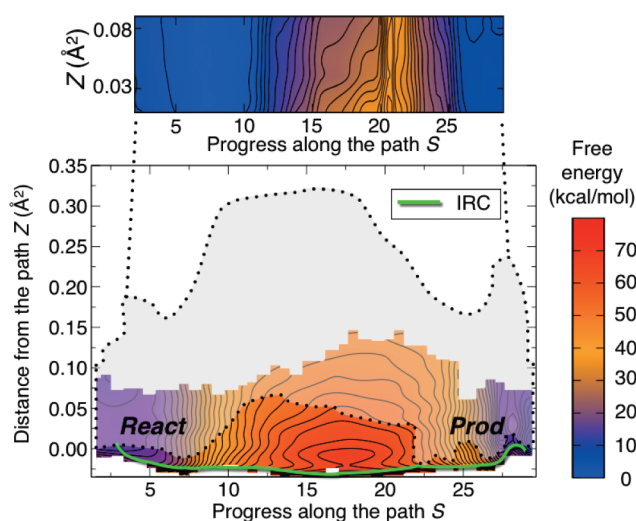


Figure 2. FES in S and Z for the first umbrella sampling run (bottom panel). The shaded gray region represents the sampled region with the second US, whose FES in the new S and Z is magnified in the upper panel. The projection of the IRC pathway on S and Z is displayed in light green. Negative values of distance Z are due to its definition,¹² as a result of a negative logarithm whose argument can be larger than one.

representative of the chemical reaction, while the remaining atoms are neglected. This choice was verified a posteriori by checking that the neglected atoms were correctly averaged out during the simulations (see the Supporting Information, SI). Then, steered-MD and US calculations, coupled with PCVs, were performed to identify the reaction pathways on the FES.

After a preliminary B3LYP/ADMP-based 1D-US run along S , we performed a more detailed bidimensional (2D) adaptive US to capture the difference between the starting IRC-based path and the simulated one (see SI). A total of 190 ps of sampling in both S and Z variables was carried out to determine a pathway of the reaction on the FES.²² The reconstruction of the 2D FES (Figure 2) was performed using the fitting procedure originally developed by Maragliano and Vanden-Eijnden.²³ In particular, we adopted the version recently modified by Monteferrante et al.²⁴ Figure 2 includes the projection of the IRC pathway in S and Z (light green line), which lies on a narrow channel at low values of Z . Only a couple of kilocalories per mole separated IRC from a wider region at lower energy, which indicated alternative pathways on the FES. To further refine the FES, we performed steered-MD along the S variable. A confining wall over Z allowed the system relative freedom to relax onto the new pathway, while preventing the system from escaping the reactive region.²⁵ A set of frames was then extracted from the steered-MD trajectory. This set represented a novel guess path for subsequent PCVs/ES calculations. Further 2D US simulations were then performed, as previously described (see ref 22 and the SI). The newly sampled region is shaded gray in Figure 2 and magnified in the upper panel. This indicates the location of the FES associated with the novel reaction path with respect to the original IRC. The resulting free energy projected over S and Z showed almost vertical isolines (upper panel of Figure 2), which indicated that the new reference path was lying in a wider, low, free energy pathway, and allowed for a meaningful 1D projection (Figure 3A). Here, we point out that currently our approach does not account for quantum effects, although they are estimated to be negligible for the reaction under study.¹⁰

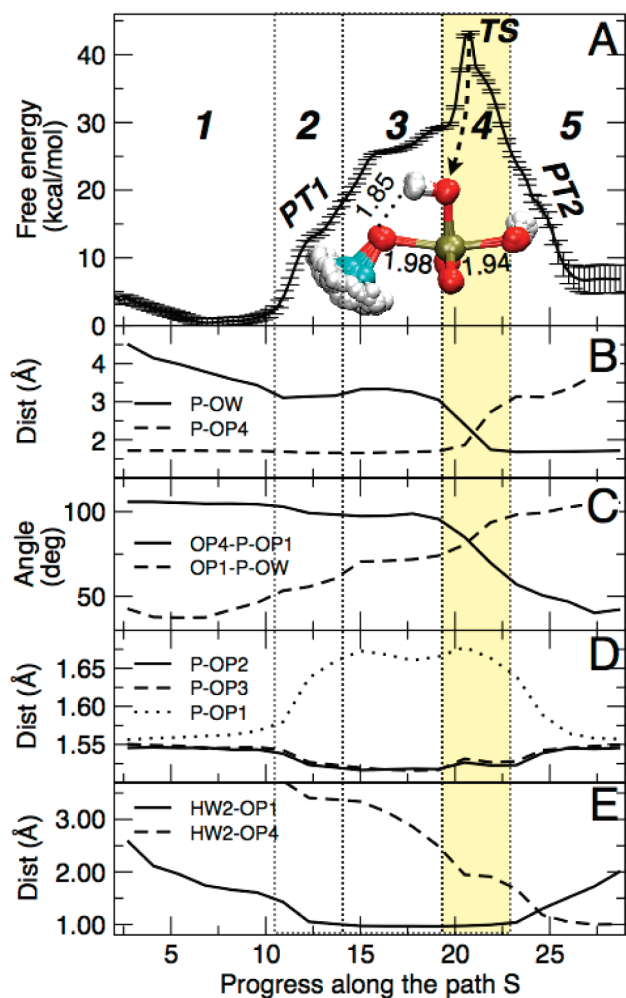


Figure 3. (A) 1D free energy profile over S (the progress of the reaction). The transition state is also displayed in ball-and-stick representation. (B–E) Behavior of the relevant structural features during the reaction. The TS region is highlighted in yellow.

The energy changes observed along the path could be ascribed to five consecutive structural events generated by different relevant components in the gradient plot (see Figure 4). When $S \in [1, 10]$ (region 1), the nucleophilic water approached the scissile phosphate, and the distance P–OW was markedly shortened without any apparent energetic penalty (Figure 3A, B). In the interval $S \in [10, 13]$ (region 2), proton HW2 was transferred from the water to OP1, which led to a barrier of ~ 10 kcal/mol on the FES (Figure 3, PT1 in panel A and E). This barrier is in agreement with that reported by Florián and Warshel.¹⁰ A decrease of the nonbridging oxygen lengths P–OP2 and P–OP3 (Figure 3D), together with a lengthening of the bridging P–OP1 bond, was also associated with this event. In principle, PT1 could also happen on OP2 or OP3. Therefore, we disclosed three degenerate pathways for this key event. Our calculations showed that this degeneracy did not significantly affect the free energy profile, since it accounted for a negligible change of ~ 0.2 kcal/mol (see the SI). At $S \in [13, 19]$ (region 3), the slope of the free energy profile changed, showing a continuous increase of 15 kcal/mol. This was mainly due to the HW2–OP1 dipole reorientation with respect to the phosphate dipole, which facilitated the nucleophilic attack of the hydroxyl ion,

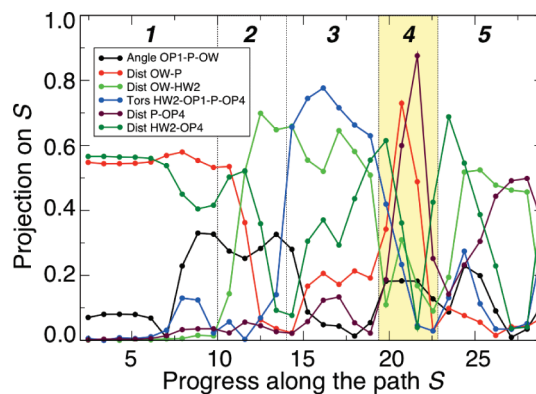


Figure 4. Absolute values of the projection of the gradients of relevant degrees of freedom over the gradient of the path variable S suitably normalized. Degrees of freedoms showing high values play a pivotal role in the reaction progress. This pattern was used to define the five regions as shown also in Figure 3. This partitioning scheme could be in principle applied to other chemical reactions.

otherwise unlikely. To further investigate the role of dipole reorientation, we carried out a gradient analysis versus the reaction variable S (see Figure 4).

This pointed to the driving components in the mean force, when moving from the reactants to the products. As far as the dipole reorientation was concerned, we could consistently observe that the torsion angle HW2–OP1–P–OP4 showed the highest contribution in region 3.

At this point, the phosphate inversion was due to the concerted change of P–OW and P–OP4 bond lengths, and OP4–P–OP1 and OP1–P–OW angles (Figure 3B,C). The associative TS geometry is represented by an “ensemble” of structures in Figure 3A. In the TS, the lengths of the forming (OW–P) and breaking (OP4–P) bonds were 1.94 Å and 1.98 Å, respectively. The geometrical features of the FES still pointed to a concerted associative mechanism. When $S \in [22.5, 29]$ (region 5), a late PT facilitated the leaving group departure (see Figure 3, PT2 in panels A and E).

The main difference with the FES-TS when compared to the PES-TS was the orientation of the proton HW2, which here pointed toward the leaving group. The HW2 PT1 and the subsequent rotation of HW2 around the P–OP1 bond is a necessary preliminary step; this key event was fundamental to inducing the nucleophilic attack of the hydroxide ion with a favorable dipole interaction. In addition, the forming and breaking bonds were ~ 0.1 Å longer in the FES-TS when compared to those identified in PES-TS. This difference was well captured in the More O’Ferrall–Jencks (MOFJ) plot, which, in Figure 5, is reported for a set of configurations along the minimum free energy pathway obtained via US and steered-MD. Remarkably, both US and steered-MD showed a less associative character for the methyl phosphate hydrolysis on the FES, compared to that identified on the PES via IRC. In particular, as shown in Figure 5, steered-MD captured the concerted nature of the mechanism, although it overemphasized the less associative character of the nucleophilic attack to the phosphate. The estimated free energy barrier was 43 (± 0.7) kcal/mol, in good agreement with the experimental value of 44–47 kcal/mol. The free energy associated with this barrier was about 6 kcal/mol less than the PES activation energy. With these simulations, we have shown that effects related to the intrinsic anharmonicity of the FES can

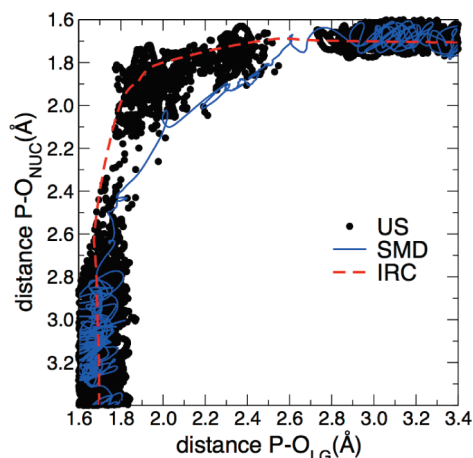


Figure 5. More O'Ferral–Jencks (MOFJ) plot for the IRC pathway. The configurations from the minimum free energy path obtained from US and steered-MD run over the low free energy pathway. P–O_{NUC} and P–O_{LG} are the distances between P and the nucleophile oxygen (OW) and the leaving group (OP4), respectively.

indeed change a reaction's TS in terms of structure and energetics with respect to a TS identified on the PES.

In summary, we have herein reported an innovative approach to the study of reaction mechanisms. Complementing traditional investigations on the PES, our approach focuses on the reaction's FES, highlighting and quantifying the smoothing effects of thermal energy on pathways identified initially on the PES. We have shown this to be true even for simple reactions, such as methyl phosphate hydrolysis. In this respect, the present methodology can be seen as extending the standard static calculations for a saddle point search toward methods that allow a direct sampling of the phase space of a chemical reaction, including its entropic effects. For instance, this approach could be very helpful in studying enzymatic reactions. In addition, it could be instrumental to better understanding linear free energy relationship experiments, which measure the change in reaction rates as a response of a different leaving or attacking group. Finally, from a mechanistic standpoint, the present study has offered a detailed understanding of the key factors that regulate the associative mechanism of a phosphoryl transfer reaction, including the proton shuttle required for a more proficient nucleophilic attack and departure of the leaving group.

■ ASSOCIATED CONTENT

S Supporting Information. DFT B3LYP structure coordinates of reagents, TSs, and products; 1D umbrella sampling results; tests over the choice of the degrees of freedom involved; and a description of the 2D umbrella sampling procedure. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: andrea.cavalli@iit.it.

■ ACKNOWLEDGMENT

Giovanni Bussi is kindly acknowledged for providing the stochastic rescaling routines. We thank IIT Integrated Computational Multiscale Platform for the computational time.

■ REFERENCES

- (1) Thrular, D. G.; Pliego, J. R., Jr. In *Continuum Solvation Models in Chemical Physics: From Theory to Application*; Mennucci, B., Cammi, R., Eds.; Wiley: Chichester, U. K., 2008; pp 338–365.
- (2) Barone, V. *J. Chem. Phys.* **2005**, *122*, 014108.
- (3) Johnson, L. N.; Lewis, R. J. *Chem. Rev.* **2001**, *101*, 2209–42.
- (4) Knowles, J. R. *Annu. Rev. Biochem.* **1980**, *49*, 877–919.
- (5) Aqvist, J.; Kolmodin, K.; Florián, J.; Warshel, A. *Chem. Biol.* **1999**, *6*, R71–R80.
- (6) Klähn, M.; Rosta, E.; Warshel, A. *J. Am. Chem. Soc.* **2006**, *128*, 15310–15323.
- (7) De Vivo, M.; Dal Peraro, M.; Klein, M. L. *J. Am. Chem. Soc.* **2008**, *130*, 10955–10962.
- (8) De Vivo, M.; Ensing, B.; Dal Peraro, M.; Gomez, G. A.; Christianson, D. W.; Klein, M. L. *J. Am. Chem. Soc.* **2007**, *129*, 387–394.
- (9) Florián, J.; Goodman, M. F.; Warshel, A. *J. Am. Chem. Soc.* **2003**, *125*, 8163–8177.
- (10) Florián, J.; Warshel, A. *J. Phys. Chem. B* **1998**, *102*, 719–734.
- (11) Kamerlin, S. C. L.; Florián, J.; Warshel, A. *ChemPhysChem* **2008**, *9*, 1767–1773.
- (12) Branduardi, D.; Gervasio, F. L.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*, 054103.
- (13) (a) Schlegel, H. B.; Millam, J. M.; Iyengar, G. A.; Voth, G. A.; Scuseria, G. E.; Daniels, A. D.; Frisch, M. J. *J. Chem. Phys.* **2001**, *114*, 9758–9763. (b) Iyengar, S. S.; Schlegel, H. B.; Millam, I. M.; Voth, G. A.; Scuseria, G. E.; Frisch, M. J. *J. Chem. Phys.* **2001**, *115*, 10291–10302.
- (14) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, 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, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision A.2; Gaussian, Inc., Wallingford CT, 2009.
- (15) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471–2474.
- (16) Bussi, G.; Donadio, D.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*, 014101.
- (17) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. Comput. Chem.* **2003**, *24*, 669–681.
- (18) Kamerlin, S. C. L.; Haranczyk, M.; Warshel, A. *Chem. Phys. Chem.* **2009**, *10*, 1125–1134.
- (19) Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; Raiteri, P.; Donadio, D.; Marinelli, F.; Pietrucci, F.; Broglia, F. A.; Parrinello, M. *Comput. Phys. Commun.* **2009**, *180*, 1961–1972.
- (20) (a) Hratchian, H. P.; Schlegel, H. B. In *Theory and Applications of Computational Chemistry: the First 40 Years*; Dykstra, C. E., Frenking, G., Kim, K. S., Scuseria, G. E., Eds.; Elsevier: Amsterdam, 2005, pp 195–249. (b) A 6-31++G** basis set was adopted in all of the present calculations. The IRC provided 200 steps in both forward and backward directions.
- (21) The reaction in question comprises three distinct events: (1) PT1 from the water molecule to the phosphate dianion, (2) rotation of the proton HW2 around the oxygen OP1, and (3) transfer of the proton HW2 to the leaving group (PT2). Therefore, we considered in PCVs representation the following key atoms: OW, HW2, HW1, P, OP1, and OP4. From the IRC trajectory, 30 equally spaced frames were selected, with an average root-mean-square deviation of 0.12 Å. This allowed us to choose a value of $\lambda = 110.0 \text{ \AA}^{-2}$. λ is the smoothing parameter for running PCVs calculations.

(22) Umbrella potentials on both S and Z spaces for a total of 506 grid points. The spring constant on the S was 1.2×10^3 kcal/mol, while that chosen for Z was 5×10^5 kcal/mol \AA^{-4} .

(23) Maragliano, L.; Vanden-Eijnden, E. *J. Chem. Phys.* **2008**, *128*, 184110.

(24) Monteferrante, M; Bonella, S.; Meloni, S.; Ciccotti, G. *Mol. Simul.* **2009**, *35*, 1116–1129.

(25) A harmonic wall over the Z variable was imposed for values higher than 0.04\AA^2 . A relative soft spring constant of 3.1 kcal/mol \AA^{-4} with a scaling factor of 0.04\AA^2 was chosen.

(26) Lad, C.; Williams, N. H.; Wolfenden, R. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 5607–5610.

An $n \log n$ Generalized Born Approximation

Ramu Anandakrishnan,[†] Mayank Daga,[†] and Alexey V. Onufriev^{*,‡}[†]Department of Computer Science and [‡]Department of Computer Science and Physics, Virginia Tech, Blacksburg, Virginia 24061, United States

ABSTRACT: Molecular dynamics (MD) simulations based on the generalized Born (GB) model of implicit solvation offer a number of important advantages over the traditional explicit solvent based simulations. Yet, in MD simulations, the GB model has not been able to reach its full potential partly due to its computational cost, which scales as $\sim n^2$, where n is the number of solute atoms. We present here an $\sim n \log n$ approximation for the generalized Born (GB) implicit solvent model. The approximation is based on the hierarchical charge partitioning (HCP) method (Anandakrishnan and Onufriev *J. Comput. Chem.* **2010**, *31*, 691–706) previously developed and tested for electrostatic computations in gas-phase and distant dependent dielectric models. The HCP uses the natural organization of biomolecular structures to partition the structures into multiple hierarchical levels of components. The charge distribution for each of these components is approximated by a much smaller number of charges. The approximate charges are then used for computing electrostatic interactions with distant components, while the full set of atomic charges are used for nearby components. To apply the HCP concept to the GB model, we define the equivalent of the effective Born radius for components. The component effective Born radius is then used in GB computations for points that are distant from the component. This HCP approximation for GB (HCP-GB) is implemented in the open source MD software, NAB in AmberTools, and tested on a set of representative biomolecular structures ranging in size from 632 atoms to ~ 3 million atoms. For this set of test structures, the HCP-GB method is 1.1–390 times faster than the GB computation without additional approximations (the reference GB computation), depending on the size of the structure. Similar to the spherical cutoff method with GB (cutoff-GB), which also scales as $\sim n \log n$, the HCP-GB is relatively simple. However, for the structures considered here, we show that the HCP-GB method is more accurate than the cutoff-GB method as measured by relative RMS error in electrostatic force compared to the reference (no cutoff) GB computation. MD simulations of four biomolecular structures on 50 ns time scales show that the backbone RMS deviation for the HCP-GB method is in reasonable agreement with the reference GB simulation. A critical difference between the cutoff-GB and HCP-GB methods is that the cutoff-GB method completely ignores interactions due to atoms beyond the cutoff distance, whereas the HCP-GB method uses an approximation for interactions due to distant atoms. Our testing suggests that completely ignoring distant interactions, as the cutoff-GB does, can lead to qualitatively incorrect results. In general, we found that the HCP-GB method reproduces key characteristics of dynamics, such as residue fluctuation, χ_1/χ_2 flips, and DNA flexibility, more accurately than the cutoff-GB method. As a practical demonstration, the HCP-GB simulation of a 348 000 atom chromatin fiber was used to refine the starting structure. Our findings suggest that the HCP-GB method is preferable to the cutoff-GB method for molecular dynamics based on pairwise implicit solvent GB models.

1. INTRODUCTION

Atomistic molecular dynamics (MD) simulations can be used to study biomolecules where experimental investigation is expensive or infeasible.^{1–6} However, the duration and system size for such simulations are limited by the computational cost of long-range electrostatic interactions, which, without further approximations, scales as $\sim N^2$ where N is the total number of atoms.^{7–10} We do not consider coarse-grained approximations here, which are in general less accurate than atomic-level approximations.¹¹

Historically, the first approximation widely used in the context of molecular dynamics (MD) to speedup the computation of long-range interactions was the spherical cutoff method. The simple spherical cutoff method treats atoms within a cutoff distance exactly while ignoring all other atoms.^{12,13} The spherical cutoff method can however produce many artifacts such as spurious forces or artificial structures around the cutoff distance.^{14–16} The particle mesh Ewald (PME) method was developed to address the shortcomings of the spherical cutoff method and has become the de facto “industry standard” for

explicit solvent MD. The PME imposes an artificial periodic boundary condition where a central cell containing the molecule of interest is assumed to be surrounded by an infinite array of images of the central cell. With this assumption, the long-range interaction, which decays slowly with distance, can be represented as the sum of two fast converging series—one in real space and the other in Fourier space.^{17–20} Another explicit solvent method, the fast multipole method,^{21–23} was tested for biomolecular simulations but has not been widely adopted, most likely due to its algorithmic complexity and instabilities caused by discontinuities inherent in the method.²⁴ In general, all three of these methods scale as $\sim N \log N$, where N is the total number of atoms in the system including the solvent atoms.

Realistic simulations require that the biomolecular structure be immersed in a solvent, typically water with ions. Implicit solvent models, such as the generalized Born (GB) approximation,

Received: July 13, 2010

Published: January 27, 2011

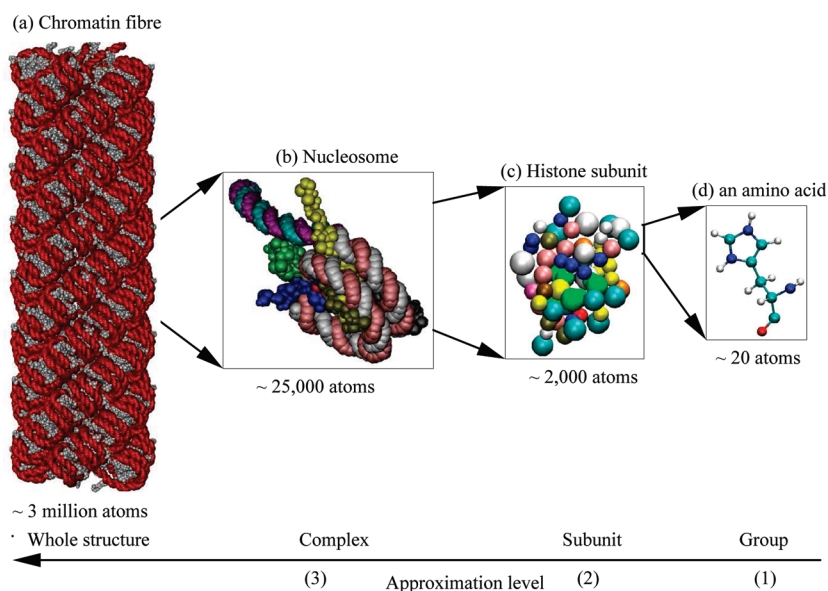


Figure 1. Example of the natural hierarchical partitioning of a chromatin fiber. (a) The fiber is made up of 100 nucleosome complexes. The individual nucleotide groups in the fiber are represented as red beads and amino acid groups as gray beads. (b) Each complex (level 3) is made up of 13 subunits with the segments of DNA linking nucleosome complexes being treated as separate subunits. A complex is shown here with each subunit represented in a different color. (c) Each subunit (level 2) is made up of 49–142 groups. The linker histone subunit is shown here with the groups colored by the type of amino acid. (d) Each group (level 1) is made up of 7–32 atoms (level 0). A histidine amino acid group is shown here with atoms represented as small spheres and covalent bonds between the atoms represented as links. The atoms are colored by the type of atom. The total fiber consists of approximately 3 million atoms. The fiber was constructed as described in Wong et al.⁵⁹ The images were rendered using VMD.⁶⁰ For clarity, only 10 of the 13 subunits are shown in a and b.

analytically represent the solvent as a continuum.^{25–50} An important benefit of implicit solvent simulations is that conformational space is sampled faster due to the reduction of solvent viscosity.^{51,52} Other benefits include instantaneous dielectric response from the solvent due to changes in solute charge state, and the elimination of “noise” in the energy landscape due to small variations in solvent structure.⁵³ Consequently, implicit solvent models are often used for applications where it is important to explore a large number of conformational states, such as for protein folding,⁵⁴ replica exchange,⁵⁵ and docking simulations.⁵⁶ However, the functional form for the most widely used practical implicit solvent model for MD, the GB model, scales as $\sim n^2$, where n represents the number of solute atoms only. ($n < N$ where N refers to the total number of atoms, including solute and solvent atoms, used for explicit solvent computations, while n refers to the number of solute atoms only, used in the implicit solvent computations.) One approach for reducing computational cost is to apply the spherical cutoff concept to the GB implicit solvent model, i.e., ignore interactions and computations involving atoms beyond a cutoff distance. We refer to this approach as the cutoff-GB method. Such an approach can reduce computational cost to $\sim n \log n$. However, the cutoff-GB may suffer from the same shortcomings as the spherical cutoff method, such as spurious forces and artificial structures around the cutoff distance. Although there are studies based on the successful use of the cutoff-GB method,^{34,57} we are not aware of a large scale systematic study that examines the effect of the cutoff on the accuracy of the GB model. To the best of our knowledge, the GB model has not been used with the PME or the fast multipole methods, most likely because the functional form of the GB model does not easily lend itself to the Ewald transformation used by the PME method or the multipole expansion used by the fast multipole method.

We present here an $\sim n \log n$ GB approximation that retains the simplicity of the cutoff-GB approximation, while in most cases being more accurate for the set of test structures considered here. Moreover, our testing demonstrates that the method presented here more accurately reproduces important characteristics of dynamics compared to the cutoff-GB method. Our approach is based on the hierarchical charge partitioning (HCP) approximation developed by us previously.⁵⁸ To approximate long-range electrostatic interactions, the HCP uses the natural organization of biomolecules into multiple hierarchical levels of components, as illustrated in Figure 1—atoms (level 0); nucleic and amino acid groups (level 1); protein, DNA, and RNA subunits (level 2); complexes of multiple subunits (level 3); and higher level structures such as fibres and membranes. The charge distribution for components above the atomic level are approximated by a much smaller number of charges. For components that are distant from the point of interest, these approximate charges are used in the computation of electrostatic interaction, while the atomic charges (level 0) are used for nearby components (Figure 2). The greater the distance from the point of interest, the larger (higher level) is the component used in the approximation of electrostatic interactions. In our previous study, we have shown that this approximation scales as $\sim n \log n$ for biomolecular structures. The HCP concept is used here to reduce the computational cost of each of the three $\sim n^2$ computations in the GB model—the computation of electrostatic vacuum energy, solvation energy, and the so-called effective Born radii—to $\sim n \log n$.

The remainder of the paper is organized as follows. In the Methods section, we briefly review the GB implicit solvent model and describe how the HCP concept is applied to the implicit solvent GB model (HCP-GB). The HCP-GB method was tested using a set of representative biomolecular structures ranging in

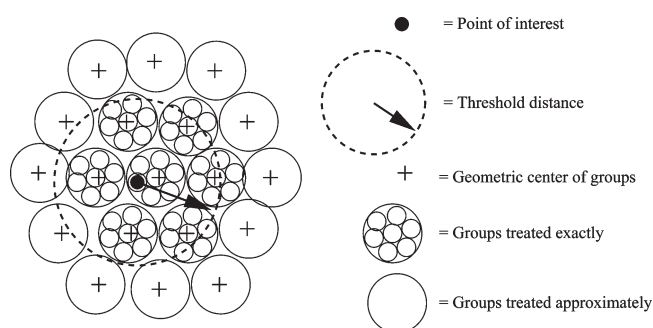


Figure 2. The HCP threshold distance. For the first level of approximation shown here, groups within the threshold distance from the point of interest are treated exactly using atomic charges (level 0), while groups beyond the threshold distance are approximated by a small number of charges (level 1). The distance to a group is computed from the point of interest to the geometric center of the group.

size from 632 atoms to 3 016 000 atoms with absolute total charge ranging from 1 to 21 424 e . The accuracy, speedup, dynamics, and conservation of momentum and energy are discussed in the Results section. In the Conclusion section, we summarize our finding and discuss the applicability of HCP-GB to practical MD problems.

2. METHODS

A short description of the GB implicit solvent model is included below, followed by a detailed description of how the HCP approximation is used to reduce the computational cost of the GB model from $\sim n^2$ to $\sim n \log n$. Also included in this section is a description of the structures and protocols used for testing the HCP-GB method.

2.1. Generalized Born (GB) Implicit Solvent Model. The electrostatic energy of a system, E^{elec} , in the presence of a solvent can be approximated by the GB implicit solvent model⁶¹ as

$$E^{\text{elec}} = E^{\text{vac}} + E^{\text{solv}} \quad (1)$$

$$E^{\text{vac}} = \sum_i^n \sum_{j>i}^n \frac{q_i q_j}{r_{ij}} \quad (2)$$

$$E^{\text{solv}} \approx -\frac{1}{2} \left(1 - \frac{1}{\epsilon_w} \right) \sum_i^n \sum_j^n \frac{q_i q_j}{[r_{ij}^2 + B_i B_j] e^{(-r_{ij}^2/4B_i B_j)}]^{1/2}} \quad (3)$$

where E^{vac} and E^{solv} are the electrostatic vacuum and solvation energy, ϵ_w is the dielectric constant of the solvent, q_i and q_j are the charges of atoms i and j , r_{ij} is the distance between the atoms, and B_i and B_j are their effective Born radii. For the purpose of this work, we consider the following Coulomb field approximation for the effective Born radius, B_i , as implemented in NAB (or Amber):⁶²

$$\frac{1}{B_i} \approx \frac{1}{R_i} - \frac{1}{4\pi} \int_{|r_{ik}|>R_i}^{\text{solute}} \frac{1}{|r_{ik}|^4} dV \quad (4)$$

where R_i is the intrinsic radius of charge i , r_{ik} is the distance from i to any point k in the solute volume, and $\int_{|r_{ik}|>R_i}^{\text{solute}} dV$ is the volume integral over the volume occupied by the solute (the cavity formed in the solvent by the solute) excluding the volume of the atom i itself.

2.2. HCP-GB—An $n \log n$ GB Approximation. Note that the computation of electrostatic vacuum energy E^{vac} in eq 2 and solvation energy E^{solv} in eq 3 both scale as $\sim n^2$. In MD software

that implements the GB implicit solvent model, such as Amber,⁶² analytical pairwise approximations for computing the effective Born radii B_i in eq 4 also scale as $\sim n^2$, unless further approximations are made. The HCP concept is used to reduce the computational cost for each of these computations to $\sim n \log n$, as described below.

2.2.1. $n \log n$ Approximation for Electrostatic Vacuum Energy. The previous HCP study⁵⁸ describes in detail the $\sim n \log n$ approximation for computing electrostatic vacuum energy E^{vac} . The key concepts from the study are summarized here. Biomolecular structures are naturally organized into multiple hierarchical levels as illustrated in Figure 1 for a chromatin fiber. Atoms are at the lowest level (level 0); groups of atoms form amino and nucleic acids (level 1); protein, DNA, and RNA chains made up of these groups form subunits (level 2); multiple subunits form complexes (level 3); and multiple complexes join together to form larger structures such as fibers and membranes. The HCP approximates atomic charges within each of the components above level 0, by a much smaller number of charges (1 or 2). For the one-charge approximation for a component, the approximate charge is placed at the “center of charge” for the component with a charge value equal to the net charge of the component. For the two-charge approximation, the two approximate charges are placed at the “center of charge” of the positive and negative charges, respectively. The center of charge is calculated in a manner similar to the center of mass when the total charge is nonzero.⁵⁸ When the total charge is zero, the component does not contribute to the approximate computation and is ignored.

The HCP then uses these approximate charges for computing electrostatic interactions beyond predefined threshold distances (Figure 2). For example, consider a structure consisting of four levels, 0–3, see Figure 3. A separate threshold distance, h_1 , h_2 , and h_3 , is defined for levels 1, 2, and 3, respectively. For complexes (level 3) farther than h_3 from the point of interest, the approximate charges for the complex are used in the computation. Otherwise, for subunits (level 2) within the complex that are farther than h_2 , the approximate charges for the subunit are used in the computation. Finally, individual atomic charges are used in the computations for charges within the level 1 threshold distance h_1 . This top-down algorithm results in $\sim n \log n$ scaling based on assumptions generally consistent with realistic biomolecular systems. Consider a hypothetical structure consisting of n atoms such that, for any given atom, there are k atoms (level 0) within the level 1 threshold distance h_1 , k groups (level 1) between h_1 and the level 2 threshold distance h_2 , k subunits (level 2) between h_2 and the level 3 threshold distance h_3 , and so on. Such a structure can be represented as a hierarchical tree with each internal node representing a component with k nodes immediately below each internal node, and with a total of n leaf nodes representing the atoms. The computational cost of the HCP algorithm for such a structure scales as $\sim n \log n$. For a more detailed description, refer to the previous HCP study.⁵⁸ This previous study also showed that for the computation of electrostatic vacuum energy, E^{vac} , in eq 2, the relatively simple HCP approximation can be comparable in accuracy to the more complex particle mesh Ewald (PME) method and more accurate than the simple spherical cutoff method.

2.2.2. $n \log n$ Approximation for Solvation Energy. To reduce the computational cost of solvation energy, E^{solv} , in eq 3 from $\sim n^2$ to $\sim n \log n$, we first define a component effective

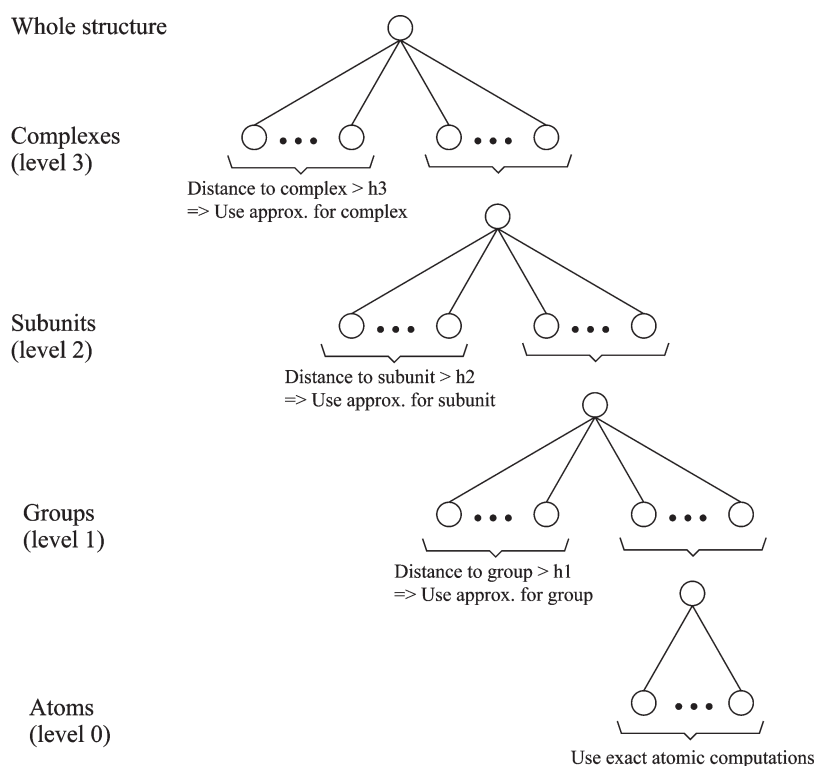


Figure 3. Illustration of the HCP approximation. Biomolecular structures are naturally organized into multiple hierarchical levels of components—complexes, subunits, groups, and atoms—as represented by the tree structure shown here. Approximations are used for computations involving distant components, while exact atomic computations are used for atoms within nearby groups. The HCP algorithm proceeds from the top level down to the lowest level to determine the level of approximation to use. The level of approximation used is determined by the distance of a component from the point of interest compared to the threshold distance for the level of the component— h_1 , h_2 , and h_3 for levels 1, 2, and 3, respectively.

Born radius, B_c for components above the atomic level. The component effective Born radii, B_c are then used to approximate the contribution of distant components to the solvation energy of atom i instead of the effective Born radii, B_j , of the individual atoms within these distant components.

To derive a simple functional form for the component effective Born radius, B_c we consider the limit of r_{ij} , $r_{ic} \rightarrow 0$, where r_{ij} is the distance from atom i to atom $j \in c$ and r_{ic} is the distance from atom i to component c . Let E_{ic}^{solv} represent the contribution of component c to the solvation energy of atom i .

$$E_{ic}^{\text{solv}} = -\frac{1}{2} \left(1 - \frac{1}{\epsilon_w}\right) \sum_{j \in c} \frac{q_i q_j}{[r_{ij}^2 + B_i B_j e^{(-r_{ij}^2/4B_i B_j)}]^{1/2}} \quad (5)$$

$$\approx -\frac{1}{2} \left(1 - \frac{1}{\epsilon_w}\right) \sum_{j \in c} \frac{q_i q_j}{[B_i B_j]^{1/2}} \quad (r_{ij} \rightarrow 0) \quad (6)$$

$$\text{Also, } E_{ic}^{\text{solv}} = -\frac{1}{2} \left(1 - \frac{1}{\epsilon_w}\right) \frac{q_i q_c}{[r_{ic}^2 + B_i B_c e^{(-r_{ic}^2/4B_i B_c)}]^{1/2}} \quad (7)$$

$$\approx -\frac{1}{2} \left(1 - \frac{1}{\epsilon_w}\right) \frac{q_i q_c}{[B_i B_c]^{1/2}} \quad (r_{ic} \rightarrow 0) \quad (8)$$

$$\Rightarrow \frac{1}{B_c} \approx \left[\frac{1}{q_c} \sum_{j \in c} \frac{q_j}{B_j^{1/2}} \right]^2 \quad (9)$$

For the one-charge HCP approximation, q_c is the net charge of the component, and the sum in the above equations is over all the atoms in the component. For the two-charge approximation, two separate component effective Born radii are computed, one for each of the two approximate charges, i.e., one for the positively charged atoms and another for the negatively charged atoms. In this case, q_c represents the total positive or negative charge and the sum is over the positively or negatively charged atoms, respectively.

In this work, we also considered two alternatives to eq 9 for component effective Born radii. One approximation, by Archontis and Simonson³⁹ developed in the context of a coarse grain model, defines the equivalent of the component effective Born radius as the harmonic average of its constituent atomic Born radii weighted by the square of atomic charges. The resulting expression is similar to eq 9 except that the constituent atomic Born radii are weighted by atomic charges. Another approach is to use the analytical approximation for effective Born radii defined by eq 10 described below, with i representing a component c instead of an atom, and $j \neq i$ replaced by $j \in c$. We examined these alternatives (results included in Appendix A.1) and found that on average the approach described above by eq 9 is more accurate, although in some specific instances, one of the other alternatives can be more accurate. We have therefore chosen to base all further analysis on eq 9 but note that future work may lead to better approximations.

2.2.3. $n \log n$ Approximation for Effective Born Radii B_i . To compute the integral in eq 4, the Coulomb field approximation for effective Born radii B_i , we will consider here one commonly

used approximation to eq 4. However, the main idea can be applied to any volume based approximation for computing effective Born radii. In the specific approximation considered here, the integral in eq 4 is computed over the volume occupied by individual atoms, ignoring overlaps between atoms and spaces between atoms that are inaccessible to the solvent.⁴⁴ For the case where the atoms i and j do not overlap, this analytical approximation computes the effective Born radius B_i as

$$\frac{1}{B_i} \approx \frac{1}{R_i} - \sum_{j \neq i} \left[\frac{R_j}{2[r_{ij}^2 - R_j^2]} - \frac{1}{4r_{ij}} \log \frac{r_{ij} + R_j}{r_{ij} - R_j} \right] \quad (10)$$

The computation of effective Born radii using the above approximation scales as $\sim n^2$. The HCP approximation can be used to reduce the computational cost of the above equation to $\sim n \log n$, as follows. We define a component radius R_c for a component c which can be used to approximate the contribution of distant components to the effective Born radius of an atom, replacing the computations involving the individual atoms within the component. Then, using the HCP approach described in section 2.2.1 above, the effective Born radius B_i for atom i can be approximated in $\sim n \log n$ computations.

To derive a simple expression for component radius R_c , we consider the limit of $r_{ij} \rightarrow r_{ic}$, where r_{ij} is the distance from atom i to atom $j \in c$, and r_{ic} is the distance from atom i to component r_{ij} . For distant components c , let B_{ic} be the contribution of the atoms $j \in c$, to the effective Born radius of i . B_{ic} can be approximated by a truncated Taylor series expansion of eq 10 as

$$\frac{1}{B_{ic}} = C \sum_{j \in c} \frac{R_j^3}{r_{ij}^4} + \text{higher order terms} \quad (11)$$

$$\approx C \sum_{j \in c} \frac{R_j^3}{r_{ic}^4} (r_{ij} \rightarrow r_{ic}) \quad (12)$$

$$\text{Also, } \frac{1}{B_{ic}} = C \frac{R_c^3}{r_{ic}^4} \quad (13)$$

$$\Rightarrow R_c^3 \approx \sum_{j \in c} R_j^3 \quad (14)$$

where C is a constant. Interestingly, R_c is the radius of a sphere with the same volume as the sum of volumes of its constituent atoms, which is what one may intuitively expect. On the basis of the simple and intuitive nature of the expression for R_c , we conjecture that the form of eq 14 is independent of the specifics of eq 10. For a distant component c , the component radius R_c is used in eq 10 in place of R_j and r_{ic} in place of r_{ij} for atoms $j \in c$. Higher level components are used in the computation of effective Born radii B_i for more distant components such that computational cost scales as $\sim n \log n$, as described in section 2.2.1 above.

2.2.4. $n \log n$ Approximation for Solvation Forces. The solvation force on an atom i , F_i^{solv} , is computed as the derivative of the solvation potential ϕ_i^{solv} using the chain rule, as follows:

$$\phi_i^{\text{solv}} = \sum_{j \neq i} \phi_{ij}^{\text{solv}} \quad (15)$$

$$\phi_{ij}^{\text{solv}} = -\frac{1}{2} \left(1 - \frac{1}{\epsilon_w} \right) \frac{q_i q_j}{[r_{ij}^2 + B_i B_j e^{(-r_{ij}^2/4B_i B_j)}]^{1/2}} \quad (16)$$

$$F_i^{\text{solv}} = \sum_{j \neq i} F_{ij}^{\text{solv}} \quad (17)$$

$$F_{ij}^{\text{solv}} = -q_i \partial \phi_{ij}^{\text{solv}} / \partial r_{ij} \quad (18)$$

$$= -\frac{1}{4} \left(1 - \frac{1}{\epsilon_w} \right) \frac{q_i q_j}{[r_{ij}^2 + B_i B_j e^{-r_{ij}^2/4B_i B_j}]^{3/2}} \times \quad (19)$$

$$\left[2r_{ij} - \frac{1}{2} r_{ij} e^{-r_{ij}^2/4B_i B_j} + e^{-r_{ij}^2/4B_i B_j} \left(1 + \frac{r_{ij}^2}{4B_i B_j} \right) \left(B_j \frac{\partial B_i}{\partial r_{ij}} + B_i \frac{\partial B_j}{\partial r_{ij}} \right) \right] \quad (20)$$

$$\frac{\partial B_i}{\partial r_{ij}} = B_i^2 \sum_{j \neq i} \left[\frac{-R_j r_{ij}}{(r_{ij}^2 - R_j^2)^2} + \frac{1}{4r_{ij}^2} \log \frac{R_j + r_{ij}}{R_j - r_{ij}} + \frac{R_j}{2r_{ij}(R_j^2 - r_{ij}^2)} \right] \quad (21)$$

where ϕ_{ij}^{solv} is the solvation potential contribution of atom j at atom i , F_{ij}^{solv} is the corresponding force contribution, ϵ_w is the dielectric constant of the solvent, r_{ij} is the distance between atoms i and j , q_i and q_j are the atomic charges, and R_i and R_j are the intrinsic radii. Here, $\partial B_i / \partial r_{ij}$ is the derivative of the effective Born radii (eq 10). For distant components, component intrinsic radii and component effective Born radii are used in the above equations instead of the atomic intrinsic radii and atomic effective Born radii. Using the HCP approach described in section 2.2.1 above, solvation forces can then be approximated in $\sim n \log n$ computations.

2.3. Test Structures and Protocols. To assess performance of the HCP-GB method in the context of molecular dynamics, we implemented the method in NAB, the open source molecular dynamics (MD) software in AmberTools v1.3.⁶³ The HCP implementation in NAB is scheduled to be released with AmberTools v1.5, for general use. In some sense, NAB is a minimal version of the production Amber MD software⁶² and is particularly well suited for experimentation unlike the highly optimized but also more complex production version. NAB however does use the same force fields and implements the same GB implicit solvent methods and options as the production Amber code. For the purpose of this study, we used the commonly used OBC GB model (IGB = 5 in Amber⁶⁵).

Performance in both accuracy and speed was evaluated relative to the reference GB computation without any additional approximations (reference GB). We also compared the HCP-GB method to the same GB implicit solvent model with a spherical cutoff (cutoff-GB). The cutoff-GB method ignores all interactions beyond a cutoff distance for the computation of electrostatic energy and effective Born radii in eqs 1–3 and 10. Our previous study⁵⁸ had compared the electrostatic vacuum energy and forces computed by the HCP method to the particle mesh Ewald (PME) explicit solvent method. However, to the best of our knowledge, the GB implicit solvent model has not been implemented for the PME method in readily available molecular dynamics software. Therefore, a similar comparison for the HCP-GB method was not performed here.

The HCP-GB method was tested on a set of eight representative biomolecular structures ranging in size from 632 atoms to 3 016 000 atoms with absolute total charge ranging from 1 to 21 424e (Table 1). The H++ server (<https://biophysics.cs.vt.edu/H++>) was used to add missing hydrogens to these structures.⁶⁴

Table 1. List of Representative Structures Used for Testing^a

structure	PDB ID	size (atoms)	charge (e)	cutoff dist (Å)	threshold dist (Å)		
					h_1	h_2	h_3
10 bp B-DNA fragment	2BNA	632	18	21	21	n/a	n/a
immunoglobulin binding domain	1BDD	726	2	15	15	n/a	n/a
ubiquitin	1UBQ	1231	1	15	15	n/a	n/a
thioredoxin	2TRX	1654	5	15	15	n/a	n/a
nucleosome core particle	1KX5	25101	133	21	21	90	n/a
microtubule sheet	<i>b</i>	158016	360	15	15	48	n/a
virus capsid	1A6C	475500	120	15	15	66	n/a
chromatin fiber	<i>c</i>	3016000	21424	21	21	90	169

^a Unless stated otherwise, the cutoff and threshold distances listed here were used for all testing. ^b The microtubule sheet was constructed as described in Wang and Nogales.⁶⁶ ^c The chromatin fiber was constructed as described in Wong et al.⁵⁹

The HCP threshold distances were chosen such that, for a given atom within a given test structure, the exact atomic computation (level 0) is used for interactions with other atoms within its own and nearest neighboring groups (level 1), as illustrated in Figure 2. To satisfy this condition, threshold distances h_l are calculated as $h_l = R_l^{\max} + 2 \times R_1^{\max}$ where l is the HCP level, R_l^{\max} is the maximum component radius at level l , and R_1^{\max} is the maximum group (level 1) radius, for a given structure. The HCP threshold distances thus calculated for each of the test structures are shown in Table 1. These are the suggested conservative defaults for these and other similar structures. The HCP-GB level 1 threshold distance for a given structure is also used as the cutoff distance for the cutoff-GB computations. Unless stated otherwise, these threshold and cutoff distances were used for all of the testing described in the Results section.

Four metrics were used to measure the accuracy of the approximate methods: relative error in electrostatic energy (relative energy error) Err^E , for vacuum, solvation, and net electrostatic energy, and relative RMS error in electrostatic force (relative force error) Err^F , calculated as

$$\text{Err}^E = |E^{\text{approx}} - E^{\text{ref}}|/E^{\text{ref}} \quad (22)$$

$$\text{Err}^F = \text{Err}^{\text{rms}}/F^{\text{avg}} \quad (23)$$

$$\text{Err}^{\text{rms}} = \left[\sum_{i=1}^n |F_i^{\text{approx}} - F_i^{\text{ref}}|^2/n \right]^{1/2} \quad (24)$$

$$F^{\text{avg}} = \sum_{i=1}^n |F_i^{\text{ref}}|/n \quad (25)$$

where E^{approx} is the energy calculated using an approximation, E^{ref} is the energy calculated using the reference GB computation without cutoffs or the use of HCP, Err^{rms} is the root-mean-square (RMS) error in force for the atoms in a given structure, F^{avg} is the average force, and F_i^{approx} and F_i^{ref} are the force on atom i calculated using the approximate and reference GB computations, respectively.

Speedup was measured as CPU time for the reference (no cutoff) GB computation divided by the CPU time for the approximation tested.^a All testing was conducted on Virginia Tech's System X computer cluster (<http://www.arc.vt.edu>) consisting of 1100 dual core 2.5 GHz PowerPC 970FX processors with 4

GB of RAM, running the Apple Mac OS X 10.3.9, and connected by 10 Gbps InfiniBand switches. Where possible, testing was performed using a single CPU (a single core of the dual core processor) to reduce the potential variability due to interprocessor communication. However, due to the large memory requirements for the neighbor list used by the cutoff-GB method, it was not possible to run the cutoff-GB computation for structures larger than 200 000 atoms using a single CPU in the test environment described above. Therefore, 16 CPUs were used for the 475 500 atom virus capsid and 128 CPUs for the 3 016 000 atom chromatin fiber. For comparison on an equal footing, the reference GB computations and the HCP-GB computations were also performed with the same number of CPUs. When multiple CPUs were used, the CPU time for the longest running CPU was used to calculate speedup. To limit the run time for the reference GB computation to a few days, speedup was calculated for 1000 iterations of MD for structures with <10 000 atoms, 100 iterations for structures with 10 000–1 000 000 atoms, and 10 iterations for the structure with >1 000 000 atoms. To make the results representative of typical simulations involving much larger numbers of iterations, the CPU time excludes the time for loading the data and initialization prior to starting the simulation. Note that the speedup may vary with the computing system characteristics, such as interprocessor communication network, number of processors used, processor architecture, memory configuration, etc. A detailed analysis of the effect of these characteristics on speedup is beyond the scope of this study, which focuses on the algorithm.

The following parameters and protocol were used for the simulations, unless otherwise stated. The threshold distances used are listed in Table 1. 6–12 van der Waals interactions for the HCP-GB were computed using only the atoms that are within the level 1 threshold distance, i.e., atoms that are treated exactly. The simulations used the Amber ff99SB force field.⁶⁷ Langevin dynamics with a collision frequency of 50 ps⁻¹ (appropriate for water) was used for temperature control, a surface-area dependent energy of 0.005 kcal/mol/Å² was added, and an inverse Debye–Huckel length of 0.125 Å⁻¹ was used to represent a 0.145 M salt concentration. A 1 fs time step was used for the simulation with the nonbonded neighbor list being updated after every step. Note that updating the nonbonded neighbor list less frequently will improve the speedup of the cutoff-GB method; however, the speedup of the HCP-GB method can also be improved similarly by updating component radii and charges less frequently. For simplicity and for

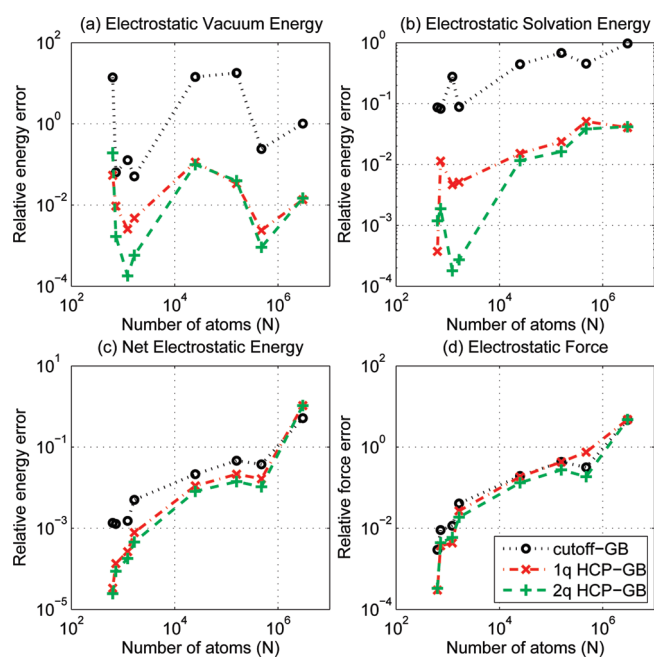


Figure 4. Accuracy of the HCP-GB and cutoff-GB methods relative to the reference GB computation without cutoffs. Accuracy is computed as the relative error in (a) vacuum, (b) solvation, and (c) net electrostatic energy, and (d) relative RMS error in electrostatic force, for the one-charge and two-charge HCP-GB and the cutoff-GB method. Connecting lines are shown to guide the eye.

comparison on an equal footing, the nonbonded neighbor list and component radii and charges are updated after every step. Default values were used for all other parameters. The simulation protocol consisted of five stages. First, the starting structure was minimized using the conjugate gradient method with a restraint weight of $5.0 \text{ kcal/mol/\AA}^2$. Next, the system was heated to 300 K over 10 ps with a restraint weight of $1.0 \text{ kcal/mol/\AA}^2$. The system was then equilibrated for 10 ps at 300 K with a restraint weight of $0.1 \text{ kcal/mol/\AA}^2$, and then for another 10 ps with a restraint weight of $0.01 \text{ kcal/mol/\AA}^2$. Finally, all restraints were removed for the production stage.

3. RESULTS AND DISCUSSION

We examined a number of characteristics of the HCP-GB method that are important for molecular dynamics—accuracy, speed, dynamics, and conservation of energy and momentum, which are discussed below.

3.1. Accuracy. Figure 4 shows the accuracy for the one- and two-charge HCP-GB methods compared to the cutoff-GB method. For the test structures considered here, the values of the two components of electrostatic interactions—vacuum and solvation energies—as calculated by the HCP-GB method are significantly more accurate than that of the cutoff-GB method (Figure 4a and b). For the net electrostatic energy and force, the relative improvement provided by the HCP-GB is significant for the smaller structures and decreases with structure size (Figure 4c and d). For the largest structure considered here—the 3 million atom chromatin fiber—the relative energy error for the HCP-GB method is slightly higher than that of the cutoff-GB method. Preliminary analysis suggests that the larger HCP-GB error for the chromatin fiber may be due to the negligible contribution of very distant components to the net energy

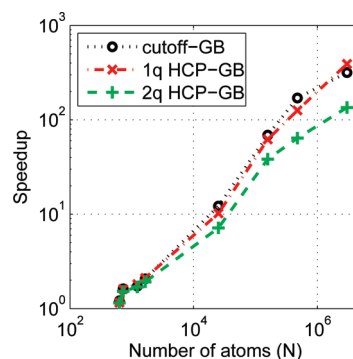


Figure 5. Speedup for the HCP-GB and cutoff-GB methods relative to the reference GB computation without cutoffs. Threshold and cutoff distances used for the different structures are listed in Table 1. Connecting lines are shown to guide the eye.

computation, even though the individual vacuum and solvation components may be large. Small errors in the estimation of these individual components can result in a large relative error in net electrostatic energy. Thus, ignoring the contribution of these distant components, as the cutoff-GB method does, may actually decrease the error in total energy as defined by the above metrics. However, as our examination of key characteristics of dynamics in section 3.6 below shows, the single point net force and net energy error metrics presented above are too crude to unambiguously differentiate between the expected performance of the cutoff-GB and HCP-GB methods in the context of molecular dynamics. For example, one can expect the cutoff scheme to neglect a roughly equal number of pairwise interactions of roughly equal magnitude but of opposite sign. The resulting cancellation of error in total electrostatic energy can be deceptive. As we shall see later in section 3.6, neglect of charge–charge interactions clearly manifests itself by producing artifacts in dynamics.

3.2. Speedup. Figure 5 shows that the speedup for the one-charge HCP-GB and cutoff-GB methods are comparable,^b while the two-charge HCP-GB is slower. Surprisingly, the cutoff-GB method is slower than the one-charge HCP-GB method for the three million atom chromatin fiber. We speculate that this is because the NAB implementation of the cutoff method does not scale well with system size due to the additional memory access required for the large neighbor list used by the method.

As noted earlier, unlike the production pmemd module of Amber 8, NAB is not highly optimized. However, on the basis of the run times for a 0.1 ns simulation of the nucleosome core particle (1KX5), compared to an equivalent simulation by Ruscio and Onufriev,⁶⁸ we estimate that NAB v1.3 is only about 1.5 times slower than the production pmemd module of Amber 8 on Virginia Tech's System X computer cluster described above.

3.3. Tradeoff between Speed and Accuracy. For a given structure, the speed and accuracy of the HCP-GB method depends primarily on two parameters: the number of charges used to approximate the components and the threshold distances. As seen in Figures 4 and 5, on the basis of net energy and force metrics, the two-charge approximation is more accurate but slower than the one-charge approximation. Figure 6 shows that increasing the threshold distance improves accuracy but reduces speed. However, as our analysis of key characteristics of dynamics (section 3.6) shows, the single-point error metrics, based on net energy or force, used above may not provide a complete measure

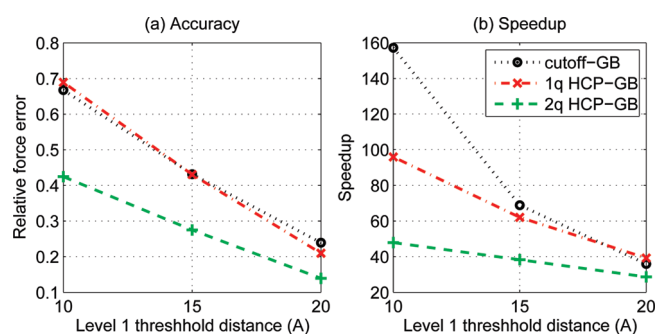


Figure 6. Tradeoff between (a) accuracy and (b) speed for the 158 016 atom microtubule structure. Cutoff and level 1 threshold distances are varied from 10 Å to 20 Å. Level 2 threshold distance is 48 Å. Connecting lines are shown to guide the eye.

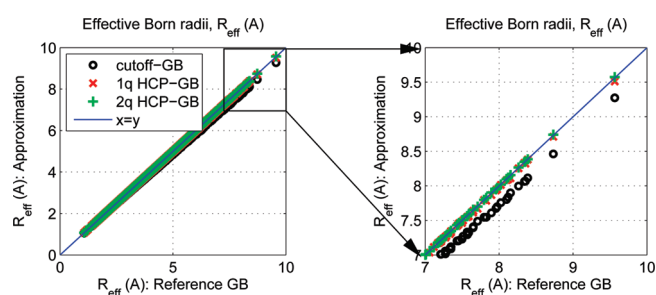


Figure 7. Accuracy of effective Born radii approximations. Effective Born radii for the 1653 atoms of thioredoxin (2TRX) are calculated using the one-charge HCP-GB, the two-charge HCP-GB, and the cutoff based (cutoff-GB) approximations and plotted against the reference GB computation without cutoffs.

of correctness in the context of molecular dynamics. The optimal choice of parameters depends on the structure and problem under consideration. For the purpose of this study, we have chosen conservative threshold distances (Table 1) such that for any given atom, atoms within its own group and immediately neighboring groups (level 1) are treated exactly, as described in section 2.3. It is possible that shorter than default threshold distances may be acceptable for specific applications, but we suggest that the decision to use shorter threshold distances be made on a case-by-case basis. For example, the two-charge HCP-GB simulation of immunoglobulin binding domain (1BDD) remains stable when the level 1 threshold distance is reduced from the recommended 15 Å to 10 Å, but the protein quickly unfolds when the threshold distance is further reduced to 5 Å.

3.4. Accuracy of the HCP Approximation for Effective Born Radii. We tested the HCP based approximation for effective Born radii on a typical structure used in this context, thioredoxin (2TRX). Figure 7 shows that for this structure, the HCP-GB approximation with a threshold distance of 15 Å is slightly more accurate than a cutoff based approximation with a 15 Å cutoff distance. The overall RMS error in effective Born radii relative to the reference GB computation without cutoffs is 0.0058 Å for the one-charge HCP-GB, 0.0017 Å for the two-charge HCP-GB, and 0.0557 Å for the cutoff-GB method.

Both the HCP-GB and the cutoff-GB introduce two sources of error into the total electrostatic energy, relative to the no-cutoff reference. One is the approximations to effective Born radii, and the other is the approximations to the electrostatic interactions.

Table 2. Relative RMS Error in Total Electrostatic Energy for Thioredoxin (2TRX) Due to Different Approximations of Effective Born Radii^a

effective Born radii approximated using	relative RMS error in total electrostatic energy	
	energy calculated with	
	cutoff	no cutoff
cutoff	0.50%	0.49%
1-q HCP	0.05%	<0.01%

^a RMS error is calculated relative to the reference GB computation. Cutoff and level 1 threshold distances of 15 Å were used for these computations.

The relative impact of these two sources is shown in Table 2. Clearly, for the spherical cutoff, the error in effective Born radii is the dominant source of error in the total electrostatic energy. The use of the HCP-GB approximation for effective Born radii can reduce this error by an order of magnitude. Whether these errors in effective Born radii will have a material impact on dynamics depends on, among other factors, the relative magnitude of the errors inherent in the approximation used to compute effective Born radii in the reference model. Nevertheless, the improvement in the accuracy of effective Born radii using the HCP approximation, compared to the cutoff approximation, comes at little or no additional cost and should therefore be used instead of the cutoff approximation.

3.5. Stability in MD Simulations. To test the stability of the HCP-GB algorithm, we ran 50 ns MD simulations of the immunoglobulin binding domain (1BDD), ubiquitin (1UBQ), a 10 base-pair fragment of B-DNA (2BNA), and thioredoxin (2TRX). Figure 8 shows the backbone RMS deviation from the crystal structure for the simulations, which are summarized in Table 3. These results suggest that the trajectory for the cutoff-GB and HCP-GB methods are generally in reasonable agreement with the reference GB simulation. For 1BDD, the one-charge HCP-GB trajectory shows RMS deviations similar to the cutoff-GB trajectory but substantially larger than the two-charge HCP-GB or the reference GB trajectories. This example emphasizes how subtle errors in charge–charge interactions can result in qualitatively different conformational dynamics. On a practical level, it suggests that the one-charge HCP-GB may not be appropriate for the simulation of small flexible structures, such as 1BDD, where small inaccuracies in the potential can lead to large structural deviations over the course of the trajectory. For such structures, we recommend the two-charge HCP-GB.

The above simulations were run with a Langevin collision frequency of 50 ps⁻¹ for thermal coupling. We also performed, for the same set of structures, 10 ns simulations with the thermal coupling reduced to 0.01 ps⁻¹ (results not shown). As expected, these simulations resulted in an enhanced sampling of conformational space, as was seen by more frequent excursions in RMS space. The weak Langevin coupling simulations were in general agreement with the simulations that used strong Langevin coupling. For example, for 1BDD, the two-charge HCP-GB and the reference GB simulations exhibited similar RMS deviations from the starting structure, while the one-charge HCP-GB and cutoff-GB resulted in much higher RMS deviations toward the end of the respective trajectories.

3.6. Detailed Characteristics of the Simulation Dynamics. An important qualitative difference between the HCP-GB

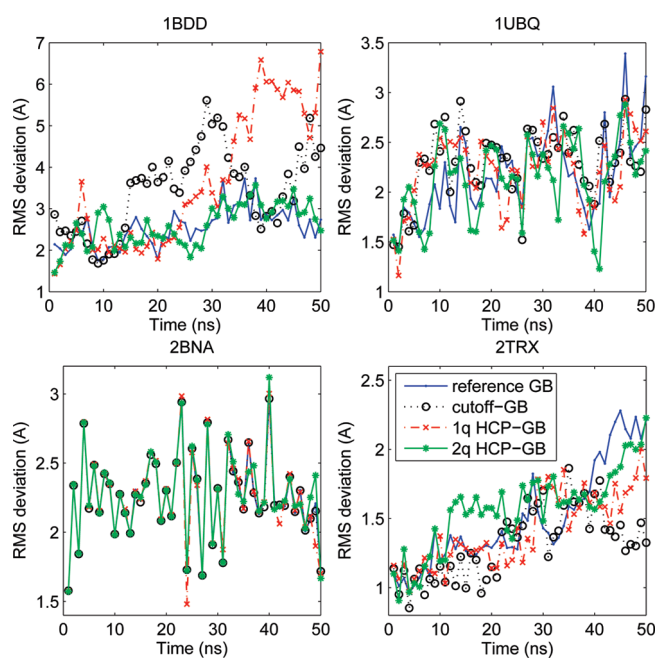


Figure 8. RMS deviation from the starting structure for 50 ns MD simulations of immunoglobulin binding domain (1BDD), ubiquitin (1UBQ), B-DNA (2BNA), and thioredoxin (2TRX) using the reference GB, cutoff-GB, and HCP-GB methods. RMS deviation is calculated for backbone heavy atoms. The trajectory is sampled every 1 ns. Connecting lines are shown to guide the eye.

Table 3. RMS Deviation from the Starting Structure for 50 ns Simulations of Immunoglobulin Binding Domain (1BDD), Ubiquitin (1UBQ), B-DNA (2BNA), and Thioredoxin (2TRX)^a

PDB ID	average RMS deviation \pm standard deviation (Å)			
	reference GB	cutoff-GB	one-charge HCP-GB	two-charge HCP-GB
1BDD	2.64 \pm 0.45	3.69 \pm 0.88	3.92 \pm 1.66	2.72 \pm 0.48
1UBQ	2.29 \pm 0.37	2.39 \pm 0.29	2.32 \pm 0.33	2.22 \pm 0.36
2BNA	2.24 \pm 0.30	2.24 \pm 0.30	2.23 \pm 0.31	2.25 \pm 0.31
2TRX	1.60 \pm 0.33	1.41 \pm 0.23	1.50 \pm 0.23	1.67 \pm 0.19

^a RMS deviation is calculated for backbone heavy atoms. The trajectory is sampled every 10 ps. Averages are for the last 40 ns of the 50 ns simulations. Standard deviation is computed as $\sqrt{[\sum_i (\text{RMS}_i - \mu)^2 / s]}$, where RMS_i is the RMS deviation for the i th sample, μ is the average RMS deviation, and s is the number of samples.

method and the cutoff-GB method is that the cutoff-GB method completely ignores the effect of all charges beyond the cutoff distance, while the HCP-GB method approximates the effect of distant charges. We believe that ignoring these distant charges can, under many circumstances, lead to qualitatively different, and incorrect, results. Consider for example the RMS fluctuation in the position of residues—a characteristic of internal dynamics of the structure. To quantify the overall difference in fluctuation for all residues compared to the reference GB simulation, we compute the RMS difference in RMS fluctuation for the 50 ns simulation of the four structures described in the Stability section above. For the structures tested here, the RMS difference in fluctuation (Table 4) indicates that on average both the one-charge

Table 4. Detailed Characteristics of Simulation Dynamics from 50 ns Simulations of Immunoglobulin Binding Domain (1BDD), Ubiquitin (1UBQ), B-DNA (2BNA), and Thioredoxin (2TRX)^a

PDB ID	cutoff-GB	one-charge	two-charge
		HCP-GB	HCP-GB
RMS difference in RMS residue fluctuations (Å)			
1BDD	0.51	0.54	0.37
1UBQ	0.24	0.20	0.15
2BNA	0.003	0.01	0.03
2TRX	0.37	0.32	0.22
average	0.28	0.26	0.19
RMS difference in distribution of χ_1 angles (% occurrence)			
1BDD	3.26	2.19	2.49
1UBQ	3.15	3.26	2.81
2TRX	3.17	3.48	3.16
average	3.19	2.98	2.82
RMS difference in distribution of χ_2 angles (% occurrence)			
1BDD	2.79	2.01	2.21
1UBQ	2.66	2.63	2.37
2TRX	2.87	2.91	3.01
average	2.77	2.52	2.53

^a RMS difference was calculated relative to the reference GB simulation. The trajectory was sampled every 10 ps. χ angles do not apply to the DNA strand 2BNA. A bin size of 10° was used for calculating the distribution of χ angles.

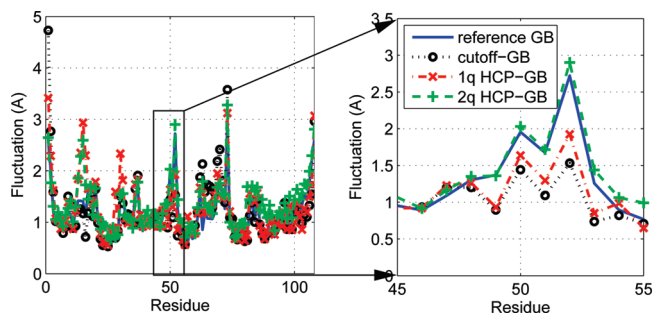


Figure 9. RMS fluctuation in residue positions for a 50 ns simulation of thioredoxin. The trajectory was sampled every 10 ps. Connecting lines are shown to guide the eye.

and two-charge HCP-GB simulations are in better agreement with the reference GB simulation than the cutoff-GB method. The differences in RMS fluctuation from the 50 ns simulation of thioredoxin are highlighted in Figure 9.

Similarly, consider the χ_1 angles for the functionally important CYS-32 of thioredoxin and THR-7 of ubiquitin (Figure 10). The χ_1 angle for CYS-32 flips between approximately -180° and $+60^\circ$ during the two-charge HCP-GB simulations as does the “correct” reference GB simulation. Whereas, for the one-charge HCP-GB and cutoff-GB methods, the angle stays at approximately -180° . And the χ_1 angle for THR-7 stays around approximately 60° during the reference GB and the HCP-GB simulations, whereas for the cutoff-GB simulation, the angle flips briefly between approximately -60° and $+60^\circ$. To quantify the overall difference in the distribution of χ_1 and χ_2 angles, we

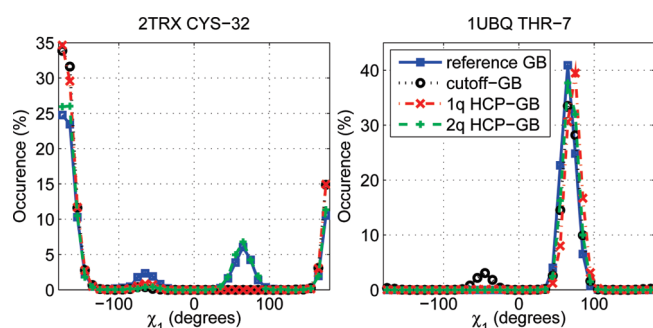


Figure 10. Distribution of χ_1 angles for the functionally important CYS-32 of thioredoxin (2TRX) and THR-7 of ubiquitin (1UBQ) from 50 ns simulations. The trajectory was sampled every 10 ps. A bin size of 10° was used for calculating the distribution of χ_1 angles. Connecting lines are shown to guide the eye.

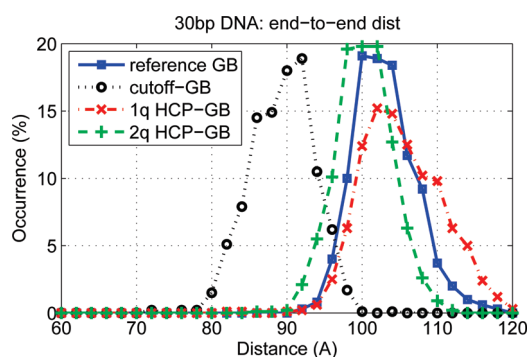


Figure 11. Distribution of distances between terminal base pairs for a 30 bp DNA strand from a 10 ns MD simulation. Trajectories were sampled every 10 ps, and a bin size of 2 Å is used for calculating the distribution of distances between terminal base pairs. Connecting lines are shown to guide the eye.

computed the RMS difference in the distribution compared to the reference GB simulation. The RMS difference in the distribution of χ_1 and χ_2 angles for 50 ns simulations of the four structures described in the “Stability” section above indicates that on average the one- and two-charge HCP-GB simulations are in better agreement with the reference GB simulation than the cutoff-GB simulations (Table 4).

To further examine the effect of ignoring distant charges we ran a 10 ns simulation of a 30 base-pair DNA strand with the same setup as described in section 2.3 above, but without the salt. Figure 11 shows the distribution of distances between terminal base pairs. The distribution shows that for the cutoff-GB method, with an average end-to-end distance of 89 Å, the structure is more flexible than for the reference GB or the one- and two-charge HCP-GB methods, with average end-to-end distances of 103, 105, and 100 Å, respectively. This difference in flexibility is most likely due to the fact that the cutoff-GB method completely ignores distant charges which contribute to the bending rigidity of the DNA chain.

We also ran a 0.3 ns simulation of the nucleosome core particle, with the same setup as described in section 2.3 above, but without the salt. When using the reference GB and the HCP-GB methods, all of the histone tails collapse onto the DNA chain within 0.1 ns, consistent with experimental observations at low salt concentrations,⁶⁹ whereas when the cutoff-GB method is used, two of the positively charged tails fail to collapse onto the

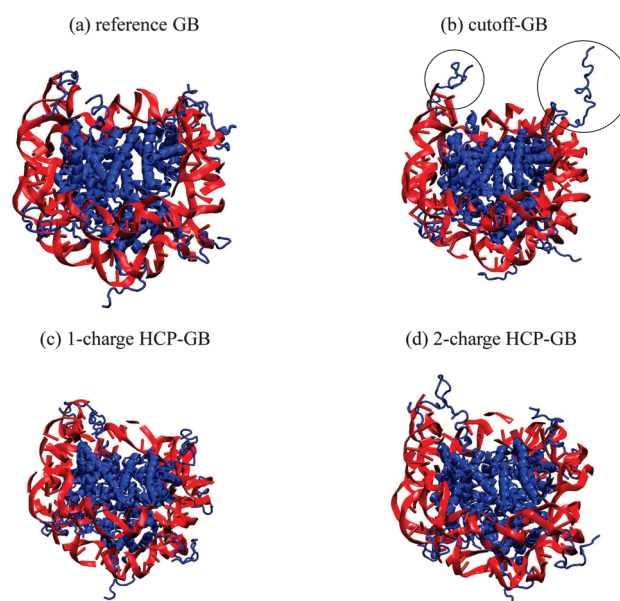


Figure 12. Nucleosome core particle after 0.1 ns simulation using (a) reference GB, (b) cutoff-GB, (c) one-charge HCP-GB, and (d) two-charge HCP-GB computations. Positively charged histone tails have collapsed onto the DNA for the reference GB and HCP-GB, whereas for the cutoff-GB, two of the tails fail to collapse.

negatively charged DNA, Figure 12. Again, this is most likely because, in the case of the cutoff-GB method, the positive charges at the ends of the histone tail do not “feel” the attraction from the highly charged DNA chain.

The above results suggest that in general the HCP-GB reproduces the dynamics of the reference GB simulation more accurately than the cutoff-GB method.

3.7. A Practical Application: Chromatin Fiber. We expect the HCP-GB to be indispensable in the modeling of large structures where the pairwise GB without further approximation is impractical. One such example is the chromatin fiber where a 348 000 atom (12 nucleosome) structure is needed at a minimum to study its functional characteristics. Such a structure can be constructed using the crystal structure for the nucleosome (1KX5) as a starting point. Multiple copies of the nucleosome can then be combined to construct the chromatin fiber, using a set of coordinate transformations described by Wong et al.⁵⁹ The coordinate transformations result in a number of severe steric clashes. A 15 ps simulation of the fiber using the two-charge HCP-GB significantly reduces the steric clashes, as seen by the large reduction in the potential energy (Figure 13). To reduce run time for this simulation, the protocol described in section 2.3 was modified to reduce the heating and equilibration stages from 10 to 2 ps.

3.8. Mitigating the Effect of Violating Newton’s Third Law. Although the HCP uses the same all-atom force field as the reference GB computation, the HCP is a multiscale model in that different levels of approximations are used for the same set of atoms depending on their distance from the point of interest. The asymmetric interactions due to the multiscale approximations can violate Newton’s third law, resulting in a residual force on the system.^{58,70} This residual force can produce an artificial center of mass motion and an overall rotation of the structure. A net residual force within a closed system causes the system as a whole to accelerate, even though there is no external force, resulting in the nonconservation of energy. Table 5 shows the net

force and torque due to the violation of Newton's third law for the HCP-GB method, on the set of test structures considered here, along with estimated center of mass displacement, rotation, and kinetic energy after 10 steps of a typical molecular dynamics simulation (10 fs). To estimate the kinetic energies, we treat the structures as rigid bodies and assume that the principal axis of rotation passes through the center of mass.

For the test structures considered here, the 3 million atom chromatin fiber represents the worst case with a linear displacement of 1×10^{-6} Å, rotation of 1×10^{-8} radians and kinetic energy of 0.09 kT after 10 steps of MD. These spurious motions are small compared to the stochastic collisions used in constant temperature simulations, which are on the order of 1 kT, and may not materially affect the dynamics of the simulation if a strong enough coupling to a thermal bath is used. However, implicit solvent MD simulations often use minimal or no viscosity to increase the sampling of conformation space. These regimes can result in a large center of mass drift, which can be inconvenient when visualizing or analyzing the trajectory. For example, for a 1 ns simulation of thioredoxin

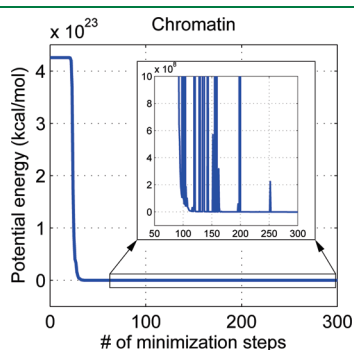


Figure 13. HCP-GB simulation of a 348 000 atom chromatin fiber. Potential energy drops rapidly as steric clashes are resolved. Connecting lines are shown to guide the eye.

using the two-charge HCP-GB method and Langevin dynamics with a collision frequency of 50 ps^{-1} , the center of mass drift is 2.24 Å, similar to the 1.81 Å drift for the reference GB simulation. Whereas with a collision frequency of 1 ps^{-1} , the center of mass drift in the HCP-GB simulation is 30.96 Å, and as much as 183.54 Å with a collision frequency of 0.1 ps^{-1} . A commonly used approach for removing center of mass drift and rotation during the course of molecular dynamics is to employ a velocity correction algorithm, e. g., the NSCM option in Amber⁶² which specifies the frequency at which center-of-mass motion is removed. We have implemented the same option in NAB. The velocity correction approach does not however correct the source of the problem—the net residual force. Moreover, a velocity correction not only eliminates the artificial motion caused by the violation of Newton's third law but also affects the random motion due to Langevin dynamics which may not be desirable in some situations. Therefore, we considered applying a force correction aimed at mitigating the effects of the third law violation. Appendix A.2 describes the two force correction approaches we considered—a molecular level and a component level force correction. By neutralizing the net residual forces, the force correction eliminates the systematic drift in the center of mass position caused by the violation of Newton's third law. For example, for the 1 ns simulation of thioredoxin using the two-charge HCP-GB method with a Langevin dynamics collision frequency of 0.1 ps^{-1} , the molecular level force correction reduces the drift from 183.54 Å to 29.49 Å, which is similar to the 29.47 Å drift for the reference GB simulation, which, of course, does not violate the third law within numerical precision of the integrator. The force correction however has several shortcomings compared to the velocity correction often used by existing MD algorithms. It causes an increase in the force error as described in Appendix A.2, while velocity correction does not affect the forces. Unlike velocity correction, the force correction only eliminates drift, not rotation. And, the force correction must be applied at every step of the simulation since it eliminates the center of

Table 5. Center-of-Mass Motion Due to Violation of Newton's Third Law for HCP-GB

structure PDB ID	residual		after 10 iterations of dynamics		
	force (kcal/mol/Å)	torque (kcal/mol)	displacement (Å)	rotation (radians)	kinetic energy (kT)
	1-q HCP-GB				
2BNA	0.003	0.38	1×10^{-8}	1×10^{-8}	6×10^{-9}
1BDD	0.22	0.73	9×10^{-7}	3×10^{-8}	4×10^{-7}
1UBQ	0.24	0.64	6×10^{-7}	1×10^{-8}	2×10^{-7}
2TRX	0.97	3.81	2×10^{-6}	4×10^{-8}	3×10^{-6}
1KXS	1.36	51.40	1×10^{-7}	3×10^{-9}	6×10^{-7}
microtubule	179.24	2603.63	3×10^{-6}	5×10^{-9}	0.001
1A6C	2×10^{-11}	6×10^{-10}	1×10^{-19}	2×10^{-22}	4×10^{-30}
chromatin	6804.00	1198829.76	6×10^{-6}	1×10^{-8}	0.09
	2-q HCP-GB				
2BNA	0.015	0.07	5×10^{-8}	2×10^{-9}	2×10^{-9}
1BDD	0.18	0.44	7×10^{-7}	2×10^{-8}	2×10^{-7}
1UBQ	0.23	0.58	6×10^{-7}	1×10^{-8}	2×10^{-7}
2TRX	0.98	4.64	2×10^{-6}	5×10^{-8}	3×10^{-6}
1KXS	2.94	57.98	3×10^{-7}	4×10^{-9}	2×10^{-6}
microtubule	59.76	598.27	1×10^{-6}	1×10^{-9}	0.0001
1A6C	2×10^{-11}	4×10^{-10}	1×10^{-19}	1×10^{-22}	3×10^{-30}
chromatin	3483.09	836570.85	3×10^{-6}	7×10^{-9}	0.03

mass acceleration (change in velocity), whereas a velocity correction can be applied less frequently since it eliminates center of mass velocity itself. For these reasons, a velocity correction, its drawbacks notwithstanding, may be preferable to the force correction for eliminating the drift and rotation caused by the violation of Newton's third law. Very preliminary testing suggests that the velocity correction may improve the stability of HCPGB simulations for small structures. For example, the RMS deviation for the one-charge HCP-GB simulation of 1BDD (Figure 8a), is in closer agreement with the reference GB simulation when the velocity correction is used (results not shown).

3.9. Mitigating the Effect of the Discontinuity at Threshold Boundaries. During the course of molecular dynamics, atoms may cross threshold boundaries that determine the level of approximation used in the computation of potentials and forces. These discontinuous changes in the level of approximation result in changes in potential that are inconsistent with the forces acting on individual atoms; i.e., the force is not equal to the derivative of the potential with respect to distance and can result in the nonconservation of energy and instability in the simulation.⁷¹ For example, 1000 steps of constant energy simulation for thioredoxin, without Langevin dynamics or surface-area dependent energy (Figure 14), show that energy is not conserved for the cutoff-GB and to a lesser extent the HCP-GB methods. However, the nonconservation of energy is much larger in the case of the cutoff-GB method. For the cutoff-GB method, the energy contribution of an atom abruptly drops to zero when the atom moves beyond the cutoff boundary, resulting in a larger change in energy compared

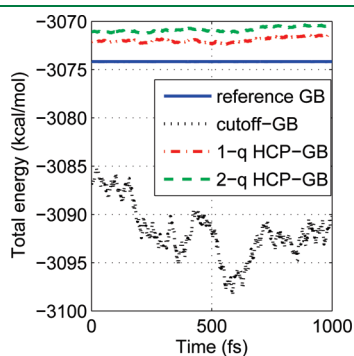


Figure 14. Total energy for 1000 steps of constant energy MD simulation for thioredoxin. The figure shows that energy is not conserved for the cutoff-GB and to a lesser extent the HCP-GB methods. Connecting lines are shown to guide the eye.

to the HCP-GB method, where the energy contribution of the atom is replaced by an approximation when the component containing the atom moves beyond the threshold boundary.

The discontinuity at threshold and cutoff boundaries can be eliminated by the use of a smoothing function.^{72,73} The smoothing function eliminates the discontinuity by gradually switching from one level of approximation to another over a short switching distance. Delle Site⁷⁰ has however shown that the smoothing function cannot in general restore the conservation of energy for multiscale methods. To eliminate the discontinuity at threshold boundaries, we adapted the smoothing function described by Loncharich and Brooks.¹⁵ The smoothing function is used to calculate the force $f(r)$ at a distance r from the point of interest inside the switching region $h < r < h + s$, where h is the threshold distance and s is the switching distance, as follows:

$$f(r) = S(r)f_{h+s} + (1 - S(r))f_h \quad (h \leq r \leq h + s) \quad (26)$$

$$S(r) = \frac{(h^2 - r^2)^2(h^2 + 2r^2 - 3(h+s)^2)}{(h^2 - (h+s)^2)^3} \quad (27)$$

where f_h and f_{h+s} are the forces due to a component computed by the HCP-GB at h and $h + s$, respectively.

The nonconservation of energy can be measured as the standard deviation in total energy, and the discontinuity in computed energy as the standard deviation in the change in total energy between consecutive steps in the MD simulation. Table 6 summarizes these metrics for the HCP-GB method with and without smoothing, the cutoff-GB method, and the reference GB computation. The table shows that, although the HCP-GB method does not conserve energy, it represents a significant improvement over the cutoff-GB method in that respect. On the other hand, although smoothing does improve energy conservation, it is comparable to extending the level 1 threshold distance to the end of the smoothing region, 18 Å in this case, as measured by the standard deviation in total energy (Table 6). The HCP-GB with smoothing does show less discontinuity than the HCP-GB with the extended threshold distance; however, the difference may not be sufficient to justify the higher computational cost of smoothing compared to simply extending the threshold distance. Note that the smoothing function can also improve the accuracy of the spherical cutoff method and has been studied previously.^{15,72,73} The results shown in the preceding subsections do not include the smoothing function for either the cutoff-GB or the HCP-GB methods.

We stress that in the case of multiscale approximations based on pairwise potentials, such as the HCP, exact energy

Table 6. Effect of the Use of the Smoothing Function on Energy Conservation and Discontinuity for 1000 steps of MD for Thioredoxin^a

standard deviation (kcal/mol)	reference GB	cutoff GB	1-q HCP-GB	2-q HCP-GB
total energy (no smoothing)	0.0012	2.8213	0.2483	0.2476
(HCP-GB with smoothing)			0.0803	0.0992
(HCP-GB with $h_1 = 18$ Å)			0.0942	0.1013
Δ total energy (no smoothing)	0.0007	0.3316	0.0228	0.0203
(HCP-GB with smoothing)			0.0033	0.0025
(HCP-GB with $h_1 = 18$ Å)			0.0062	0.0051

^a Degree of energy conservation is measured as the standard deviation in total energy and degree of discontinuity as the standard deviation in change in total energy between consecutive steps of MD simulation. The default level 1 threshold distance, h_1 , for the HCP-GB method is 15 Å. For comparison, we include the reference GB computation, the cutoff-GB method with a 15 Å cutoff distance, and the HCP-GB method with a 18 Å level 1 threshold distance. The small fluctuation in total energy for the reference GB computation is due to the finite integrator time step.

conservation cannot be achieved due to the violation of Newton's third law.

4. CONCLUSION

Implicit solvent models are routinely used where it is important to sample a large conformation space, such as for protein folding, replica exchange, and docking simulations. However, the implicit solvent model employed most extensively in molecular dynamics—the generalized Born (GB) model—scales poorly as $\sim n^2$, where n is the number of solute atoms, limiting their usefulness for long time-scale simulations or the simulation of large structures. We have presented here an $\sim n \log n$ implementation of the implicit solvent GB model based on the hierarchical charge partitioning (HCP) approximation previously developed by us. The HCP method uses the natural organization of biomolecular structures to partition the structures into multiple hierarchical levels of components such as atoms, groups (residues), subunits (chains), and complexes. The charge distribution for each of these components other than the atoms are approximated by a small (one or two) number of charges. For the computation of electrostatic interactions with distant components, the HCP uses the approximate charges while using the atomic charges for nearby components. The greater the distance from the point of interest, the larger (higher level) is the component used in the approximation. We have previously described a top-down algorithm for HCP that scales as $\sim n \log n$ for biomolecular structures.

This study extends the HCP approximation to the GB model (HCP-GB) such that both the computation of pairwise interactions and the effective Born radii scale as $\sim n \log n$.

The HCP-GB method is implemented in the open source molecular dynamics software, NAB, in AmberTools v1.3. The HCP implementation in NAB is scheduled to be released with AmberTools v1.5, for general use. The accuracy, speed, and stability of the method were then evaluated on a set of representative biomolecular structures ranging in size from 632 to ~ 3 million atoms. The performance of the HCP-GB method was compared to the spherical cutoff method with GB (cutoff-GB) where all computations, including the computation of the effective Born radii, ignore all atoms beyond a specified cutoff distance. Our results show that the HCP-GB method is more accurate, as measured by the relative RMS error in electrostatic force, than the cutoff-GB method for the structures tested. Depending on the size of the structure, the HCP-GB method was also 1.1 to 390 times faster than the reference GB computation. An analysis of 50 ns simulations of four structures—B-DNA, immunoglobulin binding domain, ubiquitin, and thioredoxin—shows that the results for the HCP-GB simulation are in reasonable agreement with the reference GB simulation without cutoffs. For the very small (726 atom) immunoglobulin binding domain protein (1BDD), the one-charge HCP-GB method exhibited RMS deviations from the crystal structure similar to the cutoff-GB and larger than the reference-GB and two-charge HCP-GB simulations. Therefore, we do not recommend the use of one-charge HCP-GB for the simulation of such small structures. However, very preliminary testing suggests that the velocity correction, described below, may improve the stability of HCPGB simulations for small structures.

There is also an important qualitative difference between the HCP-GB method and the cutoff-GB method. The cutoff-GB ignores charges beyond the cutoff distance while the HCP-GB

method approximates the influence of distance charges. Our testing suggests that this difference can have a significant impact on details of the dynamics. For example, for the 50 ns simulations of four structures, the residue flexibility and χ_1 and χ_2 angles for the cutoff-GB simulations show larger deviations from the reference GB simulation than the HCP-GB simulations. Similarly, a 10 ns simulation of a 30 base-pair DNA strand showed that the flexibility of the molecule, as measured by end-to-end distance, using the HCP-GB method was similar to that of the reference GB simulation, whereas the cutoff-GB method results showed a more flexible molecule. And a series of simulations of the nucleosome core particle showed that with the reference GB and HCP-GB methods all of the positively charged tails of the histone chains collapsed onto the negatively charged DNA, whereas two of the histone tails failed to do so with the cutoff-GB method.

Due to its multiscale nature, the HCP-GB method can violate Newton's third law, resulting in a residual center of mass force and torque. For the structures tested here, the effect of the residual force and torque is much smaller than the "noise" due to stochastic collisions used in constant temperature simulations with strong coupling to a thermal bath. However, when a weak coupling is used to increase the sampling of conformational space, the residual force and torque may cause the structure to drift and rotate, making it inconvenient for visualization and analysis. For simulations with weak coupling to a thermal bath, the center of mass motion and rotation can be eliminated by using a velocity correction. The multiscale nature of HCP-GB can also result in discontinuities at threshold boundaries, which can cause energy not to be conserved. The discontinuity and the resultant nonconservation of energy for the HCP-GB method is however much smaller than that of the cutoff-GB method. Smoothing functions can be used to reduce the discontinuities and the nonconservation of energy. However, we found that increasing the threshold distance may be a more effective way of achieving the same result.

To demonstrate a practical application of the HCP-GB method, we used it to refine a 348 000 atom chromatin fiber. The 15 ns all-atom simulation successfully resolved numerous severe steric clashes, significantly improving the quality of the starting structure.

In conclusion, the $\sim n \log n$ HCP-GB method is always faster than the $\sim n^2$ reference GB computation without additional approximations. Although the speed of the HCP-GB method is comparable to using a spherical cutoff for GB computations, which also scales as $\sim n \log n$, the HCP-GB method on average more closely reproduces key characteristics of the dynamics of the reference GB simulations. Our testing suggests that this may be because the HCP-GB method approximates the influence of distant charges, unlike the cutoff-GB method, which completely ignores them. In general, our findings suggest that compared to the cutoff-GB, the HCP-GB method may always be the preferable approach for speeding up pairwise GB computations for molecular dynamics. Where speed is critical, one can consider using the one-charge HCP-GB instead of the two-charge HCP-GB or reducing threshold distances from the recommended conservative threshold distances.

This study was intended to be a proof-of-concept of a novel method, and a number of potential improvements and optimizations remain to be studied, in particular, further optimization of the placement of approximate charges, comparison of alternate approximations for component effective Born radii, choice of parameters, comparison of velocity vs force correction, and the treatment of very distant components. Most importantly, more extensive testing is required to further define the applicability and limitations of the proposed $\sim n \log n$ GB method.

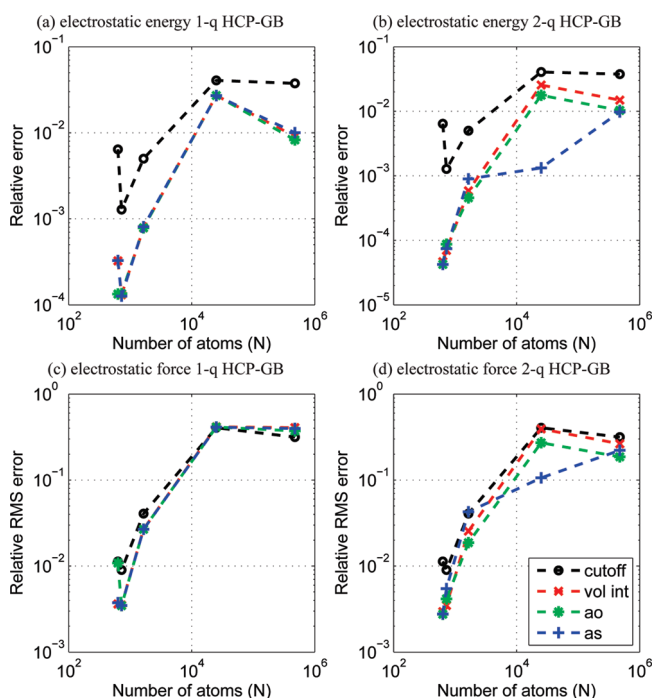


Figure 15. Comparison of four alternative methods for computing component effective Born radii showing relative error in electrostatic energy for (a) one-charge HCP-GB, (b) two-charge HCP-GB, and relative RMS error in electrostatic force for (c) one-charge HCP-GB and (d) two-charge HCP-GB. The four alternative methods are the spherical cutoff method, the volume integral (vol int) based on eq 10, Anandakrishnan–Onufriev (ao) defined by eq 9, and Archontis–Simonson (as) defined by eq 28. Cutoff distance = 15 Å. HCP threshold distance $h_1 = 15$ Å, $h_2 = 80$ Å, and $h_3 = 175$ Å. Connecting lines are shown to guide the eye.

A. APPENDIX

A.1. Component Effective Born Radii. The component effective Born radii are used to approximate the contribution of distant components to the solvation energy, as described in the Methods section. We examined three different alternatives for this approximation: the approximation defined by eq 9, the approximation defined by Archontis and Simonson,³⁹ and the volume integral approximation based on eq 10. Archontis and Simonson approximate the Born radius B_c for a component c as

$$\frac{1}{B_c} \approx \frac{1}{\sum_{i \in c} q_i^2} \sum_{i \in c} \frac{q_i^2}{B_i} \quad (28)$$

where q_i and B_i are the charges and effective Born radii, respectively, for the atoms i belonging to component c . On average, we found that the first alternative (eq 9) was most accurate, as shown in Figure 15.

A.2. Force Corrections for Neutralizing Net Residual Force. The violation of Newton's third law by the HCP-GB method, as described in section 3.8, results in a net residual force. We considered two approaches for neutralizing the net residual force—a molecular level and a component level force correction.

The molecular level approach applies a mass weighted force correction to each atom to neutralize the total residual force on the whole structure. This force correction is computed as

$$\mathbf{f}_i^{\text{corr}} = \mathbf{f}^{\text{res}} m_i / M \quad (29)$$

where $\mathbf{f}_i^{\text{corr}}$ is the force correction subtracted from the force on atom i , \mathbf{f}^{res} is the total residual force, m_i is the mass of atom i , and M is the total mass of the structure.

The component level approach is more complex. It aims to eliminate not only the net residual force but also the net residual force for each component, where the residual force for a component is the difference between the total force on a component due to all other atoms and the total force on all other atoms due to the component. In other words, it aims to restore Newton's third law at the component level. For the first level of HCP approximation, the force correction $\mathbf{f}_i^{\text{corr}}$ for an atom i belonging to component c is calculated as

$$\mathbf{f}_i^{\text{corr}} = \mathbf{f}_c^{\text{diff}} / n_c \quad (i \in c) \quad (30)$$

$$\mathbf{f}_c^{\text{diff}} = \left[\sum_{j \notin c} \sum_{i \in c} \mathbf{f}_{ji} + \sum_{k \neq c} \sum_{i \in c} \mathbf{f}_{ki} \right] + \left[\sum_{i \in c} \sum_{j \notin c} \mathbf{f}_{ij} + \sum_{j \notin c} \mathbf{f}_{cj} \right] \quad (31)$$

where $\mathbf{f}_c^{\text{diff}}$ is the difference between the total force on a component due to all other atoms and the total force on all other atoms due to the component. The first term on the right hand side of eq 31 is the force \mathbf{f}_{ji} on the atoms i belonging to component c due to other atoms j not belonging to c , computed at the atomic level (level 0). The second term is the force \mathbf{f}_{ki} on atoms i belonging to c due to other components k treated at the component level (level 1). The third term is the force \mathbf{f}_{ij} due to c on atoms j not belonging to c where the atoms i within c are treated at the atomic level (level 0). And the last term is the force \mathbf{f}_{cj} due to c on atoms j not belonging to c where c is treated at the component level (level 1). This force correction is generalized for higher levels of HCP by including terms in eq 31 for higher level HCP components. Since the individual terms on the right hand side of eq 31 are already being computed, the incremental cost of both force correction approaches scales as $\sim n$, which is $< n \log n$.

Although the above force corrections neutralize the net force due to the violation of Newton's third law, they also cause an increase in the force error. To see why, consider the case where the force error compared to the reference GB computation is approximately zero. In this case, any net force correction will result in an increase in the force error. In general, when the force error is less than half the net force correction, the net force correction will cause an increase in force error. Thus, on average, where the force error is randomly distributed, the net force correction will result in an increase in force error.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: alexey@cs.vt.edu.

■ ADDITIONAL NOTE

^a NAB, and the production Amber software, do not explicitly include a “no-cutoff” option; instead, the no-cutoff computation is performed by using the cutoff method with the cutoff distance set to a value greater than the structure size (large-cutoff), e.g., 999. This approach requires a large amount of memory for the neighbor list used by the cutoff method, even though the list is unnecessary in this case, since it always contains all of the atoms. Due to the large memory requirement, structures larger than 200 000 atoms require a larger number of processors (>128)

than was readily available in the system used. Therefore, we implemented a no-cutoff option in NAB that does not use a neighbor list. Since a neighbor list does not need to be computed, the no-cutoff option is faster than the large-cutoff approach, and the speedup results reported here are somewhat lower than what would have been obtained using the large-cutoff approach available in NAB.

^bThe average speedup for the seven structures tested here was 82× for the cutoff-GB, 85× for the one-charge HCP-GB, and 36× for the two-charge HCP-GB methods.

REFERENCES

- (1) Dodson, G. G.; Lane, D. P.; Verma, C. S. *EMBO Rep.* **2008**, *9*, 144–150.
- (2) Freddolino, P. L.; Arkhipov, A. S.; Larson, S. B.; McPherson, A.; Schulten, K. *Structure* **2006**, *14*, 437–449.
- (3) Karplus, M.; McCammon, J. A. *Nat. Struct. Biol.* **2002**, *9*, 646–652.
- (4) Karplus, M.; Kuriyan, J. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 6679–6685.
- (5) Wang, W.; Donini, O.; Reyes, C. M.; Kollman, P. A. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 211–243.
- (6) Venkateswarlu, D. *BMC Struct. Biol.* **2010**, *10*, 7.
- (7) Kumar, S.; Huang, C.; Zheng, G.; Bohm, E.; Bhatele, A.; Phillips, J. C.; Yu, H.; Kalé, L. V. *IBM J. Res. Dev.* **2008**, *52*, 177–187.
- (8) Ruscio, J. Z.; Kumar, D.; Shukla, M.; Prisant, M. G.; Murali, T. M.; Onufriev, A. V. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 9204–9209.
- (9) Shaw, D. E.; et al. *Comm. ACM* **2008**, *51*, 91–97.
- (10) Zhou, R.; Eleftheriou, M.; Hon, C. C.; Germain, R. S.; Royyuru, A. K.; Berne, B. J. *IBM J. Res. Dev.* **2008**, *52*, 19.
- (11) Ejtehadi, M. R.; Avall, S. P.; Plotkin, S. S. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 15088–15093.
- (12) Beck, D. A. C.; Armen, R. S.; Daggett, V. *Biochemistry* **2005**, *44*, 609–616.
- (13) Ruvinsky, A. M.; Vakser, I. A. *Proteins* **2008**, *70*, 1498–1505.
- (14) Mark, P.; Nilsson, L. *J. Comput. Chem.* **2002**, *23*, 1211–1219.
- (15) Loncharich, R. J.; Brooks, B. R. *Proteins* **1989**, *6*, 32–45.
- (16) Schreiber, H.; Steinhäuser, O. *Chem. Phys.* **1992**, *168*, 75–89.
- (17) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
- (18) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577–8593.
- (19) Toukmaji, A. Y.; Board, J. A. *Comput. Phys. Commun.* **1996**, *95*, 73–92.
- (20) York, D.; Yang, W. *J. Chem. Phys.* **1994**, *101*, 3298–3300.
- (21) Carrier, J.; Greengard, L.; Rokhlin, V. *SIAM J. Sci. Stat. Comp.* **1988**, *9*, 669–686.
- (22) Cai, W.; Deng, S.; Jacobs, D. *J. Chem. Phys.* **2007**, *223*, 846–864.
- (23) Lambert, C. G.; Darden, T. A.; Board, J. A., Jr. *J. Chem. Phys.* **1996**, *126*, 274–285.
- (24) Bishop, T. C.; Skeel, R. D.; Schulten, K. *J. Comput. Chem.* **1997**, *18*, 1785–1791.
- (25) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127–6129.
- (26) Bashford, D.; Case, D. A. *Annu. Rev. Phys. Chem.* **2000**, *51*, 129–152.
- (27) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *Chem. Phys. Lett.* **1995**, *246*, 122–129.
- (28) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem.* **1996**, *100*, 19824–19836.
- (29) Ghosh, A.; Rapp, C. S.; Friesner, R. A. *J. Phys. Chem. B* **1998**, *102*, 10983–10990.
- (30) Lee, M. S.; Salisbury, J.; Brooks, C. L., III. *J. Chem. Phys.* **2002**, *116*, 10606–10614.
- (31) Onufriev, A.; Bashford, D.; Case, D. A. *Proteins* **2004**, *55*, 383–394.
- (32) Tsui, V.; Case, D. *J. Am. Chem. Soc.* **2000**, *122*, 2489–2498.
- (33) Cramer, C.; Truhlar, D. *Chem. Rev.* **1999**, *99*, 2161–2200.
- (34) David, L.; Luo, R.; Gilson, M. K. *J. Comput. Chem.* **2000**, *21*, 295–309.
- (35) Im, W.; Lee, M. S.; Brooks, C. L. *J. Comput. Chem.* **2003**, *24*, 1691–1702.
- (36) Schaefer, M.; Karplus, M. *J. Phys. Chem.* **1996**, *100*, 1578–1599.
- (37) Calimet, N.; Schaefer, M.; Simonson, T. *Proteins* **2001**, *45*, 144–158.
- (38) Feig, M.; Im, W.; Brooks, C. L. *J. Chem. Phys.* **2004**, *120*, 903–911.
- (39) Archontis, G.; Simonson, T. *J. Phys. Chem. B* **2005**, *109*, 22667–22673.
- (40) Feig, M.; Brooks, C. L. *Curr. Opin. Struct. Biol.* **2004**, *14*, 217–224.
- (41) Nymeyer, H.; Garcia, A. E. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 13934–13939.
- (42) Scarsi, M.; Apostolakis, J.; Cafilisch, A. *J. Phys. Chem. A* **1997**, *101*, 8098–8106.
- (43) Dominy, B. N.; Brooks, C. L. *J. Phys. Chem. B* **1999**, *103*, 3765–3773.
- (44) Gallicchio, E.; Levy, R. M. *J. Comput. Chem.* **2004**, *25*, 479–499.
- (45) Grant, J. A.; Pickup, B. T.; Sykes, M. J.; Kitchen, C. A.; Nicholls, A. *Phys. Chem. Chem. Phys.* **2007**, *9*, 4913–4922.
- (46) Haberthür, U.; Cafilisch, A. *J. Comput. Chem.* **2007**, *29*, 701–715.
- (47) Spassov, V. Z.; Yan, L.; Szalma, S. *J. Phys. Chem. B* **2002**, *106*, 8726–8738.
- (48) Ulmschneider, M. B.; Ulmschneider, J. P.; Sansom, M. S.; Di Nola, A. *Biophys. J.* **2007**, *92*, 2338–2349.
- (49) Tanizaki, S.; Feig, M. *J. Chem. Phys.* **2005**, *122*, 124706.
- (50) Zhang, L. Y.; Gallicchio, E.; Friesner, R. A.; Levy, R. M. *J. Comput. Chem.* **2001**, *22*, 591–607.
- (51) Shen My, M. Y.; Freed, K. F. *Biophys. J.* **2002**, *82*, 1791–1808.
- (52) Rhee, Y. M.; Pande, V. S. *J. Phys. Chem. B* **2008**, *112*, 6221–6227.
- (53) Onufriev, A. V. In *Modeling Solvent Environments*; Feig, M., Ed.; Wiley-VCH: Weinheim, Germany, 2010; Chapter 6. Continuum Electrostatic Solvent Modeling with the Generalized Born Model, pp 127–165.
- (54) Zagrovic, B.; Pande, V. *J. Comput. Chem.* **2003**, *24*, 1432–1436.
- (55) Pitera, J. W.; Swope, W. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 7587–7592.
- (56) Guimarães, C. R.; Cardozo, M. *J. Chem. Inf. Model.* **2008**, *48*, 958–970.
- (57) Tsui, V.; Case, D. A. *Biopolymers* **2001**, *56*, 275–291.
- (58) Anandakrishnan, R.; Onufriev, A. V. *J. Comput. Chem.* **2010**, *31*, 691–706.
- (59) Wong, H.; Victor, J.-M.; Mozziconacci, J. *PLOS One* **2007**, *4*, 36, e877.
- (60) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33–38.
- (61) Mongan, J.; Simmerling, C.; McCammon, J. A.; Case, D. A.; Onufriev, A. *J. Chem. Theory Comput.* **2007**, *3*, 156–169.
- (62) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. *J. Comput. Chem.* **2005**, *26*, 1668–1688.
- (63) Macke, T.; Case, D. In *Molecular Modeling of Nucleic Acids*; Leontes, N. B., SantaLucia, J., Eds.; American Chemical Society: Washington, DC, 1998; Chapter: Modeling Unusual Nucleic Acid Structures, pp 379–393.
- (64) Gordon, J.; Myers, J.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A. *Nucleic Acids Res.* **2005**, *33*, 68–71.
- (65) Onufriev, A.; Bashford, D.; Case, D. *J. Phys. Chem. B* **2000**, *104*, 3712–3720.
- (66) Wang, H.-W. W.; Nogales, E. *Nature* **2005**, *435*, 911–915.

- (67) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65*, 712–725.
- (68) Ruscio, J. Z.; Onufriev, A. *Biophys. J.* **2006**, *91*, 4121–4132.
- (69) Bertin, A.; Leforestier, A.; Durand, D.; Livolant, F. *Biochemistry* **2004**, *43*, 4773–4780.
- (70) Delle Site, L. *Phys. Rev. E* **2007**, *76*, 047701.
- (71) Ensing, B.; Nielsen, S. O.; Moore, P. B.; Klein, M. L.; Parrinello, M. *J. Chem. Theory Comput.* **2007**, *3*, 1100–1105.
- (72) Leach, A. In *Molecular Modeling: Principles and Applications*; Prentice Hall: New York, 2001; Chapter 6. Computer Simulation Methods, pp 303–352.
- (73) Schlick, T. *Molecular Modeling and Simulation, an Interdisciplinary Guide*; Springer-Verlag: New York, 2002.

The Influence of Cholesterol on the Properties and Permeability of Hypericin Derivatives in Lipid Membranes

Emma S. E. Eriksson and Leif A. Eriksson*

School of Chemistry, National University of Ireland—Galway, Galway, Ireland

S Supporting Information

ABSTRACT: The promising photosensitizing properties of hypericin, a natural quinone substituted with hydroxyl and alkyl groups, have led to the proposal that it can be utilized in photodynamic therapy. Neither the detailed mechanism behind the powerful action of hypericin, arising as a result of light excitation, nor the intracellular localization and transportation of the molecule is yet fully understood. The behavior of hypericin derivatives in a pure dipalmitoylphosphatidylcholine (DPPC) lipid membrane has recently been studied theoretically by means of molecular dynamics simulations. Natural membranes however contain many important constituents—cholesterol being one of the most essential—that influence the function and structure of the membrane, and thereby also the behavior of drug molecules therein. In the present study, we investigated hypericin and its brominated derivatives in membranes containing 9 and 25 mol % cholesterol. The results show that the presence of cholesterol in the membrane affects the permeability of the hypericin molecules and does so differently for the various molecules in the two membranes. Hypericin containing one bromine was found to exhibit the lowest free energy profile for the transport process into the lipids, and also the highest permeability coefficients, indicating that this molecule displays the fastest and easiest diffusion in the membranes. All three molecules were found to accumulate most preferably close to the polar headgroup region in both membranes.

1. INTRODUCTION

1.1. Cholesterol in Lipid Membranes. Cholesterol (Figure 1) is an important compound in nature, possessing many essential properties, not only as precursor to several vitamins and hormones but also as an important constituent in biological membranes, besides phospholipids and glycolipids, in which it for example increases mechanical strength, regulates phase behavior, and reduces the passive permeability of water and other small molecules. The cholesterol molecule is made up of three groups that have all proven essential for their effect on membranes: the fused rigid steroid rings, the hydroxyl group attached to one of the rings, and the short flexible hydrocarbon chain.¹ Although intracellular synthesis of cholesterol takes place in the endoplasmic reticulum (ER) and external cholesterol is transported to the lysosomes where it is hydrolyzed, the majority of the cellular cholesterol is found in the plasma membrane,² and an equilibrium process in which cholesterol is being transported between the cell membrane and the cytosol has been proposed.³

Cholesterol is found in a wide range of concentrations in various animal membranes, normally around 20–30 mol %, but plasma membranes of some cells contain up to 50 mol %.⁴ In membranes in which the cholesterol concentration is high (>25 mol %), an additional phase exists besides the solid-ordered and liquid-disordered ones: the liquid-ordered phase.^{5,6} In this phase, the lipids are translationally disordered and conformationally ordered, i.e., a combination of the two other phases in which the lipids are either completely ordered or disordered. In the liquid-ordered state, both ordering^{7,8} and condensing^{9,10} effects caused by cholesterol are observed. Cholesterol has also been extensively studied in the controversial field of lipid rafts and its presence in those. Lipid rafts are dynamic liquid-ordered

domains made up by cholesterol, sphingolipids, and proteins important in, for example, signaling.¹¹

1.2. Computational Studies of Cholesterol-Containing Membranes. Basic initial computational studies of cholesterol/lipid membranes were performed more than two decades ago^{12–14} and have been followed by numerous more extended and detailed studies carried out by Monte Carlo (MC) and molecular dynamics (MD) simulation techniques in recent years. The improvement of computers and algorithms has enabled progression to study larger systems, including more than 1000 lipids, and the use of longer simulation time scales.^{15,16} Many properties of cholesterol-containing membranes as seen experimentally have been reproduced theoretically, thereby also enabling the evaluation of detailed properties that are difficult to observe experimentally.

It has been clearly shown, both experimentally and theoretically, that cholesterol has a crucial effect on the properties of the membrane and that the cholesterol concentration plays an important role. A wide range of concentrations of cholesterol has been included in the computational studies to cover for the occurrence in natural membranes. Cholesterol has an ordering and condensing effect, two features that are closely related and that have been observed in numerous computational studies at varying cholesterol concentrations.^{15–22} The ordering and condensing effect results in a decreased membrane surface area and thereby a reduced area per lipid. Chui et al. performed simulations of DPPC bilayers with cholesterol concentrations ranging from 4 to 50 mol % and found a linear relationship between

Received: September 15, 2010

Published: January 28, 2011

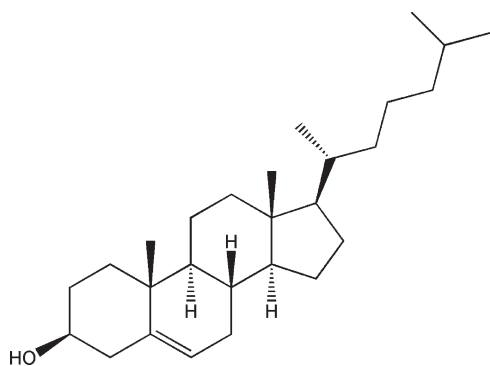


Figure 1. Chemical structure of the cholesterol molecule.

increased cholesterol concentration and a decrease in area per lipid at cholesterol concentrations in the range of 12–50 mol %.²¹ However, Tu et al. showed that a cholesterol concentration of 12.5 mol % had no significant effect on the conformation and packing of the hydrocarbon chains.²³ Smondyrev and Berkowitz observed that the cholesterol molecules exhibit a larger tilt when the concentration is low, whereas at higher concentrations, in which the hydrocarbon tails are more ordered and extended, cholesterol displays a reduced tilt.¹⁷ A reduction in electron density in the center of the DPPC bilayer was also observed with higher cholesterol concentrations as well as an increase in hydrophobic thickness,¹⁷ supported by neutron-scattering experiments on DMPC bilayers.²⁴

Martinez-Seara et al. recently published a detailed study of the unique structural functionality of cholesterol in its ability to initiate the liquid-ordered phase.²⁵ They showed that, at concentrations below ~30 mol %, the cholesterol molecules avoid locations adjacent to each other but prefer locations separated by ~1 nm, i.e., located in the second coordination shell. It was also found that triangular connections between neighboring cholesterol molecules exist, as opposed to demethylated cholesterol (missing the two off-plane methyl groups; Dchol) for which linear connections exist. The two faces of cholesterol were also studied in detail, and it was shown that ordering and condensing was less pronounced in membranes with Dchol, indicating that the off-plane methyl groups were essential for the unique properties of cholesterol.²⁶

Overall, unsaturated lipids show weaker interaction with cholesterol than saturated ones, and condensation and ordering is consequently less evident.²⁷ Solvation of saturated lipid chains occurs most preferably with the smooth α face of cholesterol, the flat side with no substituents.²⁰ In lipids containing one saturated and one unsaturated chain, cholesterol is preferably solvated by the saturated acyl chain.²⁸ For unsaturated chains, however, the β face induces higher ordering.²⁷ Even lipid chains not neighboring any cholesterol in the system²⁰ or within a radius of at least a few nanometers²⁸ were found to display higher order than in a pure lipid bilayer. The position of the double bond in unsaturated lipids has been shown to significantly influence the interactions with cholesterol regarding condensing and ordering, with the double bond located in the middle of the lipid acyl chain resulting in the smallest effects, whereas when located at the end of the chain, it resulted in larger effects.^{29,30}

Niemelä et al. recently used three large membrane systems (1024 lipids in each) with varying unsaturated palmitoylphosphatidylcholine (POPC), palmitoylsphingomyelin, and cholesterol composition in order to study lipid rafts.¹⁶ Two of the

systems represented membranes with coexistent liquid-ordered and liquid-disordered domains as a result of high cholesterol content. They confirmed that the raft domain membranes most likely influence the function of membrane proteins, mainly due to significant differences in pressure profiles compared to the liquid-disordered single phase membrane.

Water molecules penetrate deeper into the bilayer interior in cholesterol-containing membranes and can therefore form hydrogen bonds with the hydroxyl group of cholesterol.¹⁹ The number of possible hydrogen bonds also depends on the depth at which cholesterol is located, and this in turn depends on the lipid composition in the bilayer. In a DPPC bilayer, the cholesterol molecules are located deeper into the lipid interior than in a DMPC bilayer, and the number of hydrogen bonds with water is consequently reduced in the DPPC membrane.^{18,26} At low cholesterol concentrations, the number of hydrogen bonds with water is lower than at higher concentrations, as a result of the cholesterol molecules being able to locate deeper inside the DPPC bilayer.¹⁷ Pure saturated lipid membranes are however not frequently found in nature; membranes are usually composed of mixtures of different lipids, unsaturated ones being more common than saturated, and other components. However, pure saturated lipid membranes are commonly utilized as computational models mainly due to a lack of experimental data for unsaturated lipids, which make parametrization a difficult issue. The fact that most lipids have at least one saturated chain also makes the choice of saturated lipids useful. It is however worth emphasizing that the computational studies discussed herein were performed using different programs, force fields, and lipid compositions. These differences can slightly influence the results.

1.3. Permeation of Molecules in Cholesterol-Containing Membranes. Cholesterol was shown to reduce the permeability of ions and small molecules such as Na^+ , K^+ , Cl^- , and glucose through lipid membranes.³¹ It was also discovered that the permeability of water is reduced in cholesterol-containing membranes and that this depends on the cholesterol concentration.³²

There are only a few studies performed on the unique behavior of drug molecules in lipid membranes containing cholesterol, as most of the theoretical and experimental focus has been directed into studying the effect of cholesterol itself on various lipid membranes. An experimental study performed on model membranes showed that cholesterol decreases the permeability of large drug molecules mainly due to the condensing effect of cholesterol on the membrane.³³ This condensation may in particular affect large and rigid drug molecules. Molecular interactions between the drug molecule and cholesterol can also delay the permeation through the membrane, such as has been shown for small nonsteroidal anti-inflammatory drugs.³⁴ Several studies have confirmed that cholesterol strongly influences the interactions between peptides (drugs as well as endogenous compounds) and the membrane.^{35–38}

Drug interactions with cholesterol are important also because of the fact that many drugs are transported in the body by liposomes and lipoproteins. Liposomes are commonly used in the drug delivery of both hydrophilic and hydrophobic drugs, as they contain an aqueous core surrounded by a circled lipid bilayer that can contain a significant amount of cholesterol. Hydrophobic drugs that are administered into the bloodstream can bind to lipoproteins, which are the natural cholesterol transporters in the body, and can, in those, interact with the cholesterol molecules.

With the use of computational methods making it possible to reproduce many of the features of cholesterol in membranes, the behavior of small molecules such as drugs in these can be studied

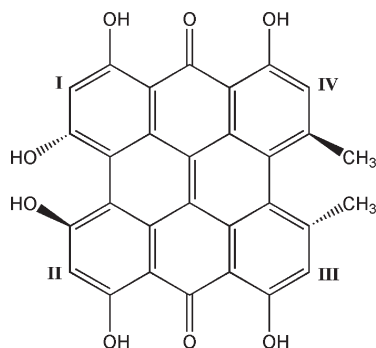


Figure 2. The hypericin molecule with numbers indicating where bromine substitution was modeled (I, Hy-Br; I–IV, Hy-4Br).

as well. Such studies can promote the design of drugs with desirable properties.

1.4. Hypericin. We have in a recent study investigated the behavior of the potent photodynamic drug hypericin (Hy; Figure 2) with no, one (Hy-Br; position I, Figure 2), and four bromines (Hy-4Br; positions I–IV, Figure 2) in a pure DPPC bilayer using molecular dynamics simulations.³⁹

Hypericin is a natural compound found in the *Hypericum* species, whose advantageous medical properties have been known for several thousand years. Besides the well-known antidepressive properties of *Hypericum* extracts,⁴⁰ it has been found that hypericin possesses antiviral^{41–46} and antitumor^{47–50} properties, as a result of the formation of reactive oxygen species (ROS) such as singlet oxygen upon light excitation, indicating that the compound could be used in photodynamic therapy (PDT). Hypericin has also been successful in the field of photophysical diagnosis of early stage tumors, as it accumulates specifically in tumor tissue from which the fluorescence of the drug can be detected.^{51–53}

The interest in making the hypericin molecule a more effective photodynamic drug has led to modifications such as halogenation. Bromination of hypericin increases the formation of ROS due to enhanced intersystem crossing from the first excited singlet state to the triplet state^{54,55} and has shown potential phototoxic activity against viruses.⁵⁶ Hypericin has been found in various cell compartments; however, the exact cellular target as well as the transport into the cell and its action of cell destruction is still to be elucidated. However, there seems to be a preference for hypericin to accumulate in lipid membranes due to its hydrophobic character, where it can initiate lipid peroxidation.^{57,58}

In a recently published study, it was proposed that cholesterol is the major reason why hypericin selectively accumulates in lipid membranes.⁵⁹ A high amount of hypericin was found to localize in raft domains rich in cholesterol rather than in less ordered regions rich in lipids. The emission spectrum suggests interactions between cholesterol and the π electrons of hypericin, resulting in effective packing of the two molecules due to the common planar structure. These results indicate that hypericin most likely can enter the cell membrane through diffusion; however, in the presence of lipoproteins such as low-density lipoproteins in the blood, these can be likely carriers of hypericin and can also assist in cell-entering.^{60–62} Lipoproteins, as the natural carriers of cholesterol in the body, are in that sense important in the aspect of possible cholesterol interactions with the drug.

In our previous molecular dynamics study of hypericin derivatives in pure DPPC lipid bilayers, we found a strong preference

for the hypericin molecules to accumulate in the bilayer, close to the polar headgroups and the interface between the lipids and water, a location that enables interactions between the hydroxyl groups of hypericin and water.³⁹ The largest gain in free energy for the transfer process of moving from water into the lipids as well as the fastest diffusion through the membrane was shown for Hy-Br, indicating that this molecule would have the highest probability to penetrate the membrane and reach the interior of the cell. Experimentally, it has been shown that halogenated drugs display larger permeability coefficients through lipid membranes.⁶³

In order to extend the previous study, we are herein including cholesterol in the membrane model. The study was performed on two cholesterol/lipid membrane systems, one containing a low concentration (9 mol %) cholesterol and one containing a higher concentration (25 mol %) cholesterol, and with the same hypericin derivatives as in the previous study (Hy, Hy-Br, and Hy-4Br; Figure 2). The present study was performed using a membrane model containing twice as many lipids and with twice as long production runs compared to the previous one.

2. COMPUTATIONAL METHODOLOGY

The molecular dynamics program GROMACS (version 4.0.4)⁶⁴ was used throughout the study, together with the united atom GROMACS force field. The membrane model used was a dipalmitoylphosphatidylcholine (DPPC) bilayer consisting of 128 lipids and 3655 water molecules that had been equilibrated⁶⁵ and simulated for 100 ns.^{66–68} This membrane model contains a larger number of lipids than the model used in our previous study; however the number of water molecules is more or less the same, leading to a thinner water phase in the present model.

The cholesterol structure was first geometry optimized using the Gaussian program⁶⁹ at the B3LYP/6-31G(d,p) level of theory. Using the coordinates obtained in the quantum optimizations, the topology of the cholesterol molecule was obtained using the PRODRG software⁷⁰ through its Web server [http://davap1.bioch.dundee.ac.uk/prodrp/], which generates topologies based on the GROMOS87 force field. Mulliken charges obtained from the optimization were assigned to the cholesterol molecule, and small charge groups with total charges close to zero were used. Thereafter, the structure was minimized by the steepest descent algorithm followed by a 100 ps equilibration simulation with a time step of 0.5 fs.

Two cholesterol/DPPC bilayers were constructed, one with 9 mol % (cholesterol/lipid ratio 12:116) and one with 25 mol % (cholesterol/lipid ratio 32:96) cholesterol, by randomly replacing lipid molecules with cholesterol in the membrane model. The same number of lipids was replaced in each monolayer. The membranes were minimized using steepest descent and equilibrated for 5 ns at 100 K. A simulated annealing simulation was then performed to increase the temperature from 100 to 500 K and then reduce it to 323 K, in steps of 50 K and 100 ps. The two membranes were equilibrated 20 ns at 323 K.

During the initial simulated annealing of the cholesterol/DPPC systems, the structures and conformations of the membranes were seriously disrupted and the increase in kinetic energy resulting from the heating generated the lipids in a highly flexible disordered state. This led to the possibility for the cholesterol molecules to move around in the membrane, both within and in-between the monolayers. In the membrane containing 32 cholesterol molecules (25 mol %), with initially 16 in each

monolayer, some cholesterol molecules moved in-between the monolayers, resulting in 14 cholesterol molecules in one monolayer and 18 in the other. In the system with 9 mol % cholesterol, the 12 cholesterol molecules stayed in their respective monolayers (six in each) during all simulations.

The geometries of the hypericin molecules were generated as outlined above for cholesterol and the topologies obtained using the PRODRG software. Mulliken atomic charges obtained from the geometry optimizations, as well as small charge groups, were assigned to the molecules. As bromine is not parametrized in the GROMOS87 force field, Lennard-Jones and ligand parameters for chlorine were used instead. For the DPPC phospholipids, a standard united atom force field was applied,⁷¹ and for water, we used the SPC model.⁷² Parameters used for the cholesterol and hypericin molecules are provided in the Supporting Information.

Six independent simulations were performed, one for each neutral hypericin derivative (Hy, Hy-Br, and Hy-4Br) in each of the two membranes. Two hypericin molecules of each derivative were inserted into the membrane model, one in the outer region of the water phase and one in the middle of the lipid phase. The systems were equilibrated for 200 ns, followed by 100 ns production runs in which the system trajectories were collected every 0.8 ps. During the equilibrations, the hypericin molecules moved into the lipids at different stages of the simulations. All simulations were performed using a time step of 2 fs. In a set of test simulations, 10 hypericin (Hy only) molecules were also studied in the two membranes. All 10 molecules were initially inserted into the water phase, and the behavior of the molecules was monitored.

In all simulations, the isothermal–isobaric ensemble (NPT) at $T = 323$ K and $p = 1$ bar was used. The temperature and pressure were held constant using a Nosé–Hoover thermostat^{73,74} with a coupling constant of 0.1 ps and a semi-isotropic Parrinello–Rahman barostat^{75,76} with a coupling constant of 1 ps. A particle mesh Ewald scheme^{77,78} was used to calculate the electrostatic interactions with a 10 Å cutoff for the real space. The same cutoff was used for van der Waals interactions (Lennard-Jones terms). Bond lengths were constrained using the LINCS algorithm.⁷⁹ Analyses were performed on the equilibration runs to verify equilibration convergence and on the production runs from which all reported data were obtained.

A potential of mean force formalism was used to calculate free energy profiles for hypericin molecules across the lipid bilayer (the direction of the z axis). The z component of the force, F_z , acting on the molecule at certain constrained distances between the molecule and the bilayer (DPPC) center of mass was collected at different positions along the z axis. The free energy for the transfer process between z_i and z_f is written as

$$\Delta G = G_{z_f} - G_{z_i} = - \int_{z_i}^{z_f} \langle F_z \rangle_z dz \quad (1)$$

where the bracket means an average over the forces collected at each constrained distance. To calculate the free energy profile for the translocation of each molecule, 34 constrained simulations were performed in which the hypericin molecule was located at positions differing by 0.1 nm along the z axis direction. The starting points for the simulations were sampled from the previous unconstrained simulations. To sample the points in the middle of the bilayer, where the molecules were never located during the unconstrained simulations, a weak force was used to push the molecule toward the bilayer middle, choosing the value

of the force so as to make the least perturbation possible on the bilayer system.

At each point in water (9 mol % cholesterol: 2.8–3.3 nm from the bilayer center; 25 mol % cholesterol: 3.1–3.3 nm from the bilayer center) equilibration was performed for at least 3 ns, followed by a production run of 4 ns. Inside the lipid bilayer (9 mol % cholesterol: 0–2.7 nm from the bilayer center; 25 mol % cholesterol: 0–3.0 nm from the bilayer center), an increase in the sampling was needed due to the slower motion of the molecules, and therefore, each point was equilibrated for at least 4.7 ns and a production run of 10 ns followed. For some of the systems with the hypericin molecules located within or close to the headgroup region, it was difficult to reach an equilibrated system due to competing interactions with the hypericins from the lipids and water. The thinner water phase in the present membrane model reduces the number of constrained simulations as the maximum distance to the bilayer center used herein was 3.3 nm, compared to 4.0 nm in our previous pure DPPC study. In the more condensed 25 mol % cholesterol membrane, the water phase is thicker than in the lower cholesterol membrane (as discussed below), and additional frames further out into the water phase could in principle have been included. However, in order to compare the results from the two membranes, the same number of frames was employed for both systems. The constrained simulations in the 25 mol % cholesterol membrane were run with the z box length fixed, whereas for the 10 mol % cholesterol membrane, this was not possible.

The force acting on the hypericin center of mass was collected at every time step during the production run. A SHAKE algorithm⁸⁰ was used to constrain the distance between the center of mass of the bilayer and the hypericin molecules (the molecules were constrained in the z direction but allowed to rotate). In the 25 mol % cholesterol membrane, in which the two monolayers contained different numbers of cholesterol molecules, the constrained hypericin molecule was located in the monolayer containing the most cholesterol (18).

The permeability is defined as the current density divided by the concentration gradient across the membrane. The procedure developed by Marrink and Berendsen⁸¹ was adopted to calculate the permeability coefficients, based on the fluctuation dissipation theorem and using the deviation of the instantaneous force, $F(z,t)$, from the average force acting on the molecule obtained during the constrained dynamics:

$$\Delta F(z,t) = F(z,t) - \langle F(z,t) \rangle \quad (2)$$

The local time-dependent friction coefficient, ξ , can be calculated from the following autocorrelation function:

$$\xi(z,t) = \langle \Delta F(z,t) \Delta F(z,0) \rangle / RT \quad (3)$$

where T is the absolute temperature and R is the gas constant. By integrating the friction coefficient, one can obtain the diffusion coefficient, D :

$$D(z) = RT / \xi(z) = (RT)^2 / \int_0^\infty \langle \Delta F(z,t) \Delta F(z,0) \rangle dt \quad (4)$$

This function was fitted to a double exponential using a nonlinear fitting procedure⁸¹ in order to integrate the autocorrelation of the force fluctuations:

$$C(t) = A_0 \exp(-t/\tau_0) + A_1 \exp(-t/\tau_1) \quad (5)$$

This translates to the molecules moving inside the lipid bilayer on two distinct time scales, corresponding to the two decay times τ_0 and τ_1 , one fast and one slow.

The permeability coefficient, P , can be calculated by integrating over the local resistances across the membrane, $R(z)$. $R(z)$ is obtained by dividing the exponential of the previously calculated free energies, $\Delta G(z)$, by the diffusion coefficients, $D(z)$:

$$1/P = \int R(z) dz = \int_{z_i}^{z_f} \frac{\exp(\Delta G(z)/kT)}{D(z)} dz \quad (6)$$

Interaction between cholesterol molecules at the two concentrations and for each different hypericin system was monitored through radial distribution functions (RDF), as were interaction between cholesterol and DPPC molecules. The data reveal that the concentration of cholesterol in the current systems is too low for the cholesterols to interact directly—instead they all form a first and even a second coordination shell of DPPC molecules, before the nearest cholesterol molecule appears (COM distances around 1 nm)—in agreement with the findings of Martinez-Seara et al.²⁵ No specific effects from the different hypericins were observed. The RDFs are provided in the Supporting Information.

3. RESULTS

The fact that cholesterol molecules were exchanged between the monolayers in the 25 mol % cholesterol bilayer during the simulated annealing is an interesting observation. Flip-flopping of cholesterol is an important natural process and has experimentally been demonstrated to occur with a half-time of <1 s.⁸² Coarse-grained MD simulations estimated the rate of possible flip-flops to be in the same range.⁸³ Umbrella sampling MD simulations and atomistic MD simulations have also been applied to study possible flip-flopping of cholesterol.⁸⁴ The most probable flip-flop path for cholesterol was calculated with a modified string method and was found to involve the cholesterol molecule first tilting and then moving to the bilayer center. The free energy barrier for the flip-flop in the DPPC bilayer was found to be higher than in diarachidonylphosphatidylcholine (DAPC), but lower than in POPC. However, spontaneous flip-flops of cholesterol have not yet been observed during MD simulation “production runs”, although for ketosterol this phenomenon has been detected in atomistic MD simulations.⁸⁵ Note that the flip-flops observed in the present study occurred during the simulated annealing, in which the temperature was significantly higher than under normal conditions, resulting in more movement in the system that gives rise to a possible exchange of molecules between the monolayers. No spontaneous flip-flops were observed during the simulations at constant temperature (323 K).

In Figure 3, we show snapshots from the production simulations of the two membranes with hypericin. In the membrane with 9 mol % cholesterol, the lipid and cholesterol molecules are free to move to a larger extent, and the molecules are tilted more than in the bilayer with higher cholesterol content. The hydrocarbon tails of both the lipids and the cholesterol molecules are in a disordered state. The condensing effect of cholesterol is clearly seen in the 25 mol % cholesterol membrane, and the hydrocarbon tails of the lipids and the cholesterol molecules are more ordered and more aligned to the bilayer normal. The increased thickness of the water phase, as a result of the decreased membrane surface area, is also clearly visible in the 25 mol %

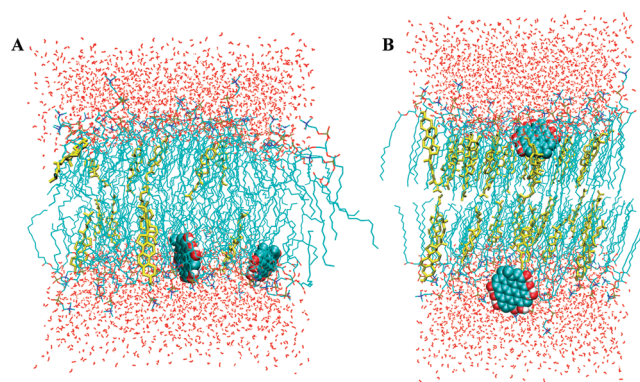


Figure 3. Snapshots from the simulations of two hypericin molecules in (A) 9 mol % and (B) 25 mol % cholesterol membranes. Cholesterol molecules are displayed in yellow.

cholesterol membrane. The condensing effect is due to the smaller size of a cholesterol molecule relative to DPPC, and consequently, with higher cholesterol content, the effect is more pronounced. The cholesterol molecule is shorter than a DPPC molecule and can therefore fit tightly in-between two DPPC molecules, often with its hydroxyl group in the level of the carbonyl group of the DPPC molecules. The condensing effect can also be displayed as the area per lipid, which decreases with an increasing concentration of cholesterol. The area per DPPC lipid was calculated by subtracting from the total area per monolayer the total area occupied by cholesterol molecules and dividing the difference by the number of lipids in one monolayer. We used the area per cholesterol molecule ($=32 \text{ \AA}^2$) obtained from X-ray diffraction experiments,⁸⁶ assuming that the cholesterol area remained constant, whereas the area per lipid was taken as an average over the last 10 ns of the equilibration of the membranes. For the 9 mol % cholesterol membrane, the area per lipid was estimated to be 59.7 \AA^2 and, in the 25 mol % cholesterol membrane, 48.1 \AA^2 in the monolayer with 18 cholesterol molecules and 50.8 \AA^2 in the monolayer with 14 cholesterol molecules. Smondryev and Berkowitz used the same approach to calculate the area per lipid in membranes with 11 and 50 mol % cholesterol, giving 58.3 \AA^2 and $44.7/46.5 \text{ \AA}^2$, respectively.¹⁷ In a pure DPPC membrane, the average area per lipid has been measured to be 61.6 \AA^2 .⁸⁷ Our generated data thus fit well into this range.

The hypericin molecules entered the membranes at different stages of the equilibration simulations. However, the same equilibration time (200 ns) was used for all six systems, and by that time all of the molecules were well inside the lipids. During the equilibrations, the molecules were free to enter any of the two monolayers, from the water or from the center of the bilayer. In the case of Hy in the 9 mol % cholesterol membrane and Hy-4Br in the 25 mol % cholesterol membrane, the two molecules ended up in the same monolayer. However, the molecules were not interacting, as the minimum distance between the two molecules during the production runs was overall larger than the van der Waals cutoff of 10 Å. Neither of the hypericin molecules crossed the center of the bilayer or returned into the water phase after entering the monolayer region.

In the 9 mol % cholesterol membrane, all three hypericin derivatives entered the lipids in regions where the local cholesterol concentration was low. The reason for this behavior might be that the local density is lower in the regions where there are no

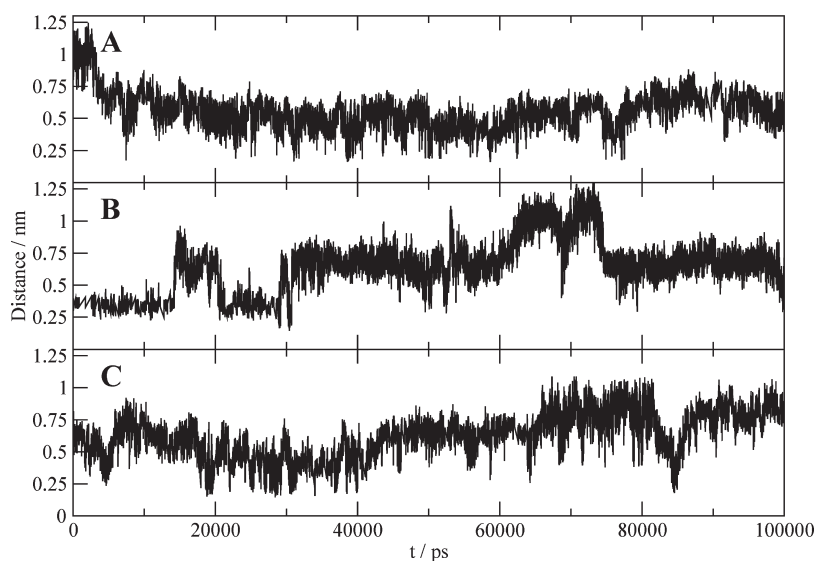


Figure 4. Minimum distance (nm) between any pair of atoms of a hypericin molecule (A, Hy; B, Hy-Br; C, Hy-4Br) and a cholesterol molecule during the production simulations in the 9 mol % cholesterol membrane.

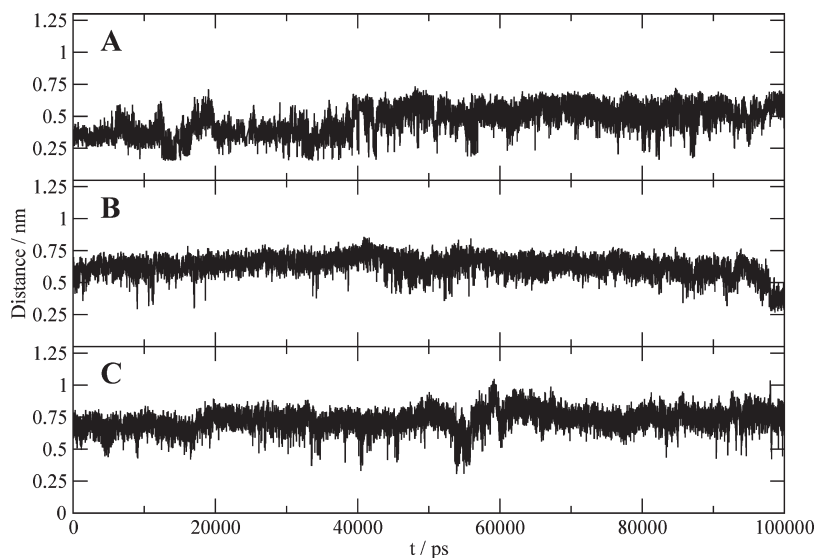


Figure 5. Minimum distance (nm) between any pair of atoms of a hypericin molecule (A, Hy; B, Hy-Br; C, Hy-4Br) and a cholesterol molecule during the production simulations in the 25 mol % cholesterol membrane.

cholesterol molecules, making it easier for the large hypericin molecules to incorporate. In the 25 mol % cholesterol membrane, this behavior was harder to determine with certainty due to the overall higher cholesterol concentration.

The minimum distance between any pair of atoms from either of the two hypericin molecules in each simulation and any of the cholesterol molecules was investigated during the production simulations (Figures 4 and 5) to explore if the hypericin molecules were able to interact with the cholesterol molecules. In the membrane with 9 mol % cholesterol, the minimum distance between Hy and cholesterol was in the range 0.25–1.25 nm, with an average distance of 0.56 nm. The distance between Hy-Br and Hy-4Br and cholesterol fluctuated more frequently than for Hy, varying between 0.2 and 1.3 nm for Hy-Br and between 0.2 and 1.1 nm for Hy-4Br, with an average distance of 0.63 nm for both molecules. In the 25 mol % cholesterol membrane, the distance

between the hypericin molecules and cholesterol displayed less variation, with distances in the range of 0.2–1.0 nm and an average for Hy of 0.46 nm; for Hy-Br, 0.64 nm; and for Hy-4Br, 0.73 nm. All average distances are summarized in Table 1. The Hy molecules are obviously the ones located closest to cholesterol in both membranes, and in the 25 mol % cholesterol membrane, the minimum distance between Hy and cholesterol is shorter than in the 9 mol % cholesterol membrane. For Hy-Br, there is no significant difference between the two membranes, whereas for Hy-4Br the minimum distance is longer in the 25 mol % cholesterol membrane, nearly 3 Å longer than for Hy in the same membrane. We emphasize, however, that the trends observed should be taken with some caution given the limited amount of data.

Density profiles for the different systems are displayed in Figure 6 and show that the probability to find the hypericin

Table 1. Average Distances (in nm) between Any Pair of Atoms of a Hypericin Molecule and a Cholesterol Molecule Inside the Two Membranes during the Production Runs

molecule	9 mol % cholesterol	25 mol % cholesterol
Hy	0.56	0.46
Hy-Br	0.63	0.64
Hy-4Br	0.63	0.73

molecules close to the polar headgroups and the interface between the lipids and water is high. This being the densest region of the bilayer, it would not be favorable for the large and inflexible hypericin molecules to accumulate if it was not for the possibility to interact with water while still being embedded in the lipids. This is discussed further in connection with the radial distribution functions below. Comparing the two membranes reveals that Hy is found overall slightly closer to the lipid/water interface in the 9 mol % cholesterol membrane. For Hy-Br, the molecule that enters the lipids from water (right peak) ends up closer to the lipid/water interface than the molecule that was initially positioned in the middle of the bilayer. This is seen in both membranes. Hy-4Br displays wider density profiles than Hy and Hy-Br, in both membranes, and moves closer to the bilayer center than any of the others. This results in less interaction with water and is supported by observations made in a pure DPPC membrane.³⁹ Concluded from the density profiles is that at least one of the two Hy and Hy-Br molecules in each simulation is located closer to the lipid/water interface in the 9 mol % cholesterol membrane. Similar density profiles have been found for smaller drug molecules such as psoralens and 5-aminolevulinic acid and derivatives thereof; however, these molecules were located slightly closer to the bilayer center.^{88,89}

In natural membranes, unsaturated lipids are important constituents, especially in the formation of raft domains. Due to the use of saturated lipids in the present study, possible interactions between the hypericin molecules and unsaturated lipids, as well as between cholesterol and unsaturated lipids was not accounted for, which could affect the partition and behavior of the molecules. It has been found that cholesterol interacts stronger with saturated lipids than unsaturated ones, resulting in more pronounced effects on the structure of the membrane, including condensing and ordering aspects.²⁷ Consequently, this affects the behavior of small molecules, such as hypericin, in the membranes. As an initial study of the behavior of hypericin derivatives in cholesterol containing membranes, and to enable comparisons with earlier work on hypericin in purely saturated bilayers, we used saturated lipids also in the present study. As a next step, in order to follow up this study, unsaturated lipids would be included for comparison.

In cholesterol-containing membranes, the lipids are more ordered and their tails more aligned to the bilayer normal, compared to in a pure DPPC membrane. This results in a larger bilayer thickness and is seen by comparing the DPPC density profile for the 25 mol % cholesterol membrane in Figure 6 with the DPPC density profiles for the pure lipid membrane.³⁹ The bilayer thickness is increased by a few Ångstroms in the cholesterol-containing membrane. In the 9 mol % cholesterol membrane, the bilayer thickness is approximately the same as in the pure DPPC membrane. However, the fact that the bilayer thickness is increased does not necessarily mean that each monolayer thickness is increased in the 25 mol % cholesterol membrane but is rather a result of the monolayers being more

separated, resulting in a decrease in density in the center of the bilayer (Figure 6). These findings are in agreement with previous observations that the hydrophobic thickness is increased and the electron density in the center of the bilayer is decreased in cholesterol-containing membranes.^{17,24} In the 9 mol % cholesterol membrane, the monolayer separation is similar to the one in the pure DPPC membrane, indicated by the higher density in the bilayer center. The increased free space in the middle of the 25 mol % cholesterol membrane could be expected to result in a reduced probability of locating the hydrophobic hypericin molecules in that region, even more than in the more compact membranes. This is supported by the calculated free energy profiles for the transport process through the membrane and is further discussed below.

A test set of simulations with 10 hypericin (Hy) molecules initially placed in the water region of the two systems provided information about interactions between the molecules (data not shown). It was clear that in the water phase the molecules tend to stack together due to the nonfavorable polar environment. Previous studies have confirmed that hypericin molecules interact with each other in solution, forming dimers when the concentration is low⁹⁰ and H aggregates of at least four molecules positioned face to face when the concentration is higher.⁹¹ It is clear from the simulations performed herein that the hypericin molecules interact strongly with each other in conformations of either two or four molecules in the water phase; however, whether dimers or H aggregates are predominantly formed is difficult to determine due to the high degree of movement in the system. The aggregation did not seem to affect the ability to enter the lipids, and the molecules remained stacked together also within the lipids, indicating that these interactions are favorable also in a nonpolar environment. It has previously been observed that the photodynamic properties of hypericin are altered by aggregation, resulting in a decreased singlet oxygen yield,⁹⁰ an issue that needs to be considered if the hypericin molecules are to be activated by light within a membrane.

Radial distribution functions between oxygen atoms on the hypericin derivatives and hydrogen atoms in the surrounding water (Figure 7A) and between polar hydrogen atoms on the hypericin derivatives and oxygen atoms in the surrounding water (Figure 7B) were calculated for both membranes. The first peak in both figures (at ~0.18 nm) corresponds to a hydrogen bond. The following peak in Figure 7A corresponds to a second hydrogen in the same water molecule or a second solvation shell, whereas the second peak in Figure 7B corresponds to a second solvation shell of water. Following this, there is an increase in amplitude of the radial distribution functions as more and more water molecules are included in shells of higher order.

The peaks corresponding to hydrogen bonds suggest that the hypericin molecules, to a greater or smaller extent, interact with water as they accumulate in the region close to the interface between the lipids and water. In a pure DPPC membrane, a clear trend was observed for the radial distribution functions, with Hy displaying the highest radial distribution functions overall, followed by Hy-Br and Hy-4Br, respectively.³⁹ Such a clear trend cannot be observed in the cholesterol-containing membranes. In the 9 mol % cholesterol membrane, Hy-Br displays the highest radial distribution function for a hydrogen bond between oxygen atoms on hypericin and hydrogen atoms in water followed by Hy and Hy-4Br, respectively. This feature can be explained by the finding that one of the Hy-Br molecules was located very close to the lipid water interface, as seen in the density profiles. However,

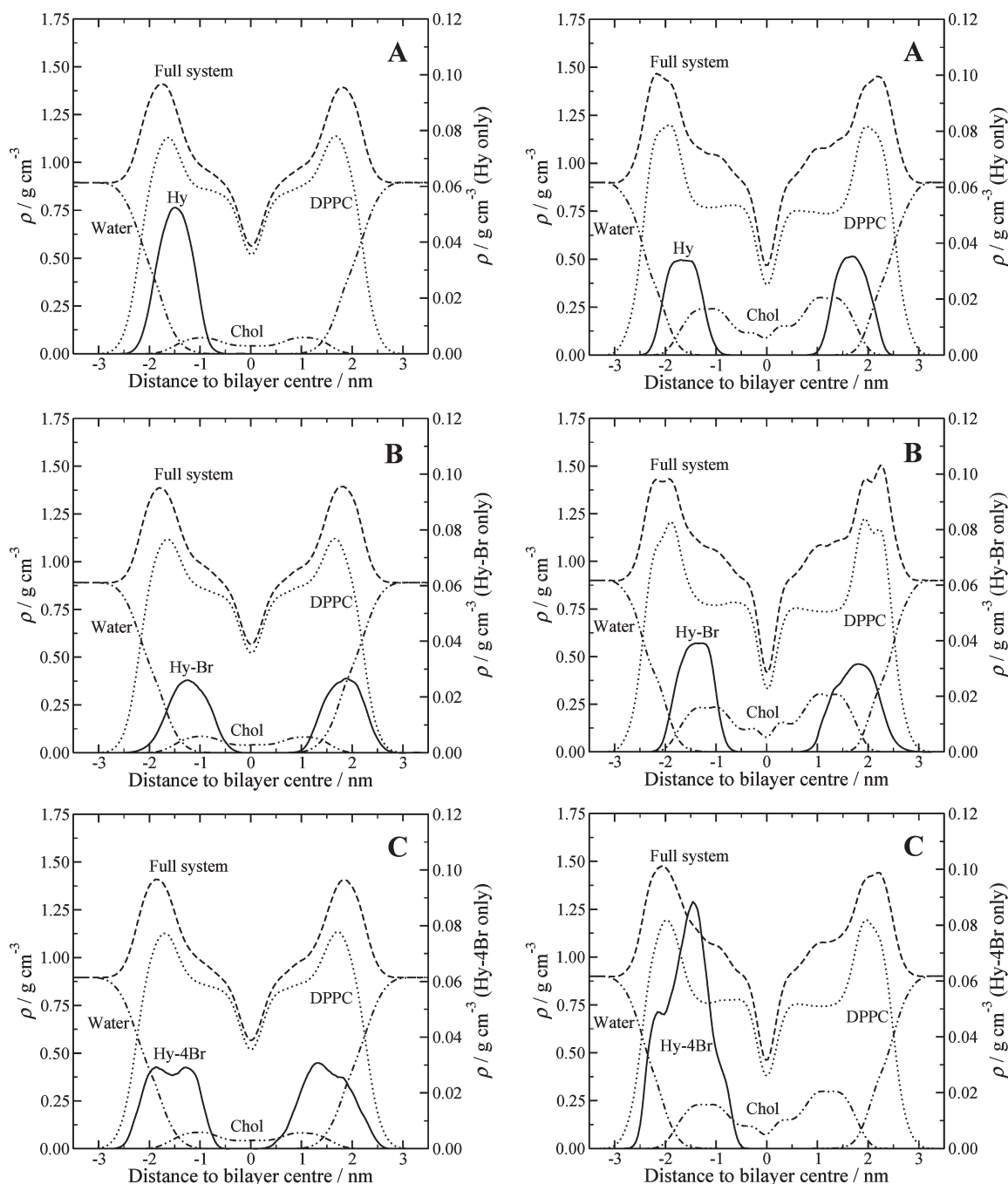


Figure 6. Density profiles for two hypericin derivatives (A, Hy; B, Hy-Br; C, Hy-4Br) in the DPPC bilayer with 9% cholesterol (left) and 25% cholesterol (right).

Hy displays the highest radial distribution function for a hydrogen bond between hydrogen atoms of hypericin and oxygen atoms in water, albeit very low, followed by even lower amplitudes for Hy-Br and Hy-4Br. In the 25 mol % cholesterol membrane, Hy displays the highest radial distribution function for a hydrogen bond between oxygen atoms on hypericin and hydrogen atoms in water followed by Hy-Br and Hy-4Br at equal amplitudes, whereas Hy and Hy-Br display the highest radial distribution functions for hydrogen bonds between hydrogen atoms of hypericin and oxygen atoms in water. Common for all membranes, with and without cholesterol, is that Hy-4Br overall displays the lowest radial distribution functions. This is explained

by the fact that this molecule was moving closer to the bilayer middle during the simulations in both membranes, hence reducing the ability to interact with water.

The mean-square displacement (MSD)⁹² reveals details about the movements of a molecule inside the bilayer and is defined by

$$\text{MSD}(t) = \langle |\vec{r}(t) - \vec{r}(0)|^2 \rangle \quad (7)$$

where $\vec{r}(0)$ and $\vec{r}(t)$ are the positions of a particle at time $t = 0$ and at a certain time t .

The brackets indicate a time average over all similar particles and over different time origins along the simulation. The Einstein

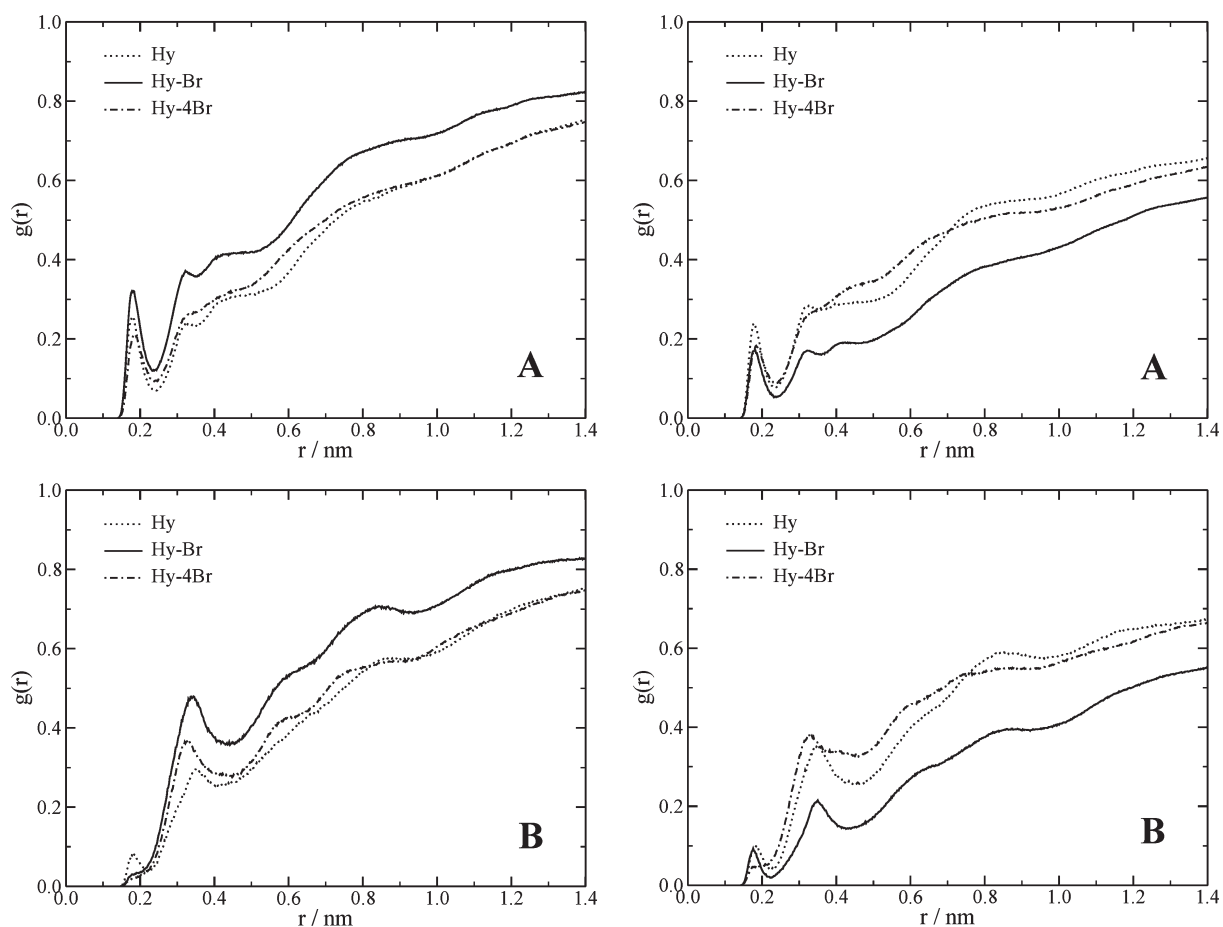


Figure 7. Radial distribution functions between (A) oxygen atoms on the hypericin derivatives and hydrogen atoms in the surrounding water and (B) polar hydrogen atoms on the hypericin derivatives and oxygen atoms in the surrounding water in the 9 mol % (left) and 25 mol % cholesterol membranes (right).

relation allows for the calculation of the diffusion coefficient, D , at sufficiently long simulation times:⁹²

$$D = \lim_{t \rightarrow \infty} \frac{1}{2dt} \langle |r_i(t) - r_i(0)|^2 \rangle \quad (8)$$

where d is the dimensionality of the space.

This way, one can obtain the MSD for the molecules moving in the bilayer plane ($d = 2$) and along the bilayer normal ($d = 1$), respectively. The MSD provides a measure of the average distance a molecule travels in the system, and the growth rate of the MSD depends on how often the molecule collides, i.e., a measure of the ease of diffusion of the molecule.

Like other molecules diffusing in confined media, the hypericin molecules never reach the Einsteinian limit of proper diffusion within the limited time of the simulation, and anomalous diffusion occurs where MSD is proportional to t^n , with $0 < n < 1$. The implication is that a direct comparison with experimental diffusion coefficients cannot be made. However, on the basis of the MSD, one can state which molecules have a higher or lower diffusive regime. The MSDs of the hypericin derivatives in the bilayer plane and along the normal of the bilayer (z direction) in the two membranes are displayed in Figure 8A and B, respectively. The addition of a bromine atom to the hypericin molecule does not significantly affect the movement in the bilayer plane in the 9 mol % cholesterol membrane, whereas the addition of four bromine atoms makes the molecule

move more easily. Despite the fact that Hy-4Br is heavier than the two other molecules, its movement in the bilayer plane is faster due to a lower propensity of forming hydrogen bonds with water, as seen above. In the 25 mol % cholesterol membrane, however, the movement in the bilayer plane is highest for Hy, followed by Hy-4Br and Hy-Br, and the movement of both Hy and Hy-4Br is considerably higher here than in the membrane of lower cholesterol concentration. This situation does not reflect the hydrogen bond capability with water, as it was shown above that Hy exhibits the highest radial distribution functions in the 25 mol % cholesterol membrane. The movement of all molecules in the bilayer plane of the 9 mol % cholesterol membrane is slower than in a pure DPPC membrane.³⁹ The movement of Hy in the bilayer plane of the 25 mol % cholesterol membrane is similar to that in a pure DPPC membrane, whereas for Hy-Br and Hy-4Br the movement is slower.

The MSDs for the molecules along the bilayer normal of the two membranes are displayed in Figure 8B. This movement is, as opposed to the movement in the bilayer plane, finite and should hence be interpreted with caution. The MSD profiles in the direction of the bilayer normal of the 9 mol % cholesterol membrane are similar, and on the same order of magnitude, as the ones obtained in a pure DPPC membrane, displaying the slowest movement for Hy and similar, yet faster, movement for Hy-Br and Hy-4Br,³⁹ and overall slightly faster than in the 25 mol % cholesterol membrane. It is clearly seen that the

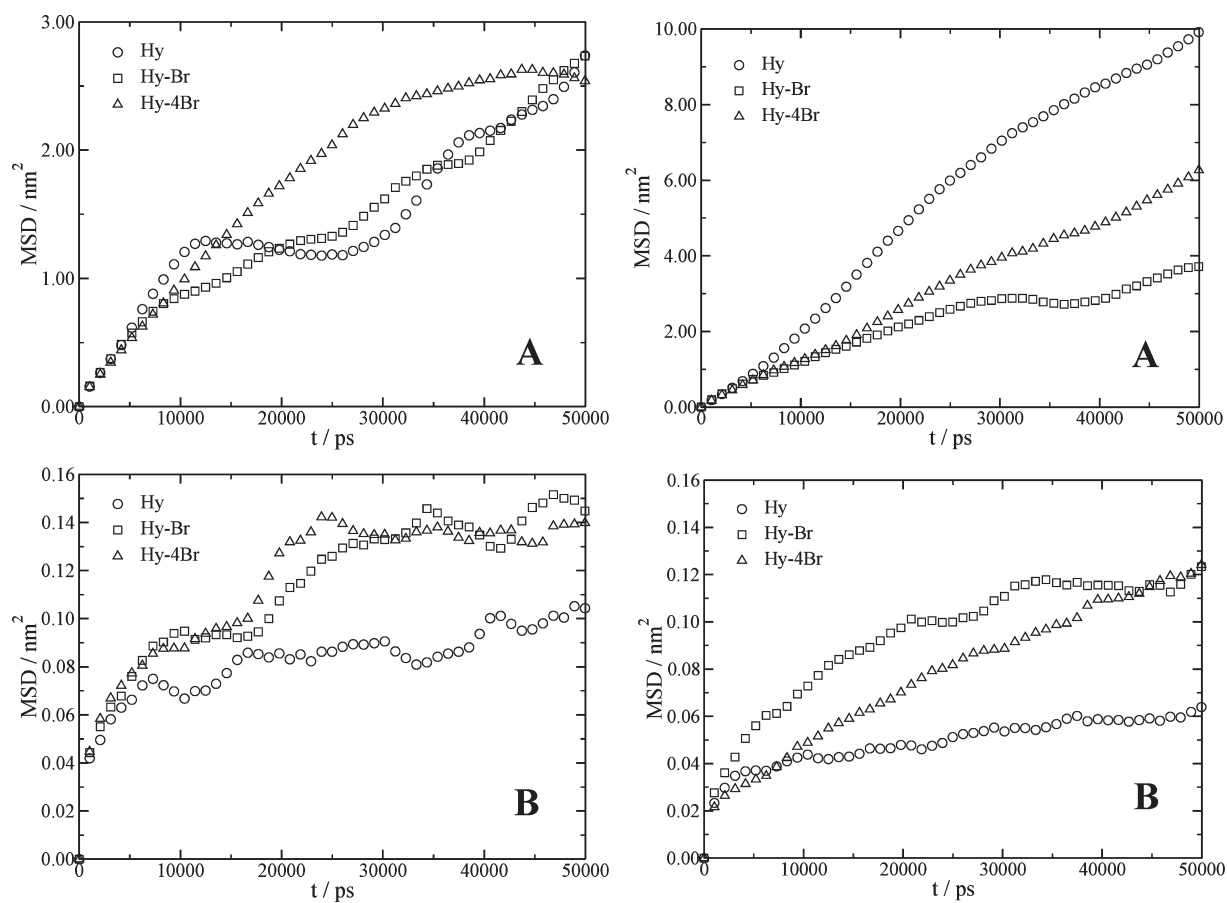


Figure 8. MSD in (A) the *xy* plane and (B) the direction normal to the bilayer in the 9 mol % (left) and 25 mol % cholesterol membrane (right). Note the different scales on the y axes in A.

movement in the bilayer plane is significantly higher than in the direction of the bilayer normal.

The local diffusion coefficients are displayed in Figure 9 as functions of the distance to the bilayer center and were obtained by integrating the fitted autocorrelation functions (eq 4). It is clear that the diffusion of the molecules is low inside the lipids, whereas in the water, the diffusion is faster, which is seen by the increase in diffusion in the far left and far right regions of the graphs in Figure 9. As we only studied the hypericin molecules at a distance of at most 3.3 nm from the bilayer center, the molecules are at this distance still interacting with the lipids, which allow for less free movement of the molecules compared to bulk water. In the pure DPPC system, in which systems at larger distances from the bilayer center (the diffusion further out in the water phase) were studied, diffusion in the water phase was at least 10 times faster than inside the lipids, and strongly dependent on the size of the molecules.³⁹ We expect diffusion coefficients on the same order of magnitude far out in the water phase of the cholesterol-containing systems as well.

In the 9 mol % cholesterol membrane, the diffusion of all three molecules inside the lipids was relatively constant, with a small decrease in diffusion for Hy and Hy-4Br and a small increase in diffusion for Hy-Br, in the very middle of the membrane. Hy shows the lowest diffusion at the bilayer center. However, in the 25 mol % cholesterol membrane, the diffusion close to the middle of the membrane is significantly increased compared to further out in the lipids and also compared to the middle of the 9 mol % cholesterol membrane. The large increase in diffusion is

due to the increased free space in the middle of the high cholesterol membrane, in which the monolayers are more separated. This free space allows for fast diffusion due to few interactions with lipid molecules. Hy displays the fastest diffusion in the middle, followed by Hy-Br and Hy-4Br. In the remaining lipid region, the diffusion is overall lower than in the membrane containing a lower amount of cholesterol.

Free energy profiles for the transport process from water and into the lipids, as a function of the distance to the bilayer center, were calculated using the potential of mean force formalism outlined above⁹³ and are displayed in Figure 10 together with the data obtained in the pure DPPC membrane. In a previous study performed using the same technique, but with only 2 ns production runs (we used 4 ns in water and 10 ns in the lipid bilayer), errors in free energy were found to be in the range of 0.7–4 kJ/mol in the bilayer middle, the region with the largest errors.⁹⁴ As the simulations performed in the present study were longer than the ones for which these errors were calculated, we expect the errors herein to be in the same range, if not smaller.

For the hypericin molecules, local minima were found in the region 1–2 nm from the bilayer center, close to the polar headgroup region where the molecules were found to accumulate during the unconstrained simulations. As the molecules continue moving across the bilayer, the free energy increases and shows a maximum in the very middle of the bilayer. In both cholesterol-containing membranes, Hy-Br shows the deepest minimum close to the polar headgroups, followed by Hy-4Br and Hy. For Hy-Br, the minimum is deeper in the 25 mol % cholesterol membrane,

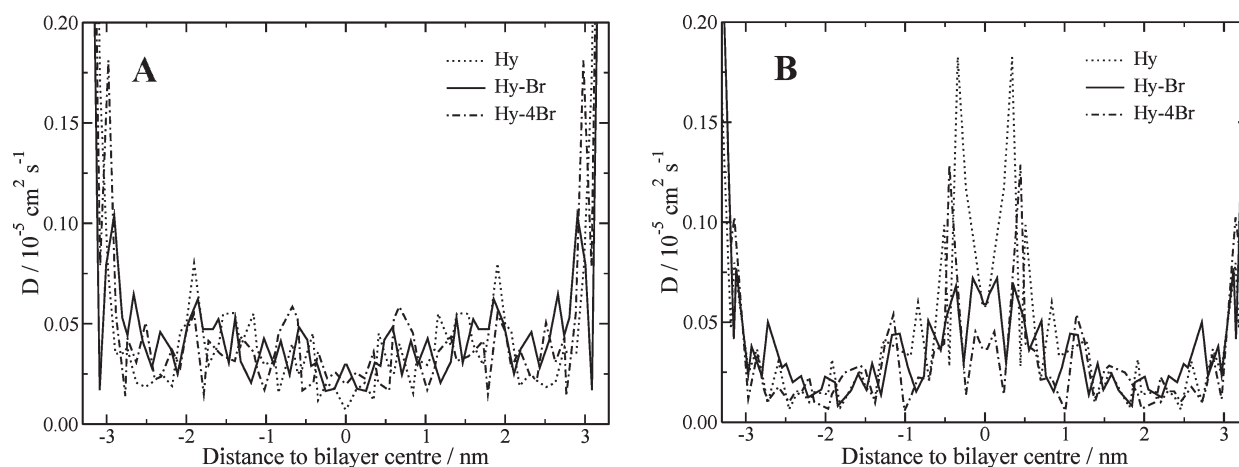


Figure 9. Local diffusion coefficients of the hypericin derivatives in (A) 9 mol % and (B) 25 mol % cholesterol membranes, as a function of the distance to the bilayer middle.

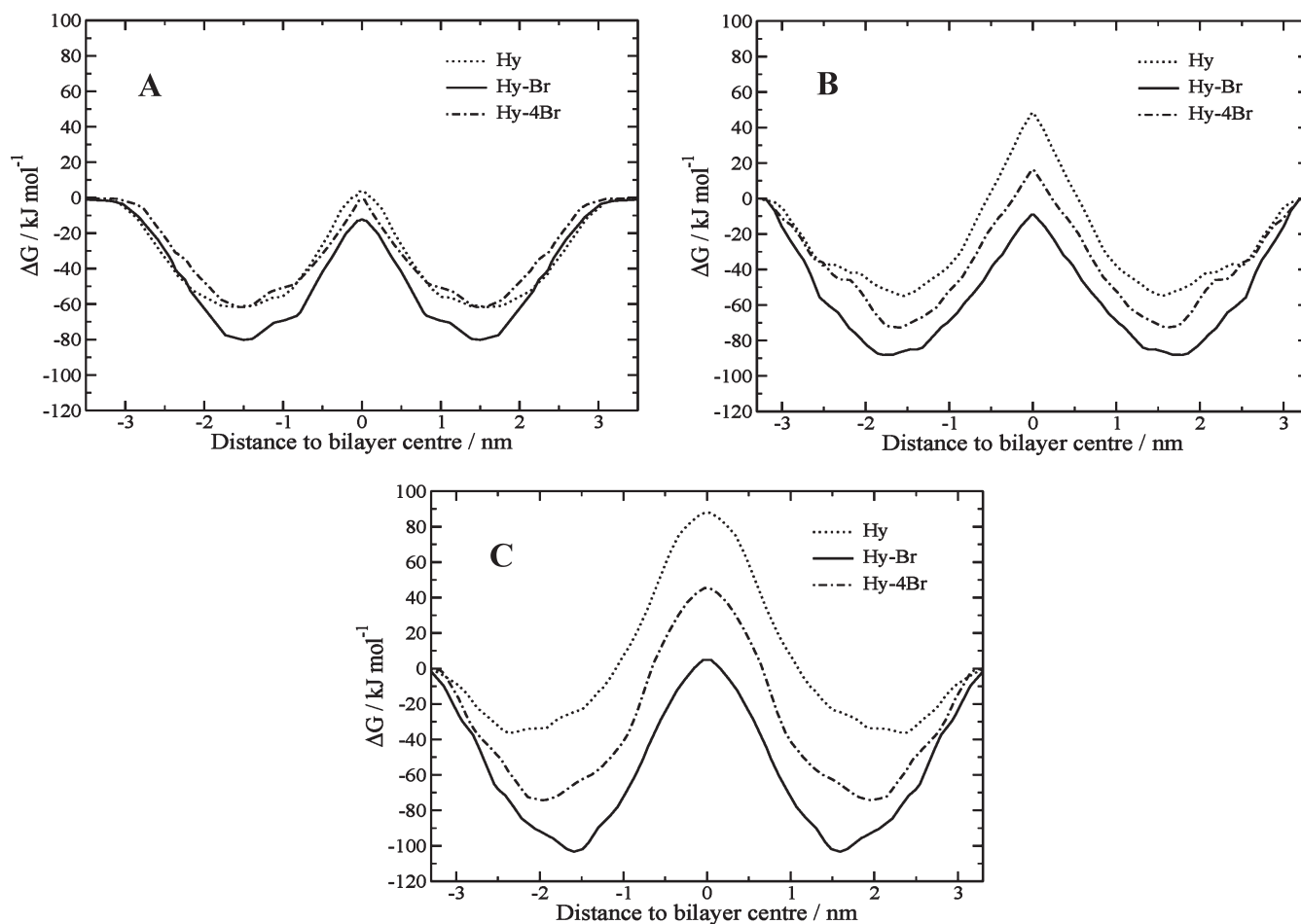


Figure 10. Free energy profiles for the hypericin derivatives inside the (A) 0 mol %, (B) 9 mol %, and (C) 25 mol % cholesterol membranes.

whereas for Hy-4Br, the minimum is more or less at the same depth in the two membranes, and for Hy, the minimum is less deep in the 25 mol % cholesterol membrane. In the middle of the bilayer, the free energy of the molecules follows the same order, with Hy-Br displaying the lowest energy. All three molecules display higher energy barriers in the high cholesterol membrane, with positive ΔG values compared to when in water. In the 9 mol %

cholesterol membrane, however, Hy-Br displays a negative overall free energy also in the bilayer center.

In the low cholesterol-containing membrane, the local minima of the molecules are found further out from the bilayer center with a larger decrease in free energy; accordingly the minimum of Hy-Br is found further toward the lipid/water interface, followed by Hy-4Br and Hy. Interestingly, in the high cholesterol

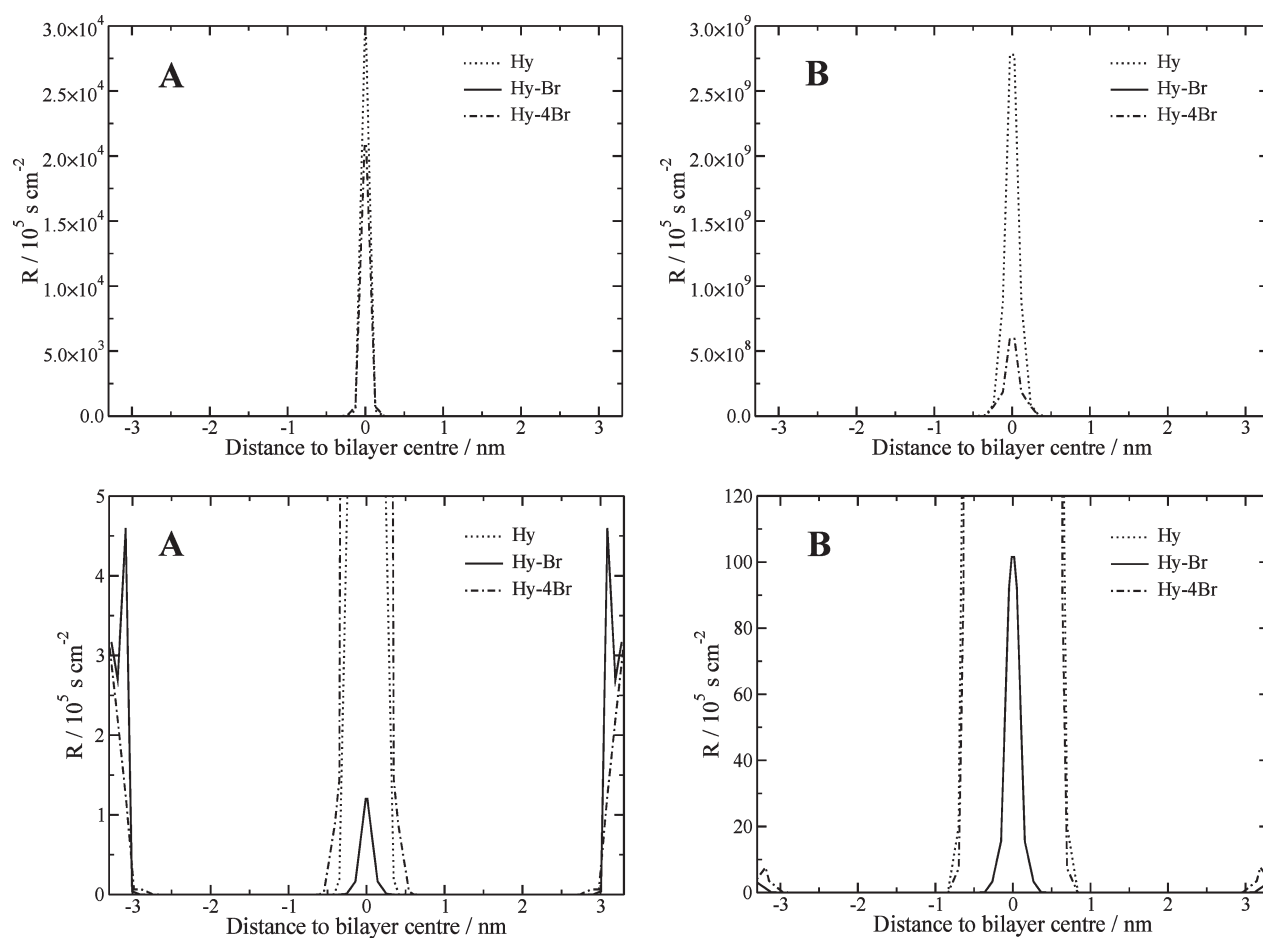


Figure 11. Local resistance profiles of the hypericin derivatives in (A) 9 mol % and (B) 25 mol % cholesterol membranes, as a function of the distance to the bilayer middle. The Hy profile was scaled down by a factor 3×10^5 in the 9 mol % cholesterol membrane and by a factor 10^6 in the 25 mol % cholesterol membrane. Magnification of the local resistance profiles of Hy-Br are displayed in the lower graphs. Note the different scales on the y axis of the graphs.

membrane, the opposite situation is observed. Here, the minimum lies closer to the bilayer center with a larger decrease in free energy, generating the minimum of Hy-Br significantly closer to the bilayer center than the two other molecules.

The free energy maxima of the molecules in the 25 mol % cholesterol membrane are clearly more rounded than the narrow maxima found with a lower cholesterol concentration. This is due to the fact that, as discussed above, the two monolayers are more separated in the membrane with higher cholesterol content. The free space in the middle of the membrane is also responsible for the larger increase in free energy in this region compared to in the membrane with a lower cholesterol content. The free space constitutes a less hydrophobic region offering less possible interactions with lipids, in which it is not favorable for the hydrophobic molecules to be located, and thereby generating a large increase in free energy for transport process into/across that region.

In the pure DPPC bilayer, Hy-Br was also found to display the lowest overall change in free energy along the bilayer normal (Figure 10A).³⁹ In the 9 mol % cholesterol membrane, Hy-Br and Hy-4Br show deeper free energy minima in the headgroup region compared to the pure DPPC membrane, whereas Hy displays an increase. In the middle of the bilayer, the free energy is increased (more positive) for all three molecules. Only Hy-Br still displays a negative free energy, however slightly less negative

than without cholesterol present. Hy-4Br, which displayed a small negative change in free energy in the middle of the pure DPPC membrane, is now on the positive side. In the 25 mol % cholesterol membrane, the same pattern is seen when comparing with the pure DPPC membrane, although with larger overall changes than in the 9 mol % cholesterol membrane. It is clear that Hy is most negatively affected by the inclusion of cholesterol in the membrane, displaying a significant increase in free energy both in the middle of the bilayer and close to the polar headgroups.

The local resistance was calculated from the local diffusion coefficients and the free energy profiles using eq 6, and the resulting profiles are displayed in Figure 11. An increase in resistance is seen for all molecules in the middle of the bilayer; however, the magnitude of this peak clearly differs for the three molecules. The resistance for Hy is significantly higher than for the two other molecules in both membranes, and therefore it was scaled down to more clearly display the profiles. The resistance for Hy-Br is considerably lower than for the two other molecules in both membranes, and its profiles can only be seen in the lower graphs of Figure 11 where we show a magnification of the low resistance region. In the 25 mol % cholesterol membrane, the resistances for the molecules are overall higher compared to when the cholesterol concentration is lower. In both membranes, the Hy and Hy-4Br resistances in the middle are considerably

Table 2. Permeability Coefficients (in cm/s) of the Hypericin Derivatives inside the Membranes without and with Cholesterol

molecule	0 mol % cholesterol ³⁹	9 mol % cholesterol	25 mol % cholesterol
Hy	4.21×10^{-4}	8.25×10^{-12}	1.63×10^{-17}
Hy-Br	4.94×10^{-3}	5.49×10^{-3}	4.12×10^{-4}
Hy-4Br	1.51×10^{-3}	3.31×10^{-6}	6.93×10^{-11}

higher than in the water phase, as is seen in the lower graphs of Figure 11. For Hy-Br, however, the resistance in the middle is higher than in water in the 25 mol % cholesterol membrane, whereas in the 9 mol % cholesterol membrane, the resistance in the middle is lower than that in water. The free energy is the major contributor to the shape of the resistance profiles, with an increase in free energy giving an increase in resistance. The lowest free energy of Hy-Br in the middle of the membrane therefore generates the lowest resistance, and the opposite for Hy. The molecules exhibit higher resistances in the middle of the cholesterol-containing membranes than in a pure DPPC membrane.³⁹

Permeability coefficients were calculated by integrating the resistance profiles across the bilayer and are displayed in Table 2. In both cholesterol-containing membranes, the permeation decreases in the order Hy-Br > Hy-4Br > Hy, in agreement with the findings in a pure DPPC membrane,³⁹ for which data are also included in Table 2. In the 25 mol % cholesterol membrane, the permeation is overall slower than when the cholesterol concentration is lower. The difference in permeation of Hy-Br in the two cholesterol-containing membranes is only 1 order of magnitude, whereas for the two other molecules, the difference is significantly larger. The permeation of Hy-Br is not considerably affected by the inclusion of cholesterol, when compared to that in a pure DPPC membrane.³⁹ In fact, the permeation in the low cholesterol membrane is slightly faster than in the pure DPPC membrane. For the two other molecules, the permeation is considerably reduced compared to the pure DPPC membrane. As mentioned above, the resistance, and thereby the permeability, strongly depends on the free energy, and with an increase in energy follows a decrease in permeation. This supports the finding of Hy-Br displaying the fastest permeation (large permeability coefficient) as it exhibits the lowest free energy, both close to the polar headgroups and in the center of the bilayer. Hy-Br was also displaying the largest movement in the *z* direction of the bilayer, as seen above. These results suggest that Hy-Br most likely would have the highest probability of penetrating the plasma membrane and reach the interior of the cell, whereas the two other molecules would be more prone to residing in the membrane and from there cause photodamage.

Experimentally, it has been shown that halogenated drugs display larger permeability coefficients through lipid membranes;⁶³ however, the permeability of the molecules can also be reduced by the addition of large and heavy substituents. This can explain the reduced permeability of tetra-brominated hypericin compared with the monobrominated one.

The finding that Hy displays the slowest diffusion in the membranes can be explained by the fact that this molecule was found to reside closer to cholesterol molecules in both membranes, and it also forms hydrogen bonds to the water phase. Even though the distance between Hy and cholesterol was never short enough to allow direct interactions, the ordering of the

lipids close to cholesterol molecules can affect the permeability of the hypericin molecule.

4. CONCLUSIONS

The behavior of hypericin, a natural compound possessing photodynamic properties, and its mono- and tetra-brominated derivatives was studied in DPPC lipid membranes containing 9 and 25 mol % cholesterol, respectively, by means of molecular dynamics simulations. The three molecules were found to accumulate in the region close to the polar headgroups of the lipids, close to the interface between the lipids and water. This location enables interactions between the hydroxyl groups of the hypericin molecules and water, supported by hydrogen bonds found in the radial distribution functions. Overall, the hypericin molecules in the 25 mol % cholesterol membrane were found slightly closer to the bilayer middle than in the low cholesterol membrane. With a high amount of cholesterol present in the membrane, the two monolayers are slightly separated and allow for fast diffusion in that region. However, this region also offers fewer lipid interactions with the hydrophobic hypericin molecules, making it a nonfavorable location to reside in, indicated by the large increase in free energy for the transport process of the molecules into that region. Close to the polar headgroups, where the hypericin molecules were found to accumulate, the free energy profiles showed local minima. The size of these minima was dependent on the molecule and the amount of cholesterol in the membrane. The permeability coefficients were overall lower in the high cholesterol membrane. Hy-Br displayed the largest decrease in free energy both in the middle of the bilayer and close to the headgroups, as well as the highest permeability coefficients in both membranes. This indicates that this molecule would have the highest probability to penetrate a plasma membrane, independently of the cholesterol concentration, and reach the interior of a cell. The two other molecules, in particular Hy, exhibit significantly lower permeability and are expected to be found within the membrane.

■ ASSOCIATED CONTENT

S Supporting Information. Cholesterol topology, hypericin topology, Hy-Br topology, Hy-4Br topology, evolution of the area per lipid during the last 10 ns of the equilibration of the cholesterol containing membranes, and RDFs for DPPC-Chol and Chol-Chol center-of-mass during production runs (100 ns) for each membrane and each hypericin system. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: leif.eriksson@nuigalway.ie.

■ ACKNOWLEDGMENT

The National University of Ireland—Galway is gratefully acknowledged for financial support. We also acknowledge Dr. Daniel dos Santos (Univ. of Lisbon) for valuable discussions.

■ REFERENCES

- (1) De Kruff, B.; De Greef, W. J.; Van Eyk, R. V. W.; Demel, R. A.; Van Deene, L. *Biochim. Biophys. Acta* **1973**, *298*, 479–499.
- (2) Lange, Y.; Swaisgood, M. H.; Ramos, B. V.; Steck, T. L. *J. Biol. Chem.* **1989**, *264*, 3786–3793.

- (3) Bretscher, M. S.; Munro, S. *Science* **1993**, *261*, 1280–1281.
- (4) Sackmann, E. In *Structure and dynamics of membranes*; Lipowsky, R., Sackmann, E., Eds.; Elsevier: Amsterdam, 1995; pp 1–62.
- (5) Ipsen, J. H.; Karlström, G.; Mouritsen, O. G.; Wennerström, H.; Zuckermann, M. J. *Biochim. Biophys. Acta* **1987**, *905*, 162–172.
- (6) Vist, M. R.; Davis, J. H. *Biochemistry* **1990**, *29*, 451–464.
- (7) Boggs, J. M.; Hsia, J. C. *Biochim. Biophys. Acta* **1972**, *290*, 32–42.
- (8) Oldfield, E.; Meadows, M.; Rice, D.; Jacobs, R. *Biochemistry* **1978**, *17*, 2727–2740.
- (9) Marsh, D.; Smith, I. C. P. *Biochim. Biophys. Acta* **1973**, *298*, 133–144.
- (10) Yeagle, P. L. *Biochim. Biophys. Acta* **1985**, *822*, 267–287.
- (11) Simons, K.; Toomre, D. *Nat. Rev. Mol. Cell Biol.* **2000**, *1*, 31–39.
- (12) Scott, H. L.; Kalaskar, S. *Biochemistry* **1989**, *28*, 3687–3691.
- (13) Scott, H. L. *Biophys. J.* **1991**, *59*, 445–455.
- (14) Edholm, O.; Nyberg, A. M. *Biophys. J.* **1992**, *63*, 1081–1089.
- (15) Hofsäuss, C.; Lindahl, E.; Edholm, O. *Biophys. J.* **2003**, *84*, 2192–2206.
- (16) Niemelä, P. S.; Ollila, S.; Hyvönen, M. T.; Karttunen, M.; Vattulainen, I. *PLoS Comput. Biol.* **2007**, *3*, 304–312.
- (17) Smondryev, A. M.; Berkowitz, M. L. *Biophys. J.* **1999**, *77*, 2075–2089.
- (18) Pasenkiewicz-Gierula, M.; Róg, T.; Kitamura, K.; Kusumi, A. *Biophys. J.* **2000**, *78*, 1376–1389.
- (19) Chiu, S. W.; Jakobsson, E.; Scott, H. L. *J. Chem. Phys.* **2001**, *114*, 5435–5443.
- (20) Róg, T.; Pasenkiewicz-Gierula, M. *Biophys. J.* **2001**, *81*, 2190–2202.
- (21) Chiu, S. W.; Jakobsson, E.; Mashl, R. J.; Scott, H. L. *Biophys. J.* **2002**, *83*, 1842–1853.
- (22) Falck, E.; Patra, M.; Karttunen, M.; Hyvönen, M. T.; Vattulainen, I. *Biophys. J.* **2004**, *87*, 1076–1091.
- (23) Tu, K. C.; Klein, M. L.; Tobias, D. J. *Biophys. J.* **1998**, *75*, 2147–2156.
- (24) Douliez, J. P.; Leonard, A.; Dufour, E. J. *J. Phys. Chem.* **1996**, *100*, 18450–18457.
- (25) Martinez-Seara, H.; Róg, T.; Karttunen, M.; Vattulainen, I.; Reigada, R. *PLoS One* paper: e11162, pages 1–11.
- (26) Róg, T.; Pasenkiewicz-Gierula, M.; Vattulainen, I.; Karttunen, M. *Biophys. J.* **2007**, *92*, 3346–3357.
- (27) Róg, T.; Pasenkiewicz-Gierula, M. *Biochimie* **2006**, *88*, 449–460.
- (28) Pitman, M. C.; Suits, F.; MacKerell, A. D.; Feller, S. E. *Biochemistry* **2004**, *43*, 15318–15328.
- (29) Martinez-Seara, H.; Róg, T.; Pasenkiewicz-Gierula, M.; Vattulainen, I.; Karttunen, M.; Reigada, R. *J. Phys. Chem. B* **2007**, *111*, 11162–11168.
- (30) Martinez-Seara, H.; Róg, T.; Pasenkiewicz-Gierula, M.; Vattulainen, I.; Karttunen, M.; Reigada, R. *Biophys. J.* **2008**, *95*, 3295–3305.
- (31) Papahadjopoulos, D.; Nir, S.; Ohki, S. *Biochim. Biophys. Acta* **1972**, *266*, 561–583.
- (32) Finkelstein, A.; Cass, A. *Nature* **1967**, *216*, 717–718.
- (33) Zhao, L. Y.; Feng, S. S. *J. Colloid Interface Sci.* **2006**, *300*, 314–326.
- (34) Gere-Paszti, E.; Farkas, O.; Prodan, M.; Forgacs, E. *Chromatographia* **2003**, *57*, 599–604.
- (35) Söderlund, T.; Lehtonen, J. Y. A.; Kinnunen, P. K. *J. Mol. Pharmacol.* **1999**, *55*, 32–38.
- (36) Ghannam, M. M.; Mady, M. M.; Khalil, W. A. *Biophys. Chem.* **1999**, *80*, 31–40.
- (37) Prenner, E. J.; Lewis, R.; Jelokhani-Niaraki, M.; Hodges, R. S.; McElhaney, R. N. *Biochim. Biophys. Acta Biomembr.* **2001**, *1510*, 83–92.
- (38) Ji, S. R.; Wu, Y.; Sui, S. F. *J. Biol. Chem.* **2002**, *277*, 6273–6279.
- (39) Eriksson, E. S. E.; dos Santos, D. J. V. A.; Guedes, R. C.; Eriksson, L. A. *J. Chem. Theory Comput.* **2009**, *5*, 3139–3149.
- (40) Linde, K.; Mulrow, C. D.; Berner, M.; Egger, M. *Cochrane Database Syst. Rev.* **2005**, *73*.
- (41) Meruelo, D.; Lavie, G.; Lavie, D. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 5230–5234.
- (42) Lopezbazzocchi, I.; Hudson, J. B.; Towers, G. H. N. *Photochem. Photobiol.* **1991**, *54*, 95–98.
- (43) Hudson, J. B.; Lopezbazzocchi, I.; Towers, G. H. N. *Antivir. Res.* **1991**, *15*, 101–112.
- (44) Degar, S.; Prince, A. M.; Pascual, D.; Lavie, G.; Levin, B.; Mazur, Y.; Lavie, D.; Ehrlich, L. S.; Carter, C.; Meruelo, D. *AIDS Res. Hum. Retrovir.* **1992**, *8*, 1929–1936.
- (45) Moraleda, G.; Wu, T. T.; Jilbert, A. R.; Aldrich, C. E.; Condreay, L. D.; Larsen, S. H.; Tang, J. C.; Colacino, J. M.; Mason, W. S. *Antiviral Res.* **1993**, *20*, 235–247.
- (46) Lenard, J.; Rabson, A.; Vanderoef, R. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 158–162.
- (47) Thomas, C.; Pardini, R. S. *Photochem. Photobiol.* **1992**, *55*, 831–837.
- (48) Andreoni, A.; Colasanti, A.; Colasanti, P.; Mastrocinque, M.; Riccio, P.; Roberti, G. *Photochem. Photobiol.* **1994**, *59*, 529–533.
- (49) VanderWerf, Q. M.; Saxton, R. E.; Chang, A.; Horton, D.; Paiva, M. B.; Anderson, J.; Foote, C.; Soudant, J.; Mathey, A.; Castro, D. J. *Laryngoscope* **1996**, *106*, 479–483.
- (50) Liu, C. D.; Kwan, D.; Saxton, R. E.; McFadden, D. W. *J. Surg. Res.* **2000**, *93*, 137–143.
- (51) D’Hallewin, M. A.; De Witte, P. A.; Waelkens, E.; Merlevede, W.; Baert, L. *J. Urol.* **2000**, *164*, 349–351.
- (52) D’Hallewin, M. A.; Kamuhabwa, A. R.; Roskams, T.; De Witte, P. A. M.; Baert, L. *BJU Int.* **2002**, *89*, 760–763.
- (53) Pytel, A.; Schmeller, N. *Urology* **2002**, *59*, 216–219.
- (54) Delaey, E.; Zupko, I.; Chen, B.; Derycke, A.; Van Laar, F.; De Vos, D.; De Witte, P. *Int. J. Oncol.* **2003**, *23*, 519–524.
- (55) Guedes, R. C.; Eriksson, L. A. *J. Photochem. Photobiol. Chem.* **2006**, *178*, 41–49.
- (56) Hudson, J. B.; Delaey, E.; de Witte, P. A. *Photochem. Photobiol.* **1999**, *70*, 820–822.
- (57) Senthil, V.; Jones, L. R.; Senthil, K.; Grossweiner, L. I. *Photochem. Photobiol.* **1994**, *59*, 40–47.
- (58) Chaloupka, R.; Obsil, T.; Plasek, J.; Sureau, F. *Biochim. Biophys. Acta Biomembr.* **1999**, *1418*, 39–47.
- (59) Ho, Y. F.; Wu, M. H.; Cheng, B. H.; Chen, Y. W.; Shih, M. C. *Biochim. Biophys. Acta Biomembr.* **2009**, *1788*, 1287–1295.
- (60) Chen, B.; Xu, Y.; Roskams, T.; Delaey, E.; Agostinis, P.; Vandenheede, J. R.; de Witte, P. *Int. J. Canc.* **2001**, *93*, 275–282.
- (61) Kascakova, S.; Refregiers, M.; Jancura, D.; Sureau, F.; Maurizot, J. C.; Miskovsky, P. *Photochem. Photobiol.* **2005**, *81*, 1395–1403.
- (62) Mukherjee, P.; Adhikary, R.; Halder, M.; Petrich, J. W.; Miskovsky, P. *Photochem. Photobiol.* **2008**, *84*, 706–712.
- (63) Gerebtzoff, G.; Li-Blatter, X.; Fischer, H.; Frentzel, A.; Seelig, A. *ChemBioChem* **2004**, *5*, 676–684.
- (64) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theor. Comput.* **2008**, *4*, 435–447.
- (65) Tieleman, D. P.; Berendsen, H. J. C. *J. Chem. Phys.* **1996**, *105*, 4871–4880.
- (66) Patra, M.; Karttunen, M.; Hyvönen, M. T.; Falck, E.; Lindqvist, P.; Vattulainen, I. *Biophys. J.* **2003**, *84*, 3636–3645.
- (67) Patra, M.; Karttunen, M.; Hyvönen, M. T.; Falck, E.; Vattulainen, I. *J. Phys. Chem. B* **2004**, *108*, 4485–4494.
- (68) Biological Physics & Soft Condensed Matter Group of Prof. Mikko Karttunen, Dept. of Applied Mathematics, University of Western Ontario, SoftSimu - Downloads Zone. <http://www.apmaths.uwo.ca/~mkarttu/downloads.shtml>, File: dppc-128–100.pdb (accessed Jan 2009).
- (69) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.;

Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, Revision B.02; Gaussian, Inc.: Pittsburgh, PA, 2003.

(70) Schuttelkopf, A. W.; van Aalten, D. M. *Acta Crystallogr., Sect. D* **2004**, *60*, 1355–1363.

(71) Lavie, G.; Mazur, Y.; Lavie, D.; Prince, A. M.; Pascual, D.; Liebes, L.; Levin, B.; Meruelo, D. *Transfusion* **1995**, *35*, 392–400.

(72) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces*; Reidel Publishing Company: Dordrecht, The Netherlands, 1981.

(73) Nose, S. *Mol. Phys.* **1984**, *52*, 255–268.

(74) Hoover, W. G. *Phys. Rev.* **1985**, *31*, 1695–1697.

(75) Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981**, *52*, 7182–7190.

(76) Nose, S.; Klein, M. L. *Phys. Rev. Lett.* **1983**, *50*, 1207–1210.

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

(78) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577–8593.

(79) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. J. *Comput. Chem.* **1997**, *18*, 1463–1472.

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

(81) Marrink, S. J.; Berendsen, H. J. C. *J. Phys. Chem.* **1994**, *98*, 4155–4168.

(82) Steck, T. L.; Ye, J.; Lange, Y. *Biophys. J.* **2002**, *83*, 2118–2125.

(83) Bennett, W. F. D.; MacCallum, J. L.; Hinner, M. J.; Marrink, S. J.; Tieleman, D. P. *J. Am. Chem. Soc.* **2009**, *131*, 12714–12720.

(84) Jo, S.; Rui, H. A.; Lim, J. B.; Klauda, J. B.; Im, W. *J. Phys. Chem. B* **2010**, *114*, 13342–13348.

(85) Róg, T.; Stimson, L. M.; Pasenkiewicz-Gierula, M.; Vattulainen, I.; Karttunen, M. *J. Phys. Chem. B* **2008**, *112*, 1946–1952.

(86) Engelman, D. M.; Rothman, J. E. *J. Biol. Chem.* **1972**, *247*, 3694–3697.

(87) Smondyrev, A. M.; Berkowitz, M. L. *J. Comput. Chem.* **1999**, *20*, 531–545.

(88) dos Santos, D. J. V. A.; Eriksson, L. A. *Biophys. J.* **2006**, *91*, 2464–2474.

(89) Erdtman, E.; dos Santos, D. J. V. A.; Löfgren, L.; Eriksson, L. A. *Chem. Phys. Lett.* **2008**, *463*, 178–182.

(90) Burel, L.; Jardon, P. *J. Chim. Phys. Phys.* **1996**, *93*, 300–316.

(91) Falk, H.; Meyer, J. *Monatsh. Chem.* **1994**, *125*, 753–762.

(92) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press: Oxford, U. K., 1990.

(93) Paci, E.; Ciccotti, G.; Ferrario, M.; Kapral, R. *Chem. Phys. Lett.* **1991**, *176*, 581–587.

(94) Bemporad, D.; Essex, J. W.; Luttmann, C. *J. Phys. Chem. B* **2004**, *108*, 4875–4884.

Extracting Realistic Kinetics of Rare Activated Processes from Accelerated Molecular Dynamics Using Kramers' Theory

Urmi Doshi and Donald Hamelberg*

Department of Chemistry and The Center for Biotechnology and Drug Design, Georgia State University, Atlanta, Georgia 30302-4098, United States

ABSTRACT: The *cis*–*trans* isomerization of peptide bonds is very slow, occurring in hundreds of seconds. Kinetic studies of such processes using straightforward molecular dynamics are currently not possible. Here, we use Kramers' rate theory in the high friction regime in combination with accelerated molecular dynamics in explicit solvent to successfully retrieve the normal rate of *cis* to *trans* switching in the glycyl–prolyl dipeptide. Our approach bypasses the time-reweighting problem of the hyperdynamics scheme, wherein the addition of the bias potential alters the transition state regions and avoids an accurate estimation of kinetics. By performing accelerated molecular dynamics at a few different levels of acceleration, the rate of isomerization is enhanced as much as 10^{10} to 10^{11} times. Remarkably, the normal rates obtained by simply extrapolating to zero bias are within an order of experimental estimates. This provides validation from a kinetic standpoint of the ω torsional parameters of the AMBER force field that were recently revised by matching to experimentally measured equilibrium properties. We also provide a comparative analysis of the performance of the widely used water models, i.e., TIP3P and SPC/E, in estimating the kinetics of *cis*–*trans* isomerization. Furthermore, we show that the dynamic properties of bulk water can be corrected by adjusting the collision frequency in a Langevin thermostat, which then allows for better reproduction of *cis*–*trans* isomerization kinetics and a closer agreement of rates between experiments and simulations.

INTRODUCTION

Atomistic molecular dynamic (MD) simulations carried out in explicit solvent can provide the closest comparison to a realistic picture of a biomolecular process. In most cases, the main objectives of such simulations are observation of stochastic dynamics to understand the detailed reaction mechanisms and characterization of the system both thermodynamically and kinetically. These include adequate representation of relevant conformational states and reliable estimation of rate coefficients. However, it is very challenging to meet these objectives for biomolecular processes for a number of reasons. First, there are computational limitations to simulating the enormous number of degrees of freedom involved. Second, in spite of significant improvement, the empirical force fields employed for biomolecules are not exact. Third, while most biomolecular processes of interest occur on time scales of several microseconds or longer, standard MD is currently limited to only nanoseconds. Even with advancements in parallel computing, direct MD has become accessible to a maximum of a few microseconds. This time scale problem eventually gives rise to the sampling issue even for the simplest of molecules and can become quite overwhelming, especially for activated processes that have energy barriers significantly higher than $k_B T$. Since biomolecular events are stochastic in nature, one needs to generate multiple trajectories that start from the reactant and take different reaction paths and times to reach the product. Also, in order to obtain reasonable kinetics, it is necessary to have a good sampling of reaction paths. But for long time-scale processes, it is not possible to observe a sufficient number of barrier-crossing events during the length of a typical brute-force MD and hence accurately calculate kinetic rate constants.

Traditionally, the time scale problem is approached using transition state theory (TST)^{1,2} and the related reactive flux method.³ Using the prevalent constrained sampling methods, one can obtain the free energy profile along the reaction coordinate determined in advance.^{4–6} Without the need to perform dynamics in long simulations and observe the actual events, the rates can be simply estimated from the free energy difference between the transition and the stable states (i.e., $k^{\text{TST}} = (k_B T)/h \exp(-\Delta F^\ddagger/(k_B T))$). The TST rates are not accurate, usually overestimated, because the effects of recrossing the barrier region are not accounted for and $(k_B T)/h$, which may be a reasonable prefactor for small organic molecules, is highly inappropriate for biomolecules. For long time-scale events, the reactive flux method provides a correction over TST rates (i.e., $k^{\text{RF}} = \kappa k^{\text{TST}}$) from short-time behaviors. Instead of an energy well, trajectories are initiated in the transition state region and monitored whether they reach the product/reactant well. From several such trajectories, one can evaluate the transition velocity at the transition point and the transmission coefficient κ , which gives the fraction of the trajectories that actually make it to an energy well without recrossing the transition state region. The rate coefficients can then be calculated from these quantities and the relative probability distribution of the transition and the ground states obtained from equilibrium simulations. In the past, when computational resources were limited to observe complete activated events, such approaches were useful to study isomerization of side-chain dihedrals in proteins and estimate rate constants.^{7,8} However, the results of the reactive flux method depend strongly

Received: September 21, 2010

Published: January 27, 2011

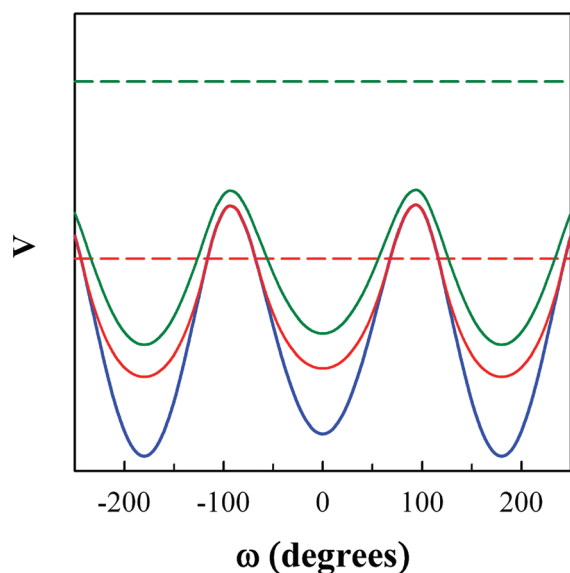


Figure 1. Schematic representation of original potential and bias potentials at two different values of boost energy E . The blue curve depicts a typical torsional potential along the ω dihedral. The potential has a minimum around 0° representing the *cis* state, and the *trans* states populate the regions around $+$ and -180° due to periodicity. The red and green curves show the modified potentials corresponding to two different values of E (red and green dotted lines).

on the choice of the transition point and the reaction coordinate. For biomolecules with hyper-dimensional potential energy landscape, it is very challenging to locate the precise transition point and choose the correct reaction coordinate that is often a collective variable involving more than one order parameter. Besides, the reactive flux method may not be applicable for certain activated bioprocesses that have broad transition state regions because the trajectories originating from there may take a very long time to reach the ground states and thereby not remain short-time.

Recent efforts toward obtaining a detailed picture of kinetic behavior of biomolecules from simulation studies have resulted in several improved methodologies and algorithms.⁹ These include transition path sampling,^{10,11} transition interface sampling,¹² partial path transition interface sampling,¹³ ensemble dynamics in combination with distributed computing,¹⁴ milestoning,¹⁵ Markovian formalism,¹⁶ string method,¹⁷ forward flux sampling,^{18–20} and hyperdynamics.²¹ With the majority of these techniques, one need not impose a reaction coordinate, but relevant order parameters that clearly demarcate the stable states must be extracted from the simulations. The calculation of rate constants, in most cases, is not trivial and typically involves gathering a vast number of short trajectories very many times from various intermediate hyperplanes between the initial and final states. Large-scale motions such as the folding of peptides^{22–24} and small proteins¹⁴ as well as allosteric transitions in oligomeric proteins²⁵ have been investigated with explicit solvent simulations. These processes have relatively lower but rough energy barriers and occur on the microsecond time scale. On the other hand, kinetic studies of more localized motions such as the isomerization of backbone dihedrals with large activation barriers that take place in nanoseconds to milliseconds have also been implemented in simple dipeptides.^{26,27} But compared to rotations around side-chain χ and backbone ϕ/ψ

torsions that involve single bonds, the *cis*–*trans* isomerization of the ω dihedral angle, i.e., rotation around the peptide bond with pseudo-double-bond character, has energy barriers in the range of 30 – $35 k_B T$ (~ 20 kcal/mol) and usually occurs on the time scale range of several hundred seconds.²⁸ Computational kinetic studies of such extremely slow processes have so far been possible only with methods based on Voter's hyperdynamics scheme.^{21,29} These studies aimed at obtaining either the relative rates of *cis*–*trans* isomerization in different peptides^{30,31} or absolute rates in implicit solvent.³² However, computing absolute kinetic rate constants from simulations in explicit water, which are comparable to experimental ones, has never been attempted.

In the hyperdynamics approach,^{21,33} which is based on TST, the original potential is raised by adding a bias potential only in the regions of energy basins but not the transition states, such that the barriers are lowered and the rate of escape from an energy basin is significantly increased. Thus, sampling of conformations on the modified potential is enhanced. The original equilibrium properties of the system can be retrieved after simply correcting for the bias. Similarly, reweighting the actual times of accelerated events yields the true transition rate. The advantage of this approach is that no prior knowledge of the reaction coordinate or potential energy minima or maxima (i.e., barriers) is required, and therefore it is possible to carry out unconstrained simulations in which one can actually observe the entire event without any discontinuity. In Voter's hyperdynamics scheme, the transition state regions are identified during the course of the simulation, which can become computationally very intensive and impractical for larger biomolecules. In the accelerated MD treatment proposed by Hamelberg et al.,²⁹ a novel expression for the bias potential is used, which eliminates the step of dynamically identifying transition states and increases the applicability of the method to large biological systems.³⁴ Due to improved sampling of conformational space, the accelerated MD approach has been shown to capture long time scale protein dynamics and therefore better reproduce NMR observables such as residual dipolar couplings, scalar J couplings,³⁵ and very recently also chemical shifts.³⁶ The bias potential $\Delta V(\mathbf{r}) = (E - V(\mathbf{r}))^2 / (\alpha + (E - V(\mathbf{r})))$ is continuous, non-negative, and added to the original potential $V(\mathbf{r})$ only when $V(\mathbf{r})$ falls lower than a preset boost energy E . Here, the parameter α and the boost energy E determine the level of acceleration of MD. However, the introduction of this bias potential should not alter the transition state regions if one desires original kinetics by reweighting time. It is very hard to fulfill this condition for hyper-dimensional surfaces since there are many saddle points that represent the transition state. This problem is illustrated in Figure 1, which shows a schematic representation of the original torsional potential projected along the peptide bond ω dihedral. Also shown are the modified potentials using two different values of E but the same α . When E is set below the transition state regions (red dotted line), the modified potential is not altered around the saddle points (i.e., $\Delta V(\mathbf{r}) = 0$). In the second case, when E is set to a much higher value (green dotted line), $\Delta V(\mathbf{r})$ does not remain nonzero at the transition state regions. Thus, reweighting time to obtain true kinetics will not be valid in the second case.

For most biomolecular events, although the second case will usually be encountered, the retrieval of equilibrium properties at conditions of zero bias will not be affected. However, estimation of true kinetic rate constants from accelerated MD will not be accurate. To overcome this problem, we recently proposed a new approach³⁷ in which accelerated MD is combined with Kramers'

rate theory^{38,39} in the overdamped regime. We successfully demonstrated that it is possible to retrieve the true rate of *cis*–*trans* isomerization in a simple model system of a peptide bond (*N*-methylacetamide) by carrying out simulations at various levels of acceleration.³⁷

The advantage of this method is that there is no need to reweight time, and hence, one need not worry about whether transition state regions are modified or not based on the choice of parameters that control the extent of accelerations.³⁷ From a series of accelerated MD simulations where E was kept constant and α was varied, the rates of escape from the *cis* well were obtained directly without any approximations, i.e., by observing the actual dynamics of the peptide bond with several barrier-crossing events. Using Kramers' rate theory in the overdamped limit, the true rates were then extrapolated to conditions of zero bias/acceleration (i.e., when $\alpha \rightarrow \infty$). Alternatively, the *cis* \rightarrow *trans* rate could be estimated from a single accelerated MD simulation. By solving the Smoluchowski equation, the effective diffusion coefficient on the modified landscape projected onto the ω dihedral could be calculated directly. If diffusion coefficients remained nearly constant with the changes in α , one could then use the effective diffusion coefficient on the modified landscape and the reweighted free energy profile to estimate the true *cis* \rightarrow *trans* rate from Kramers' theory. In our previous studies,³⁷ we validated our method by obtaining a close agreement between the true rates estimated from a series of accelerated MD and rates calculated from a long unbiased MD simulation. To make *cis*–*trans* isomerization accessible to straightforward MD, we neglected the solvent degrees of freedom and considerably lowered the potential energy barrier between the *cis* and the *trans* isomers in an artificial manner (i.e., by modifying the default AMBER parameters for the ω torsional force constants). Both accelerated and normal MD were then performed on this new potential in the absence of solvent.

In this work, we apply the same approach to obtain realistic kinetics of *cis*–*trans* isomerization in a simple but biologically relevant dipeptide solvated in explicit water. We compare the performance of the commonly used rigid nonpolarizable models of water, i.e., TIP3P⁴⁰ and SPC/E,⁴¹ in describing the *cis*–*trans* isomerization kinetics. Recently, we carried out the reparameterization of AMBER ω torsions by reproducing experimentally derived *cis*–*trans* equilibria (the value of the ω angle defines the *cis* and the *trans* states) and barrier heights from the accelerated MD of several model peptides.⁴² The present study will give us an opportunity to test whether the new improved parameters obtained from an equilibrium point of view will also be valid for the calculation of rate constants that are comparable to experimental estimates.

RESULTS AND DISCUSSION

From the biological point of view, studying the process of *cis*–*trans* isomerization, especially of prolyl peptidyl bonds, is very crucial for understanding the mechanisms of a number of biomolecular switches⁴³ and the ubiquitously found *cis*–*trans* isomerases.⁴⁴ Moreover, the distinctly slow phase in the folding of several proteins is found to be due to *cis*–*trans* isomerization.⁴⁵ We therefore chose the glycyl–prolyl peptide (Figure 2) to study the isomerization kinetics of the prolyl–peptidyl bond. Also, for this peptide, there is availability of experimental *cis*–*trans* isomerization rates²⁸ against which we could verify the true rates resulting from our accelerated MD studies.

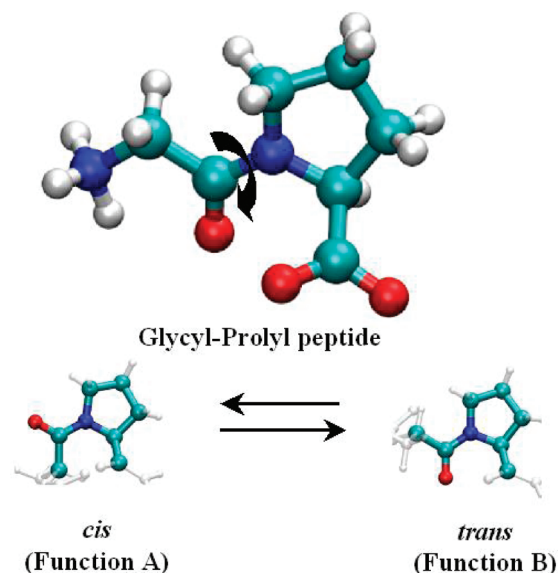


Figure 2. Glycyl–proline, the model system to study *cis*–*trans* isomerization kinetics of ω dihedral preceding the Pro residue. The lower panel shows that different functions are associated with the *cis* and the *trans* isomers in biomolecular switches.

At first, we simulated the solvated peptide with accelerated MD without any temperature regulation; i.e., we carried out microcanonical (NVE) runs for 1 μ s. For larger systems, the microcanonical ensemble is similar to the canonical one, and energy is conserved for runs of few hundred nanoseconds as the system acts as its own bath. However, for a small system such as the one used in this study (with a total of only \sim 1600–1900 atoms) and ran for very long times, the NVE ensemble resulted in very large drifts in temperature. The fluctuations in the potential and kinetic energy were not adequately dissipated such that a steady decrease in temperature was observed across a 1 μ s trajectory for each level of acceleration. Therefore, it was necessary to maintain the temperature during the course of MD runs using a thermostat. The temperature is determined by the average kinetic energy of the system, which in turn depends on the velocities of the particles. Since a thermostat maintains the temperature by either rescaling or modifying the velocities of the particles, its use affects the dynamics of the system and therefore is not recommended while studying the kinetics of a process. In a Langevin thermostat, the temperature is regulated by adjusting the velocity and kinetic energy of each particle.⁴⁶ This is achieved by adding a frictional force (γv) proportional to the velocity of each particle (i.e., each atom of the system) and a random white noise that follows a Gaussian distribution with zero mean and variance being a function of mass of the particle, frictional coefficient γ (or collision frequency), and the desired temperature. Collisions between the particles cause friction and thereby dampen the dynamics of the system. In the case of *cis*–*trans* isomerization where the barriers are very large, there is a clear separation of time scales. Therefore, the perturbation caused by the use of a Langevin thermostat may not significantly affect the kinetics of *cis*–*trans* isomerization as much as it may to systems with intrinsically much lower barriers and diffusive dynamics. It has been shown earlier from NVE simulations that the commonly used water models do not reproduce the dynamic properties of bulk water, namely, the self-diffusivity of water.⁴⁷ The self-diffusion coefficient of water in TIP3P is much larger than that from experiments, suggesting that dynamics of

Table 1. Self-Diffusion Coefficients of Different Water Models Obtained from MD Simulations

water model	thermostat	collision frequency, γ (ps^{-1})	self-diffusion coefficient of water calculated from MD simulations at 298.15 K ($\times 10^{-5} \text{ cm}^2/\text{s}$)
TIP3P	Langevin	1	5.067
TIP3P	Langevin	9	2.281
SPC/E	Langevin	1	2.299
TIP3P	Berendsen		4.557
SPC/E	Berendsen		2.338
experiments			2.300

biomolecules solvated in TIP3P water will be much faster. The benefit of using a Langevin thermostat is that one can fine-tune γ to calibrate the self-diffusion coefficient of water against experimentally measured values. One can then investigate biomolecular dynamics with this preset value of γ . We therefore calculated the self-diffusion coefficients $D_{\text{H}_2\text{O}}$ of TIP3P and SPC/E water from the Einstein relation⁴⁸ ($\lim_{t \rightarrow \infty} \langle |\mathbf{r}(t_0 + t) - \mathbf{r}(t_0)|^2 \rangle = 6D_{\text{H}_2\text{O}}t$, where \mathbf{r} is the position at time t of each oxygen and hydrogen atom and the brackets represent an average over all water atoms and initial time points t_0); i.e., from the slope of the plots of mean square displacement vs time in the region of 150–400 ps. A collision frequency of 9 ps^{-1} using TIP3P matched the computational (this work) estimate of water's self-diffusion constant to the empirical one. The Berendsen thermostat⁴⁹ rescales the atomic velocities to control the temperature to the desired value. In doing so, the conditions of the canonical ensemble are destroyed. Furthermore, velocity rescaling has been shown to predict water's self-diffusivity with significantly larger variation.⁴⁷ As can be seen from Table 1, the dynamics in TIP3P water with a Langevin ($\gamma = 1 \text{ ps}^{-1}$) or Berendsen thermostat is much faster than that expected from experiments. In spite of this fact, we carried out the simulations of *cis*–*trans* isomerization in all five combinations listed in Table 1 to investigate the effects of these commonly used water models and temperature-control protocols on peptide dynamics.

For each combination of water model and thermostat, we performed accelerated MD at five different values of α (20, 25, 30, 35, and 40 kcal/mol) and kept E constant. Further, for each value of α , simulations were carried out for as long as $1 \mu\text{s}$. We recorded the value of the ω angle preceding the Pro (Figure 1) at every time step (i.e., 2 fs) and calculated the number of transitions from the *cis* to the *trans* wells in either direction. We also measured the time the peptide spends in the *cis* well before every instance it transitions into the *trans* configuration and thus obtained a distribution of dwell times, $p(\tau)$. As done in earlier works,^{31,37} we then calculated the probability of survival, $S(t)$, in the *cis* well for time t and longer from $S(t) = \int_t^{+\infty} p(\tau) dt$. Figure 3 shows the survival probabilities at five different levels of acceleration of *cis*–*trans* isomerization in TIP3P water and a Langevin thermostat with $\gamma = 9 \text{ ps}^{-1}$.

The rates of isomerization from *cis* to *trans* were obtained from single exponential fits to the slow phases. The amplitudes of the fast phases, which represent the recrossings, reduce as the barriers become larger and relatively sharper with increasing α . According to Kramers' rate theory^{38,39} in one dimension, the rate of escape from a well over a barrier of height F_b in the

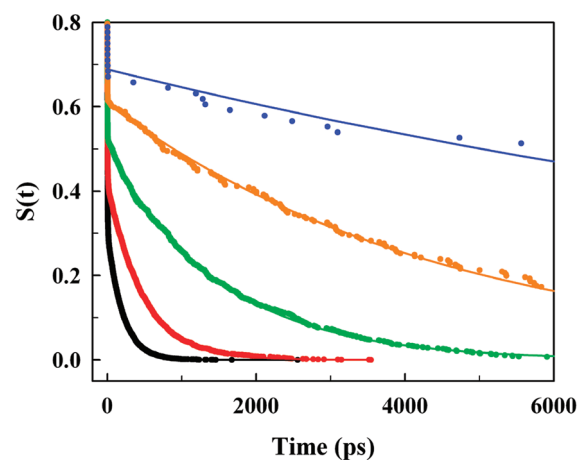


Figure 3. Decay of the probability of survival in the *cis* well for different levels of acceleration, i.e., $\alpha = 20$ kcal/mol (black), $\alpha = 25$ kcal/mol (red), $\alpha = 30$ kcal/mol (green), $\alpha = 35$ kcal/mol (orange), and $\alpha = 40$ kcal/mol (blue) in TIP3P water using a Langevin thermostat with $\gamma = 9 \text{ ps}^{-1}$. Boost energy, E , is set to 64 kcal/mol. Solid lines are mono-exponential fits to the slow phases.

overdamped limit is given by $k = (\omega_b \omega_c) / (2\pi \xi) \exp(-F_b / (k_B T))$, where ξ is the frictional rate and frequencies ω_c and ω_b are respectively the curvatures of the well and the barrier regions that are assumed to be roughly parabolic. The barrier height F_b varies with α as $F_b = F_0 - q / (\alpha + p)$,^{31,37} where F_0 is the free energy barrier height for the *cis* to *trans* transition on the reweighted (i.e., unbiased) one-dimensional free energy profile. p and q are constants from the fit of F_b versus α and are related to $\langle E - V(\mathbf{r}) \rangle$ and $\langle (E - V(\mathbf{r}))^2 \rangle$, respectively. Here $\langle \dots \rangle$ denotes an ensemble average over the configuration space. Substituting $\xi = (k_B T) / D$, where D represents the apparent diffusivity on a flat potential, $F_b = F_0 - q / (\alpha + p)$, expressing ω_b and ω_c in units of $(k_B T)^{1/2} / (\text{deg})$ and F_0 , α , q , and p in units of $k_B T$, and taking the natural log on both sides results in the linear relation: $\ln k_{c \rightarrow t}^\alpha / (\omega_b^\alpha \omega_c^\alpha) = \ln D^\alpha / (2\pi) - F_0 + q / (\alpha + p)$. Here, the subscript α denotes that the rates, curvatures, and apparent diffusion coefficients vary with the acceleration factor α .

We calculated the rate of (bidirectional) transition $k_{c \rightarrow t}^\alpha$ from the *cis* to the *trans* wells from accelerated MD simulations using five different values of α and the curvatures ω_c^α and ω_b^α from the respective unweighted free energy profiles. As done earlier,³⁷ assuming D^α to be constant, we then obtained the normal rate constants from the intercept of the plot of $\ln k_{c \rightarrow t}^\alpha / (\omega_b^\alpha \omega_c^\alpha)$ vs $q / (\alpha + p)$ and the curvatures $\omega_c^{\alpha \rightarrow \infty}$ and $\omega_b^{\alpha \rightarrow \infty}$ measured from the reweighted free energy profile along the ω angle. Figure 4 shows the plot of $\ln k_{c \rightarrow t}^\alpha / (\omega_b^\alpha \omega_c^\alpha)$ vs $q / (\alpha + p)$ for simulations performed in SPC/E and TIP3P water with a Langevin thermostat. For the range of α considered here, the peptide dynamics (i.e., rate of isomerization) was accelerated by a factor of 10^{10} to 10^{11} as compared to normal rates.

From $k_{c \rightarrow t}^\alpha$ obtained from simulations and the effective rate of escape, k_S^α , from the *cis* well calculated by solving Smoluchowski's diffusion equation on unweighted free energy profiles corresponding to different values of α , we estimated the apparent diffusion coefficients as $D^\alpha = k_{c \rightarrow t}^\alpha / k_S^\alpha$. For each water model, D^α did not change significantly with α ; i.e., $D_{\text{SPC/E}} = 9.8 (\pm 1.5) \times 10^{13} \text{ deg}^2/\text{s}$; $D_{\text{TIP3P}(\gamma=1)} = 12.2 (\pm 2.9) \times 10^{13} \text{ deg}^2/\text{s}$; $D_{\text{TIP3P}(\gamma=9)} = 8.3 (\pm 1.9) \times 10^{13} \text{ deg}^2/\text{s}$. This suggested that instead of a series one could perform only a single accelerated

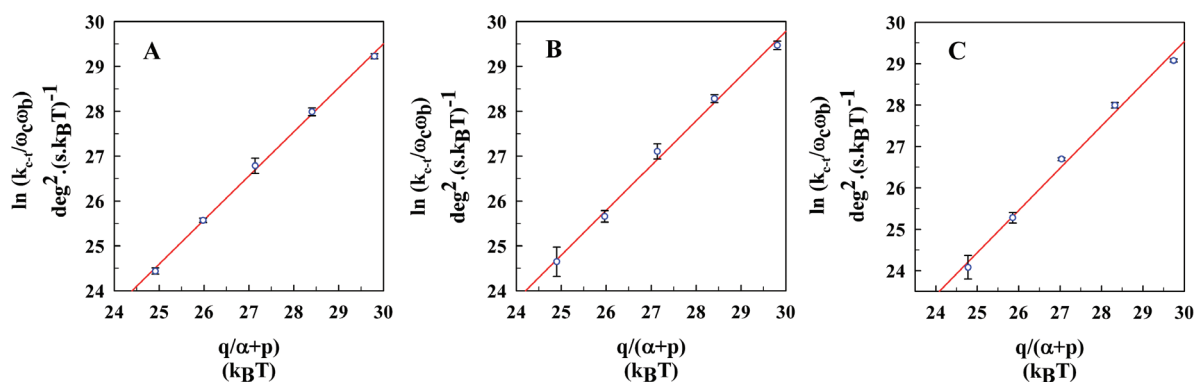


Figure 4. Kramers' plot, i.e., $\ln k_{c \rightarrow t}^{\alpha} / (\omega_b^{\alpha} \omega_c^{\alpha})$ vs $q / (\alpha + p)$ obtained from accelerated MD simulations using a Langevin thermostat in (A) SPC/E ($\gamma = 1 \text{ ps}^{-1}$), (B) TIP3P ($\gamma = 1 \text{ ps}^{-1}$), and (C) TIP3P ($\gamma = 9 \text{ ps}^{-1}$). Blue circles are data points from simulations at different values of α , and red lines are linear fits to data. Error bars obtained from three independent $1 \mu\text{s}$ runs are shown in black.

Table 2. Comparison of Rates of *cis* to *trans* Isomerization Obtained from Experiments and Simulations in Different Water Models and Thermostats^a

water model	thermostat	collision frequency, γ (ps^{-1})	normal rate constant (s^{-1})
TIP3P	Langevin	1	0.0137 ± 0.0093
TIP3P	Langevin	9	0.0054 ± 0.0047
SPC/E	Langevin	1	0.0134 ± 0.0045
TIP3P	Berendsen		0.0238 ± 0.0128
SPC/E	Berendsen		0.0208 ± 0.0109
experiments			0.0022 ± 0.0006

^a The normal rate constants and their errors listed above are calculated from three independent runs of $1 \mu\text{s}$ for each set of water model and thermostat.

MD and calculate the apparent diffusion coefficient on the unweighted free energy profile. In the case of a glycyL-prolyl peptide plugging, $D \approx 1e^{14} \text{ deg}^2/\text{s}$ and $\omega_c^{\alpha \rightarrow \infty}$, $\omega_b^{\alpha \rightarrow \infty}$, and F_b from the normal free energy profile in Kramers' rate equation yielded the rate of *cis-trans* isomerization (e.g., $\sim 0.001 \text{ s}^{-1}$ for TIP3P ($\gamma = 9 \text{ ps}^{-1}$)) very similar to those obtained from a series of accelerated MD simulations.

The average normal *cis* to *trans* rate constants corresponding to zero bias for each combination of water model and thermostat are listed in Table 2. Overall, the rates were overestimated by only an order of magnitude as compared to that measured from pD-jump experiments.²⁸ Such close agreement even with the use of an empirical fixed-charged force field was quite encouraging. This clearly showed that the newly revised AMBER parameters for ω torsions obtained from matching experimental barrier heights were appropriate also for reproducing the kinetics of *cis-trans* isomerization. Although a Berendsen thermostat is usually not the choice for temperature regulation and for providing a canonical distribution, in this case of *cis-trans* isomerization, it gives kinetic rates similar to those calculated using a Langevin thermostat. The rates of isomerization were slightly slower from simulations in a Berendsen thermostat but still within an order of magnitude of those from experiments.

Molecular dynamics simulations in the TIP3P water model and Langevin thermostat with a collision frequency of 9 ps^{-1} yielded the best agreement with experimental estimates of *cis-trans* isomerization rates. The calculated rates differed from

experimental ones only by a factor of 2, which was excellent considering the differences in conditions under which experiments [i.e., pD-jump in deuterated water in combination with ^{13}C or ^1H NMR²⁸] and computational simulations are performed. This suggested that the rates obtained from simulations could be matched more closely to experimentally measured ones by correcting the self-diffusivity of water. Thus, while using empirical force fields and imperfect water models for biomolecular simulation one could consider γ as a kinetic parameter and fine-tune it to obtain better agreement with experiments.

Both TIP3P and SPC/E water in combination with a Langevin thermostat ($\gamma = 1 \text{ ps}^{-1}$) gave very similar results (i.e., faster kinetics) in spite of the differences in the self-diffusion coefficients of water calculated with these models. When the collision frequency was increased to 9 ps^{-1} in TIP3P water, the kinetics became slower and much closer to experiments. This could be explained using the relation of diffusion coefficient and energetic roughness on a rough and effective 1-D potential proposed by Zwanzig:⁵⁰ $D = D_0 \exp[-(\varepsilon/k_B T)^\theta]$. Here, D_0 is the diffusion coefficient on the smooth potential and ε is the roughness, which follows a random distribution reflected in $\theta = 2$. SPC/E water is rougher than TIP3P [$\varepsilon_{\text{SPC/E}} = 1.01$, $\varepsilon_{\text{TIP3P}} = 0.87^{51}$], which was estimated from the temperature-dependent part of the above equation that predominantly contributes to the differences in the effective diffusion coefficients at lower temperatures. At ambient temperature, i.e., 298 K, the effects of roughness of the two water models (coming from the exponential quantity in the above equation) already accounts for a factor of 2 (i.e., $D_{\text{TIP3P}} \approx 2D_{\text{SPC/E}}$). The temperature-independent part, D_0 , seems more to be controlled by the frictional drag, which can be adjusted by the collision frequency in the Langevin-thermostat setup. Hence, the same values of γ for TIP3P and SPC/E yielded similar normal rate constants for *cis* to *trans* transitions.

CONCLUSIONS

Accurate estimation of rate constants from accelerated MD requires no alteration of potential in the transition state regions upon addition of the bias. However, for biomolecular processes, the above condition is often not satisfied. To circumvent this problem, we applied Kramers' rate theory in the high friction regime to the results of accelerated MD to obtain normal rate constants corresponding to dynamics on the original potential. We have successfully demonstrated our approach by simulating

the kinetics of *cis* to *trans* isomerization of a Pro-containing dipeptide in explicit water. By varying the extent of acceleration in accelerated MD, the speed of *cis*–*trans* isomerization was increased from hundreds of seconds to tens of nanoseconds. This allowed us to actually observe the *cis* to *trans* transitions several times and estimate normal rates from linear extrapolation to conditions of zero bias. The normal rates obtained from accelerated MD were in notable agreement with those from experiments, given the use of empirical force fields and water models. For the first time, our work kinetically validated the AMBER torsional parameters. We have further suggested that the kinetics of activated processes could be better reproduced by correcting the self-diffusivity of water models regularly used in biosimulations. This could be achieved by fine-tuning the collision frequency in a Langevin thermostat. The present work therefore provides the encouragement to investigate the kinetics of activated processes in larger and more complex systems in explicit water.

COMPUTATIONAL DETAILS

Using MD simulations, we investigated the *cis*–*trans* isomerization in the zwitterionic form of the glycyl–proline dipeptide. We used the AMBER 10.0 suite of programs⁵² with the modified version⁵³ of the parm99 (ff99SB) force field⁵⁴ and the revised parameter for the ω torsional angles (i.e., $V_2/2 (X-C-N-X) = 14$ kcal/mol).⁴² With the *xleap* program, the peptide was solvated in a cubic periodic box of dimensions 28 Å and filled with either ~600 TIP3P⁴⁰ or ~500 SPC/E⁴¹ water molecules, which were placed up to 9 Å away from the peptide. The system was equilibrated in the NPT ensemble at 1 bar of pressure and a temperature of 298.15 K. The SHAKE⁵⁵ algorithm (with a tolerance of 0.0001) was used to constrain bonds involving hydrogen atoms, and particle mesh Ewald⁵⁶ summation was used for long-range interactions. Short-range nonbonded interactions were calculated with a cutoff of 9 Å with the nonbonded pair list updated every 20 fs. The *pmemd* module of AMBER 10.0 was modified to carry out production runs with accelerated MD. Throughout the simulations, the temperature was regulated with a Langevin⁴⁶ or Berendsen⁴⁹ thermostat while the reference pressure was maintained by coupling to an external bath using a coupling constant of 1 ps. A time step of 2 fs was used to integrate Newton's equations of motions. For accelerated MD, the boost energy E was set to 50 kcal/mol above the average total torsional energy of ~14 kcal/mol. Three independent production MD runs were carried out for 1 μ s (a total of 3 μ s) for each of the five different values of α (i.e., 20, 25, 30, 35, and 40 kcal/mol). This was done for each combination of water model and thermostat, summing up the total simulation time to $3 \times 5 = 15 \mu$ s. From each production run at each level of acceleration, $k_{c \rightarrow t}^\alpha$ was obtained and the normal rate constant was calculated from a linear regression of $\ln k_{c \rightarrow t}^\alpha / (\omega_b^\alpha \omega_c^\alpha)$ vs $q/(\alpha + p)$ where the average values of p and q were used. Table 2 lists the average rates with errors obtained from the above procedure applied to all three runs. In Figure 4, the average $\ln k_{c \rightarrow t}^\alpha / (\omega_b^\alpha \omega_c^\alpha)$ and its errors obtained from three runs are shown.

For the calculation of the self-diffusion coefficient of water, the simulations were carried out with a periodic box of dimensions $19 \times 19 \times 19 \text{ \AA}^3$ and filled with either 362 TIP3P or 364 SPC/E water molecules. After reaching the density of 0.9855 g/cm³ with equilibration, the MD run was carried out for 1 ns at 298.15 K using a Langevin thermostat. For SPC/E, the collision frequency

γ was set to 1 ps^{-1} , whereas for TIP3P, a γ of 1 as well as 9 ps^{-1} was used in two different simulations.

AUTHOR INFORMATION

Corresponding Author

*Tel.: +1 404 413 5564. Fax: +1 404 413 5505. E-mail: dhamelberg@gsu.edu.

ACKNOWLEDGMENT

This work is supported in part by the National Science Foundation CAREER MCB-0953061 and Georgia Cancer Coalition. This work was also supported by Georgia State's IBM System p5 supercomputer, acquired through a partnership of the Southeastern Universities Research Association and IBM, supporting the SURAgri initiative.

REFERENCES

- (1) Eyring, H. *J. Chem. Phys.* **1935**, *3*, 107–115.
- (2) Laidler, K. J.; King, M. C. *J. Phys. Chem.* **1983**, *87*, 2657–2664.
- (3) Chandler, D. *J. Chem. Phys.* **1978**, *68*, 2959–2970.
- (4) Grubmuller, H. *Phys. Rev. E* **1995**, *52*, 2893–2906.
- (5) Huber, T.; Torda, A. E.; Van Gunsteren, W. F. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 695–708.
- (6) Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187–199.
- (7) Ghosh, I.; McCammon, J. A. *Biophys. J.* **1987**, *51*.
- (8) Northrup, S. H.; Pear, S. M.; Lee, C. Y.; McCammon, J. A.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 4035–4039.
- (9) Elber, R. *Curr. Opin. Struct. Biol.* **2005**, *15*, 151–6.
- (10) Dellago, C.; Bolhuis, P. G.; Csajka, F. S.; Chandler, D. *J. Chem. Phys.* **1998**, *108*, 1964–1977.
- (11) Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L. *Annu. Rev. Phys. Chem.* **2002**, *53*, 291–318.
- (12) Van Erp, T. S.; Moroni, D.; Bolhuis, P. G. *J. Chem. Phys.* **2003**, *118*, 7762–7774.
- (13) Moroni, D.; Bolhuis, P. G.; van Erp, T. S. *J. Chem. Phys.* **2004**, *120*, 4055–65.
- (14) Pande, V. S.; Baker, I.; Chapman, J.; Elmer, S. P.; Khaliq, S.; Larson, S. M.; Rhee, Y. M.; Shirts, M. R.; Snow, C. D.; Sorin, E. J.; Zagrovic, B. *Biopolymers* **2003**, *68*, 91–109.
- (15) Faradjian, A. K.; Elber, R. *J. Chem. Phys.* **2004**, *120*, 10880–9.
- (16) Swope, W. C.; Pitera, J. W.; Suits, F. *J. Phys. Chem. B* **2004**, *108*, 6571–6581.
- (17) Weinan, E.; Weiqing, R.; Vanden-Eijnden, E. *J. Phys. Chem. B* **2005**, *109*, 6688–6693.
- (18) Allen, R. J.; Warren, P. B.; ten Wolde, P. R. *Phys. Rev. Lett.* **2005**, *94*.
- (19) Allen, R. J.; Frenkel, D.; ten Wolde, P. R. *J. Chem. Phys.* **2006**, *124*, 024102.
- (20) Borrero, E. E.; Escobedo, F. A. *J. Phys. Chem. B* **2009**, *113*, 6434–6445.
- (21) Voter, A. F. *J. Chem. Phys.* **1997**, *106*, 4665–4677.
- (22) Bolhuis, P. G. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 12129–12134.
- (23) Kuczera, K.; Jas, G. S.; Elber, R. *J. Phys. Chem. A* **2009**, *113*, 7461–7473.
- (24) Swope, W. C.; Pitera, J. W.; Suits, F.; Pitman, M.; Eleftheriou, M.; Fitch, B. G.; Germain, R. S.; Rayshubski, A.; Ward, T. J. C.; Zheshkov, Y.; Zhou, R. *J. Phys. Chem. B* **2004**, *108*, 6582–6594.
- (25) Elber, R. *Biophys. J.* **2007**, *92*, L85–L87.
- (26) Bolhuis, P. G.; Dellago, C.; Chandler, D. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 5877–5882.
- (27) De Oliveira, C. A. F.; Hamelberg, D.; McCammon, J. A. *J. Chem. Phys.* **2007**, *127*.
- (28) Grathwohl, C.; Wuthrich, K. *Biopolymers* **1981**, *20*, 2623–2633.

- (29) Hamelberg, D.; Mongan, J.; McCammon, J. A. *J. Chem. Phys.* **2004**, *120*, 11919–11929.
- (30) Hamelberg, D.; McCammon, J. A. *J. Am. Chem. Soc.* **2005**, *127*, 13778–13779.
- (31) Hamelberg, D.; Shen, T.; McCammon, J. A. *J. Am. Chem. Soc.* **2005**, *127*, 1969–1974.
- (32) Hamelberg, D.; Shen, T.; McCammon, J. A. *J. Chem. Phys.* **2005**, *122*, 241103.
- (33) Voter, A. F. *Phys. Rev. Lett.* **1997**, *78*, 3908–3911.
- (34) Hamelberg, D.; McCammon, A. *J. Am. Chem. Soc.* **2009**, *131*, 147–152.
- (35) Markwick, P. R. L.; Bouvignies, G.; Blackledge, M. *J. Am. Chem. Soc.* **2007**, *129*, 4724–4730.
- (36) Markwick, P. R. L.; Cervantes, C. F.; Abel, B. L.; Komives, E. A.; Blackledge, M.; McCammon, J. A. *J. Am. Chem. Soc.* **2010**, *132*, 1220–1221.
- (37) Xin, Y.; Doshi, U.; Hamelberg, D. *J. Chem. Phys.* **2010**, *132*, 224101.
- (38) Hanggi, P.; Talkner, P.; Borkovec, M. *Rev. Mod. Phys.* **1990**, *62*, 251–342.
- (39) Kramers, H. A. *Physica (Utrecht)* **1940**, *7*, 284–304.
- (40) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (41) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269–6271.
- (42) Doshi, U.; Hamelberg, D. *J. Phys. Chem. B* **2009**, *113*, 16590–16595.
- (43) Lu, K. P.; Finn, G.; Lee, T. H.; Nicholson, L. K. *Nat. Chem. Biol.* **2007**, *3*, 619–29.
- (44) Fanghanel, J.; Fischer, G. *Front. Biosci.* **2004**, *9*, 3453–78.
- (45) Brandts, J. F.; Halvorson, H. R.; Brennan, M. *Biochemistry* **1975**, *14*, 4953–63.
- (46) Adelman, S. A.; Doll, J. D. *J. Chem. Phys.* **1976**, *64*, 2375–2388.
- (47) Mark, P.; Nilsson, L. *J. Phys. Chem. A* **2001**, *105*, 9954–9960.
- (48) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford Science Publications: New York, 2007.
- (49) Berendsen, H. J. C.; Postma, W. F.; Van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- (50) Zwanzig, R. *Proc. Natl. Acad. Sci. U. S. A.* **1988**, *85*, 2029–30.
- (51) Johnson, Q.; Doshi, U.; Shen, T.; Hamelberg, D. *J. Chem. Theory Comput.* **2010**, *6*, 2591–2597.
- (52) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, M.; Walker, R. C.; Zhang, W.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman, P. A. *AMBER 10*; University of California: San Francisco, CA, 2008.
- (53) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65*, 712–25.
- (54) 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.
- (55) Ryckaert, J.; Cicotti, G.; Berendsen, H. J. *Comp. Phys. Phys.* **1977**, *23*, 327–341.
- (56) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089–10092.

Exploiting Configurational Freezing in Nonequilibrium Monte Carlo Simulations

Paolo Nicolini,[†] Diego Frezzato,[‡] and Riccardo Chelli^{*,†,¶}

[†]Dipartimento di Chimica, Università di Firenze, Via della Lastruccia 3, I-50019 Sesto Fiorentino, Italy

[‡]Dipartimento di Scienze Chimiche, Università di Padova, Via Marzolo 1, I-35131 Padova, Italy

[¶]European Laboratory for Nonlinear Spectroscopy (LENS), Via Nello Carrara 1, I-50019 Sesto Fiorentino, Italy

ABSTRACT: To achieve acceptable accuracy in fast-switching free energy estimates by Jarzynski equality [*Phys. Rev. Lett.* **1997**, *78*, 2690] or Crooks fluctuation theorem [*J. Stat. Phys.* **1998**, *90*, 1481], it is often necessary to realize a large number of externally driven trajectories. This is basically due to inefficient calculation of path-ensemble averages arising from the work dissipated during the nonequilibrium paths. We propose a computational technique, addressed to Monte Carlo simulations, to improve free energy estimates by lowering the dissipated work. The method is inspired by the dynamical freezing approach, recently developed in the context of molecular dynamics simulations [*Phys. Rev. E* **2009**, *80*, 041124]. The idea is to limit the configurational sampling to particles of a well-established region of the sample (namely, the region where dissipation is supposed to occur), while leaving fixed (frozen) the other particles. Therefore, the method, called *configurational freezing*, is based on the reasonable assumption that dissipation is a local phenomenon in single-molecule nonequilibrium processes, a statement which is satisfied by most processes, including folding of biopolymers, molecular docking, alchemical transformations, etc. At variance with standard simulations, in configurational freezing simulations the computational cost is not correlated with the size of the whole system, but rather with that of the reaction site. The method is illustrated in two examples, i.e., the calculation of the water to methane relative hydration free energy and the calculation of the potential of mean force of two methane molecules in water solution as a function of their distance.

1. INTRODUCTION

In the framework of methods devised for estimating free energy differences, an interesting scenario has been disclosed by two nonequilibrium work relations, the Jarzynski equality¹ (JE) and the Crooks fluctuation theorem^{2,3} (CFT). These theorems relate the free energy difference between two thermodynamic states to the external work performed in an ensemble of realizations switching the system between such states. The switching procedure is accomplished through an external control parameter correlated with some collective coordinate of the system (e.g., an interatomic distance, a torsional angle, the morphing coordinate in alchemical transformations, etc.). In the past decade, JE and CFT have found wide application in biophysical research to investigate the mechanical and elastic single-molecule properties of important biological molecules such as DNA,⁴ RNA,^{5,6} and proteins.⁷

In computer experiments, the control parameter acts on the system by means of an externally driven force. Computer simulations based on this type of technique are typically known as steered molecular dynamics^{8,9} (MD) or steered Monte Carlo^{10,11} (MC) simulations. Specifically, in order to estimate free energy differences using JE, one equilibrium simulation must first be performed by constraining the control parameter to the value of one thermodynamic state. During this simulation, microstates (atomic coordinates and momenta in MD simulations and atomic coordinates in MC simulations) are recorded at regular time intervals. In a second phase, steered MD/MC simulations are performed starting from the saved microstates, thus realizing nonequilibrium paths where the control parameter is driven to the value of the second thermodynamic state with the time schedule common to all realizations. In

CFT-based calculations, the protocol described above must also be applied to the backward direction of the process. Clearly, the schedule of the control parameter must be reverse in time.

In such calculations, it is of basic importance to improve the path-sampling efficiency, which globally increases by lowering the work dissipated during the realizations. To this aim, several approaches have been developed. They include biased path sampling,^{12–15} generation of non-Hamiltonian equations of motion,¹⁶ and optimal protocol strategies.¹⁷ A significant speed-up was also obtained by MD simulations with large time steps.¹⁸ A limitation of the methods listed above lies in their drastic dependence on the sample size. Considering that single-molecule fast-switching experiments perturb the system in a well-defined region of space around the reaction site, most molecules far from this site (typically the solvent molecules) may not be involved in the dissipation process, thus persisting in their state of equilibrium. This implies that any relevant worsening should be observed in the performances of nonequilibrium methods if these molecules were not accounted for in sampling schemes, simply because their effect on dissipation is negligible. As an example, we may think of the calculation of the binding free energy of two solvated molecules. In such a case, the dynamics of the solvent molecules far from the center of mass of the two target molecules is irrelevant for dissipation. Another example could be the calculation of the potential of mean force^{19,20} (PMF) related to processes of molecular traffic in transport proteins. Here, the reaction/dissipation site can be localized

Received: October 4, 2010

Published: February 09, 2011

around the protein channel. In spite of the fact that molecules far from the reaction site *do not affect* dissipation significantly, they *do affect* the overall cost of the computation because, in any case, all interparticle forces must be calculated to evolve the system. In this regard, a methodology for improving fast-switching free energy estimates in large systems via JE or CFT was recently proposed in the context of MD simulations by some of us.²¹ The strategy is based on the dynamical freezing of a subset of particles which are supposed to not be involved in the dissipation, while leaving the particles near the reaction site dynamically active. This freezing is accomplished with a synchronous scaling of the masses and velocities of the involved particles by keeping their individual kinetic energy unchanged. Such an approach allows a less frequent calculation of the forces between dynamically frozen particles, with a significant lowering of the simulation cost. By updating the list of particles belonging to dynamically frozen and unfrozen regions at regular time intervals, it has been shown that it is possible to design algorithms independent of the dynamical evolution of the system.

In this article, we extend the idea of dynamical freezing to MC simulations. Since in MC simulations the sampling is configurational rather than dynamical, we term the method *configurational freezing* (CF). In the context of nonequilibrium MC simulations, configurational freezing is synonymous with zero probability of selecting a particle for a trial move. In particular, a particle can be chosen for a trial move only if its distance from the reaction site is smaller than a prior established value, whereas selection of particles that do not meet this condition is skipped. Once a particle is selected, the trial moves that leave the particle within the threshold distance from the reaction site are accepted with a probability that satisfies the detailed balance condition, the other trial moves being rejected. With respect to standard MC simulations, CF allows enrichment of the sampling in the dissipation region. It is evident that CF does not preserve the ergodicity, because sampling occurs only in a limited region of the phase space. On the other hand, ergodicity is not a requirement for JE and CFT.² However, it is important to stress that the initial microstates must be sampled at equilibrium, and therefore the preliminary simulation(s), aimed at generating such microstates, must be carried out without applying CF.

In fast-switching CF simulations, the arbitrariness of sampling restricted phase-space regions does not introduce any approximation in the method (in fact, as we will show, the validity of CFT and JE is preserved). Actually, the tunability of CF in terms of the sampled phase space is used to optimize the calculation, i.e., to reduce the number of realizations needed to get accurate free energy estimates using nonequilibrium work theorems. This means that CF could be safely applied even if most particles entering in dissipation mechanisms were ruled out from the sampling. This aspect can be better understood considering the limit case of CF where all of the particles not correlated to the change of the control parameter are frozen. The ensuing algorithm would lead to switching realizations infinitely fast, and hence the JE applied to the instantaneous work samples would correspond to the known free energy perturbation method.²² If the initial microstates of the instantaneous switching realizations were picked randomly from equilibrium simulations of the two end states, then the CFT²³ would be equivalent to the Bennett method.²⁴

As a test case, we report the calculation of the water to methane relative hydration free energy using alchemical transformations. This system has been widely used to assess the performances of various free energy methods.^{25,26} In particular, we estimate the simulation speedup by using CF in comparison

to the standard fast-switching approach, where the phase-space sampling of the whole system is performed. It is clear that CF is well suited to such a type of system, because the reaction/dissipation site is easily identified as the volume of the sphere centered on the molecule subject to the alchemical transformation. The only variable to be defined in the algorithm is the radius of the sphere which, while not affecting the validity of the nonequilibrium work theorems, does affect the overall efficiency of the method.

As a further example, we report on the calculation of the PMF of two methane molecules in water solution as a function of their distance. Also this system has been used as a benchmark in molecular modeling.^{27–29} This case is however slightly more complex, because a high-dissipation region cannot be identified easily. A possible strategy could be to choose the center of mass of the methane molecules as the center of a sphere within which solvent molecules are not frozen. This approach would not differ conceptually from the previous case. However, following ref 21, we have opted to define the reaction/dissipation region as the one corresponding to the union of the spheres centered on the methane molecules.

The outline of the article follows. In section 2.1, we report the demonstration of the JE and CFT in the context of MC simulations.² In section 2.2, we present and justify the CF algorithm. Technical details on the MC simulations and on the systems are given in section 3, while the simulation results are reported and discussed in section 4. Concluding remarks can be found in section 5.

2. THEORY

2.1. Jarzynski Equality and Crooks Fluctuation Theorem in Monte Carlo Sampling Schemes. Since CF is addressed to the application of JE and CFT in MC simulations, here we report the demonstration of JE and CFT proposed by Crooks² for such a type of sampling schemes. Actually, the basic assumptions are rather general, i.e., that the dynamics of the system is Markovian and microscopically reversible (principle of detailed balance). The former condition ensures that the system is memory-less, while the latter ensures that the system is time reversible and that the equilibrium probability of the microstates is canonically distributed.

Suppose that the system is canonical with temperature T . At step t , the microstate of the system is defined by the vector x_t specifying the positions of all particles and by the externally controlled parameter λ_t . In MC simulations, x_t and λ_t evolve with discrete steps. Therefore, the driven evolution of the system from the initial microstate, $\{x_0, \lambda_0\}$, to the final microstate, $\{x_\tau, \lambda_\tau\}$, can be represented as a sequence of microstates where moves of the particles and of the control parameter are made on alternate steps as

$$\Gamma_F \equiv \{x_0, \lambda_0\} \rightarrow \{x_0, \lambda_1\} \rightarrow \{x_1, \lambda_1\} \dots \{x_{\tau-1}, \lambda_{\tau-1}\} \rightarrow \{x_\tau, \lambda_\tau\} \quad (1)$$

The previous sequence of microstates defines completely a driven trajectory during which work is performed on the system and heat is exchanged between the system and heat reservoir. In particular, work is done when the control parameter is moved at a fixed configuration, while heat is exchanged when the particle positions evolve at a fixed control parameter. The total work performed on the system, W , and the total heat exchanged with

the reservoir, Q , are

$$W = \sum_{t=0}^{\tau-1} [E(x_t, \lambda_{t+1}) - E(x_t, \lambda_t)] \quad (2)$$

$$Q = \sum_{t=1}^{\tau} [E(x_t, \lambda_t) - E(x_{t-1}, \lambda_t)] \quad (3)$$

where $E(x_t, \lambda_t)$ is the energy of the microstate $\{x_t, \lambda_t\}$. The reverse path, Γ_R , conjugated to the forward path specified in eq 1 is simply defined by the reverse sequence of microstates. For such a path, work and heat are the negative of the forward path direction. If we assume that the evolution of the system is Markovian, then the probability of making a transition between two states depends only on the state of the system at the initial step. Thus, the probability of following the forward path Γ_F through phase space given the initial state, $\{x_0, \lambda_0\}$, and the sequence of the control parameter, $\{\lambda_0, \lambda_1, \dots, \lambda_\tau\}$, can be split into a product of conditional probabilities as follows:

$$P[\Gamma_F|x_0, \lambda_0] = \prod_{t=1}^{\tau} P[x_t - 1, \lambda_t \rightarrow x_t, \lambda_t | x_{t-1}, \lambda_t] \quad (4)$$

In the previous equation, we have implicitly considered that the moves of the control parameter are performed with a probability equal to one. In an analogous way, we can define the probability of the reverse path Γ_R as

$$P[\Gamma_R|x_\tau, \lambda_\tau] = \prod_{t=1}^{\tau} P[x_{t-1}, \lambda_t \leftarrow x_t, \lambda_t | x_t, \lambda_t] \quad (5)$$

The single steps are assumed to be microscopically reversible and therefore obey the detailed balance condition for all fixed values of λ .

$$\frac{P[x_t - 1, \lambda_t \rightarrow x_t, \lambda_t | x_{t-1}, \lambda_t]}{P[x_t - 1, \lambda_t \leftarrow x_t, \lambda_t | x_t, \lambda_t]} = \frac{e^{-\beta E(x_t, \lambda_t)}}{e^{-\beta E(x_{t-1}, \lambda_t)}} \quad (6)$$

where $\beta^{-1} = k_B T$, with k_B being Boltzmann's constant. Equations 3–6 allow writing the ratio between $P[\Gamma_F|x_0, \lambda_0]$ and $P[\Gamma_R|x_\tau, \lambda_\tau]$ as

$$\frac{P[\Gamma_F|x_0, \lambda_0]}{P[\Gamma_R|x_\tau, \lambda_\tau]} = \prod_{t=1}^{\tau} \frac{e^{-\beta E(x_t, \lambda_t)}}{e^{-\beta E(x_{t-1}, \lambda_t)}} = e^{-\beta Q} \quad (7)$$

If we also specify that both forward and reverse paths start from equilibrium distributions, then the ratio between the total probabilities of the Γ_F and Γ_R trajectories becomes

$$\frac{P(x_0, \lambda_0)P[\Gamma_F|x_0, \lambda_0]}{P(x_\tau, \lambda_\tau)P[\Gamma_R|x_\tau, \lambda_\tau]} = e^{\beta[W - F(\lambda_\tau) + F(\lambda_0)]} \quad (8)$$

where $F(\lambda)$ is the Helmholtz free energy given the value λ of the control parameter and W is the work done on the system in the forward path (eq 2). In going from eq 7 to eq 8, we have considered the expression of the probability for a canonical distribution

$$P(x_t, \lambda_t) = \frac{e^{-\beta E(x_t, \lambda_t)}}{e^{-\beta F(\lambda_t)}} \quad (9)$$

and the first law of thermodynamics, $E(x_\tau, \lambda_\tau) - E(x_0, \lambda_0) = W + Q$ (this is simply obtained by summing eq 2 to eq 3). Equation 8 is the known expression of CFT relating the ratio of the probabilities of observing a driven path and its reverse path to the Helmholtz free energy difference between the two states and to the work W

performed on the system during the forward path. In the case of constant-pressure constant-temperature conditions, the Helmholtz free energy is replaced by the Gibbs free energy.³⁰

Given a fixed sequence of the external control parameter, e.g., $\{\lambda_0, \lambda_1, \dots, \lambda_\tau\}$, an infinite number of paths of the type of eq 1 can be observed. Indicating with $P(x_0, \lambda_0)P[\Gamma_F|x_0, \lambda_0]$ the probability of observing a given path Γ_F (see eq 8) and with W the associated work (eq 2), we can write the path-ensemble average of the function $\exp(-\beta W)$ as

$$\langle e^{-\beta W} \rangle_{\text{all}\Gamma_F} = \sum_{\text{all}\Gamma_F} \exp(-\beta W) P(x_0, \lambda_0) P[\Gamma_F|x_0, \lambda_0] \quad (10)$$

This average over the forward paths can be changed to an average over the reverse conjugated paths using eq 8

$$\begin{aligned} \langle e^{-\beta W} \rangle_{\text{all}\Gamma_F} &= \sum_{\text{all}\Gamma_R} e^{-\beta \Delta F} P(x_\tau, \lambda_\tau) P[\Gamma_R|x_\tau, \lambda_\tau] \\ &= e^{-\beta \Delta F} \end{aligned} \quad (11)$$

where $\Delta F = F(\lambda_\tau) - F(\lambda_0)$ and the sum runs over all Γ_R paths conjugated to Γ_F . The last step follows because the free energy difference is path independent, and because probabilities are normalized. Equation 11 is the JE for the forward direction of the process. Analogous derivation leads to a very similar expression for the reverse direction.

2.2. The Configurational Freezing Algorithm. In this section, we first illustrate the basic aspects of the CF algorithm and describe its implementation in MC simulations. Then we provide a theoretical justification of the method to the light of the requirements needed for JE and CFT to be valid (see section 2.1):

- (1) First of all, the initial microstates of the driven realizations are generated. As stated in the Introduction, these microstates can be picked at regular step intervals³¹ from an equilibrium MC simulation fixing the control parameter, λ , to the value of a state of interest (λ_0 in section 2.1). In order to make realizations in the backward direction, another equilibrium MC simulation must be performed fixing λ to the value of the second state (λ_τ in section 2.1). CF does not enter into play at this stage, and hence the generation of the initial microstates is made as in standard fast-switching numerical experiments.
- (2) Once an appropriate number of initial microstates is obtained, we must define a region of space where the external perturbation, arising from the driven change of the control parameter, is supposed to be localized. During steered MC simulations, only the particles found in this region will be selected for trial moves, and any attempted move bringing the particles out of the established region will be rejected systematically. Therefore, the aim of the algorithm is to relax the system exactly where perturbation occurs, thus lowering the global dissipated work. Note that accurate identification of such a region is not strictly necessary, since the validity of JE and CFT is not conditioned to it. On the other hand, selecting a “good mobility region” would lead to a computational gain because, for a given number of MC trial moves, the accuracy of CFT and JE in recovering free energy differences increases by lowering the dissipated work. We could summarize the idea by stating that in fast-switching single-molecule experiments the motion of atoms far from the reaction site is immaterial.

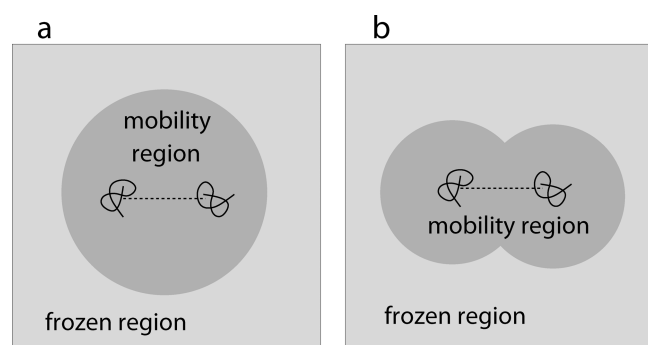


Figure 1. Schematic representation of configurational freezing applied to the binding process of two molecules in a fluid. *Panel a:* Fixed mobility-sphere approach. The mobility region (in dark gray) is kept fixed during the driven simulations. *Panel b:* multiple mobility-sphere approach. The mobility region corresponds to the union of the mobility spheres centered on the molecules. For the sake of clarity, in both cases, the solvent particles are represented as a continuum (light gray plus dark gray).

The simplest way to define the mobility region is to identify a sphere, held fixed during the system evolution, including most particles (atoms, molecules, molecular fragments, etc.) involved in the driven process. If the radius of the sphere is large enough, then we may reasonably be confident that an effective mobility region has been chosen. However, the radius cannot be very large for not including too many particles unaffected by the external perturbation. In analogy with the dynamical freezing technique,²¹ we term this sphere the “mobility sphere”. A schematic illustration of this approach applied to the binding process of two molecules in a fluid is represented in Figure 1a.

Although the previous scheme can be implemented easily in MC programs, it does not guarantee good efficiency, especially when the high-dissipation region has no spherical symmetry. Typical examples are the unfolding and refolding processes of biopolymers. A more suitable criterion for treating asymmetric driven processes is to select more particles as centers of mobility spheres. Thus, the overall mobility region corresponds to the union of the single mobility spheres. For instance, some atoms of the involved biopolymer could be chosen as centers of mobility spheres in driven folding processes. At variance with the criterion illustrated above, this choice implies that the mobility region can change in shape and size during the system evolution,³² because the particles associated with the mobility spheres can in principle be moved. Therefore, the list of mobile and frozen particles must be updated often during the dynamics, thus making the current scheme computationally more expensive. However, for very large systems and strongly asymmetric driven processes, efficiency can be largely recovered because the mobility regions are localized there where dissipation occurs. A schematic representation of the multiple mobility-sphere approach applied to the binding process of two molecules in a fluid is shown in Figure 1b.

- (3) Before starting forward and backward realizations, the driving sequence of the control parameter, $\{\lambda_0, \lambda_1, \dots, \lambda_\tau\}$, and the sequence of MC moves that realize the free evolution of the system (the step $\{x_{i-1}, \lambda_i\} \rightarrow \{x_i, \lambda_i\}$ in the notation of section 2.1) are established. The latter may occur with an arbitrary scheme such as single-particle moves, a sequence of single-particle moves, or even

cluster moves. As already stated, the sampling criteria are that (i) only the particles in the mobility spheres are subject to trial moves and, among these trial moves, (ii) only the ones leaving the particles inside the mobility spheres are accepted with a probability that preserves the detailed balance, the other moves being rejected. Note that, as we will prove later, the latter condition is not independent, but derives from the combination of the former and the detailed balance. If necessary, after each MC move, be it a λ or a system move, the list of particles inside the mobility spheres is updated.

In order to justify CF, it should be proved that the algorithm preserves the detailed balance condition, i.e., the microreversibility, and that the system evolution is Markovian. These arguments will be addressed in the remaining part of the current section. We denote the probability of generating a microstate j from a microstate i as α_{ij} (stochastic matrix). By definition, both microstates are characterized by the same value of λ . Therefore, their canonical probabilities, indicated here as p_i and p_j , are given by eq 9 with identical free energy $F(\lambda)$. Without a loss of generality, suppose that the microstates i and j differ only for the position of one particle, say the particle n , or, in other words that a MC move involves the change of position of only one particle. The matrix element α_{ij} can be decomposed as the product of two terms, i.e., the probability of selecting the particle n and the conditional probability of generating the move of the particle n from its position in the microstate i , r_i , to its position in the microstate j , r_j , given the particle n has been selected:

$$\alpha_{ij} = P_{\text{sel}}[n] P_{\text{move}}[r_i \rightarrow r_j | n] \quad (12)$$

Typically, the new position r_j is randomly picked around the original position r_i , and $P_{\text{move}}[r_i \rightarrow r_j | n]$ takes an unknown (but constant) value that we do not need to determine. In standard MC sampling, the probability $P_{\text{sel}}[n]$ is simply the reciprocal of the number of particles, which implies that α_{ij} is symmetric.

Since CF and standard Metropolis MC schemes^{33,34} differ only for the stochastic matrix (which in both cases depends only on the current microstate), the Markov condition is trivially satisfied. The proof that microreversibility is preserved in CF is simply based on the fact that the CF acceptance ratio is just derived by imposing the detailed balance condition, which reads as follows

$$p_i \alpha_{ij} \text{acc}(i \rightarrow j) = p_j \alpha_{ji} \text{acc}(j \rightarrow i) \quad (13)$$

where $\text{acc}(i \rightarrow j)$ and $\text{acc}(j \rightarrow i)$ are the probabilities of accepting the trial moves $i \rightarrow j$ and $j \rightarrow i$, respectively. The Metropolis solution³⁴ for the acceptance ratio is $\text{acc}(i \rightarrow j) = \min(1, p_j/p_i)$.

For simplicity, from now on, we denote the position of the particle n with *in* if it belongs to the mobility region and with *out* if it does not belong to the mobility region. In CF, the probability $P_{\text{move}}[r_i \rightarrow r_j | n]$ does not differ from the standard algorithm, whereas

$$\begin{aligned} P_{\text{sel}}[n] &= N_{\text{in}}^{-1} && \text{if the particle } n \text{ is in} \\ P_{\text{sel}}[n] &= 0 && \text{if the particle } n \text{ is out} \end{aligned} \quad (14)$$

where N_{in} is the number of particles inside the mobility spheres. In principle, we can identify four types of MC moves: *in* \rightarrow *in*, *in* \rightarrow *out*, *out* \rightarrow *in*, and *out* \rightarrow *out*. Actually, the *out* \rightarrow *in* and *out* \rightarrow *out* moves are not generated because of the second condition of eq 14. On the basis of eqs 12–14 and considering that

$P_{\text{move}}[r_i \rightarrow r_j | n] = P_{\text{move}}[r_j \rightarrow r_i | n]$ by construction, the detailed balance for the $in \rightarrow in$ and $in \rightarrow out$ moves can be written respectively as

$$\begin{aligned} p_i \text{acc}(i_{\text{in}} \rightarrow j_{\text{in}}) &= p_j \text{acc}(j_{\text{in}} \rightarrow i_{\text{in}}) \\ p_i N_{\text{in}}^{-1} P_{\text{move}}[x_i \rightarrow x_j | n] \text{acc}(i_{\text{in}} \rightarrow j_{\text{out}}) &= 0 \end{aligned} \quad (15)$$

Solutions of the previous equations are

$$\begin{aligned} \text{acc}(i_{\text{in}} \rightarrow j_{\text{in}}) &= \min(1, p_j / p_i) \\ \text{acc}(i_{\text{in}} \rightarrow j_{\text{out}}) &= 0 \end{aligned} \quad (16)$$

As anticipated, eq 16 implies that MC moves which take one particle out of the mobility region are never accepted. In conclusion, the use of the acceptance criteria of eq 16 in steered MC simulations ensures that sufficient conditions for CFT to be valid are met. A direct consequence of eq 14 is that CF does not preserve ergodicity. On the other hand, in order to apply CFT and JE, the ergodicity may not be satisfied during the λ -switching simulations (see section 2.1). Rather, ergodicity is a necessary requirement when producing the initial microstates from equilibrium simulations. In fact, as remarked at the beginning of this section (see point 1 above), CF must not be applied in the preparatory simulations.

We stress again that the choice of the reference particles (centers of mobility spheres) is not a decisive aspect of the method, but rather a key point to obtain an effective simulation speedup. Basically, the choice should be made on the basis of physical statements, identifying where the control parameter acts on the system, under the reasonable assumption that dissipation is a local phenomenon in nonequilibrium processes. Only if we can localize such a region, we can apply CF with the hopes of an effective and size-dependent computational gain.

Before concluding the illustration of the method, it is worthwhile to note the correlation existing between CF and an early sampling method, i.e., preferential sampling.³⁵ This technique was devised to enhance equilibrium sampling in well-established regions of space, typically regions around the solute. To this aim, a high-mobility region is first identified. A parameter p defines how often we wish to move the *out* particles relative to the *in* ones: p lies between 0 and 1, values close to 0 corresponding to much more frequent moves of the *in* particles. A MC move consists of the following steps: (1) a particle is chosen at random; (2) if it is *in*, a trial move is made; (3) if it is *out*, a random number is generated uniformly between 0 and 1. If p is greater than the random number, then a trial move is made. If not, then we return to step 1 without accumulating any averages. Trial moves are accepted with a probability dependent on p as follows:

$$\begin{aligned} \text{acc}(i_{\text{in}} \rightarrow j_{\text{in}}) &= \min(1, p_j / p_i) \\ \text{acc}(i_{\text{in}} \rightarrow j_{\text{out}}) &= \min(1, A p_j / p_i) \\ \text{acc}(i_{\text{out}} \rightarrow j_{\text{in}}) &= \min(1, B p_j / p_i) \\ \text{acc}(i_{\text{out}} \rightarrow j_{\text{out}}) &= \min(1, p_j / p_i), \end{aligned} \quad (17)$$

with A and B being

$$\begin{aligned} A &= p \{1 + (p - 1) / [pN + (1 - p)N_{\text{in}}]\}^{-1} \\ B &= p^{-1} \{1 - (p - 1) / [pN + (1 - p)N_{\text{in}}]\}^{-1} \end{aligned} \quad (18)$$

where N is the total number of particles and N_{in} is the number of *in* particles in the microstate i . By selecting only *in* particles (CF condition), namely, by setting $p = 0$, the acceptance ratios of

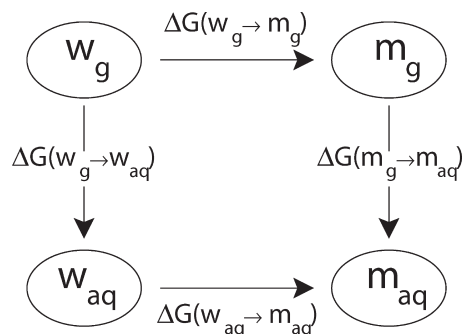


Figure 2. Thermodynamic cycle used to calculate the water to methane relative hydration free energy.

eq 17 take the following forms:

$$\begin{aligned} \text{acc}(i_{\text{in}} \rightarrow j_{\text{in}}) &= \min(1, p_j / p_i) \\ \text{acc}(i_{\text{in}} \rightarrow j_{\text{out}}) &= 0 \\ \text{acc}(i_{\text{out}} \rightarrow j_{\text{in}}) &= 1 \\ \text{acc}(i_{\text{out}} \rightarrow j_{\text{out}}) &= \min(1, p_j / p_i) \end{aligned} \quad (19)$$

In such a case, since we never select *out* particles, the last two statements of eq 19 are immaterial. The CF algorithm is thus recovered.

We point out that any variety of preferential sampling, such as those based either on molecular jumps from high- to low-mobility regions and vice versa³⁵ (the method described succinctly above) or on intermolecular distance criteria,³⁶ may also be employed to enhance sampling around the reaction site in nonequilibrium MC simulations. In this case, however, a downgrading of the performances with respect to CF is expected because MC moves of particles far from the perturbed region, and hence pretty ineffective for dissipation, can always be realized.

3. COMPUTATIONAL DETAILS

3.1. Water to Methane Relative Hydration Free Energy.

The calculation of the water to methane relative hydration free energy, $\Delta\Delta G_{\text{hyd}}$ (or $\Delta\Delta F_{\text{hyd}}$), has been taken often as a benchmark for comparing free energy methods.^{25,26} $\Delta\Delta G_{\text{hyd}}$ can be computed through the thermodynamic cycle shown in Figure 2. Given the free energy differences defined in the figure, the water to methane relative hydration free energy is

$$\begin{aligned} \Delta\Delta G_{\text{hyd}} &= \Delta G(w_g \rightarrow w_{\text{aq}}) - \Delta G(m_g \rightarrow m_{\text{aq}}) \\ &= \Delta G(w_g \rightarrow m_g) - \Delta G(w_{\text{aq}} \rightarrow m_{\text{aq}}) \end{aligned} \quad (20)$$

where w_{aq} and w_g indicate solvated and gaseous water, while m_{aq} and m_g indicate solvated and gaseous methane. The free energy difference between the systems w_{aq} and m_{aq} , $\Delta G(w_{\text{aq}} \rightarrow m_{\text{aq}})$, can be written as the sum of two contributions:

$$\Delta G(w_{\text{aq}} \rightarrow m_{\text{aq}}) = \Delta G_g(w_{\text{aq}} \rightarrow m_{\text{aq}}) + \Delta G_{\text{pert}}(w_{\text{aq}} \rightarrow m_{\text{aq}}) \quad (21)$$

where $\Delta G_g(w_{\text{aq}} \rightarrow m_{\text{aq}})$ is the free energy needed to morph water into methane in the gas phase by taking the molecular structures as in water solution, while $\Delta G_{\text{pert}}(w_{\text{aq}} \rightarrow m_{\text{aq}})$ is the solvent-phase perturbation free energy difference between water and methane. In our case, since a rigid water model and a united-atom methane model have been used, $\Delta G_g(w_{\text{aq}} \rightarrow m_{\text{aq}})$ equals $\Delta G(w_g \rightarrow m_g)$.

This leads to the equality

$$\begin{aligned}\Delta\Delta G_{\text{hyd}} &= -\Delta G_{\text{pert}}(w_{\text{aq}} \rightarrow m_{\text{aq}}) \\ &= \Delta G_{\text{pert}}(m_{\text{aq}} \rightarrow w_{\text{aq}})\end{aligned}\quad (22)$$

Thus, the water to methane relative hydration free energy can be calculated from a single calculation of the solvent-phase perturbation free energy and compared to the experimental value of $-34.57 \text{ kJ mol}^{-1}$ resulting from subtracting the hydration free energy of methane,³⁷ 8.09 kJ mol^{-1} , from the hydration free energy of water,³⁷ $-26.48 \text{ kJ mol}^{-1}$. However, in order to prevent comparisons that are biased from inaccurate modeling of the system (adopted potential models, underlying approximations in the simulation method, etc.), we found it more appropriate to compare $\Delta\Delta G_{\text{hyd}}$ estimated from nonequilibrium methods to that obtained from another well-established methodology such as thermodynamic integration.¹⁹

At variance with previous studies^{25,26} where potential and structural parameters were varied to morph water into methane or vice versa, we employ only a change of potential parameters. In particular, we adopt a morphing energy function dependent on the reaction coordinate λ as follows:

$$E(\lambda) = E_1 + \lambda(E_0 - E_1) \quad (23)$$

where E_0 is the energy of the w_{aq} system and E_1 is the energy of the m_{aq} system. On the basis of eqs 22 and 23, we define $\Delta\Delta G_{\text{hyd}} = G(\lambda = 1) - G(\lambda = 0)$, where $G(\lambda)$ is the Gibbs free energy of the “hybrid” thermodynamic state characterized by the energy function of eq 23. To calculate E_1 , the methane united atom has been placed on the oxygen site of water.

As stated above, the reference value of $\Delta\Delta G_{\text{hyd}}$ has been computed by thermodynamic integration. Various equilibrium MC simulations have been performed with a fixed λ value (from $\lambda = 0$ to $\lambda = 1$ in λ steps of 0.04 for a total of 26 simulations). For each simulation, the sample was first equilibrated in 165 Msteps (from now on, we will use Msteps and Ksteps to denote 10^6 and 10^3 trial moves, respectively). Equilibration was verified by monitoring the energy and volume of the system. For each MC simulation, the average value of the derivative of $E(\lambda)$ with respect to λ , $\langle \partial E(\lambda) / \partial \lambda \rangle_\lambda = \langle E_0 - E_1 \rangle_\lambda$, has been calculated over 150 Msteps. Given this set of averages, the water to methane relative hydration free energy is found by solving numerically the integral:

$$\Delta\Delta G_{\text{hyd}} = \int_0^1 \left\langle \frac{\partial E(\lambda)}{\partial \lambda} \right\rangle_\lambda d\lambda \quad (24)$$

The error in thermodynamic integration has been determined by calculating, for each fixed- λ simulation, batch averages of the quantity $\langle \partial E(\lambda) / \partial \lambda \rangle_\lambda$ (for a total of 300 averages per λ value). Each average has been evaluated over 0.5 Msteps. The standard error of these averages has then been integrated across the entire λ coordinate to yield the free energy error.²⁵ Note that this error analysis will overestimate the error but should be a sufficiently sensitive reference to allow comparison between the various algorithms.

The equilibrium simulations performed at $\lambda = 0$ and $\lambda = 1$ have also been employed to store the microstates used for CF and standard steered MC simulations (2000 microstates for each value of λ). Forward and backward nonequilibrium simulations have been performed using the energy function of eq 23, changing λ by a constant quantity every single-molecule trial move. By using the standard algorithm (no CF), various series of

nonequilibrium simulations have been carried out differing in the number of steps (200, 100, 50, 25, 20, 10, and 5 Ksteps). In CF simulations, one mobility sphere centered on the oxygen/methane site of the morphed molecule is adopted. In these simulations, the numbers of steps reported above have been employed with a mobility-sphere radius, R , of 1 nm. Other series of CF simulations have been performed using 100 Ksteps and different mobility-sphere radii ($R = 0.25, 0.30, 0.35, 0.40, 0.45, 0.58, 0.685, 0.85, 1.00, 1.20, \text{ and } 1.50 \text{ nm}$).

The water molecule is described by a rigid TIP4P model,³⁸ while the united atom model of ref 39 has been used for methane. In both models, short-range repulsive and dispersive forces are accounted for by a Lennard-Jones potential with mixed terms by Lorentz–Berthelot rules. Electrostatic interactions are calculated using the standard Coulomb law. A cutoff radius of 1.5 nm is used to switch off the interatomic potential energies. The simulation sample is made of one solute molecule subject to morphing and 1678 solvent (water) molecules. Constant-pressure (0.1 MPa), constant-temperature (298 K) MC simulations have been performed using a cubic box with standard periodic boundary conditions. The solute and volume moves are both attempted with a probability of 10^{-3} . Solute and solvent moves consist of rigid-body translations and rotations, with a maximum translation of 0.02 nm and a maximum rotation of 5° for the solvent and 10° for the solute. The volume moves changed the volume of the simulation box by a maximum of 0.4 nm^3 .

3.2. Potential of Mean Force of a Methane Dimer in Water Solution. For this type of experiment, most technical details (potential models, thermodynamic conditions, criteria for MC moves, etc.) are given in section 3.1. Here, we only report the basic differences. The system is made of two methane molecules and 1678 water molecules. The PMF is calculated along the methane–methane intermolecular distance from 0.3 to 0.7 nm. In order to eliminate the Jacobian contribution from the PMF, MC simulations have been realized by displacing the two methane molecules along a fixed direction. The reference free energy profile has been calculated using finite-difference thermodynamic integration.⁴⁰ Specifically, simulations have been run at various (fixed) z values, z_1, z_2, \dots, z_h , where $z_1 = 0.3 \text{ nm}$, $z_h = 0.7 \text{ nm}$, $z_{i+1} - z_i = 0.01 \text{ nm}$, and hence $h = 41$. These parameters allow good phase space overlap between neighboring simulations. The property accumulated during each simulation is the free energy gradient $\partial G / \partial z$. The free energy gradients can be approximated numerically by the finite difference $(\Delta G / \Delta z)_z$, where $\Delta G = G(z + \Delta z) - G(z)$ can be found using the Zwanzig formula²²

$$\Delta G = -\beta^{-1} \ln \langle e^{-\beta[E(z + \Delta z) - E(z)]} \rangle_z \quad (25)$$

where the average is calculated over the ensemble generated by a simulation at fixed z (the so-called reference state) whose energy is $E(z)$. The energy $E(z + \Delta z)$ is related to the system perturbed by a Δz increment whose size is 10^{-4} nm . The PMF as a function of z is then found by integrating over the measured gradients

$$G(z) = \int_{z_1}^z \left(\frac{\Delta G}{\Delta z} \right)_z dz \quad (26)$$

In the case of nonequilibrium numerical experiments, the initial microstates for the forward and backward realizations have been picked from equilibrium MC simulations realized with $z = z_1$ and $z = z_h$, respectively. Therefore, the forward direction is assumed to correspond to a breaking up of the methane dimer

($z_1 \rightarrow z_h$), whereas the backward direction corresponds to an approach of the molecules ($z_1 \leftarrow z_h$). A total of 2000 initial microstates have been produced for each direction of the process. We remark that, in these calculations, the control parameter corresponds to the collective coordinate z , which is the methane–methane distance fixed along a given direction of the space. The length of the steered MC simulations is 5.04 Msteps. The control parameter z is moved every 1 step by a quantity of 1.333×10^{-4} nm. In CF steered MC simulations, two mobility spheres centered on the methane molecules have been considered. On the basis of the symmetry of the problem, the mobility spheres have been assumed to be of equal size. The list of *in* water molecules is updated after each move of the control parameter z and after each volume move. Tests with various values of the mobility-sphere radius ($R = 0.6, 0.8, 1.0, 1.2,$ and 1.5 nm) have been realized, including $R = \infty$ corresponding to standard type simulations.

4. RESULTS

4.1. Water to Methane Relative Hydration Free Energy. In applications of CFT to free energy calculations, neither eq 8 nor its extended form involving work distribution functions (see eq 20 of ref 3) are of actual practical use, the former because the probabilities of conjugated twin trajectories cannot be known, the latter because free energy differences can only be determined resorting to somehow arbitrary procedures.⁹ An effective CFT-based method to estimate free energy differences was developed by Shirts and co-workers exploiting maximum likelihood arguments.²³ The resulting equation, that we report below (eq 27), is equivalent to that proposed earlier by Bennett in a different context.²⁴ Denoting the work measurements in the forward direction ($\lambda = 0 \rightarrow \lambda = 1$) as $W_1^{(F)}, W_2^{(F)}, \dots, W_{n_F}^{(F)}$ and those in the backward direction ($\lambda = 1 \rightarrow \lambda = 0$) as $W_1^{(B)}, W_2^{(B)}, \dots, W_{n_B}^{(B)}$, the formula reads as follows:

$$\sum_{i=1}^{n_F} \left[1 + \frac{n_F}{n_B} e^{\beta(W_i^{(F)} - \Delta G)} \right]^{-1} = \sum_{j=1}^{n_B} \left[1 + \frac{n_B}{n_F} e^{\beta(W_j^{(B)} + \Delta G)} \right]^{-1} \quad (27)$$

where $\Delta G = G(1) - G(0) \equiv \Delta\Delta G_{\text{hyd}}$. ΔG is obtained by solving eq 27 iteratively. Shirts and co-workers also proposed a way of evaluating the variance of ΔG from maximum likelihood methods, by correcting the estimate in the case of the restriction from the fixed probability of the forward and backward work measurements to the fixed number of forward and backward work measurements. The variance of ΔG is²³

$$\sigma^2 = -\frac{1}{\beta^2} \left(\frac{1}{n_F} + \frac{1}{n_B} \right) + \frac{2}{\beta^2} \left\{ \sum_{i=1}^{n_F} [1 + \cosh(w_i^{(F)})]^{-1} + \sum_{j=1}^{n_B} [1 + \cosh(w_j^{(B)})]^{-1} \right\}^{-1} \quad (28)$$

where $w_i^{(F)} = \beta W_i^{(F)} - \beta \Delta G - \ln(n_B/n_F)$ and $w_j^{(B)} = \beta W_j^{(B)} + \beta \Delta G + \ln(n_B/n_F)$. The quantity σ^2 can be calculated once ΔG is recovered from eq 27.

Estimates of the water to methane relative hydration free energy obtained from eq 27 using both standard (from now on ST) and CF methods are reported in Figure 3 for various series of

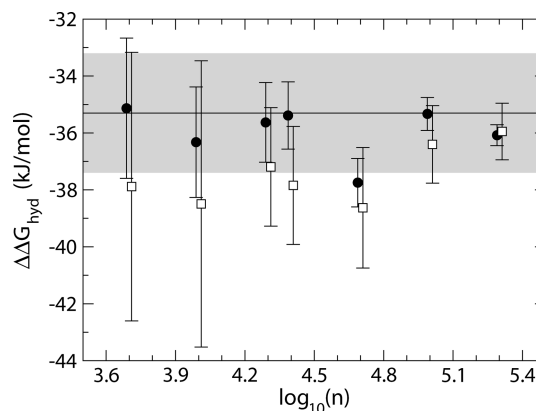


Figure 3. Water to methane relative hydration free energy, $\Delta\Delta G_{\text{hyd}}$ (from eq 27), as a function of $\log_{10}(n)$, where n is the number of steps in steered MC simulations. Full circles, $\Delta\Delta G_{\text{hyd}}$ from CF simulations; open squares, $\Delta\Delta G_{\text{hyd}}$ from ST simulations. The error bars correspond to $\pm\sigma$ (from eq 28). The data are slightly shifted along the abscissa to allow a better visualization of the error bars. Solid horizontal line: $\Delta\Delta G_{\text{hyd}}$ from thermodynamic integration (shaded area bounds the error).

steered MC simulations differing in the number of steps. The result from thermodynamic integration is also reported in the figure. The radius of the mobility sphere adopted in CF calculations is 1 nm. The error in thermodynamic integration data has been calculated as described in section 3.1, while the error bars for nonequilibrium methods correspond to $\pm\sigma$ (eq 28). There are two most important differences between the CF and ST methods emerging from Figure 3. First, an overall closer agreement with thermodynamic integration data is observed for the CF method. In this respect, it is also significant that the error band related to the reference free energy (shaded area in Figure 3) encompasses six out of seven CF free energy estimates, against only three estimates obtained from ST simulations. In the second instance, the σ value computed from CF simulations is systematically smaller than that recovered from the ST approach. Simple statistical arguments explain why the error increases by decreasing the number of simulation steps. From the radius of the mobility sphere ($R = 1$ nm) and from the average size of the simulated system,⁴¹ the number of frozen molecules is estimated to be, on average, slightly smaller than 92%. This implies that the number of trial moves of the molecules inside the mobility-sphere region (i.e., around the reaction site) is, on average, 12 times greater in CF than in ST simulations. The more efficient sampling around the reaction site obtained with CF leads to more accurate free energy estimates for all series of steered MC simulations. Consistently, we note that, regardless of the simulation length, the error resulting from the ST method is roughly double compared to that from CF. An increase of the former is even expected for larger simulation boxes.

In this context, we define computational gain as the ratio $n_{\text{st}}/n_{\text{cf}}$ between the number of simulation steps needed with ST and CF to get the same free energy error σ . In Figure 4a, we report σ (taken from Figure 3) as a function of the number of simulation steps. Both methods, CF and ST, show a quite regular trend. As remarked above, the error increases monotonically by decreasing the number of steps. To estimate the computational gain, the two sets of data have been fitted with arbitrary functions of n_{st} and n_{cf} , namely, $\sigma = an_{\text{st}}^b$ for the ST data and $\sigma = a'n_{\text{cf}}^{b'}$ for the CF data, where the fitting parameters are $a = 227.885 \text{ kJ mol}^{-1}$, $b = -0.44501$, $a' = 218.397 \text{ kJ mol}^{-1}$, and $b' = -0.51886$. From these functions, it is

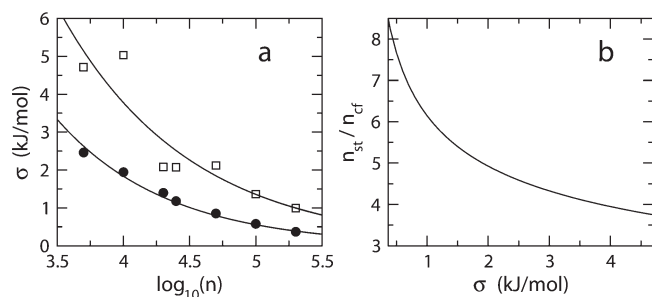


Figure 4. Error on the water to methane relative hydration free energy. *Panel a:* σ (from eq 28) as a function of $\log_{10}(n)$, where n is the number of steps in steered MC simulations. Full circles, data from CF simulations; open squares, data from ST simulations; solid lines, fitted curves (see text for details). *Panel b:* computational gain, n_{st}/n_{cf} (from eq 29), as a function of σ .

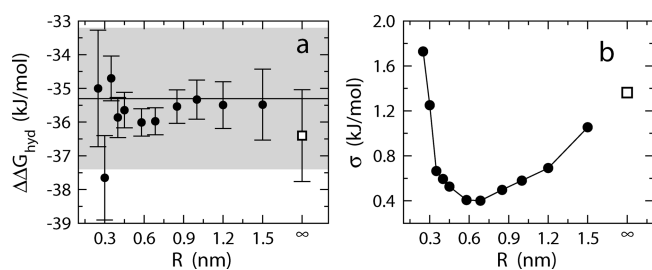


Figure 5. *Panel a:* Water to methane relative hydration free energy, $\Delta\Delta G_{hyd}$ (from eq 27), as a function of the radius of the mobility sphere, R . All data refer to steered MC simulations of 100 Ksteps. Full circles, $\Delta\Delta G_{hyd}$ from CF simulations; open square, $\Delta\Delta G_{hyd}$ from ST simulations. The error bars correspond to $\pm\sigma$ (from eq 28). Solid line: $\Delta\Delta G_{hyd}$ from thermodynamic integration (shaded area bounds the error). *Panel b:* σ as a function of R . Full circles, data from CF simulations; open square, data from ST simulations. The line is drawn as a guide for the eyes.

possible to recover the ratio n_{st}/n_{cf} as a function of σ

$$\frac{n_{st}}{n_{cf}} = \frac{(\sigma/a)^{1/b}}{(\sigma/a')^{1/b'}} \quad (29)$$

Equation 29 is drawn in Figure 4b limiting the σ range to that observed in our numerical experiments. The computational gain is consistently above 3.5, and for more accurate free energy estimates (low values of σ) it can overtake 7. It should be noted, however, that computational gain could be even greater by increasing the size of the simulation box. In fact, for a given number of simulation steps, the performance of CF is almost unaffected by changes in box size, whereas in ST simulations the MC moves would be equally distributed through the sample with evident worsening of the sampling close to the reaction site.

The choice of the radius of the mobility sphere(s), R , though arbitrary in CF, is a key point to improve the efficiency of the method. The free energy estimates along with the error bars obtained by using 100-Kstep-long simulations are reported in Figure 5a for various values of R . For comparison, the result from the ST simulation is also shown in the figure. It can be noted that CF outperforms ST for all R values above a threshold of about 0.3 nm. In principle, there are two opposite effects correlated with a change of R . By decreasing R , the sampling can be more and more focused on the reaction site, namely, on the particles

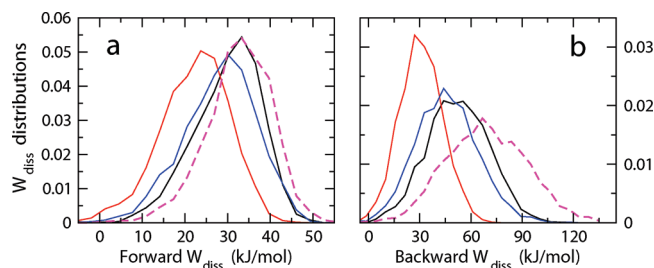


Figure 6. *Panel a:* Normalized distribution functions of the work dissipated in the forward realizations of the alchemical process (morphing methane into water). *Panel b:* Normalized distribution functions of the work dissipated in the backward realizations of the alchemical process (morphing water into methane). Black, red, and blue lines are related to CF simulations using $R = 1.5, 0.685,$ and 0.35 nm, respectively. Dashed lines are related to ST simulations. All data refer to steered MC simulations of 100 Ksteps.

mostly involved in dissipation. Although this fact would seem globally positive for improving the performances of CF, on the other hand, we must consider that small R values may exclude important (for dissipation) portions of the system from sampling. To understand the latter aspect, it is useful to think to the limit case $R = 0$, which means that all solvent molecules are configurationally frozen. In such a calculation, only the volume and the particles bearing the mobility sphere(s) (here, the molecule subject to morphing) would evolve in time. Clearly, since most of the system particles are not allowed to relax, dissipation may be very large with an ensuing increase of the error. These observations call for the existence of an optimal R value for which the two effects balance to eventually give a minimum free energy error. In our system, such a minimum lies between $R = 0.58$ and $R = 0.685$ nm, as can be observed in Figure 5b, where we show σ as a function of R (data taken from Figure 5a). It is encouraging that the free energy error is not strongly sensitive to R changes for distances around the first solvation shell.²⁵ It should however be remarked that this may depend on both the chemical nature of the system and the kind of nonequilibrium experiment (number of steps, type of collective coordinate correlated with the control parameter, etc.). In spite of this, we may state with reasonable confidence that CF outperforms ST if the mobility region encompasses the first solvation shell at least.

The picture emerging from the previous discussion calls into play dissipation arguments to explain the dependence of the CF efficiency by the size of the mobility region, as well as the better performances of CF with respect to ST at a fixed number of steps. To quantify this idea, in Figure 6 we report the distribution functions of the work dissipated in the forward and backward realizations of the alchemical process. The dissipated work for the forward and backward realizations is calculated as $W_{diss}^{(F)} = W^{(F)} - \Delta G$ and $W_{diss}^{(B)} = W^{(B)} + \Delta G$, respectively, where $W^{(F)}$, $W^{(B)}$, and ΔG are defined in eq 27. Specifically, the ST method is compared to various CF-based calculations differing in the radius of the mobility sphere, R . For the sake of clarity, only three representative radii are reported in the figure: the greatest one ($R = 1.5$ nm), one of the smallest ($R = 0.35$ nm), and the radius which provides the minimum error ($R = 0.685$ nm). As expected, the work distributions in the forward and backward directions of the process are very different. However, in terms of dissipation, both directions provide the same information. Overall, in agreement with the hypothesis formulated above, the algorithms with

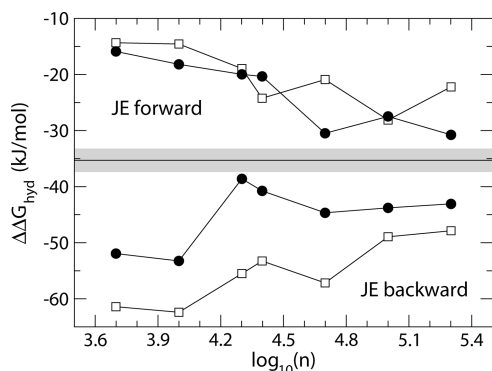


Figure 7. Water to methane relative hydration free energy, $\Delta\Delta G_{\text{hyd}}$ (from eq 30, *JE forward* in the panel; from eq 31, *JE backward* in the panel), as a function of $\log_{10}(n)$, where n is the number of steps in steered MC simulations. Full circles, $\Delta\Delta G_{\text{hyd}}$ from CF simulations; open squares, $\Delta\Delta G_{\text{hyd}}$ from ST simulations. Solid horizontal line: $\Delta\Delta G_{\text{hyd}}$ from thermodynamic integration (shaded area bounds the error). The lines are drawn as a guide for the eyes.

greater mean dissipation also provide greater error (compare Figure 5 to Figure 6). Accordingly, in Figure 6 we observe that CF with $R = 0.685$ nm gives the lowest dissipation (and hence the lowest error), while the ST method furnishes the greatest value for both error and dissipation. However, we must remark that, from the physical point of view, the error is only indirectly correlated with the mean dissipation. It rather depends on the sampling efficiency of low-work values which, on the other hand, improves by lowering the mean dissipated work (see Figure 6).

Further insights into the performances of CF can be gained by comparing JE free energy estimates obtained with ST and CF methods. Using the notation introduced in eq 27, ΔG is estimated from separate work measurements in the forward and backward directions of the process as follows:

$$e^{-\beta\Delta G} = n_{\text{F}}^{-1} \sum_{i=1}^{n_{\text{F}}} e^{-\beta W_i^{(\text{F})}} \quad (30)$$

$$e^{\beta\Delta G} = n_{\text{B}}^{-1} \sum_{j=1}^{n_{\text{B}}} e^{-\beta W_j^{(\text{B})}} \quad (31)$$

The error related to eqs 30 and 31 can be calculated using many batch path-ensemble averages,⁴² which would be too computer-time-demanding for us. However, even the simple comparison of free energy estimates as a function of the simulation steps gives useful and clear indications. Results are shown in Figure 7 along with thermodynamic integration data. As expected, the free energy estimates from eqs 30 and 31 are globally worse than those from eq 27 (compare the data with Figure 3). In fact, it is well-known that JE free energy estimates are strongly biased due to work exponential averaging.^{42–46} In our context, it is however important to observe that, globally, the free energy difference obtained with CF agrees with thermodynamic integration better than that obtained from the ST method. This is more evident in backward (from water to methane) than in the forward direction, where in only two cases ST outperforms CF. These tests show that the known problem of JE arising from biasing in exponential averages may be alleviated if reversibility is enhanced by using CF.

4.2. Potential of Mean Force of a Methane Dimer in Water Solution. In this section, we report on the PMF of a methane dimer in water solution as a function of the methane–methane

distance calculated with thermodynamic integration, ST and CF methods. Three types of PMF estimators have been considered, namely, the JE in forward and backward directions, $G_{\text{F}}(z)$ and $G_{\text{B}}(z)$, and the bidirectional estimator proposed in ref 47. The JE free energy profiles are

$$G_{\text{F}}(z) = -\beta^{-1} \ln \langle e^{-\beta W_{\text{F}}(z)} \rangle_{\text{F}} \quad (32)$$

$$G_{\text{B}}(z) = -\beta^{-1} \ln \langle e^{-\beta W_{\text{B}}(z)} \rangle_{\text{B}} \quad (33)$$

where $W_{\text{F}}(z)$ is the work done on the system to switch the control parameter from z_a to z during the forward realizations (with $z_a = 0.3$ nm) and $W_{\text{B}}(z)$ is the work done on the system to switch the control parameter from z_b to z during the backward realizations (with $z_b = 0.7$ nm). The symbols $\langle \dots \rangle_{\text{F}}$ and $\langle \dots \rangle_{\text{B}}$ indicate path-ensemble averages over the forward and backward realizations (see also eqs 30 and 31). Note that $G_{\text{F}}(z)$ and $G_{\text{B}}(z)$ are free energies with respect to reference states corresponding to $z = z_a$ and $z = z_b$, respectively. The bidirectional PMF estimator,⁴⁷ $G_{\text{FB}}(z)$, is based on eqs 32 and 33 and on the estimate of the free energy difference between the end states, $\Delta G_{ab} = G(z_b) - G(z_a)$, calculated using eq 27:

$$G_{\text{FB}}(z) = -\beta^{-1} \ln \langle e^{-\beta G_{\text{F}}(z)} + e^{-\beta[\Delta G_{ab} + G_{\text{B}}(z)]} \rangle \quad (34)$$

In Figure 8, we show $G_{\text{F}}(z)$, $G_{\text{B}}(z)$, and $G_{\text{FB}}(z)$ along with $G_{\text{TI}}(z)$ calculated from thermodynamic integration as described in section 3.2. $G_{\text{TI}}(z)$ agrees well with previous studies.^{29,48} For the sake of clarity, only three representative CF free energy profiles, obtained with $R = 0.6$, 1.0, and 1.5 nm, have been reported in the figure. Significant differences in the performances of ST and CF methods are observed. Regardless of the value of R , the PMFs obtained from CF are in better agreement with $G_{\text{TI}}(z)$ than those obtained from ST. Moreover, we note that the performances of CF may depend significantly on the radius of the mobility spheres. For the present system, the radius of 1.0 nm appears to provide the globally smallest deviation from $G_{\text{TI}}(z)$. However, for a more quantitative comparison, we need to determine somehow the overall deviation of the estimated PMFs from the reference. It is known that different PMF estimates such as $G_{\text{F}}(z)$, $G_{\text{B}}(z)$, and $G_{\text{FB}}(z)$ may differ from $G_{\text{TI}}(z)$ for an arbitrary constant, say q . Therefore, in order to make consistent comparisons, we have determined q using a least-squares procedure. Specifically, the constant q to be added to, e.g., $G_{\text{F}}(z)$, is calculated by solving the equation $\partial \eta / \partial q = 0$, where η is the root-mean-square deviation of $G_{\text{F}}(z)$ from $G_{\text{TI}}(z)$:

$$\eta = \sqrt{P^{-1} \sum_{i=1}^P [q + G_{\text{F}}(z_i) - G_{\text{TI}}(z_i)]^2} \quad (35)$$

The resolution for z employed in eq 35 is 0.01 nm, so that $P = 41$. Note that the error on $G_{\text{F}}(z)$ and $G_{\text{B}}(z)$ obtained using eq 35 is underestimated, because in work exponential averages biasing is not uniformly distributed through the space of the z coordinate (small biasing is typical of initial z points, while large biasing is typical of final z points⁴⁹). The minimum η values obtained from ST- and CF-based PMF estimators are reported as a function of R in Figure 9. It is interesting to note that the JE and the bidirectional estimator, which is based on the CFT, give comparable accuracy. This unexpected behavior is quite interesting because bidirectional methods are known to be in general more accurate than JE.^{47,50–56} However, a comparative analysis of PMF estimators is not the aim of the article, and therefore we will

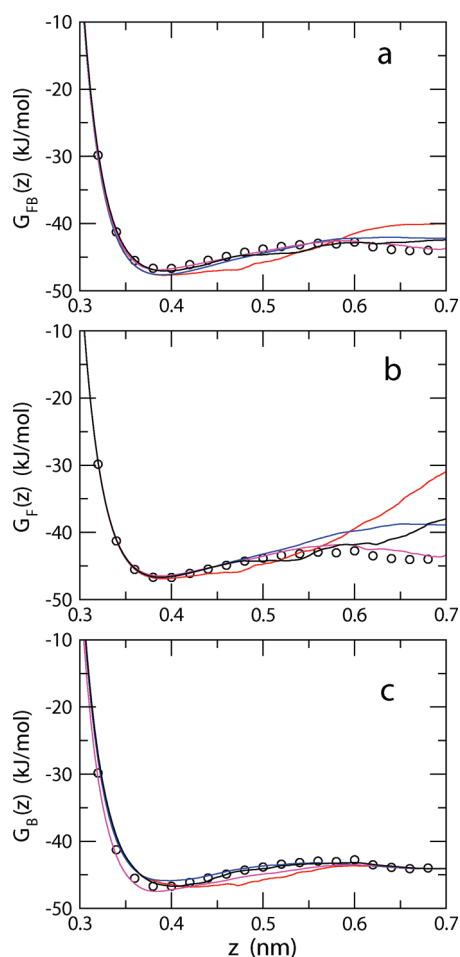


Figure 8. PMF of a methane dimer in water solution calculated using nonequilibrium methods as a function of the methane–methane distance, z . *Panel a:* $G_{\text{FB}}(z)$ (from eq 34). *Panel b:* $G_{\text{F}}(z)$ (from eq 32). *Panel c:* $G_{\text{B}}(z)$ (from eq 33). Black lines, data from CF simulations using $R = 1.5$ nm; magenta lines, data from CF simulations using $R = 1.0$ nm; blue lines, data from CF simulations using $R = 0.6$ nm; red lines, data from ST simulations; open circles, data from thermodynamic integration. Note that the additive arbitrary constants for $G_{\text{FB}}(z)$ are chosen so as to minimize the root-mean-square deviation between $G_{\text{FB}}(z)$ and $G_{\text{TI}}(z)$ (see eq 35), while the constants for $G_{\text{F}}(z)$ and $G_{\text{B}}(z)$ are taken to set $G_{\text{F}}(z_a) = G_{\text{TI}}(z_a)$ and $G_{\text{B}}(z_b) = G_{\text{TI}}(z_b)$, respectively.

not discuss this aspect further. Rather, we are interested in establishing a comparison between ST and CF approaches once the PMF estimator is given. In this respect, we note that the overall features observed for the water to methane relative hydration free energy (see section 4.1) are also found in the present system. In fact, as observed above, CF outperforms ST systematically (see η in Figure 9). This is particularly evident considering the JE estimator in the forward direction. In such a case, the η value calculated using ST exceeds the smallest and largest values of η obtained from CF by about 3.5 and 2.0 kJ mol^{-1} , respectively. Also the free energy difference $G(0.7) - G(0.39)$ calculated using forward JE with the ST method (where 0.39 nm is the methane–methane distance corresponding to the minimum free energy) is very large in comparison to the CF outcomes (see Figure 8b).

Another interesting feature of Figure 9 is the almost flat behavior of η , extending over a quite large interval of R values. As in the alchemical transformation case study, a minimum is however observed in the curve error vs R , where the error is

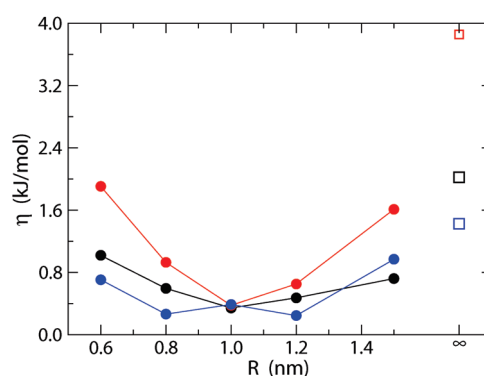


Figure 9. Minimum root-mean-square deviation, η (from eq 35), of methane dimer PMFs from the reference PMF (from thermodynamic integration) as a function of the radius of the mobility spheres, R . Black circles: error on $G_{\text{FB}}(z)$ estimated from CF simulations; red circles, error on $G_{\text{F}}(z)$ estimated from CF simulations; blue circles, error on $G_{\text{B}}(z)$ estimated from CF simulations; black square, error on $G_{\text{FB}}(z)$ estimated from ST simulations; red square, error on $G_{\text{F}}(z)$ estimated from ST simulations; blue square, error on $G_{\text{B}}(z)$ estimated from ST simulations. The lines are drawn as a guide for the eyes.

Table 1. Free Energy Difference ΔG_{ab} (from eq 27) and σ Error (from eq 28) for the Methane Dimer in Solution Calculated Using the ST Method and CF with Various Values of the Mobility-Sphere Radius, R (in units of nm)^a

simulation algorithm	ΔG_{ab}	σ
ST	39.42	0.61
CF ($R = 0.6$)	40.70	0.37
CF ($R = 0.8$)	41.86	0.21
CF ($R = 1.0$)	42.43	0.19
CF ($R = 1.2$)	42.98	0.24
CF ($R = 1.5$)	42.86	0.33
TI	44.1	3.4

^a The free energy difference by thermodynamic integration (TI) and the related error have been calculated as described in the text. All quantities are in units of kJ mol^{-1} .

represented here by η . This minimum lies at about $R = 1.0$ nm, against $R \approx 0.6$ – 0.7 nm of the alchemical transformation. The origin of this optimal mobility-sphere radius has been discussed in section 4.1. Here, we remark on the significant difference between the two optimal radii, to be ascribed, probably, to the extent of the external perturbation. It is obvious that in the current experiments the chemical environment undergoes a large perturbation during the realizations because the escorted dynamics of the methane molecules breaks down their own first solvation shell (at least). In the alchemical transformation, the structural rearrangement of the solvent around the “hybrid” water–methane molecule is evidently less relevant because morphing water into methane and vice versa does not imply a strong change in molecular volume. In some sense, the difference observed in the optimal R is a consequence of the different degree of reversibility of the two processes. This fact is indirectly supported by the number of steps needed to get reasonable convergence of the free energy estimates. While in alchemical transformations even a few simulation steps are sufficient to get satisfactory accuracy, thousands of steps are instead necessary for computing the binding free energy of a methane dimer via nonequilibrium pulling experiments.

Information complementary to η is gained by the variance of $\Delta G_{ab} = G(z_b) - G(z_a)$ calculated from eq 28 on the basis of maximum likelihood arguments. The data obtained from the CF and ST methods are reported in Table 1. For comparison, results from thermodynamic integration are also reported in the table (in this case, the error has been calculated as described in section 3.1, namely, using 300 batch averages per λ value, each simulation being of 0.5 Msteps). The similarity of σ (Table 1) and η (Figure 9) with changing R , especially the position of the minimum at about $R = 1$ nm, is remarkable. In agreement with the discussion above, the CF appears to outperform the ST method systematically.

5. CONCLUSIONS

The computational method we have presented, called *configurational freezing*, is devised to improve fast-switching free energy estimates by JE or CFT-based free energy estimators in the framework of Monte Carlo simulations. By modifying the sampling criteria in steered Monte Carlo simulations, we are able, for a fixed number of trial moves, to yield nonequilibrium trajectories where dissipation is significantly reduced with respect to a normal simulation. This dissipation decrease ultimately leads to more accurate free energy estimates. Specifically, we realize driven paths where sampling is localized around the reaction site. Therefore, the method is based on the reasonable assumption that dissipation is a local phenomenon in single-molecule nonequilibrium processes. This is expected in many processes such as the folding of biopolymers, molecular docking, alchemical transformations, etc. However, we point out that this assumption is not necessary for validating nonequilibrium work theorems (JE and CFT) but rather a statement which provides physical grounds to the efficiency of the method.

The major shortcoming of configurational freezing lies in the fact that the choice of the mobility region (the region where sampling occurs) is basically left to the chemical intuition of the researcher. The general criterion is that the mobility region must encompass the particles on which dissipation is localized. This choice, though it doesn't affect the validity of nonequilibrium work theorems, is crucial to obtain an effective computational gain. For simple systems such as the water–methane alchemical transformation treated here, the mobility region can be defined straightforwardly. In more complex systems and processes such as protein folding, dissipation cannot be localized easily. For these cases, we propose a multiple mobility-sphere approach where prior selected atoms bring mobility spheres and the overall mobility region results from the union of the single mobility spheres. This methodology works satisfactorily for our case study, i.e., the calculation of the potential of mean force of two methane molecules in water solution as a function of their distance. The self-adaptive property of the multiple mobility-sphere approach allows for the treatment of a variety of problems in addition to those illustrated in the present study and should become determinant in realizing single-molecule pulling simulations of complex biological systems.

Finally, since configurational freezing does not alter the algorithms usually employed in steered Monte Carlo simulations, it is prone to being combined with other approaches proposed for improving the efficiency of free energy estimates.

AUTHOR INFORMATION

Corresponding Author

*To whom correspondence should be addressed E-mail: riccardo.chelli@unifi.it.

ACKNOWLEDGMENT

We thank Giorgio F. Signorini, Simone Marsili and Piero Procacci (Department of Chemistry, University of Firenze, Italy) for stimulating discussions and Benedetta Morini (Department of Energetics, University of Firenze, Italy) and Antonio Bernini (Department of Systems and Informatics, University of Firenze, Italy) for insightful suggestions. We are also grateful to Gianfranco Lauria (LENS, Firenze, Italy) for technical support. This work has been supported by the European Union contract RII3-CT-2003-506350.

REFERENCES

- (1) Jarzynski, C. *Phys. Rev. Lett.* **1997**, *78*, 2690.
- (2) Crooks, G. E. *J. Stat. Phys.* **1998**, *90*, 1481.
- (3) Crooks, G. E. *Phys. Rev. E* **2000**, *61*, 2361.
- (4) Bustamante, C.; Smith, S. B.; Liphardt, J.; Smith, D. *Curr. Opin. Struct. Biol.* **2000**, *10*, 279.
- (5) Liphardt, J.; Dumont, S.; Smith, S. B.; Tinoco, I., Jr.; Bustamante, C. *Science* **2002**, *296*, 1832.
- (6) Collin, D.; Ritort, F.; Jarzynski, C.; Smith, S. B.; Tinoco, I.; Bustamante, C. *Nature* **2005**, *437*, 231.
- (7) Bornschlöggl, T.; Woehlke, G.; Rief, M. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 6992.
- (8) Park, S.; Schulten, K. *J. Chem. Phys.* **2004**, *120*, 5946.
- (9) Procacci, P.; Marsili, S.; Barducci, A.; Signorini, G. F.; Chelli, R. *J. Chem. Phys.* **2006**, *125*, 164101.
- (10) Chatelain, C. *J. Stat. Mech.* **2007**, P04011.
- (11) Mitternacht, S.; Luccioli, S.; Torcini, A.; Imparato, A.; Irbäck, A. *Biophys. J.* **2009**, *96*, 429.
- (12) Ytreberg, F. M.; Zuckerman, D. M. *J. Chem. Phys.* **2004**, *120*, 10876.
- (13) Sun, S. X. *J. Chem. Phys.* **2003**, *118*, 5769.
- (14) Geissler, P. L.; Dellago, C. *J. Phys. Chem. B* **2004**, *108*, 6667.
- (15) Wu, D.; Kofke, D. A. *J. Chem. Phys.* **2005**, *122*, 204104.
- (16) Vaikuntanathan, S.; Jarzynski, C. *Phys. Rev. Lett.* **2008**, *100*, 190601.
- (17) Schmiedl, T.; Seifert, U. *Phys. Rev. Lett.* **2007**, *98*, 108301.
- (18) Lechner, W.; Oberhofer, H.; Dellago, C.; Geissler, P. L. *J. Chem. Phys.* **2006**, *124*, 044113.
- (19) Kirkwood, J. G. *J. Chem. Phys.* **1935**, *3*, 300.
- (20) McQuarrie, D. A. *Statistical Mechanics*; Harper Collins Publishers: New York, 1976.
- (21) Nicolini, P.; Chelli, R. *Phys. Rev. E* **2009**, *80*, 041124.
- (22) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420.
- (23) Shirts, M. R.; Bair, E.; Hooker, G.; Pande, V. S. *Phys. Rev. Lett.* **2003**, *91*, 140601.
- (24) Bennett, C. H. *J. Comput. Phys.* **1976**, *22*, 245.
- (25) Woods, C. J.; Essex, J. W.; King, M. A. *J. Phys. Chem. B* **2003**, *107*, 13703.
- (26) Cossins, B. P.; Foucher, S.; Edge, C. M.; Essex, J. W. *J. Phys. Chem. B* **2009**, *113*, 5508.
- (27) Luzhkov, V. B. *Chem. Phys. Lett.* **2008**, *452*, 72.
- (28) Jorgensen, W. L.; Buckner, J. K.; Boudon, S.; Tirado-Rives, J. *J. Chem. Phys.* **1988**, *89*, 3742.
- (29) Rank, J. A.; Baker, D. *Protein Sci.* **1997**, *6*, 347.
- (30) Chelli, R.; Marsili, S.; Barducci, A.; Procacci, P. *Phys. Rev. E* **2007**, *75*, 050101(R).
- (31) The interval should be large enough to ensure uncorrelation between microstates.
- (32) The overall size of the mobility region may vary because mobility spheres can overlap.
- (33) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. N.; Teller, E. *J. Chem. Phys.* **1953**, *21*, 1087.
- (34) Frenkel, D.; Smit, B. *Understanding Molecular Simulations: From Algorithms to Applications*; Academic Press: San Diego, CA, 2002.
- (35) Owicki, J. C.; Scheraga, H. A. *Chem. Phys. Lett.* **1977**, *47*, 600.

- (36) Owicki, J. C. *Computer modeling of matter*; Lycos, P., Eds.; American Chemical Society: Washington, DC, 1978.
- (37) Ben-Naim, A.; Marcus, Y. *J. Chem. Phys.* **1984**, *81*, 2016.
- (38) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
- (39) Jorgensen, W. L.; Madura, J. D.; Swenson, C. J. *J. Am. Chem. Soc.* **1984**, *106*, 6638.
- (40) Mezei, M. *J. Chem. Phys.* **1987**, *86*, 7084.
- (41) The mean volume calculated from water-into-water and methane-into-water samples is 50.134 nm³.
- (42) Gore, J.; Ritort, F.; Bustamante, C. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 12564.
- (43) Lu, N.; Kofke, D. A. *J. Chem. Phys.* **2001**, *114*, 7303.
- (44) Wu, D.; Kofke, D. A. *Phys. Rev. E* **2004**, *70*, 066702.
- (45) Wu, D.; Kofke, D. A. *J. Chem. Phys.* **2004**, *121*, 8742.
- (46) Wu, D.; Kofke, D. A. *J. Chem. Phys.* **2005**, *123*, 054103.
- (47) Chelli, R.; Procacci, P. *Phys. Chem. Chem. Phys.* **2009**, *11*, 1152.
- (48) Rank, J. A.; Baker, D. *Biophys. Chem.* **1998**, *71*, 199.
- (49) An alternative choice to quantify the error η in the case of $G_F(z)$ and $G_B(z)$ could be to fix q in eq 35 such that $G_F(z_a) = G_{TF}(z_a)$ and $G_B(z_b) = G_{TB}(z_b)$, respectively.
- (50) Shirts, M. R.; Pande, V. S. *J. Chem. Phys.* **2005**, *122*, 144107.
- (51) Chelli, R.; Marsili, S.; Procacci, P. *Phys. Rev. E* **2008**, *77*, 031104.
- (52) Minh, D. D. L.; Adib, A. B. *Phys. Rev. Lett.* **2008**, *100*, 180602.
- (53) Lu, N.; Singh, J. K.; Kofke, D. A. *J. Chem. Phys.* **2003**, *118*, 2977.
- (54) Lu, N.; Kofke, D. A.; Woolf, T. B. *J. Comput. Chem.* **2004**, *25*, 28.
- (55) Lu, N.; Woolf, T. B.; Kofke, D. A. *Phys. Rev. E* **2004**, *69*, 057702.
- (56) Hahn, A. M.; Then, H. *Phys. Rev. E* **2010**, *81*, 041117.

Exciton/Charge-Transfer Electronic Couplings in Organic Semiconductors

Seth Difley and Troy Van Voorhis*

Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139-4307, United States

ABSTRACT: Charge transfer (CT) states and excitons are important in energy conversion processes that occur in organic light emitting devices (OLEDs) and organic solar cells. An *ab initio* density functional theory (DFT) method for obtaining CT–exciton electronic couplings between CT states and excitons is presented. This method is applied to two organic heterodimers to obtain their CT–exciton coupling and adiabatic energy surfaces near their CT–exciton diabatic surface crossings. The results show that the new method provides a new window into the role of CT states in exciton–exciton transitions within organic semiconductors.

INTRODUCTION

Organic semiconductors (OSCs) hold promise as low-cost solar cells^{1–8} and as versatile, flexible, high-contrast display technologies^{9–13} that are amenable to cost-effective large-scale production. Several of the challenges to improving the energy conversion efficiencies of these devices include maximizing absorption (or emission) efficiencies, electron and hole transport, and charge collection (or charge recombination).^{14,15} A detailed understanding of how these device properties are related to OSC materials and device architecture is important for guiding the design of semiconductor technologies. In the present study, we look closely at electronic couplings, which play a valuable role in providing this understanding.

One reason we might be interested in these couplings is that spatially localized excitons and long-range CT states play crucial roles in OSCs. Figure 1 illustrates the interplay of these states at a generic PV interface between OSC materials A and B. The first several localized singlet and triplet excitons are presented for each PV material. A nonlocal CT state that involves both A and B is also shown. Figure 1 shows two electronic state pathways that may be involved in free carrier generation. In the first step of the solar cell's operation cycle, the A–B system is photoexcited to a singlet excitonic state localized on either A or B. After some time, the system can undergo a series of relaxations from one excitonic state to another excitonic state and eventually relax to the CT state. Finally, the CT state can undergo charge separation to form free electron and hole charge carriers.¹⁶ These charge carriers can drift toward the electrodes to produce the desired current. The overall rate of carrier generation depends on the rate of each step. In optimizing these devices, we wish to choose materials and device morphologies that maximize the rate of desirable relaxation mechanisms while minimizing loss mechanism rates.¹⁴ A useful tool for estimating electronic transition rates between states *a* and *b* is the Marcus rate expression:

$$k_{ab} = \frac{2\pi}{\hbar} |H_{ab}|^2 \frac{1}{\sqrt{4\pi\lambda k_B T}} \exp\left(\frac{-(\lambda + \Delta G^\circ)^2}{4\lambda k_B T}\right) \quad (1)$$

Here, ΔG° is the driving force, λ is the reorganization energy, and $H_{ab} = \langle \psi_a | H | \psi_b \rangle$ is the coupling between states described by the

wave functions ψ_a and ψ_b and Hamiltonian H . Importantly, we note that the rate is proportional to the square modulus of the coupling. In this way, the coupling governs the relative magnitude of transition rates between states with similar driving forces and reorganization energies.

Another reason couplings are important is for the conceptual study of OSC electronic transition mechanisms. To begin this discussion, it is convenient to describe electronic energy surfaces as either being diabatic or adiabatic. Given nonorthogonal diabatic states with energies H_{aa} and H_{bb} and the coupling between these states H_{ab} , a generalized eigenvalue problem (eq 2) can be solved to obtain the corresponding adiabatic states $\psi^\pm \equiv d_a^\pm \psi_a + d_b^\pm \psi_b$ and energy surfaces ϵ^\pm :

$$\begin{pmatrix} H_{aa} & H_{ab} \\ H_{ba} & H_{bb} \end{pmatrix} \begin{pmatrix} d_a^\pm \\ d_b^\pm \end{pmatrix} = \epsilon_{\text{ad}}^\pm \begin{pmatrix} S_{aa} & S_{ab} \\ S_{ba} & S_{bb} \end{pmatrix} \begin{pmatrix} d_a^\pm \\ d_b^\pm \end{pmatrix} \quad (2)$$

An important distinction between diabatic and adiabatic states is that diabatic states are characterized by uniform electronic character as one moves along an arbitrary nuclear coordinate. Thus, if a state has ionic (covalent) character at one point on its diabatic surface, it will have ionic (covalent) character at every point on that diabatic surface. In contrast, adiabatic states, which result from rigorously applying the Born–Oppenheimer approximation, may have varied electronic character at different points on the same adiabatic energy surface. Another way to conceptualize the difference between diabatic and adiabatic states is that diabatic surfaces can energetically cross each other, while adiabatic surfaces instead undergo an “avoided crossing” near the diabatic crossings (Figure 2).

The interplay between diabatic and adiabatic states is important for understanding electronic transitions.¹⁷ That is, neither diabatic nor adiabatic states can alone describe all mechanisms. For example, suppose in Figure 2 that the system has been excited onto the upper adiabatic energy surface at point 2. By following the diabatic state, the system relaxes to the lower adiabatic surface at point 3. Meanwhile, starting from point 1, the

Received: September 5, 2010

Published: February 22, 2011

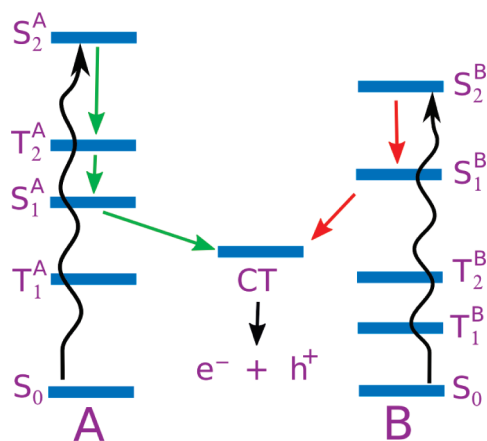


Figure 1. Two electron transfer pathways in an organic photovoltaic material. The spatial location (molecule A or B) of the localized excitons is denoted by superscript. The CT state is spread over both molecules.

system must follow the adiabat to get to the “product” at point 3. To put it another way, in Figure 2, the system behaves diabatically when optically activated and adiabatically when thermally activated. The relative magnitude of these rates is governed by the electronic coupling, making it an important quantity for the mechanistic description of OSCs.

In this study, we present an *ab initio* method for obtaining electronic couplings in which the CT states are generated by constrained DFT (CDFT)¹⁸ and the excitons are generated by time dependent density functional theory (TDDFT).¹⁹ This effort represents an expansion of previous work that described a method for obtaining electronic couplings between pairs of CDFT states.²⁰ In addition to being *ab initio*, the approach described in this study takes advantage of the balance between accuracy and computational tractability that the TDDFT and CDFT methods offer for excited states. After having demonstrated the method’s utility for constructing the adiabatic states of two specific and relevant organic dimers, we will briefly discuss implications of our coupling results for OSC electron and hole transport.

METHODS

Linear Response TDDFT. Linear response time-independent density function theory (TDDFT) is a successful method for obtaining excited state properties based solely on the response of the electron density.¹⁹ The central object in linear response is the transition density for the $i \rightarrow f$ excitation:

$$\rho_{i \rightarrow f}(r) = \langle \Psi_i | \delta(\hat{r} - r) | \Psi_f \rangle \quad (3)$$

Together with the transition energy, $\omega_{i \rightarrow f}$, $\rho_{i \rightarrow f}$ contains all of the information needed to determine the intensity of the transition under an arbitrary field. In TDDFT, the transition density is expanded in terms of products of the occupied (ϕ_j) and unoccupied (ϕ_b) Kohn–Sham (KS) orbitals as²¹

$$\rho_{i \rightarrow f}(r) = \sum_{jb} (X_{jb} \phi_b(r) \phi_j^*(r) + Y_{jb} \phi_j(r) \phi_b^*(r)) \quad (4)$$

The X and Y amplitudes that appear in this expansion can be thought of, roughly, as the amplitudes for $j \rightarrow b$ excitation and $b \rightarrow j$ de-excitation in the given transition. These amplitudes are

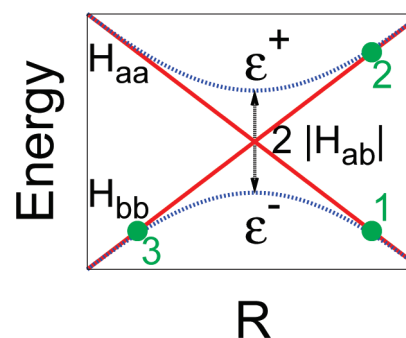


Figure 2. Cartoon of adiabatic (dashed curves labeled H_{aa} and H_{bb}) and diabatic (solid curves labeled ϵ^\pm) states at the crossing of the diabatic states as a function of some nuclear coordinate R . The coupling H_{ab} is half of the separation between the adiabatic states at the crossing. Points 1 and 3 are connected by an adiabatic state, while points 2 and 3 are connected by a diabatic state.

easily calculated using any of a number of existing implementations of the linear response TDDFT equations.

While KS-TDDFT relies on a single determinant ansatz for the ground state and only explicitly considers single excitation/de-excitation terms in computing the response, it can be proven that TDDFT gives the exact $\omega_{i \rightarrow f}$ and $\rho_{i \rightarrow j}$ if the exact exchange–correlation kernel, f_{xc} , is used.²² Unfortunately, the exact kernel is unknown, so one resorts to one of various approximations in order to apply TDDFT to molecules. For example, one commonly assumes that f_{xc} is frequency independent (the adiabatic approximation) and/or local in space (local density approximation). One common weakness of the vast majority of these commonly used functionals is that they do not treat charge transfer excitations on the same footing with localized valence excitations.²³ Typically, the CT states are far too low in energy—by an electronvolt or more in some cases—leading to very poor energy landscapes.²⁴ In practice, this problem can be softened by the use of range-separated hybrid functionals.^{25,26} By design, these functionals treat long-range CT excitations correctly, but this comes at the expense of also systematically raising valence excitation energies.²⁷ Within our group, we have explored the alternative possibility of treating CT states with constrained DFT (as described below) and using TDDFT for only the valence exciton states.

Before moving on to discuss constrained DFT, we make one note about how we will use TDDFT. In order to compute the coupling, we will need a surrogate wave function, Φ^{ex} , for the TDDFT exciton. A simple ansatz for Φ^{ex} forces the transition density between Φ^{ex} and the KS determinant to be equal to the TDDFT transition density:

$$\begin{aligned} \langle \Phi_0 | \delta(\hat{r} - r) | \Phi^{\text{ex}} \rangle &\equiv \rho_{i \rightarrow f}(r) \\ &= \sum_{jb} (X_{jb} \phi_b(r) \phi_j^*(r) + Y_{jb} \phi_j(r) \phi_b^*(r)) \end{aligned} \quad (5)$$

This serves as an implicit definition of Φ^{ex} . We note that this *will not* give us the exact excited state wave function any more than the Kohn–Sham determinant gives us the exact ground state wave function. Rather, this prescription gives us an approximate wave function that preserves an important physical property of the true system: the transition density. If we restrict our

attention to single excitations from the KS reference, it is easily verified that

$$|\Phi^{\text{ex}}\rangle = \sum_{jb} (X_{jb} + Y_{jb}) |\Phi_j^b\rangle \equiv \sum_{jb} C_{jb} |\Phi_j^b\rangle \quad (6)$$

where Φ_j^b denotes the KS single determinant where the j th occupied orbital has been replaced by the b th unoccupied orbital. Similar manipulations have been performed previously in order to associate a wave function with a TDDFT transition.²⁸

For the exact density functional, TDDFT states are rigorously adiabatic states because they obey the Born–Oppenheimer approximation. However, for commonly used approximate functionals, we typically observe that the TDDFT states behave as diabatic or diabatic-like states. Although there are TDDFT states that involve a single molecule (e.g., a Frenkel exciton) and states that involve more than one molecule (e.g., CT between a pair), these two types of states are generally energetically well separated, which essentially ensures that they will behave diabatically. Identifying the character of a given TDDFT state can be done by attachment/detachment density analysis.^{29,30} If the attachment density is confined to the same molecule as the detachment density, the state is identified as an exciton. On the other hand, if the attachment density for a TDDFT state is on a different molecule than the detachment density, the TDDFT state is identified as a CT state. This attachment/detachment analysis can be conducted at several points along a given TDDFT surface to confirm that the electronic character remains consistent from one location on the surface to another. We also note that since excitons are localized on single monomers (i.e., they are Frenkel-type excitons), exciton energies obtained for dimer systems at large monomer–monomer separations are expected to be essentially the same as exciton energies obtained for a single monomer. Therefore, dimer TDDFT energies that do not connect to a sum of monomer TDDFT energies at large separation are typically identified as CT states.

Constrained DFT for Diabatic States. Constrained density functional theory (CDFT) has been shown to be a reliable, inexpensive method for obtaining long-range CT state energies. The details of this approach have been presented elsewhere.^{18,20,31–33} Here, we briefly review CDFT and illustrate the use of this computational tool as it pertains to obtaining electronic couplings.

In the CDFT formalism, we build constraints of the form

$$\sum_{\sigma} \int w_c^{\sigma}(\mathbf{r}) \rho^{\sigma}(\mathbf{r}) \, d\mathbf{r} = N_c \quad (7)$$

where the sum is over spins such that $\sigma = \alpha$ or β , c is the constrained region of the system, w_c is a weighting function that corresponds to the constrained property, and N_c is the expectation value of the constrained property. Equation 7 is then combined as a Lagrange multiplier constraint with the Kohn–Sham energy functional $E[\rho]$ to generate a new functional

$$W[\rho, \{V_c\}] = E[\rho] + \sum_c^m V_c \left(\sum_{\sigma} \int w_c^{\sigma}(\mathbf{r}) \rho^{\sigma}(\mathbf{r}) \, d\mathbf{r} - N_c \right) \quad (8)$$

where the c th Lagrange multiplier is V_c , and there are m constraints. W is then made stationary with respect to ρ and V_c .

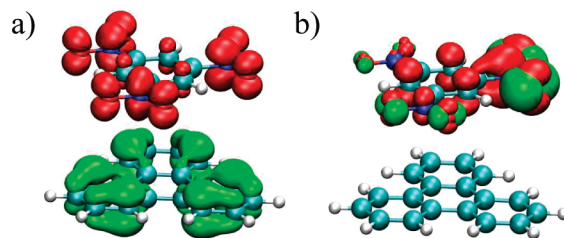


Figure 3. Attachment/detachment density plots for triphenylene:1,3,5-trinitrobenzene illustrating (a) nonlocal CT-like and (b) localized exciton-like electron densities. Red (green) regions have excess (deficient) density compared to the ground state.

By this procedure, we obtain the energy $E(\rho)$ as a natural function of the expectation value N_c . In the present study of electronic couplings, spin polarized CT states are generated by applying both charge and spin constraints via eq 7. A charge constraint is applied that forces the donor (acceptor) molecule to have an excess charge of $+1$ (-1). A concurrent constraint on the net spin forces the donor and acceptor, respectively, to have an excess spin of $\pm 1/2$. Importantly for the present study, applying these constraints produces CT states that are rigorously diabatic.

Electronic Couplings between TDDFT and CDFT States.

The electronic couplings whose properties are the focus of this study are

$$H_{ab} = \langle \psi^{\text{CT}} | H | \psi^{\text{ex}} \rangle \quad (9)$$

where ψ^{CT} and ψ^{ex} are the wave functions corresponding to the CT state and exciton and H is the electronic Hamiltonian. In particular, we are interested in computing the electronic coupling between CT states obtained by CDFT and excitons obtained by TDDFT. To do this, we adapt the constrained approach that has been successfully demonstrated for obtaining couplings between CT and neutral states.²⁰

In the constrained approach to electronic couplings, we use Kohn–Sham determinants to approximate the true wave function. This allows us to write the coupling matrix element (eq 9) in terms of a single electronic density. Following this approach, we obtain²⁰

$$H_{ab} = (E_{\text{CT}} + V_c^{\text{CT}} N_c) \langle \Phi^{\text{CT}} | \Phi^{\text{ex}} \rangle - V_c^{\text{CT}} \langle \Phi^{\text{CT}} | w_c | \Phi^{\text{ex}} \rangle \quad (10)$$

Here, E_{CT} is the energy of $|\Phi^{\text{CT}}\rangle$, $\langle \Phi^{\text{CT}} | \Phi^{\text{ex}} \rangle$ is the CT–exciton overlap, V_c^{CT} is the CT state’s constraining potential (eq 8), and $\langle \Phi^{\text{CT}} | w_c | \Phi^{\text{ex}} \rangle$ is the CT–exciton matrix element of the one-body weight operator w_c . This result makes the reasonable assertion that the electronic coupling depends on both the magnitude of the orbital overlap and the strength of the potential that was used to create the CT state.

For the true density functional, $\langle \Phi^{\text{CT}} | H | \Phi^{\text{ex}} \rangle$ is the complex conjugate of $\langle \Phi^{\text{ex}} | H | \Phi^{\text{CT}} \rangle$. However, applying the same logic as above, we find $H_{ba} = \langle \Phi^{\text{ex}} | H | \Phi^{\text{CT}} \rangle = E_{\text{ex}} \langle \Phi^{\text{ex}} | \Phi^{\text{CT}} \rangle - V_c^{\text{ex}} \langle \Phi^{\text{ex}} | w_c | \Phi^{\text{CT}} \rangle$, where V_c^{ex} is the constraining potential corresponding to the exciton. For the approximate functionals that are commonly used, these two expressions are not equivalent, so the Hermiticity condition is not fulfilled. To satisfy Hermiticity, we choose the electronic coupling to be the average of H_{ab} and H_{ba} . This average is reasonable because H_{ab} overestimates the electronic coupling when H_{ba} underestimates the coupling, and vice versa.

Equation 10 reduces the problem of computing the coupling to obtaining a zero-body overlap and one-body weight matrix element. In order to obtain a reasonable approximation to these matrix elements for excitation–CT coupling, we note that both CDFT and TDDFT states can be expressed in terms of Slater determinants. Therefore, the coupling in eq 10 can be computed as a sum of zero- and one-body matrix elements of Slater determinants (eqs 11 and 12).³⁴

$$\langle \Phi^{\text{CT}} | \Phi^{\text{ex}} \rangle = \sum_{ia} \langle \Phi^{\text{CT}} | \Phi_{ia}^{\text{ex}} \rangle C_{ia} \quad (11)$$

$$\langle \Phi^{\text{CT}} | w_c | \Phi^{\text{ex}} \rangle = \sum_{ia} \langle \Phi^{\text{CT}} | w_c | \Phi_{ia}^{\text{ex}} \rangle C_{ia} \quad (12)$$

In these equations, we have used the fact that the approximate TDDFT states can be written as sums of Slater determinants (eq 6). Computing each $\langle \Phi^{\text{CT}} | \Phi_{ia}^{\text{ex}} \rangle$ and $\langle \Phi^{\text{CT}} | H | \Phi_{ia}^{\text{ex}} \rangle$ term has an N^3 computational complexity, where N is the number of electrons in the system.³⁴ Furthermore, the sums in eqs 11 and 12 are over-occupied and virtual orbitals, both of which scale with the number of electrons in the system. Therefore, the total complexity of computing $\langle \Phi^{\text{CT}} | \Phi^{\text{ex}} \rangle$ or $\langle \Phi^{\text{CT}} | H | \Phi^{\text{ex}} \rangle$ is on the order of $N^3 \times N^2 = N^5$. With this complexity scaling, computing the coupling of even medium-sized molecular systems becomes intractable.

To decrease the complexity scaling of the coupling calculation, we use a Thouless rotation³⁵ to re-express the TDDFT excited states (eq 6) as a sum of two Slater determinants. In particular, we define $\phi_i(\pm\varepsilon) \equiv \phi_i \pm \varepsilon \sum_a C_i^a \phi_a$ where ε is small, ϕ_i is the i th occupied Kohn–Sham orbital, and ϕ_a is the a th virtual Kohn–Sham orbital. That is, we construct new orbitals $\phi_i(\pm\varepsilon)$ that mix small amounts of the virtual orbitals with each the i th occupied orbitals. From these constructed orbitals, we build a pair of Slater determinants $\Phi(\pm\varepsilon) \equiv |\phi_1(\pm\varepsilon) \phi_2(\pm\varepsilon) \phi_3(\pm\varepsilon) \dots| = |\phi_1 \phi_2 \phi_3 \dots| \pm \varepsilon \sum_{ia} C_i^a \phi_i^a + O(\varepsilon^2)$. Using these definitions, the TDDFT state becomes

$$|\Phi^{\text{ex}}\rangle = \sum_{ia} C_i^a \Phi_i^a = \lim_{\varepsilon \rightarrow 0} \left(\frac{\Phi(+\varepsilon) - \Phi(-\varepsilon)}{2\varepsilon} \right) \quad (13)$$

With the TDDFT state expressed in this two-determinant form, the matrix element of eq 11 has the manageable computational complexity of $O(N^3)$.

TDDFT states $|\Phi^{\text{ex}}\rangle$ are generally not orthogonal to the CDFT states $|\Phi^{\text{CT}}\rangle$ because the two states are eigenstates of different Hamiltonians. For this reason, we apply an orthogonalization step to put the couplings we obtain here on the same footing as couplings obtained by other methods. We define $S_{ij} = \langle \Psi^i | S | \Psi^j \rangle$ and weight matrix $w_c^{ij} = \langle \Psi^i | w_c | \Psi^j \rangle$, and solve for the generalized eigenstates of the constraint function

$$\begin{pmatrix} w_c^{aa} & w_c^{ab} \\ w_c^{ba} & w_c^{bb} \end{pmatrix} \begin{pmatrix} x_a^n \\ x_b^n \end{pmatrix} = n \begin{pmatrix} S_{aa} & S_{ab} \\ S_{ba} & S_{bb} \end{pmatrix} \begin{pmatrix} x_a^n \\ x_b^n \end{pmatrix}$$

where \mathbf{X}^n is an eigenvector of \mathbf{W}_c and n is its eigenvalue. By construction, the eigenstates of w_c are orthonormal and localized. We therefore transform the Hamiltonian to the eigenbasis of w_c via $\tilde{H} = \mathbf{X}^\dagger \mathbf{H} \mathbf{X}$, and the appropriate (orthogonal) coupling is then given by the off-diagonal element \tilde{H}_{ab} .

A variety of methods have been developed for obtaining electronic couplings. In particular, if adiabatic energy surfaces

are available, the coupling at the avoided crossing can be identified as one-half of the minimum energy separation between the adiabatic surfaces (Figure 2).³⁶ For obtaining couplings away from the avoided crossing, Mulliken–Hush methods³⁷ are often applied. A number of other empirical and semiempirical approaches for obtaining couplings between various types of electronic states have also been developed.^{38–43} The present approach is specialized in that it predicts couplings between two classes of states for which adiabatic energies are not easily obtained.

Computational Details. In this paper, we use the 3-21G basis set, B3LYP hybrid density functional, DFT, CDFT, and full linear response TDDFT as implemented in Q-Chem.⁴⁴ The basis set is intentionally small to speed up the calculations. Since this is a validation study and our conclusions are largely qualitative, we do not anticipate that a larger basis would change the picture significantly. Becke weights⁴⁵ are used in the constrained population analysis. Attachment/detachment analysis²⁹ is used to obtain the electronic spatial character of the TDDFT states.

Diabatic energy surfaces for the chosen dimers are produced by making the monomer planes parallel, scanning along the separation distance between the monomer planes, and obtaining TDDFT and CDFT states for each separation distance. For the heterodimers studied, the CDFT constraints were chosen to obtain the lowest energy CT state.

RESULTS

Triphenylene:1,3,5-Trinitrobenzene. As a first illustration of the TDDFT/CDFT coupling method, we chose a dimer consisting of triphenylene (1,3,5-trinitrobenzene) as the donor (acceptor). The small size of these molecules allows a straightforward search for CT–exciton intersections and demonstrates many of the issues that arise in obtaining the electronic couplings of long-range organic dimers.

The attachment/detachment density plots in Figure 3 show the qualitative difference between exciton and long-range CT-like states obtained by TDDFT for triphenylene:1,3,5-trinitrobenzene. In the analysis that follows, we will focus our attention on the TDDFT states that are manifested by localized densities such as in Figure 3b.

Figure 4 presents the first several singlet TDDFT states and the lowest lying CDFT state of triphenylene:1,3,5-trinitrobenzene. By attachment/detachment analysis, we find that the lowest nine states have CT-like electronic character as in the left pane of Figure 3, while the higher lying TDDFT states shown in Figure 4 have localized excitonic electronic character (Figure 3b). It is known that for many density functionals such as B3LYP, CT-like states generated by TDDFT have erroneously low energies.²⁴ Consequently, the lowest nine TDDFT states (red curves) in Figure 4 do not correspond to experimentally observable excitations, and we will thus attempt to disregard these states in what follows. Meanwhile, the higher-lying singlet excitations represented by the green curves in Figure 4 are excitons and are expected to correspond to fluorescence absorption spectra and form the exciton states of interest.

Excitons are localized on monomers, so they should not change much in energy as the monomer–monomer separation increases. This expectation that exciton energies will remain nearly constant with respect to separation distance provides a diagnostic for distinguishing excitons from TDDFT CT-like states that compliments attachment/detachment analysis. For

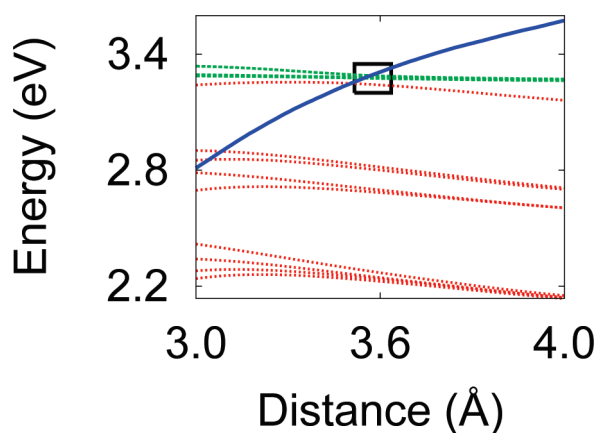


Figure 4. Diabatic energy surfaces for TDDFT excitons (dashed green curves), TDDFT CT-like states (dotted red curves), and a CDFT CT state (solid blue curve) for triphenylene:1,3,5-trinitrobenzene as a function of monomer–monomer separation distance. The inset rectangle encloses crossings of the CT state with three TDDFT excitons and one CT-like TDDFT state.

this particular dimer, we find that the exciton TDDFT states in Figure 4 remain at nearly constant energies with respect to the separation distance, but that the CT-like states erroneously decrease in energy as the monomer–monomer separation distance is increased. Meanwhile, CT states are characterized by charge separation between the two monomers. Thus, one would expect CT state energies to increase as the monomers are separated due to the attractive $1/r$ Coulombic potential between the CT state's separated charges. Indeed, we find that the CT state generated by CDFT has a positive slope over the entire range presented in Figure 4. This points out precisely why we use CDFT to obtain CT states and TDDFT to obtain excitons: CDFT correctly describes CT states but knows nothing of the TDDFT excitons.

Since we will treat the TDDFT excitons as diabatic-like states, it is important that they have consistent electronic character as we track along the monomer–monomer separation coordinate. We observe in Figure 4 that all but one of the CT-like TDDFT states are separated in energy from the excitons. Only the highest lying CT-like TDDFT state ever approaches the three lowest lying excitons S_1 , S_2 , and S_3 , and even then only at separations less than 3.7 Å. The attachment/detachment densities of S_1 , S_2 , and S_3 were inspected near 3.5 Å. S_1 and S_2 were found to have localized densities in this monomer–monomer separation range. S_3 is also primarily localized over the entire range presented in Figure 4, only showing a small amount of charge separation near 3.5 Å.

The triphenylene:1,3,5-trinitrobenzene CDFT state intersects three TDDFT states in the inset rectangle of Figure 4. We computed couplings \tilde{H}_{ab} between the CT state and these three excitons in the region of the crossings. Figure 5 presents the resulting coupling magnitudes. We observe that the couplings are on the order of 1–7 meV and that $\tilde{H}_{CT,S_3} > \tilde{H}_{CT,S_2} > \tilde{H}_{CT,S_1}$. Therefore, if the reorganization energies and driving forces are similar, we expect transitions between S_3 and the CT state to occur more easily than transitions between S_2 or S_1 and the CT state (eq 1). Another observation is that the couplings tend toward zero for large monomer–monomer separations. This reflects the decreasing orbital overlap between the exciton and CT state. Additionally, we note that although the

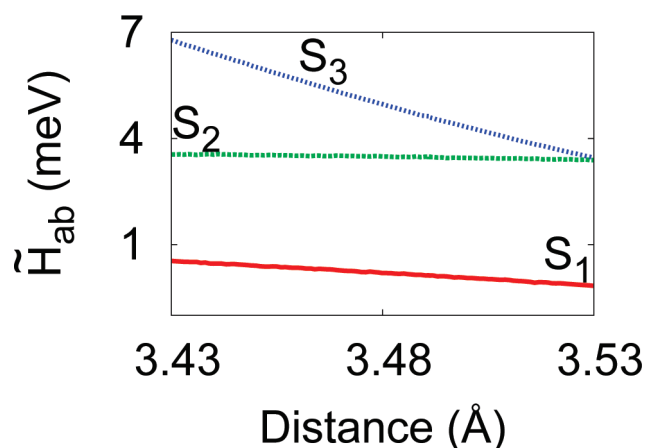


Figure 5. CT-exciton coupling magnitudes \tilde{H}_{ab} for triphenylene:1,3,5-trinitrobenzene at the diabatic state crossings in Figure 4 as a function of the monomer–monomer separation distance. Labels indicate which exciton is coupled to the CT state. We find that the couplings tend toward zero at large separations.

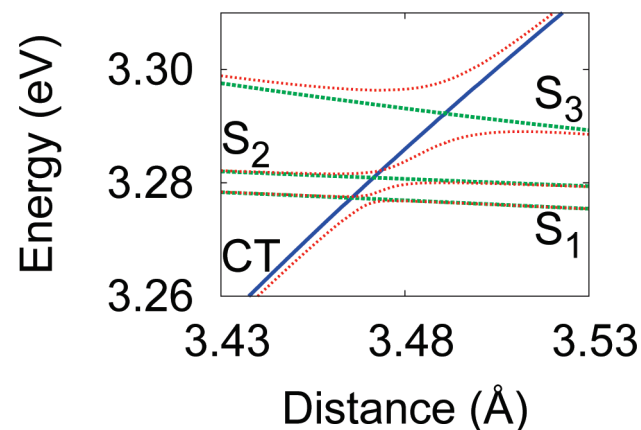


Figure 6. Diabatic exciton states (labeled green dashed curves), CT state (labeled solid blue curve), and adiabatic states (dotted red curves) of triphenylene:1,3,5-trinitrobenzene at the intersections of the CT state with S_1 , S_2 , and S_3 .

attachment/detachment density of S_3 shows mild charge separation near 3.5 Å, the magnitudes shown in Figure 5 are consistently small, as would be expected for couplings between exciton-like TDDFT states and CT states. It is therefore reasonable to treat S_1 , S_2 , and S_3 as diabatic states.

The four adiabatic states that result from solving eq 2 for the CT, S_1 , S_2 , and S_3 diabatic states and couplings are shown in Figure 6. We observe that the adiabatic states avoid each other where the diabatic states intersect. Also, the magnitude of the avoided crossing is directly related to the associated coupling magnitude. That is, the adiabatic states near the CT– S_1 (S_3) crossing most narrowly (strongly) avoid each other because the coupling between these states is small (large). Meanwhile, for regions on the energy surfaces far from avoided crossings, the adiabatic states are almost identical to the diabatic states. Importantly, Figure 6 provides a concrete pathway for a nonadiabatic transition in an organic heterodimer. For example, suppose that the triphenylene:1,3,5-trinitrobenzene dimer is initially excited to the highest-lying exciton in Figure 6, and consider how it might

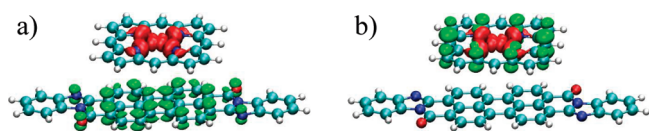


Figure 7. Attachment/detachment density plots for Zn-porphyrin:PTCBI illustrating (a) nonlocal CT-like and (b) localized exciton-like TDDFT states. Red (green) regions have excess (deficient) density compared to the ground state.

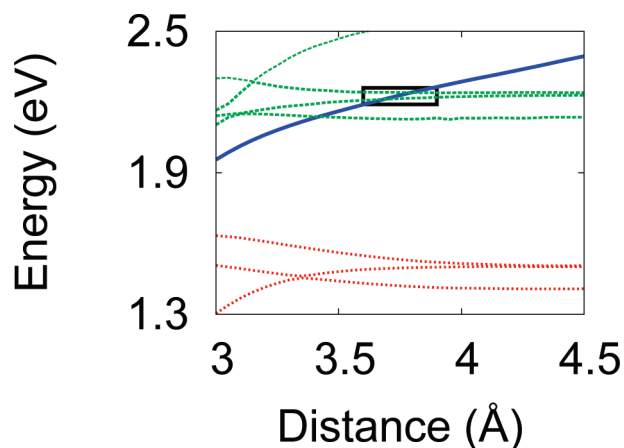


Figure 8. Diabatic energy surfaces for TDDFT excitons (dashed green curves), TDDFT CT-like states (dotted red curves), and a CDFT state (solid blue curve) for Zn-porphyrin:PTCBI as a function of monomer-monomer separation distance. The inset rectangle encloses crossings of the CT state with two TDDFT excitons. We see that the localized TDDFT states are energetically separated from the CT-like TDDFT states.

generate trapped charge carriers. In the diabatic picture, S_3 can transition to the CT state (at around 3.5 Å) and directly relax by dragging the two monomers closer together, trapping the electron and hole. Describing the same mechanism in the adiabatic picture would require starting in the fourth adiabat and making a rapid succession of nonadiabatic jumps ($4 \rightarrow 2 \rightarrow 3 \rightarrow 1$). This is not to say that this particular mechanism is operative in this particular dimer, merely that a mechanism like this is much easier to describe with the diabatic coupling than with traditional adiabatic states. We note that a similar mechanism ($S_3 \rightarrow \text{CT} \rightarrow S_1$) could be used to describe nonradiative relaxation between different bright exciton states mediated by the dark CT state.

Zn-Porphyrin:PTCBI. We have seen that the triphenylene:1,3,5-trinitrobenzene dimer provides an interesting technical demonstration of the constrained coupling method. Let us now study the CT-exciton couplings and resulting adiabatic states of a dimer composed of two organic dyes commonly used in organic semiconductors. PTCBI (3,4,9,10-perylenetetracarboxylic-bis-benzimidazole) is an organic dye often used as an electron acceptor in OSCs.^{1,46,47} It absorbs in the 450–800 nm range with absorption maxima near 525 and 700 nm.⁴⁸ Meanwhile, Zn-porphyrin is commonly used in dye-sensitized solar cells⁴⁹ and in porphyrin-fullerene solar cells.⁵⁰ Porphyrins have an absorption onset near 450 nm⁵¹ and have an important role in photosynthetic systems.^{52,53}

As for triphenylene:1,3,5-trinitrobenzene (Figure 3), we use attachment/detachment analysis to identify TDDFT states with

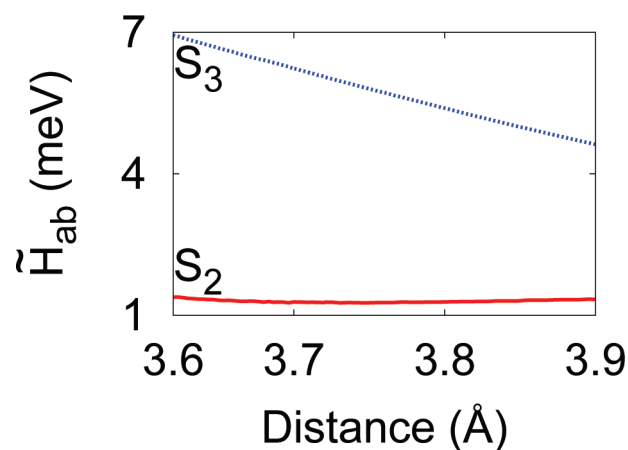


Figure 9. Coupling magnitudes near the CT- S_2 and CT- S_3 intersections labeled by the coupled exciton. We find that the CT- S_2 coupling is small over the entire range, and that the CT- S_3 couplings tends toward zero at large separations.

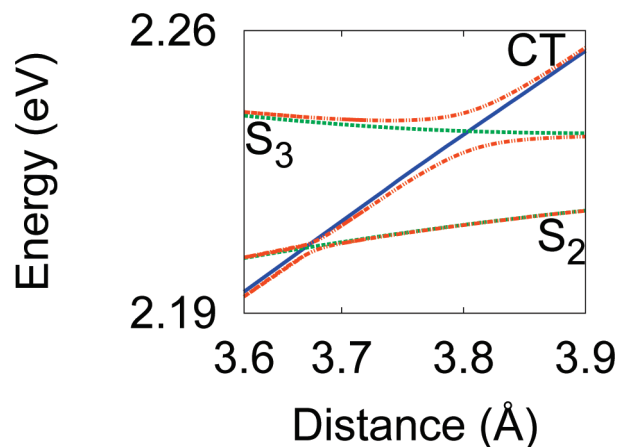


Figure 10. Diabatic exciton states (labeled green dashed curves), CT state (labeled solid blue curve), and adiabatic states (dotted red curves) of Zn-porphyrin:PTCBI at the intersections of the CT state with S_2 and S_3 .

excitonic character. Figure 7 contains representative CT-like and excitonic TDDFT densities.

Figure 8 presents the first several singlet TDDFT states and the lowest lying CDFT state of Zn-porphyrin:PTCBI. As for triphenylene:1,3,5-trinitrobenzene (Figure 4), we find that the CDFT state has a positive slope for the entire range inspected. By attachment/detachment analysis, the three TDDFT states below 1.7 eV are identified as CT-like. That is, the lowest singlet exciton states appear above 2.1 eV. Unlike for triphenylene:1,3,5-trinitrobenzene, there is an energetic delineation between the CT-like TDDFT states and the excitons, leading to clearly diabatic-like states.

In Figure 8, the CT state intersects three TDDFT states. Figure 9 presents coupling magnitudes \tilde{H}_{ab} for the upper two of these intersections. As in triphenylene:1,3,5-trinitrobenzene (Figure 5), we observe that the couplings are on the order of 0–7 meV and tend toward zero for large monomer-monomer separations. We note that the CT- S_3 coupling is much larger than the CT- S_2 coupling. Thus, by eq 1, we might expect more facile transitions between S_3 and the CT state than between S_2 and the CT state.

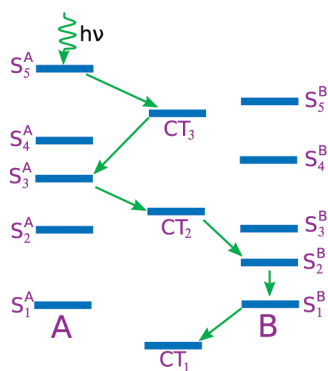


Figure 11. Illustration of CT-state mediated exciton–exciton transitions ($S_n \rightarrow CT \rightarrow S_{n-1}$).

Figure 10 presents adiabatic and diabatic CT and exciton energy surfaces in the region of the CT– S_2 and CT– S_3 intersections. We observe that the adiabatic states avoid each other where the diabatic states intersect. The avoided crossing magnitudes correspond to their associated coupling magnitude so that the adiabatic states near the CT– S_2 (S_3) crossing narrowly (strongly) avoid each other. Meanwhile, for regions on the energy surfaces far from avoided crossings, the adiabatic states are almost identical to the diabatic states. As for triphenylene:1,3,5-trinitrobenzene, we see in Figure 10 that exciton relaxation in Zn–porphyrin:PTCBI can be mediated by CT states. Given the roles of PTCBI and Zn–porphyrin as commonly used semiconductor devices, these mechanistic details about their nonadiabatic transitions are of particular interest for guiding the design of advanced solar cells and light-emitting devices.

CONCLUSIONS

We have presented an *ab initio* method for obtaining the electronic couplings between excitons and CT states in organic molecules. The utility of this method has been demonstrated by applying it to the study of the adiabatic and diabatic states and nonadiabatic transitions of two organic dimers. These results provide conceptual details of the mechanisms that allow transitions between CT states and excitons, which is an integral step in the efficient function of organic solar cells and light-emitting devices. In particular, these results show how CT states can play an important role in mediating exciton–exciton transitions (Figure 11) and conversion of excitons to free carriers.

These calculations show that it is possible to properly couple the lowest-lying CT state to a manifold of exciton states using CDFT and TDDFT in concert. Moving forward, one would like to extend this method in a number of ways. First, it would be nice if CT states other than the lowest CT state could be treated—this would allow us to characterize ultrafast relaxation involving higher-lying CT states. Second, there is a significant amount of user input that goes into these calculations—most notably the user must identify the Frenkel-like exciton states from TDDFT amidst a sea of spurious CT states. Ideally, this screening process would be automatic. For example, one could restrict the TDDFT calculation *a priori* to include only localized excitations. This would eliminate the need to screen the states manually and could potentially speed up the TDDFT calculations significantly. Finally, the calculations presented in this work have been conducted in the gas phase. Future efforts to compute these

CT–exciton couplings will use condensed phase methods such as QM/MM^{54,55} and implicit solvation models⁵⁶ that simulate effects due to bulk polarization and nuclear heterogeneity. These bulk calculations provide reorganization energies and driving forces that may be combined with the electronic couplings to provide estimates of OSC transition rates.

AUTHOR INFORMATION

Corresponding Author

*E-mail: tvan@mit.edu.

ACKNOWLEDGMENT

T.V. gratefully acknowledges support from the DOE (DE-FG02-07ER46474) and a Packard Fellowship.

REFERENCES

- (1) Tang, C. W. *Appl. Phys. Lett.* **1986**, *48*, 183–185.
- (2) Li, G.; Shrotriya, V.; Huang, J. S.; Yao, Y.; Moriarty, T.; Emery, K.; Yang, Y. *Nat. Mater.* **2005**, *4*, 864–868.
- (3) Xue, J.; Rand, B. P.; Uchida, S.; Forrest, S. R. *Adv. Mater.* **2005**, *17*, 66–70.
- (4) Park, S. H.; Roy, A.; Beaupre, S.; Cho, S.; Coates, N.; Moon, J. S.; Moses, D.; Leclerc, M.; Lee, K.; Heeger, A. J. *Nat. Photonics* **2009**, *3*, 297–302.
- (5) Kim, J. Y.; Lee, K.; Coates, N. E.; Moses, D.; Nguyen, T.; Dante, M.; Heeger, A. J. *Science* **2007**, *317*, 222–225.
- (6) Koster, L. J. A.; Mihailetschi, V. D.; Blom, P. W. M. *Appl. Phys. Lett.* **2006**, *88*, 093511.
- (7) Rand, B. P.; Burk, D. P.; Forrest, S. R. *Phys. Rev. B* **2007**, *75*, 115327.
- (8) Riede, M.; Mueller, T.; Tress, W.; Schueppel, R.; Leo, K. *Nanotechnology* **2008**, *19*, 424001.
- (9) Helfirch, W.; Scheinder, W. G. *Phys. Rev. Lett.* **1965**, *14*, 229–231.
- (10) Pope, M.; Kallmann, H. P.; Magnante, P. *J. Chem. Phys.* **1963**, *38*, 2042–2043.
- (11) van Slyke, S. A.; Tang, C. W. *Appl. Phys. Lett.* **1987**, *51*, 913–915.
- (12) Burroughes, J. H.; Bradley, D. C. C.; Brown, A. R.; Marks, R. N.; Mackay, K.; Friend, R. H.; Burns, P. L.; Holmes, A. B. *Nature* **1990**, *347*, 539–541.
- (13) Baldo, M. A.; Thompson, M. E.; Forrest, S. R. *Nature* **2000**, *403*, 750–753.
- (14) Heremans, P.; Cheyons, D.; Rand, B. *Acc. Chem. Res.* **2009**, *42*, 1740–1747.
- (15) Nelson, J.; Kwiatkowski, J. J.; Kirkpatrick, J.; Frost, J. M. *Acc. Chem. Res.* **2009**, *42*, 1768–1778.
- (16) Zhu, X.-Y.; Yang, Q.; Muntwiler, M. *Acc. Chem. Res.* **2009**, *42*, 1779–1787.
- (17) Brédas, J.-L.; Beljonne, D.; Coropceanu, V.; Cornil, J. *Chem. Rev.* **2004**, *104*, 4971–5004.
- (18) Wu, Q.; Voorhis, T. V. *Phys. Rev. A* **2005**, *72*, 024502.
- (19) Gross, E. K. U.; Dobson, J. F.; Petersilka, M. *Top. Curr. Chem.* **1996**, *181*, 81.
- (20) Wu, Q.; Voorhis, T. V. *J. Chem. Phys.* **2006**, *125*, 164105.
- (21) Furche, F. *J. Chem. Phys.* **2001**, *114*, S982–S992.
- (22) Runge, E.; Gross, E. K. U. *Phys. Rev. Lett.* **1984**, *52*, 997.
- (23) Maitra, N. T. *J. Chem. Phys.* **2005**, *122*, 234104.
- (24) Dreuw, A.; Head-Gordon, M. *J. Am. Chem. Soc.* **2004**, *126*, 4007–4016.
- (25) Vydrov, O. A.; Scuseria, G. E. *J. Chem. Phys.* **2006**, *125*, 234109.
- (26) Yanai, T.; Tew, D. P.; Handy, N. C. *Chem. Phys. Lett.* **2004**, *393*, 51–57.
- (27) Jacquemin, D.; Perpète, E. A.; Scuseria, G. E.; Ciofini, I.; Adamo, C. *J. Chem. Theory Comput.* **2008**, *4*, 123–135.

- (28) Wong, C. Y.; Curutchet, C.; Tretiak, S.; Scholes, G. D. *J. Chem. Phys.* **2009**, *130*, 081104.
- (29) Head-Gordon, M.; Graña, A. M.; Maurice, D.; White, C. A. *J. Chem. Phys.* **1995**, *99*, 14261–14270.
- (30) Martin, R. L. *J. Chem. Phys.* **2002**, *118*, 4775–4777.
- (31) Rudra, I.; Wu, Q.; Voorhis, T. V. *J. Chem. Phys.* **2006**, *124*, 24103.
- (32) Wu, Q.; Voorhis, T. V. *J. Phys. Chem. A* **2006**, *110*, 9212–9218.
- (33) Wu, Q.; Van Voorhis, T. *J. Chem. Theory Comput.* **2006**, *2*, 765–774.
- (34) Löwdin, P.-O. *Phys. Rev.* **1955**, *97*, 1474–1489.
- (35) Thouless, D. J. *Nucl. Phys.* **1960**, *21*, 225–232.
- (36) Coropceanu, V.; Cornil, J.; da Silva Filho, D.; Olivier, Y.; Silbey, R.; Brédas, J.-L. *Chem. Rev.* **2007**, *107*, 926–952.
- (37) Cave, R. J.; Newton, M. D. *Chem. Phys. Lett.* **1996**, *249*, 15–19.
- (38) Larsson, S. *J. Am. Chem. Soc.* **1981**, *103*, 4034–4040.
- (39) Kawatsu, T.; Coropceanu, V.; Ye, A.; Brédas, J.-L. *J. Chem. Phys. C* **2008**, *112*, 3429–3433.
- (40) Mančal, T.; Valkunas, L.; Fleming, G. R. *Chem. Phys. Lett.* **2006**, *432*, 301–305.
- (41) Newton, M. D. *Chem. Rev.* **1991**, *91*, 767–792.
- (42) Prytkova, T. R.; Kurnikov, I. V.; Beratan, D. N. *J. Phys. Chem. B* **2005**, *109*, 1618–1625.
- (43) Van Voorhis, T.; Kowalczyk, T.; Kaduk, B.; Wang, L. P.; Cheng, C. L.; Wu, Q. *Annu. Rev. Phys. Chem.* **2010**, *61*, 149–170.
- (44) Kong, J.; et al. *J. Comput. Chem.* **2000**, *21*, 1532–1548.
- (45) Becke, A. D. *J. Chem. Phys.* **1988**, *88*, 2547.
- (46) Peumans, P.; Bulović, V.; Forrest, S. R. *Appl. Phys. Lett.* **2000**, *76*, 2650–2652.
- (47) Kim, I.; Haverinen, H. M.; Wang, Z.; Madakuni, S.; Kim, Y.; Li, J.; Jabbour, G. E. *Chem. Mater.* **2009**, *21*, 4256–4260.
- (48) Triyana, K.; Yasuda, T.; Katsuhiko, F.; Tsutsui, T. *Thin Solid Films* **2005**, *477*, 198–202.
- (49) Schaafsma, T. J. *Sol. Energy Mater. Sol. Cells* **1995**, *38*, 349–351.
- (50) Vilmercati, P.; Castellarin Cudia, C.; Larciprete, R.; Cepek, C.; Zampieri, G.; Sangaletti, L.; Pagliara, S.; Verdini, A.; Cossaro, A.; Floreano, L.; Morgante, A.; Petaccia, L.; Lizzit, S.; Battocchio, C.; Polzonetti, G.; Goldoni, A. *Surf. Sci.* **2006**, *600*, 4018–4023.
- (51) Scandola, F.; Chiorboli, C.; Prodi, A.; Iengo, E.; Alessio, E. *Coord. Chem. Rev.* **2006**, *250*, 1471–1496.
- (52) Fischer, H.; Wenderoth, H. *Annalen* **1940**, *545*, 140–147.
- (53) Woodward, R. B.; et al. *J. Am. Chem. Soc.* **1960**, *82*, 3800–3802.
- (54) Difley, S.; Wang, L.-P.; Yeganeh, S.; Yost, S. R.; Van Voorhis, T. *Acc. Chem. Res.* **2010**, *43*, 995–1004.
- (55) Aqvist, J.; Warshel, A. *Chem. Rev.* **1993**, *93*, 2523–2544.
- (56) Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, *99*, 2161–2200.

Evaluating the Performance of DFT Functionals in Assessing the Interaction Energy and Ground-State Charge Transfer of Donor/Acceptor Complexes: Tetrathiafulvalene–Tetracyanoquinodimethane (TTF–TCNQ) as a Model Case

Gjergji Sini,[†] John S. Sears, and Jean-Luc Brédas*

School of Chemistry and Biochemistry and Center for Organic Photonics and Electronics, Georgia Institute of Technology, Atlanta, Georgia 30332-0400, United States

ABSTRACT: We have evaluated the performance of several density functional theory (DFT) functionals for the description of the ground-state electronic structure and charge transfer in donor/acceptor complexes. The tetrathiafulvalene–tetracyanoquinodimethane (TTF–TCNQ) complex has been considered as a model test case. Hybrid functionals have been chosen together with recently proposed long-range corrected functionals (ω B97X, ω B97X-D, LRC- ω PBEh, and LC- ω PBE) in order to assess the sensitivity of the results to the treatment and magnitude of exact exchange. The results show an approximately linear dependence of the ground-state charge transfer with the HOMO_{TTF}–LUMO_{TCNQ} energy gap, which in turn depends linearly on the percentage of exact exchange in the functional. The reliability of ground-state charge transfer values calculated in the framework of a mono-determinantal DFT approach was also examined.

INTRODUCTION

Since the initial works by Mulliken et al.,¹ donor–acceptor (charge-transfer) complexes have long been of interest. Recently, they have attracted increased attention due to their role in organic opto-electronic devices and in particular in organic solar cells² and organic light-emitting diodes.³ The performance of organic photovoltaic devices primarily depends on the electronic structure at the interface between an electron donor (D) component and an acceptor (A) component in both the ground and excited states. In the excited state, the ability of an exciton present either on the donor or on the acceptor to dissociate into a charge-transfer (D^+/A^-) state determines the extent of the generated photocurrent; in the ground state, the nature of the electronic interactions between donor and acceptor directly impacts the charge recombination ($D^+/A^- \rightarrow D/A$) process⁴ as well as the reverse electrical current in the dark.⁵ In organic light-emitting diodes, ground-state charge transfer between a charge-transport molecule and a molecular dopant that form a D/A complex is observed in a number of instances to greatly facilitate charge injection from an electrode into the charge-transport material.^{3,6} We focus here on the ground state of D/A complexes (while the description of the electronic structure in the lowest charge-transfer excited state will be addressed in a forthcoming work).

We are particularly interested in assessing the reliability of DFT methodologies to evaluate the interaction energy between the D and A molecules and the amount of charge transfer in the ground electronic state. DFT methods have been largely used for the study of such issues⁷ but intrinsically lead to a stabilization of the virtual molecular orbitals (MOs) with respect to the occupied MOs.⁸ This is due to (i) the fact that the virtual orbital energies in DFT are defined with respect to an N -electron problem vs $N + 1$ in Hartree–Fock (HF) theory and (ii) the self-interaction error⁹ that induces, for example, spurious long-range stabilization upon

dissociation,¹⁰ overestimates charge-transfer properties,¹¹ and results in substantial errors for long-range charge-transfer excitations.¹² The inclusion of some amount of exact HF exchange within the exchange functional is one approach often used to remedy these problems;¹³ however, such “hybrid” functionals have been found to fail for the description of long-range interactions. Other methods¹⁴ make use of the error function (erf) in order to separate the (r^{-1}) Coulomb operator into a short-range and a long-range region, using for example hybrid meta-GGA functionals for the description of the short-range interactions (the first term on the rhs of eq 1) and HF exchange for the long-range interactions (second term):

$$\frac{1}{r_{12}} = \frac{1 - \text{erf}(\omega r_{12})}{r_{12}} + \frac{\text{erf}(\omega r_{12})}{r_{12}} \quad (1)$$

The use of such functionals, referred to as ω functionals, has led in general to important improvements, specifically in the description of charge-transfer (CT) excited states.¹⁵ The ω parameter in eq 1 defines the extent of the short- and long-range regions and is typically taken as a constant optimized for a given functional. However, it has been demonstrated that the delineation between short- and long-range components (optimal ω) is a function of the electronic density of the system under study.¹⁶ In order to take this effect into account, Stein et al.¹⁵ recently proposed a simple method for optimizing the ω value for each particular system, leading to much better results, for example, in the case of CT excited-state calculations.

In the limit of the complete transfer of one electron from donor to acceptor, the ground-state wave function becomes that of an open-shell singlet and, as a result, is not well described by

Received: September 24, 2010

Published: January 18, 2011

single-reference methods. The reliability of the ground-state charge-transfer values computed within the confines of DFT has been examined previously.¹⁷ Geskin et al.^{17b} have underlined that, in the limit of weak coupling (vanishing wave function overlap) between donor and acceptor, integer charge-transfer values should be expected; i.e., the charge transfer should be either near zero or near unity. This points consequently to the necessity of a multi-configurational description. Although spin-unrestricted DFT approaches may provide a qualitatively correct description of the charge transfer in such cases,^{17b} this almost certainly results in a spin-polarized description of the electronic wave function. When considering single-reference HF and spin-restricted DFT approaches for weakly coupled complexes, the ground-state charge transfer was shown to vary in a continuous manner as a function of the (donor) HOMO–(acceptor) LUMO energy gap, while the charge-transfer values computed with multireference methods provided the expected 0-to-1 stepwise evolution.^{17b}

The same description applies in the intermediate coupling regime, as shown by Avilov et al.^{17a} in the case of a tetrathiafulvalene–tetracyanoquinodimethane (TTF–TCNQ) complex. However, these authors pointed out that, in instances where the charge-transfer values are small (lower than $\sim 0.2e$), the CASSCF and monodeterminantal RHF calculations actually give similar partial charge-transfer values, with a gradual increase in the amount of charge transferred as the $\text{HOMO}_{\text{TTF}}\text{--LUMO}_{\text{TCNQ}}$ energy gap decreases. As for the strong coupling regime, single molecule electron transport measurements¹⁸ and calculations¹⁹ have shown that the amount of charge transferred is expected to vary continuously with the voltage between electrodes. Thus, the applicability of single-reference methods for describing D/A charge transfer depends upon the strength of the electronic coupling; in addition, when single-reference DFT methods are indeed applicable, the reliability of the charge-transfer values computed with DFT methods depends as well on the choice of the functional.

The aim of our study is to analyze the performance of various DFT functionals for the description of the charge transfer in the ground state of donor–acceptor complexes. As a test case, we have chosen the TTF–TCNQ complex, as it represents a D/A complex that has been largely studied previously and for which partial charge transfer is expected; TCNQ derivatives are also widely used as molecular dopants in organic light-emitting diodes.³ The DFT functionals were chosen to include hybrid functionals with a low percentage of HF exchange (B3LYP,^{13,20} TPSSH,²¹ B97-1,²² B97-2,²³ and PBE0²⁴), hybrid functionals with a high percentage of HF exchange (BMK,²⁵ BHandH,²⁶ BHandHLYP,²⁷ M05-2X,²⁸ M06-2X,²⁹ M06-HF²⁹), and some of the more recent ω functionals (ω B97X³⁰ and ω B97X-D,³¹ as well as LRC- ω PBEh³² and LC- ω PBE³³).

THEORETICAL METHODS

Computations were performed with the Gaussian 09,³⁴ QChem 3.2,³⁵ and Molpro 2009³⁶ packages. The isolated monomers of TTF and TCNQ were completely optimized (RMS gradient 10^{-3}) at the M05-2X/6-311G** level of theory. Geometries for the TTF–TCNQ complex were constructed in cofacial and parallel-displaced configurations (employing the frozen-monomer approximation) by varying the distance between the molecular centers along the z axis and (in the case of the parallel-displaced configurations) the y axis, see Figure 1. A fixed z distance (3.45 Å) between the molecular planes was chosen for

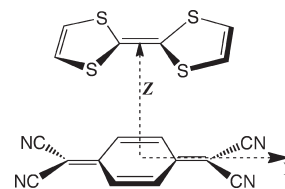


Figure 1. Tetrathiafulvalene–tetracyanoquinodimethane model complex in the cofacial geometry. The two molecular planes are parallel to one another.

the parallel-displaced configurations. Single-point computations were performed at the level of second-order Møller–Plesset perturbation theory (MP2),³⁷ spin-component-scaled (SCS)-MP2,³⁸ and DFT levels of theory employing a 6-311++G** basis and various functionals as described above. All of the functionals have been applied as defined in the literature and implemented in their respective programs, with the exception of ω B97X. For the ω B97X functional, results are presented for two ω values: $\omega = 0.3 \text{ a}_0^{-1}$ represents the standard value given in the literature, while $\omega = 0.23 \text{ a}_0^{-1}$ has been optimized by employing the procedure by Stein et al.¹⁵ and using the standard 6-31G* Pople basis set.

In order to assess the reliability of the single-reference approaches for describing the D/A complex, complete active-space self-consistent field (CASSCF) computations and broken-symmetry unrestricted DFT computations were also performed. The active space for CASSCF computations included up to eight electrons in seven MOs. The unrestricted DFT results will not be presented, as they collapse to those obtained from the restricted DFT formalism. The interaction energies were calculated with respect to the isolated monomers and were corrected for the basis-set superposition error (BSSE) by the counterpoise correction method of Boys and Bernardi.³⁹ For each method, charge transfer in the ground state is computed on the basis of the natural population analysis (NPA)⁴⁰ charges.

The benchmark level considered in this work is SCS-MP2, which has been shown to reduce the overbinding of the conventional MP2 method. Given the complexity of our system and the large basis set considered here, we were not able to produce CCSD(T) values for the interaction energies. However, a comparison of CCSD(T) and SCS-MP2 computations in a smaller 6-31G* basis gives very good agreement between these approaches. Couplings between the frontier orbitals of TTF and TCNQ have been calculated according to the approach described by Valeev et al.⁴¹ with the corresponding matrix elements evaluated with Gaussian 09.

RESULTS AND DISCUSSION

1. Frontier Orbital Analysis. On the basis of calculations on the isolated molecules, the $\text{HOMO}_{\text{TTF}}\text{--LUMO}_{\text{TCNQ}}$ energy gaps, defined as $E[\text{LUMO}_{\text{TCNQ}}] - E[\text{HOMO}_{\text{TTF}}]$, are presented in Table 1 for the various functionals and are plotted in Figure 2 along with the frontier orbital energies themselves (HOMO_{TTF} , $\text{HOMO}-1_{\text{TTF}}$, and $\text{LUMO}_{\text{TCNQ}}$) as a function of the percentage of HF exchange in the functional (% HF). Both the HOMO_{TTF} and $\text{HOMO}-1_{\text{TTF}}$ have been included in the inset of Figure 2, as both are relevant to our discussion of the charge transfer, a point that will be made more clear in the discussion below. Figure 2 shows a linear increase in the $\text{HOMO}_{\text{TTF}}\text{--LUMO}_{\text{TCNQ}}$ gap as a function of the percentage in HF exchange.

Table 1. Fraction of HF Exchange, Computed HOMO_{TTF}–LUMO_{TCNQ} Gaps, and Computed Ground-State Charge Transfer (both at the SCS-MP2 geometry and the minimum-energy geometry for the functional) for Cofacial and Parallel-Displaced Structures of the TTF–TCNQ Model Complex for Various Functionals

method	% HF	$\Delta(L-H)^a$ (eV)	q^{GS} (TTF)			
			cofacial ^b (SCS-MP2)	cofacial ^c (min)	parallel-displaced ^d (SCS-MP2)	parallel-displaced ^c (min)
TPSSh	10	−0.802	0.081		0.308	
B3LYP	20	−0.324	0.066	0.019	0.285	0.199
B97-1	21	−0.260	0.071	0.050	0.281	0.234
PBE0	25	−0.051	0.058	0.042	0.260	0.233
BMK	42	0.816	0.053	0.059	0.206	0.216
BHandH	50	1.386	0.054	0.111	0.160	0.242
M06-2X	54	1.514	0.047	0.074	0.153	0.213
M05-2X	56	1.532	0.049	0.081	0.166	0.232
LRC- ω PBEh	79.9 ^e	2.966	0.040	0.027	0.132	0.091
ω B97X-D	80.9 ^e	3.020	0.042	0.053	0.129	0.157
ω B97X-0.23	83.4 ^e	3.156	0.041		0.118	
ω B97X-0.3	92.3 ^e	3.646	0.037	0.037	0.083	0.083
LC- ω PBE	97.7 ^e	3.941				
M06-HF	100	4.139	0.035	0.076	0.063	0.142
HF		4.843	0.026		0.045	
CASSCF			0.034		0.039	
MP2			0.052	0.052	0.095	0.095

^a $\Delta E(LUMO_{TCNQ}-HOMO_{TTF})$ for the isolated molecules. ^b Cofacial geometry at an intermolecular separation of 3.45 Å, corresponding to the SCS-MP2 minimum. ^c Cofacial or parallel displaced (at fixed 3.45 Å interplanar distance) geometry corresponding to the minimum-energy geometry for the functional. ^d Parallel displacement of 3.00 Å at a fixed intermolecular distance of 3.45 Å, corresponding to the SCS-MP2 minimum. ^e Effective values (see text for details).

The HOMO–1_{TTF}–LUMO_{TCNQ} gap shows a nearly identical progression (as can be expected from the plots of the orbital energies in the inset).

The M06-HF functional, which contains 100% HF exchange, provides the greatest energy gap value among all DFT functionals, slightly over 4 eV. It is topped only by the HF value itself, 4.87 eV; however, as HF virtual orbitals are constructed on the basis of an $N + 1$ -electron system, the HOMO–LUMO gaps from HF are expected to be overestimated. At the other extreme, the TPSSh functional (that contains just 10% HF exchange) gives a “negative” gap of −0.802 eV; along with the other low HF-exchange functionals (B3LYP, B97-1, and PBE0), it provides an unphysical picture of the frontier energy levels since the LUMO_{TCNQ} energy is calculated to be lower than the HOMO_{TTF} energy for the isolated molecules. Such results are an artifact of these methods, as the experimental gas-phase values for the ionization potential of TTF and the (exothermic) electron affinity of TCNQ are 6.7 eV⁴² and 2.8 eV,⁴³ respectively (which gives an energy difference of 3.9 eV and suggests that only a partial charge transfer should be expected when the complex forms). A preliminary conclusion based on the frontier-orbital analysis is that DFT methods including a low percentage of HF exchange will be inappropriate for describing complexes such as TTF–TCNQ.

Given that a positive value for the HOMO_{TTF}–LUMO_{TCNQ} gap should be expected, the functionals containing more than 40–50% HF exchange (i.e., BMK, BHandH, M05-2X, M06-2X, and M06-HF) and the ω functionals provide for at least a reasonably physical description. Comparing the HOMO_{TTF}–LUMO_{TCNQ} gaps from Table 1 to the difference between the experimental IP value for TTF and the EA value for

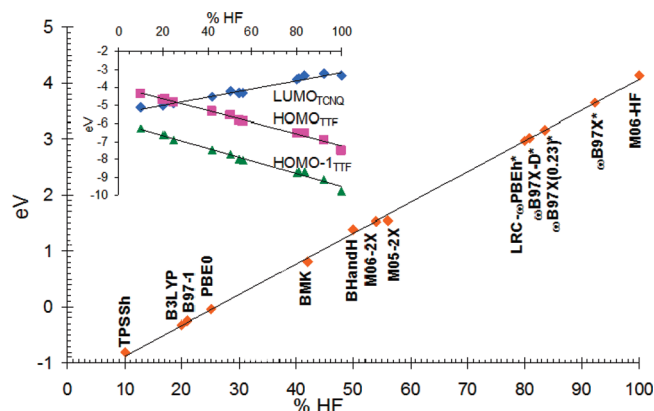


Figure 2. Computed HOMO_{TTF}–LUMO_{TCNQ} gap (orange diamonds). [Inset] Computed HOMO_{TTF}, HOMO–1_{TTF}, and LUMO_{TCNQ} energies as a function of the percentage of HF exchange. The % HF for ω functionals are effective values (see text for details).

TCNQ (3.9 eV), it is found that the ω B97X-0.3, LC- ω PBE, and M06-HF functionals give reliable results.

With regard to the ω functionals, it is not possible to assign explicit values for the percentage of HF exchange. However, it is reasonable to expect that the ω functionals with 100% HF exchange at long range contain a greater extent of HF exchange than BHandH, M05-2X, and M06-2X functionals, but a lower extent than M06-HF, as the latter only has HF exchange. Larger ω values result in a greater amount of exact exchange (being evaluated through the second term on the rhs of eq 1). The LRC- ω PBEh and ω B97X-D functionals with $\omega = 0.2 a_0^{-1}$ consequently contain a lesser degree of long-range HF exchange than

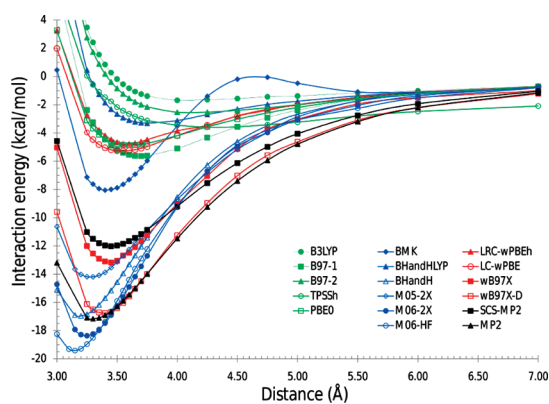


Figure 3. Potential energy surfaces for the cofacial model complex of TTF-TCNQ computed at various levels of theory with a 6-311++G** basis. All values are corrected for BSSE as described in the text.

ω B97X-0.23 ($\omega = 0.23 a_0^{-1}$), which in turn contains less HF exchange than ω B97X-0.3 ($\omega = 0.3 a_0^{-1}$) or LC- ω PBE ($\omega = 0.4 a_0^{-1}$). Using the HOMO_{TTF}-LUMO_{TCNQ} gaps and the percentage of exact exchange for the standard functionals (results from Table 1), we have used a simple linear regression to approximate the relationship between the computed gaps and the percentage of exact exchange, see Figure 2. This relation has been employed to assign *effective* percentages of exact exchange to the ω functionals; the values are calculated to range between 80 and 97% (Table 1), which indicates the significance of the long-range exact exchange in these functionals. It is worth noting that, while remaining greater than $\sim 60\%$ HF, these effective values could be different for other systems, as the contribution from HF exchange is a function of the electronic density.¹⁶

2. Interaction Energies. The ground-state charge transfer will depend heavily upon the relative geometry and orientation of the D/A molecules in the complex, falling off quickly with distance (due to the exponential decay of the overlap) and being impacted strongly by symmetry considerations (as discussed below). We thus turn our attention to the description of the ground-state potential energy surfaces. The potential energy surfaces provided by several functionals for the cofacial structure can be found in Figure 3. The SCS-MP2 values (taken here as a reference) predict a minimum of 12.02 kcal mol⁻¹ (as compared to 17.16 kcal mol⁻¹ at the MP2 level) at a separation of 3.45 Å between the molecular planes (3.25 Å at the MP2 level). The MP2 approach thus appears to substantially overbind for this system. As mentioned previously, CCSD(T) and SCS-MP2 computations in the smaller 6-31G* basis predict similar binding energies (7.24 and 8.16 kcal mol⁻¹, respectively), which are much smaller than the MP2 results for this basis set (11.75 kcal mol⁻¹).

Turning to the results for the DFT approaches collected in Figure 3, a rough separation into two groups of functionals can be observed. In general, only the long-range corrected functionals (in red) and functionals containing a high admixture of exact HF exchange (in blue) provide potential energy surfaces and interaction energies in overall good semiquantitative or qualitative agreement with the SCS-MP2 results. The best performance is shown by the ω B97X functional (whose results differ from the SCS-MP2 results by ~ 1 kcal mol⁻¹), while the M05-2X, BHandH, ω B97X-D, M06-2X, and M06-HF functionals overestimate the interaction energy by $\sim 2-7$ kcal mol⁻¹. The spurious maximum observed with the BMK functional (which remains even with an unrestricted approach) is surprising. On the other

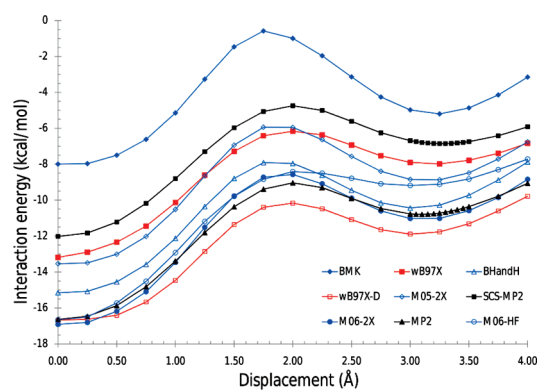


Figure 4. Potential energy surfaces for the parallel-displaced model complex of TTF-TCNQ (with a fixed interplanar separation of 3.45 Å) computed at various levels of theory and with a 6-311++G** basis. All values are corrected for BSSE as described in the text.

hand, all of the traditional functionals containing less than $\sim 30\%$ HF exchange (in green) provide an overall very poor description of the interaction energies with respect to the benchmark results, underbinding by more than ~ 7 kcal mol⁻¹.

While there are many differences between the two groups of functionals, we will focus here on the correlation with the amount of exact exchange in the functionals (see also Table 1). It is well-known that the inclusion of exact exchange is important for correcting the self-interaction error effects, which in turn reduces the delocalization error.^{11b} Accordingly, the standard hybrid functionals containing a small fraction of HF exchange (such as TPSSh, B3LYP, and PBE0) show an overall poor description of the binding. While high admixtures of exact exchange appear to be required for the description of the binding in the TTF-TCNQ model complex, the differences between, for instance, the BHandH and BHandHLYP results (both containing 50% HF exchange) indicate the importance of the specific form of the exchange functional as well. Similar observations can be made for the ω functionals. The ω B97X functional (containing up to 100% exact exchange at long range) provides the best performance, which underlines the importance of the long-range exact exchange. However, ω B97X-D (containing also up to 100% exact exchange at long range) overestimates the interaction energy, while LRC- ω PBEh and LC- ω PBE underestimate by more than 7 kcal mol⁻¹ in comparison to SCS-MP2. These observations point clearly to the importance of a careful balance between the long-range exact exchange and the quality of the short-range exchange and correlation.

It is worth mentioning that the ω values in the considered LR-corrected functionals vary between 0.2 and 0.4 a_0^{-1} ; this corresponds to approximately $1/\omega \sim 3-5 a_0$ and indicates that the separation between the short-range and long-range regions is located around 2-3 Å. The TTF-TCNQ intermolecular separation lies just beyond this range (3.45 Å), as a consequence, the complex appears to be qualitatively correctly described by both middle-range functionals (M06-2X and M05-2X) as well as long-range corrected functionals.

The potential energy surfaces for the parallel-displaced configurations (at a fixed intermolecular distance of 3.45 Å) are included in Figure 4 for MP2, SCS-MP2, and a subset of the DFT functionals investigated. A local minimum (some 5 kcal mol⁻¹ above the cofacial configuration) is predicted by all approaches for a parallel displacement around 3.25 Å. Geometry optimizations

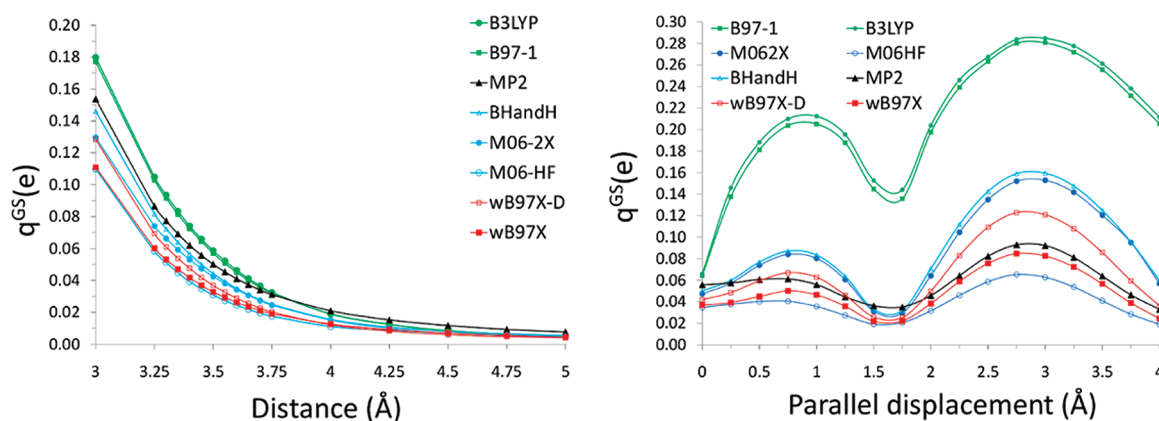


Figure 5. NPA charge transferred in the ground state (q^{GS}) calculated with the different functionals, plotted either as a function of (left panel) the interplanar distance in the cofacial configuration or (right panel) the horizontal displacement at a fixed interplanar distance of 3.45 Å for the TTF–TCNQ model system.

starting from these points give real minima without imaginary frequencies; the calculated interaction energies on the order of 7–12 kcal mol⁻¹ are consistent with the presence of complexes with such configurations in polar solvents, as observed experimentally by Tomkiewicz et al.⁴⁴

The overall performance of ω B97X remains quite good for the description of the parallel-displaced structure, where the results can be seen to strongly parallel the SCS-MP2 results. For the case of the Minnesota functionals, M05-2X provides interaction energies that are overall closer to the SCS-MP2 results. Both M05-2X and M06-2X (each containing slightly more than 50% HF exchange) predict a slightly deeper well for the second minimum (at a parallel displacement of approximately 3.0 Å), while M06-HF predicts a much shallower second minimum. All of the functionals presented contain a high admixture of HF exchange and provide a qualitatively correct description of the parallel-displaced potential energy surface.

3. Ground-State Charge Transfer. In D/A systems, the ground-state wave function (Ψ^{GS}) can be expressed as a linear combination of the neutral donor and acceptor components ($\Psi^{\text{D,A}}$) with a small admixture of a charge-transfer state (Ψ^{CT}):

$$\Psi^{\text{GS}} = \Psi^{\text{D,A}} + c\Psi^{\text{CT}} \quad (2)$$

The mixing of $\Psi^{\text{D,A}}$ and Ψ^{CT} in this expression is proportional to the coupling (V) between the covalent and ionic states and inversely proportional to their energy difference (ΔE). When the wave function in eq 2 is taken as a first-order corrected wave function in a perturbative development, the coefficient describing the contribution of the charge-transfer state is given by $c = V/\Delta E$, and the amount of charge transferred in the ground state can be expressed as^{17a} $q \propto (V/\Delta E)^2$. Thus, in the framework of a monodeterminantal description of the ground-state wave function where the charge transfer can be considered as an electron jump from an occupied frontier orbital (HOMO or HOMO-1) of TTF to the LUMO of TCNQ, the energy difference between these orbitals as well as their electronic coupling are the key parameters defining the amount of charge transferred.

The electronic coupling between the orbitals is very sensitive to the relative orientations/positions of the molecules forming the complex⁴⁵ and dominates the evolution of charge transfer with increasing intermolecular distance due to the exponential decrease in wave function overlap. This can be seen when examining the left panel of Figure 5, where q^{GS} is depicted as a

function of intermolecular distance in the cofacial configuration. The amount of charge transferred clearly falls off quickly with increasing separation (decreasing coupling) between the TTF and TCNQ molecules. While all functionals depict this feature, the functionals containing small amounts of HF exchange (B3LYP and B97-1) provide q^{GS} values that fall off more quickly in the intermediate region than for MP2 or the LRC and meta-GGA functionals.

The dependence of q^{GS} on intermolecular separation can also be clearly observed from the results in Table 1. At the SCS-MP2 minimum (3.45 Å), both B3LYP and B97-1 functionals (that contain similar percentages of HF exchange) provide similar descriptions of q^{GS} in the cofacial arrangement and values somewhat larger than for BHandH. However, when comparing the minimum-energy cofacial structures for each functional, the much poorer performance of B3LYP and BHandH in describing the intermolecular separation manifests in a rather poor description of q^{GS} as well. B3LYP predicts a much larger intermolecular separation (see Figure 3) and consequently a much smaller q^{GS} , while BHandH strongly overbinds and presents a very short intermolecular distance and thus a greatly exaggerated q^{GS} value. The overall more reasonable description of the PES afforded by B97-1 provides for q^{GS} values (at the respective minima) that are very close to the SCS-MP2 values.

With respect to the electronic coupling, the significance of the wave function overlap is clearly delineated when comparing the cofacial and parallel-displaced conformations of the TTF–TCNQ complex. The HOMO_{TTF}, HOMO-1_{TTF}, and LUMO_{TCNQ} orbitals (depicted in Figure 6) are all relevant for understanding the ground-state charge transfer in TTF–TCNQ. Both the HOMO-1_{TTF} and LUMO_{TCNQ} orbitals possess *ungerade* symmetry, while the HOMO_{TTF} is of *gerade* symmetry. In the cofacial conformation, the wave function overlap (and thus the electronic coupling) between HOMO_{TTF} and LUMO_{TCNQ} vanishes for symmetry reasons, leaving the HOMO-1_{TTF}–LUMO_{TCNQ} coupling (and energy difference) as the leading contributor to the charge transfer.

This effect can also be observed from the q^{GS} values depicted for the parallel-displaced configuration in Figure 5 (right panel). For a given parallel-displaced geometry, the HOMO_{TTF}–LUMO_{TCNQ} overlap (and the corresponding electronic coupling) depends on the relative positions of the HOMO_{TTF} and LUMO_{TCNQ} nodal surfaces (Figure 6). The minimum observed

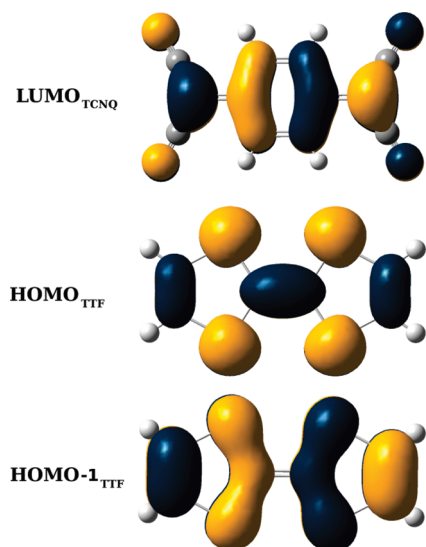


Figure 6. Sketch of the $\text{HOMO-1}_{\text{TTF}}$, HOMO_{TTF} , and $\text{LUMO}_{\text{TCNQ}}$ orbitals.

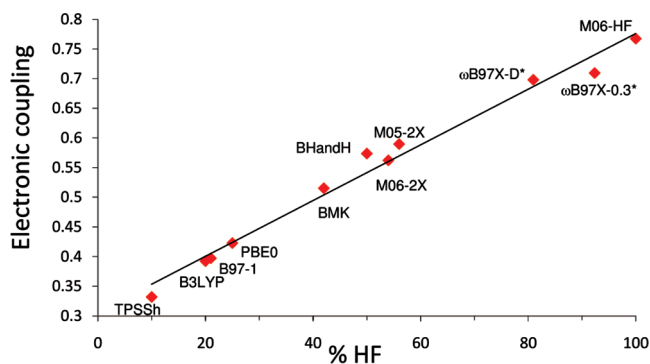


Figure 7. Electronic coupling (eV) between $\text{HOMO-1}_{\text{TTF}}$ and $\text{LUMO}_{\text{TCNQ}}$ for the cofacial TTF-TCNQ model system (3.45 Å intermolecular distance) calculated with various functionals and plotted as a function of % HF exchange.

in the q^{GS} values for parallel displacements around 1.75 Å and the maximum observed for displacements of ca. 3.0 Å are consistent with the $\text{HOMO}_{\text{TTF}}-\text{LUMO}_{\text{TCNQ}}$ couplings of 0.075 and 0.708 eV calculated at these geometries, respectively (absolute values calculated with the ωB97X functional). The decrease in the q^{GS} values with parallel displacements around 1.5–1.75 Å is driven by a decreased $\text{HOMO-1}_{\text{TTF}}-\text{LUMO}_{\text{TCNQ}}$ coupling and a mere increase in the $\text{HOMO}_{\text{TTF}}-\text{LUMO}_{\text{TCNQ}}$ coupling (which can be ascribed to the small $\text{HOMO}_{\text{TTF}}-\text{LUMO}_{\text{TCNQ}}$ overlap value of 0.0057 at 1.75 Å). The importance of employing functionals with a high admixture of HF exchange can be seen when considering the q^{GS} values for the parallel-displaced structures in the right panel of Figure 5, where both B3LYP and B97-1 greatly overestimate the contribution from the $\text{HOMO}_{\text{TTF}}-\text{LUMO}_{\text{TCNQ}}$ coupling and, as a result, the amount of charge transferred.

Thus, changes in geometry and conformation do strongly impact the electronic coupling and charge transfer; also, simple symmetry arguments can explain vanishingly small charge-transfer values. However, for a given geometry, the electronic coupling can still vary largely as a function of the choice of DFT functional. This is demonstrated to be the case for the cofacial geometry, see

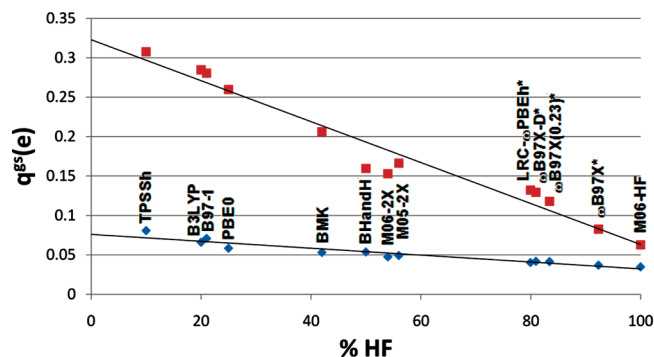


Figure 8. NPA charge transferred in the ground state (q^{GS}) calculated with the different functionals, plotted as a function of the HF exchange for the cofacial (blue) and parallel-displaced (red) configurations of the TTF-TCNQ model complex at a fixed intermolecular distance of 3.45 Å.

Figure 7, where the electronic coupling between $\text{HOMO-1}_{\text{TTF}}$ and $\text{LUMO}_{\text{TCNQ}}$ is calculated to vary by a factor of 2 among the various functionals. The computed couplings follow a nearly linear dependence upon the amount of exchange, increasing by a factor of ~ 2.3 on going from the TPSSh functional (10% HF) to the M06-HF functional (100% HF).

The NPA charges in the ground state (q^{GS}) for both the cofacial configuration and the parallel-displaced configuration of 3.0 Å are included in Table 1 and plotted in Figure 8. The MP2 calculations, which are generally used for the evaluation of net atomic or fragment charges, give values that are between those from HF and TPSSh. For the cofacial geometry at an intermolecular distance of 3.45 Å (absolute minimum in the potential energy surface), the ground-state charge transfer varies by only $0.05e$ among the functionals with a maximum value of $0.08e$ for TPSSh. As underlined above, the small values calculated in the case of the cofacial geometry are a consequence of the symmetries of the HOMO_{TTF} and $\text{LUMO}_{\text{TCNQ}}$ orbitals (Figure 6), which leads to a vanishing electronic coupling; on the other hand, in spite of the rather strong coupling between $\text{HOMO-1}_{\text{TTF}}$ and $\text{LUMO}_{\text{TCNQ}}$, there exists a substantial energetic difference between these orbitals (roughly 2 eV larger than the $\text{HOMO}-\text{LUMO}$ gap, Figure 2).

This situation is very different in the parallel-displaced configuration where the amount of charge transferred increases in some instances to over $0.3e$, see Figure 8. It is interesting to note that, while the coupling increases (in some cases by more than a factor of 2), the amount of charge transferred decreases almost linearly with the admixture of exchange. Recalling that $q \propto (V/\Delta E)^2$, it can be readily understood that the dominant factor is the large (~ 5 eV) increase in the $\text{HOMO}_{\text{TTF}}-\text{LUMO}_{\text{TCNQ}}$ ($\text{HOMO-1}_{\text{TTF}}-\text{LUMO}_{\text{TCNQ}}$) energy splitting in the denominator with increased % HF.

4. Reliability of DFT Charge-Transfer Values. In order to assess the reliability of the results obtained in the framework of the monodeterminantal approach, CASSCF computations were performed for both cofacial and parallel-displaced (corresponding to the greater q^{GS}) geometries. The active space employed includes eight electrons in seven molecular orbitals, although the HOMO_{TTF} , $\text{HOMO-1}_{\text{TTF}}$, and $\text{LUMO}_{\text{TCNQ}}$ orbitals provide the dominant contributions to the charge transfer. In both geometries, the calculated CASSCF NPA charges (Table 1) are lower than $0.1e$ and the coefficients for the charge-transfer configurations in the CASSCF wave function are vanishingly

small. Thus, the wave function for the complex is strongly dominated by the closed-shell electronic configuration (and not the open-shell singlet configuration that would be expected in the limit of full transfer of an electron). Therefore, single-reference approaches such as those employed here should provide a reliable description of the system. These results underline that a small partial charge transfer should be expected in the TTF–TCNQ complex, which appears to be consistent with the experimental observation⁴⁴ that only a limited amount ($\sim 5\%$) of TCNQ anions is detected in a TTF–TCNQ mixture in solution in a highly polar solvent such as acetonitrile (dielectric constant of 36.64) as well as with the values computed from DFT functionals with a high degree of HF exchange.

With regard to the DFT approaches, the electronic coupling values presented in Figure 7 (which vary between ~ 0.3 and 0.8 eV) demonstrate the relatively strong character of the electronic coupling in this system. By analogy with the single molecule electron transport measurements¹⁸ and theoretical results,¹⁹ the amount of charge transferred in this case is expected to vary continuously with the HOMO–LUMO gap, which means that the mono- and multideterminantal methods should give qualitatively similar descriptions. However, the poor description of the HOMO–LUMO gap provided by functionals containing only a small percentage of HF exchange (i.e., TPSSh, B97-1, B3LYP, and PBE0) results in an overestimation of the q^{GS} values. On the other hand, the functionals containing more than 50% HF exchange (i.e., BHandH, M052X, M062X, M06HF, and the ω functionals) give q^{GS} values lower than $0.2e$; this is in qualitative agreement with the CASSCF calculations (reported by Avilov et al.^{17a} and in this study) but also with the experimental observations.^{44,46} However, when considering the minimum-energy geometries (Table 1, last column), the meta-GGA functionals are found to provide q^{GS} values that are slightly greater than $0.2e$ and in lesser agreement with the MP2 value of $0.095e$, while the best agreement is found for the LRC- ω PBEh and ω B97X functionals. Again, the functionals containing a low %HF exchange overestimate the q^{GS} values.

We note that the CASSCF q^{GS} values shown in Table 1 are similar to those from HF, indicating the need for dynamical electron correlation missing from both the HF and CASSCF treatments. Thus, when considering the MP2 value of $0.095e$ (which is generally considered here as a benchmark level), it can be concluded that the ω B97X functional provides reliable results for both ω values. This is also consistent with the comparison made above between the computed HOMO–LUMO gaps and the ~ 3.9 eV experimental difference between IP(TTF) and EA(TCNQ), for which the ω B97X functional also provides very good agreement.

CONCLUSIONS

DFT calculations with a number of functionals have been carried out on a TTF–TCNQ model complex in order to evaluate their description of ground-state charge transfer. We have assessed the performance of the functionals for describing quantities such as the HOMO_{TTF}–LUMO_{TCNQ} energy difference, the binding energy in the ground state, the strength of the electronic coupling between frontier molecular orbitals, and the extent of charge transfer in the ground state.

The inclusion of large amounts (more than 60%) of long-range HF exchange in the functionals seems to be crucial in order to obtain reliable results for all quantities. Among the functionals

considered in this study, the long-range corrected functional ω B97X is found to give reliable results for the ensemble of the calculated properties. Standard hybrid functionals containing only a small percentage of HF exchange are found to be inadequate for properties directly related to the HOMO_{TTF}–LUMO_{TCNQ} energy gap.

For a fixed geometry, a linear relationship is observed between the degree of charge transfer, the HOMO_{TTF}–LUMO_{TCNQ} gap, and the percentage of HF exchange. Both the electronic coupling and the amount of charge transferred are shown to vary largely among the various functionals as well as with changes in the relative positions of the molecules. This highlights the importance of choosing a method capable of describing the potential energy surface as well as the relevant electronic couplings. The role of exchange in the evaluation of electronic couplings is the focus of ongoing work in our group.

AUTHOR INFORMATION

Corresponding Author

*E-mail: jean-luc.bredas@chemistry.gatech.edu.

Notes

[†]On leave from Laboratoire de Physicochimie des Polymères et des Interfaces (LPPI), Université de Cergy-Pontoise, 5 mail Gay-Lussac, F-95031 Cergy-Pontoise Cedex, France.

ACKNOWLEDGMENT

This work has been supported by the Center for Advanced Molecular Photovoltaics, Award No. KUS-C1-015-21, made by King Abdullah University of Science and Technology (KAUST); the Georgia Research Alliance; the STC Program of the National Science Foundation under Award DMR-0120967; and the University of Cergy-Pontoise, France.

REFERENCES

- (1) (a) Mulliken, R. S. *J. Phys. Chem.* **1952**, *56*, 801–822. (b) Mulliken, R. S. *J. Chim. Phys., Chim. Biol.* **1964**, *61*, 20–38. (c) Mulliken, R. S.; Person, W. B. *Annu. Rev. Phys. Chem.* **1962**, *13*, 107–126. (d) Orgel, L. E.; Mulliken, R. S. *J. Am. Chem. Soc.* **1957**, *79*, 4839–4846.
- (2) (a) Kippelen, B.; Bredas, J. L. *Energy Environ. Sci.* **2009**, *2*, 251–261. (b) Bredas, J. L.; Norton, J. E.; Cornil, J.; Coropceanu, V. *Acc. Chem. Res.* **2009**, *42*, 1691–1699.
- (3) Walzer, K.; Maennig, B.; Pfeiffer, M.; Leo, K. *Chem. Rev.* **2007**, *107*, 1233–1271.
- (4) Koster, L. J. A.; Smits, E. C. P.; Mihaiilechi, V. D.; Blom, P. W. M. *Phys. Rev. B* **2005**, *72*.
- (5) (a) Potscavage, W. J.; Sharma, A.; Kippelen, B. *Acc. Chem. Res.* **2009**, *42*, 1758–1767. (b) Potscavage, W. J.; Yoo, S.; Kippelen, B. *Appl. Phys. Lett.* **2008**, *93*, 193308. (c) Rand, B. P.; Burk, D. P.; Forrest, S. R. *Phys. Rev. B* **2007**, *75*, 115327. (d) Waldauf, C.; Scharber, M. C.; Schilinsky, P.; Hauch, J. A.; Brabec, C. J. *J. Appl. Phys.* **2006**, *99*, 104503.
- (6) Qi, Y. B.; Sajoto, T.; Barlow, S.; Kim, E. G.; Bredas, J. L.; Marder, S. R.; Kahn, A., *J. Am. Chem. Soc.* **2009**, *131*, 12530.
- (7) (a) Hamel, S.; Duffy, P.; Casida, M. E.; Salahub, D. R. *J. Electron Spectrosc. Relat. Phenom.* **2002**, *123*, 345–363. (b) Shankar, R.; Senthilkumar, K.; Kollandaivel, P. *Int. J. Quantum Chem.* **2009**, *109*, 764–771. (c) Zhang, G.; Musgrave, C. B. *J. Phys. Chem. A* **2007**, *111*, 1554–1561.
- (8) (a) Hua, X. L.; Chen, X. J.; Goddard, W. A. *Phys. Rev. B* **1997**, *55*, 16103–16109. (b) Manby, F. R.; Knowles, P. J. *J. Chem. Phys.* **2000**, *112*, 7002–7007.
- (9) Perdew, J. P.; Zunger, A. *Phys. Rev. B* **1981**, *23*, 5048–5079.
- (10) (a) Ruzsinszky, A.; Perdew, J. P.; Csonka, G. I.; Vydrov, O. A.; Scuseria, G. E. *J. Chem. Phys.* **2007**, *126*. (b) Braidia, B.; Hiberty, P. C.;

- Savin, A. J. *Phys. Chem. A* **1998**, *102*, 7872–7877. (c) Bally, T.; Sastry, G. N. *J. Phys. Chem. A* **1997**, *101*, 7923–7925.
- (11) (a) Kummel, S.; Kronik, L.; Perdew, J. P. *Phys. Rev. Lett.* **2004**, *93*. (b) Mori-Sanchez, P.; Wu, Q.; Yang, W. T. *J. Chem. Phys.* **2003**, *119*, 11001–11004. (c) van Gisbergen, S. J. A.; Schipper, P. R. T.; Gritsenko, O. V.; Baerends, E. J.; Snijders, J. G.; Champagne, B.; Kirtman, B. *Phys. Rev. Lett.* **1999**, *83*, 694–697.
- (12) (a) Dreuw, A.; Head-Gordon, M. *J. Am. Chem. Soc.* **2004**, *126*, 4007–4016. (b) Dreuw, A.; Weisman, J. L.; Head-Gordon, M. *J. Chem. Phys.* **2003**, *119*, 2943–2946.
- (13) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- (14) (a) Kamiya, M.; Tsuneda, T.; Hirao, K. *J. Chem. Phys.* **2002**, *117*, 6010–6015. (b) Adamson, R. D.; Dombroski, J. P.; Gill, P. M. W. *J. Comput. Chem.* **1999**, *20*, 921–927. (c) Leininger, T.; Stoll, H.; Werner, H. J.; Savin, A. *Chem. Phys. Lett.* **1997**, *275*, 151–160.
- (15) Stein, T.; Kronik, L.; Baer, R. *J. Am. Chem. Soc.* **2009**, *131*, 2818.
- (16) Baer, R.; Neuhauser, D. *Phys. Rev. Lett.* **2005**, *94*, 043002.
- (17) (a) Avilov, I.; Geskin, V.; Cornil, J. *Adv. Funct. Mater.* **2009**, *19*, 624–633. (b) Geskin, V.; Stadler, R.; Cornil, J. *Phys. Rev. B* **2009**, *80*, 085411.
- (18) Danilov, A.; Kubatkin, S.; Kafanov, S.; Hedegard, P.; Stuhr-Hansen, N.; Moth-Poulsen, K.; Bjornholm, T. *Nano Lett.* **2008**, *8*, 1–5.
- (19) Toher, C.; Filippetti, A.; Sanvito, S.; Burke, K. *Phys. Rev. Lett.* **2005**, *95*, 146402.
- (20) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.
- (21) Tao, J. M.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys. Rev. Lett.* **2003**, *91*, 146401.
- (22) Hamprecht, F. A.; Cohen, A. J.; Tozer, D. J.; Handy, N. C. *J. Chem. Phys.* **1998**, *109*, 6264–6271.
- (23) Wilson, P. J.; Bradley, T. J.; Tozer, D. J. *J. Chem. Phys.* **2001**, *115*, 9233–9242.
- (24) (a) Ernzerhof, M.; Scuseria, G. E. *J. Chem. Phys.* **1999**, *110*, 5029–5036. (b) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158–6170.
- (25) Boese, A. D.; Martin, J. M. L. *J. Chem. Phys.* **2004**, *121*, 3405–3416.
- (26) As implemented in Gaussian 09.
- (27) As implemented in Gaussian 09.
- (28) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2*, 364–382.
- (29) Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157–167.
- (30) Chai, J. D.; Head-Gordon, M. *J. Chem. Phys.* **2008**, *128*, 084106.
- (31) Chai, J. D.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.
- (32) Rohrdanz, M. A.; Martins, K. M.; Herbert, J. M. *J. Chem. Phys.* **2009**, *130*, 054112.
- (33) (a) Vydrov, O. A.; Heyd, J.; Krukau, A. V.; Scuseria, G. E. *J. Chem. Phys.* **2006**, *125*, 074106. (b) Vydrov, O. A.; Scuseria, G. E. *J. Chem. Phys.* **2006**, *125*, 234109.
- (34) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B. *Gaussian 09*, revision A.1; Gaussian, Inc.: Pittsburgh, PA, 2009.
- (35) Kong, J.; White, C. A.; Krylov, A. I.; Sherrill, D.; Adamson, R. D.; Furlani, T. R.; Lee, M. S.; Lee, A. M.; Gwaltney, S. R.; Adams, T. R.; Ochsenfeld, C.; Gilbert, A. T. B.; Kedziora, G. S.; Rassolov, V. A.; Maurice, D. R.; Nair, N.; Shao, Y. H.; Besley, N. A.; Maslen, P. E.; Dombroski, J. P.; Daschel, H.; Zhang, W. M.; Korambath, P. P.; Baker, J.; Byrd, E. F. C.; Van Voorhis, T.; Oumi, M.; Hirata, S.; Hsu, C. P.; Ishikawa, N.; Florian, J.; Warshel, A.; Johnson, B. G.; Gill, P. M. W.; Head-Gordon, M.; Pople, J. A. *J. Comput. Chem.* **2000**, *21*, 1532–1548.
- (36) Werner, H.-J.; Knowles, P. J.; Lindh, R.; Manby, F. R.; Schütz, M.; Celani, P.; Korona, T.; Mitrushenkov, A.; Rauhut, G.; Adler, T. B.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Goll, E.; Hampel, C.; Hetzer, G.; Hrenar, T.; Knizia, G.; Köppl, C.; Liu, Y.; Lloyd, A. W.; Mata, R. A.; May, A. J.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklass, A.; Palmieri, P.; Pflüger, K.; Pitzer, R.; Reiher, M.; Schumann, U.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T.; Wang, M.; Wolf, A. *MOLPRO*, version 2009.1; University College Cardiff Consultants Limited: Cardiff, U. K., 2009.
- (37) Moller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 0618–0622.
- (38) Grimme, S. *J. Chem. Phys.* **2003**, *118*, 9095–9102.
- (39) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.
- (40) Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 735–746.
- (41) Valeev, E. F.; Coropceanu, V.; da Silva, D. A.; Salman, S.; Bredas, J. L. *J. Am. Chem. Soc.* **2006**, *128*, 9882–9886.
- (42) Lichtenberger, D. L.; Johnston, R. L.; Hinkelmann, K.; Suzuki, T.; Wudl, F. *J. Am. Chem. Soc.* **1990**, *112*, 3302–3307.
- (43) Compton, R. N.; Cooper, C. D. *J. Chem. Phys.* **1977**, *66*, 4325–4329.
- (44) Tomkiewi, Y.; Torrance, J. B.; Scott, B. A.; Green, D. C. *J. Chem. Phys.* **1974**, *60*, 5111–5112.
- (45) Bredas, J. L.; Calbert, J. P.; da Silva, D. A.; Cornil, J. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 5804–5809.
- (46) (a) Yuge, R.; Miyazaki, A.; Enoki, T.; Tamada, K.; Nakamura, F.; Hara, M. *J. Phys. Chem. B* **2002**, *106*, 6894–6901. (b) Liao, J. Y.; Ho, K. C. *Sens. Actuators, B*, **2008**, *130*, 343–350.

Self-Consistent Field and Polarizable Continuum Model: A New Strategy of Solution for the Coupled Equations

Filippo Lipparini,^{*,†} Giovanni Scalmani,[‡] Benedetta Mennucci,[§] and Michael J. Frisch[‡]

[†]Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy

[‡]Gaussian, Inc., 340 Quinpiac Street Building 40, Wallingford, Connecticut 06492, United States

[§]Dipartimento di Chimica e Chimica Industriale, Università di Pisa, Via Risorgimento 35, 56126 Pisa, Italy

 Supporting Information

ABSTRACT: We present a new strategy for the solution of the self-consistent field (SCF) equations when solvent effects are included by means of the polarizable continuum model (PCM). By exploiting the recently introduced variational formalism of the PCM (VPCM), we are able to recast the self-consistent reaction field problem as an energy functional of both electronic and polarization degrees of freedom. The variational minimization of such a functional leads to the free energy of the solvated molecule at a given geometry. In this contribution we describe an effective procedure and its implementation to achieve the solution of such a variational problem. Moreover, we present numerical evidence that the new approach is superior to the traditional one in terms of performance, especially when a relatively inexpensive semiempirical method is used to describe medium- and large-size solutes.

1. INTRODUCTION

The solution of the independent particle problem by means of a self-consistent field (SCF) approach is probably the most common type of calculation performed today in the field of computational chemistry. Indeed, while they differ in the definition of the effective one-electron Hamiltonian, methods like Hartree–Fock (HF), the Kohn–Sham (KS) formulation of density functional theory (DFT), and many less accurate and computationally less expensive semiempirical methods^{1–3} share the SCF procedure as the approach to solve the underlying nonlinear eigenvalue equations. Moreover, SCF-level calculations are so widely used because they still represent a good compromise between computational cost and accuracy of the results, while the maximum size of the systems that can be studied using off-the-shelf computers increases steadily with the progress of both hardware and software technologies.

Recent efforts to extend the parametrization of semiempirical methods beyond the second row of the periodic table have been fairly successful and have led to the introduction of the PM6 method.³ These developments have the potential of making semiempirical methods a rather attractive alternative (or complement) to *ab initio* techniques, as they drastically increase the size of the molecular systems that can be approached, while keeping the computational cost manageable and still being able to reproduce nonclassical effects. In fact, despite their use of parametrization, semiempirical methods remain quantum mechanical (QM) in their principles.

The range of applicability of any method involving the SCF procedure has been further extended by their generalization to the treatment of molecules in the condensed phase. Among the models developed to account for the effects of the microscopic environment on molecular structure, properties and processes, the polarizable continuum model^{4–6} (PCM) is one of the most successful, thanks

to its generality and its versatility. According to this model, the solvent is described by means of a structureless polarizable dielectric medium. This leads to a dramatic reduction in the number of degrees of freedom as compared to explicit solvent models, while preserving a sufficiently accurate description of the interaction between solute and solvent. In addition, despite the apparently oversimplistic description of the solvent, the PCM model has been successfully applied, with very limited increase in complexity, to less uniform environments,⁷ such as anisotropic solvents, ionic solutions, interfaces between two different liquids, membranes, etc.

The polarization field (the reaction field) of the dielectric is represented by means of an apparent surface charge (ASC) distribution, located at the boundary between the dielectric medium and an empty cavity hosting the solute molecule (the molecular cavity). Over the years, various strategies have been proposed and discussed that would turn the PCM formal setup into a robust and efficient computational tool.^{8–10} In the most recent formulation, the ASC is represented with a continuous charge distribution,^{11,12} and it is expanded in a basis set of spherical Gaussian functions. The expansion coefficients, which have the physical meaning of surface polarization charges, are obtained as the solution of a linear system of equations whose dimension is given by the number of surface elements used to discretize the surface of the molecular cavity. The right-hand side of the PCM linear system depends on the solute structure and the charge distribution through its electrostatic potentials. In the case of polarizable solutes, i.e., when a QM method is used to describe the solute's density, the overall problem becomes nonlinear, and the optimal (in the variational sense) mutual polarization of both the solute and the solvent must be achieved.

Received: October 15, 2010

Published: January 24, 2011

From the formal point of view, the inclusion of the effect of the environment in the QM description of the solute is not a difficult task, as the SCF nonlinear equations maintain the same structure and the one-electron effective Hamiltonian (the Fock operator) is simply augmented by a term that represents the polarization field. The latter, in turn, is updated at each SCF cycle using the new solute's density and solving the PCM problem to evaluate the ASC. The computational machinery needed to perform an SCF calculation in the gas phase or in solution is substantially the same in the sense that no new integrals or other difficult to compute quantities are required to form the additional solvent response term in the Fock operator.¹²

On the other hand, the relative cost of the solution of the SCF equation in the gas phase and in solution may vary enormously. The cost of solving the PCM equations depends on the number of surface elements used to discretize the solute–solvent interface, and the latter is fixed once the solute geometry is known. However, the cost of computing the gas-phase Fock operator and the few additional quantities involving both solute and solvent (e.g., the solute's electrostatic potential at the surface elements) can vary dramatically depending on the underlying choice of the QM method. The inclusion of solvent effects according to the PCM model adds very little to the overall cost of an accurate DFT calculation on a small or medium system, using a large basis set; on the contrary, when a cheap semiempirical Hamiltonian is used to describe a solute containing several hundred atoms, the cost of PCM quickly becomes the most expensive step. This is a remarkable disadvantage and prevents the application of PCM to study large and complex systems in solution using semiempirical methods, while the same calculations in the gas phase can be performed using very little computational resources.

Recently, a variational formulation of the PCM has been introduced.¹³ In this new scheme, hereafter called VPCM, the polarization charges are not defined as the solution of a linear system of equations but rather as the location within the space of the polarization degrees of freedom where a suitable free energy functional is minimum. In other words, the PCM equations are recast as a minimization problem. A free energy functional is defined which depends on both the structural and electronic degrees of freedom of the solute and on the solvent polarization degrees of freedom. When variationally minimized with respect to all these parameters, the PCM energy functional provides the free energy of the polarized solute at its equilibrium geometry in solution.

There are many application under consideration for the VPCM free energy functional such as: (i) simultaneous optimization of geometry, electronic density, and polarization charges to reduce the overall computational cost of a SCF level calculation; (ii) evaluation of the free energy in solution in the context of post-SCF methods, where the reaction field introduces a coupling among the various quantities that contribute to the one-particle density matrix, without resorting to very expensive iterative procedures;^{14,15} (iii) application of PCM in connection with extended Lagrangian ab initio molecular dynamic methods¹⁶ (e.g., Car–Parrinello¹⁷ or atom-centered density matrix propagation, ADMP).^{18,19}

In ref 13 a scheme to simultaneously optimize both the molecular geometry and the polarization charges has been proposed and applied successfully to solutes described by a molecular mechanics (MM) function. In this contribution, we describe a similar approach for the simultaneous optimization of the solute electronic density and the PCM polarization charges

within the SCF procedure, which we shall use, to maintain consistency, the same VPCM acronym. By defining the proper free energy functional of both the density matrix and the polarization charges and carrying out a variational minimization with respect to both classes of parameters, we show that the functional minimum corresponds to the correct free energy in solution for the chosen form of the one-electron Hamiltonian and for a given solute geometry. Within this scheme it is possible to achieve an optimal partition of computational resources between the calculation of the gas phase and the solvent-related terms in the Fock operator, and we demonstrate that this is particularly important when the cost of the gas-phase terms is small like, e.g., when a semiempirical Hamiltonian is used. In practical terms, we show how to exploit standard direct inversion in the iterative subspace (DIIS)^{20–22} extrapolation techniques, which are commonly used to improve SCF convergence, to accelerate the minimization of the free energy functional with respect to both the density matrix and the PCM polarization charges.

This paper is organized as follows. In Section 2 we shortly recall the basic formalism of the PCM model, both in its standard form and according to the new variational scheme. A free energy functional suitable for the simultaneous solution of the SCF and PCM equations is defined, and the associated working equations are derived. Lastly, the implementation of the coupled minimization algorithm is discussed. In Section 3 we report the results of some illustrative calculations including, in particular, semiempirical single point calculations and geometry optimizations.

2. THEORY

2.1. Formalism. The PCM model solves the Poisson equation in the presence of a dielectric medium outside an empty cavity C which hosts the solute. The surface $\Gamma = \partial C$ is the boundary of this cavity and represents the interface between solute and solvent. The electrostatic potential $\varphi(\mathbf{r})$ is the solution of Poisson's equation:

$$\nabla \cdot [\varepsilon(\mathbf{r}) \nabla \varphi(\mathbf{r})] = -4\pi\rho(\mathbf{r}) \quad (1)$$

where $\rho(\mathbf{r})$ is the (nuclear and electronic) charge density of the solute. The dielectric constant function assumes, for a homogeneous solvent, the simple form:

$$\varepsilon(\mathbf{r}) = \begin{cases} 1 & \mathbf{r} \in C \\ \varepsilon & \mathbf{r} \in C \end{cases} \quad (2)$$

where ε is the macroscopic dielectric permittivity of the solvent. Using eq 2 and the appropriate boundary conditions, the problem in eq 1 is solved, and the polarization of the medium is represented by an ASC density $\sigma(\mathbf{s})$ with $\mathbf{s} \in \Gamma$. The ASC density is the solution of an integral equation whose form varies according to which member of the PCM family of models is being used. The integral equation formalism PCM (IEFPCM)^{23–25} (hereafter simply called “PCM”) is the model of choice because of its broader applicability to all values of ε and because it represents the optimal compromise between accuracy and formal complexity. The integral equation for the PCM model reads

$$\left(\frac{\varepsilon + 1}{\varepsilon - 1} \hat{\mathcal{J}} - \frac{1}{2\pi} \hat{\mathcal{D}} \right) \hat{\mathcal{J}} \sigma(\mathbf{s}) = - \left(\hat{\mathcal{J}} - \frac{1}{2\pi} \hat{\mathcal{D}} \right) \Phi(\mathbf{s}) \quad (3)$$

where $\Phi(\mathbf{r})$ is the solute's electrostatic potential, $\hat{\mathcal{J}}$ is the identity operator, while $\hat{\mathcal{J}}$ and $\hat{\mathcal{D}}$ (together with its adjoint $\hat{\mathcal{D}}^*$) are components of the so-called Calderon projector.²⁶ Their expressions are defined as

$$\hat{\mathcal{J}}\sigma(\mathbf{s}) = \int_{\Gamma} \frac{\sigma(\mathbf{s}')}{|\mathbf{s} - \mathbf{s}'|} d^2\mathbf{s}' \quad (4)$$

$$\hat{\mathcal{D}}^*\sigma(\mathbf{s}) = \int_{\Gamma} \left\{ \frac{\partial}{\partial \hat{\mathbf{n}}_s} \frac{1}{|\mathbf{s} - \mathbf{s}'|} \right\} \sigma(\mathbf{s}') d^2\mathbf{s}' \quad (5)$$

$$\hat{\mathcal{D}}\sigma(\mathbf{s}) = \int_{\Gamma} \left\{ \frac{\partial}{\partial \hat{\mathbf{n}}_{s'}} \frac{1}{|\mathbf{s} - \mathbf{s}'|} \right\} \sigma(\mathbf{s}') d^2\mathbf{s}' \quad (6)$$

where $\hat{\mathbf{n}}_s$ is the outward normal direction to the surface at point $\mathbf{s} \in \Gamma$. Taking the limit for a large dielectric constant, the conductor-like PCM^{27–29} (CPCM) model is recovered, which is still very accurate in the case of fairly polar solvents, and it has the advantage of being simpler than the more general dielectric model as the CPCM integral equation reads

$$\hat{\mathcal{J}}\sigma(\mathbf{s}) = -f(\epsilon)\Phi(\mathbf{s}) \quad (7)$$

where $f(\epsilon) = (\epsilon - 1)/\epsilon$.

Practical applications of all PCM models require a discrete representation of the ASC density over the solute–solvent interface. Recently, a CSC formalism,¹² which allows for a robust and smooth definition of all the discretized quantities, has been described. The ASC is expanded in a basis of spherical Gaussian functions:

$$\sigma(\mathbf{r}) = \sum_i \frac{q_i}{a_i} \phi_i(\mathbf{r}; \mathbf{s}_i, \zeta_i) \quad (8)$$

where a_i is the area of the i -th surface element, q_i is the corresponding expansion coefficient of σ (which has the physical dimensions of a charge), and \mathbf{s}_i is the representative point of the i -th surface element. By defining

$$\mathbf{R}_\epsilon = \frac{\epsilon + 1}{\epsilon - 1} \mathbf{1} - \frac{1}{2\pi} \mathbf{D}\mathbf{A} \quad (9)$$

$$\mathbf{T}_\epsilon = \mathbf{R}_\epsilon \mathbf{S} \quad (10)$$

and

$$\mathbf{R}_\infty = \lim_{\epsilon \rightarrow \infty} \mathbf{R}_\epsilon = \mathbf{1} - \frac{1}{2\pi} \mathbf{D}\mathbf{A} \quad (11)$$

the discretized IEFPCM equation reads

$$\mathbf{T}_\epsilon \mathbf{q} = -\mathbf{R}_\infty \mathbf{V} \quad (12)$$

where $\mathbf{V}_i = \Phi(\mathbf{s}_i)$.

As discussed in detail in the first VPCM paper,¹³ the PCM problem can be recast in a variational fashion (VPCM) by defining a suitable functional, whose minimization corresponds to the solution of eq 12. This functional assumes the following form:

$$\mathcal{G}(\tilde{\mathbf{q}}) = \frac{1}{2} \tilde{\mathbf{q}}^\dagger \tilde{\mathbf{Y}} \tilde{\mathbf{q}} + \tilde{\mathbf{q}}^\dagger \tilde{\mathbf{V}} = \frac{1}{2} \tilde{\mathbf{q}}^\dagger \mathbf{Y} \tilde{\mathbf{q}} + \tilde{\mathbf{q}}^\dagger \tilde{\mathbf{V}} \quad (13)$$

where $\tilde{\mathbf{q}} = \mathbf{R}_\infty^\dagger \mathbf{q}$, $\tilde{\mathbf{V}} = \mathbf{R}_\infty \mathbf{V}$, and

$$\left(\frac{\mathbf{T}_\epsilon \mathbf{R}_\infty^\dagger + \mathbf{R}_\infty \mathbf{T}_\epsilon^\dagger}{2} \right) = \frac{\mathbf{Y} + \mathbf{Y}^\dagger}{2} = \tilde{\mathbf{Y}} \quad (14)$$

The minimum of the functional in eq 13 with respect to the $\tilde{\mathbf{q}}$ charges is obtained by setting to zero its first derivative:

$$\frac{\partial \mathcal{G}(\tilde{\mathbf{q}})}{\partial \tilde{\mathbf{q}}} = \tilde{\mathbf{Y}} \tilde{\mathbf{q}} + \tilde{\mathbf{V}} = 0 \quad (15)$$

which is a condition equivalent to the solution of eq 12, as it can be seen by substituting in eq 15 the definitions of the appropriate quantities.

The only molecular property of the solute that appears in the expression of the free energy functional eq 13 is the transformed electrostatic potential $\tilde{\mathbf{V}}$ produced by the solute at the surface Γ . This is defined as

$$\begin{aligned} \mathbf{V}_i &= \sum_{\mu\nu} \left[\frac{\int d^3\mathbf{r} d^3\mathbf{r}' \phi_i(\mathbf{r}; \mathbf{s}_i, \zeta_i) \chi_\mu(\mathbf{r}') \chi_\nu(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \right] \mathbf{P}_{\mu\nu} \\ &= \sum_{\mu\nu} \mathbf{V}_{\mu\nu} \mathbf{P}_{\mu\nu} \end{aligned} \quad (16)$$

where $\mathbf{P} = \mathbf{C}\mathbf{C}^\dagger$ is the electronic density matrix and \mathbf{C} is the matrix of the molecular orbitals coefficients in the atomic basis set. More precisely eq 16 represents only the electronic term in the electrostatic potential, while the nuclear contribution is constant during the SCF procedure. By means of eq 13, it is possible to define a functional of both the density matrix \mathbf{P} and the transformed polarization charges $\tilde{\mathbf{q}}$ as

$$\mathcal{G}(\mathbf{P}, \tilde{\mathbf{q}}) = \langle \mathbf{h}\mathbf{P} \rangle + \frac{1}{2} \langle \mathbf{G}(\mathbf{P})\mathbf{P} \rangle + \frac{1}{2} \tilde{\mathbf{q}}^\dagger \mathbf{Y} \tilde{\mathbf{q}} + \tilde{\mathbf{q}}^\dagger \langle \tilde{\mathbf{V}}\mathbf{P} \rangle \quad (17)$$

where $\langle \dots \rangle$ denotes the matrix trace, while \mathbf{h} and $\mathbf{G}(\mathbf{P})$ are the usual one- and two-electron matrices. The last term on the right-hand side of eq 17 expresses the explicit dependence of the electrostatic potential on the electronic density \mathbf{P} . By minimizing eq 17 with respect to both \mathbf{P} , which is to be constrained to be idempotent, the following coupled equations are obtained

$$\begin{cases} \mathbf{F}(\mathbf{P}, \tilde{\mathbf{q}})\mathbf{C} = \mathbf{S}\mathbf{C}\epsilon \\ \tilde{\mathbf{Y}} \tilde{\mathbf{q}} = -\tilde{\mathbf{V}}(\mathbf{P}) \end{cases} \quad (18)$$

where

$$\mathbf{F}_{\mu\nu} = \mathbf{h}_{\mu\nu} + \mathbf{G}_{\mu\nu}(\mathbf{P}) + \tilde{\mathbf{V}}_{\mu\nu}^\dagger \tilde{\mathbf{q}}$$

is the Fock operator, including the solute–solvent interaction term. The first line in eq 18 is the well-known Roothan equation, which is nonlinear in the density matrix \mathbf{P} and is coupled to the second set of equations by the dependence of the Fock operator on the PCM charges $\tilde{\mathbf{q}}$. On the other hand, the PCM equations are coupled to the Roothan equations through the dependence of the transformed electrostatic potential on the density matrix.

2.2. Implementation. The convergence of the SCF procedure can be greatly improved by means of Pulay's DIIS^{20,21} extrapolation technique. Indeed, an extrapolation technique is usually required to ensure the convergence of the solution of the Roothan equations, since a straightforward repeated diagonalization is typically very slowly converging or not converging at all.

The idea underlying the DIIS approach is rather simple. Provided a sequence $\{x_j\}$ of points obtained by means of an iterative procedure to minimize a function, an estimate x^{ext} of the solution is obtained by an affine combination of the points with coefficients chosen to minimize a suitable error, for example, the gradient of the function itself. This leads to a linear system of

$N + 1$ equations, where N is the number of points, which can be solved by matrix inversion as N is usually a small number.

When DIIS is applied to help the convergence of the iterative SCF problem, the error vector is built from the commutator of the Fock matrix and the density matrix:

$$\mathbf{e}_i = [\mathbf{F}^i, \mathbf{P}^i]$$

which vanishes when the SCF equation is satisfied. However, it has been proven²² that using only the commutator as a measure of the DIIS error does not guarantee convergence and, in particular, does not guarantee convergence on the electronic ground state, since excited-state solutions are also characterized by vanishing commutators. Including the energy²² as an additional source of error (EDIIS) has been proven to be successful in guiding the SCF toward the ground-state solution.

We point out that if the Fock matrix is linear with respect to the density, as in the case of Hartree–Fock and most semiempirical methods

$$\mathbf{G}(\mathbf{P}^{\text{ext}}) = \tilde{\mathbf{G}}^{\text{ext}}$$

which allows for the extrapolation of the Fock matrix instead of the density, while extrapolating the density will require an additional Fock matrix formation at each SCF cycle. The above relation does not hold exactly in the case of DFT, where the exchange–correlation term in the Fock matrix is not linear in the density; however, it is possible to proceed in the same way. Typically, this will make the convergence of the SCF procedure somewhat slower but not enough to make the procedure involving the proper extrapolation of the density computationally competitive, as the latter requires two Fock matrix formation per cycle.

In our simultaneous optimization approach, we solve the coupled problem, eq 18, with the following procedure. Starting from a guess of both the density matrix \mathbf{P} and the charges $\tilde{\mathbf{q}}$, an update step is performed on both set of variables. The density update step is the standard Fock matrix diagonalization, which provides a new set of molecular orbital coefficients \mathbf{C} , while for the charges a scaled steepest-descent step is computed. We have chosen such a simple update step for the polarization charges because of its better performance when used in connection of the DIIS extrapolation (vide infra). Update methods that require the knowledge of two or more points, such as conjugate gradient or quasi-Newton methods associated with line search, have instead shown a much worse behavior when coupled with the DIIS extrapolation.

In order to compute the update step on the charges $\tilde{\mathbf{q}}$, the product between the relevant PCM matrix and a vector of charges must be evaluated. In the case of the general dielectric model, the $\tilde{\mathbf{Y}}$ matrix is used, while the \mathbf{S} matrix is required for CPCM. Depending on the available memory, the matrix–vector product may be computed on the fly, possibly exploiting a linear scaling technique, such as the fast multipole methods,^{9,30} or the matrix can be fit in memory, and the matrix–vector product carried out with quadratic cost. While the on-the-fly evaluation of the matrix–vector product by the CPCM matrix is very efficient, the $\tilde{\mathbf{Y}}$ matrix is less easily tractable as its formation involves two matrix–matrix multiplications, while the direct evaluation of the matrix–vector products requires four multiplications by \mathbf{S} and two multiplications by \mathbf{D} . Typically, if the available memory allows for the efficient formation of $\tilde{\mathbf{Y}}$, then the matrix is computed, stored in memory or on disk, and reused whenever possible. The availability of the $\tilde{\mathbf{Y}}$ matrix has the additional advantage of providing the exact diagonal elements. These are

used to scale the gradient of functional eq 17 with respect to the charges and therefore compute exactly its minimum along the steepest-descent direction, which has a closed expression in the case of a quadratic form.

From the update step on the density \mathbf{P} we get a new density and therefore a new electrostatic potential:

$$\mathbf{P}^i, \tilde{\mathbf{V}}^i \rightarrow \mathbf{P}^{i+1}, \tilde{\mathbf{V}}^{i+1}$$

while from the update step on the charges we obtain a new set of charges

$$\tilde{\mathbf{q}}^i \rightarrow \tilde{\mathbf{q}}^{i+1}$$

The DIIS extrapolation is then performed for both sets of variables at the same time. As error vector, we choose the sum of the usual DIIS error (i.e., the Fock matrix–density matrix commutator) and the gradient of functional eq 17 with respect to the charges $\tilde{\mathbf{q}}$. The two components of the error vectors are weighed by a suitable constant factor, so that

$$\mathbf{B}_{i,j} = w_F \langle \mathbf{e}_F^i, \mathbf{e}_F^j \rangle + w_Q \langle \mathbf{e}_Q^i, \mathbf{e}_Q^j \rangle \quad (19)$$

where

$$\mathbf{e}_F^i = [\mathbf{F}^i, \mathbf{P}^i], \quad \mathbf{e}_Q^i = \tilde{\mathbf{Y}}\tilde{\mathbf{q}}^i + \tilde{\mathbf{V}}^i$$

In eq 19, \mathbf{B} is the $(N + 1)$ square Pulay's matrix for $i, j \leq N$, and $\langle \cdot, \cdot \rangle$ denotes the canonical scalar product. For EDIIS we simply use the corresponding values of the free energy functional as in eq 17.

In our implementation the two-electron component of the Fock matrix and the PCM gradient are extrapolated. In fact, we note that the functional in eq 17 is the sum of a linear term involving the one-electron matrix \mathbf{h} , whose derivative with respect to the density matrix is a constant and of three quadratic terms in the density and the charges. In other words, the quadratic terms in the free energy functional are bilinear with respect to the variables \mathbf{P} and $\tilde{\mathbf{q}}$, and therefore we can extrapolate the two gradients, i.e., the two-electron part of the Fock matrix and the PCM gradient, thanks to their linearity. More in detail, the extrapolation step gives

$$\begin{aligned} \mathbf{F}_{\mu\nu}^{i+1, \text{ext}} &= \mathbf{h}_{\mu\nu} + \mathbf{G}_{\mu\nu}(\mathbf{P})^{i+1, \text{ext}} + \tilde{\mathbf{V}}_{\mu\nu}^+ \tilde{\mathbf{q}}^{i+1, \text{ext}} \\ \mathbf{g}^{i+1, \text{ext}} &= \tilde{\mathbf{Y}}\tilde{\mathbf{q}}^{i+1, \text{ext}} + \tilde{\mathbf{V}}^{i+1, \text{ext}} \end{aligned}$$

where

$$\tilde{\mathbf{V}}^{i+1, \text{ext}} = \langle \tilde{\mathbf{V}}\mathbf{P}^{i+1, \text{ext}} \rangle$$

and we obtain a consistent and coherent extrapolation of both the PCM contribution to the Fock matrix, which depends on the extrapolated charges $\tilde{\mathbf{q}}^{i+1, \text{ext}}$ and the gradient of the free energy functional with respect to the charges, which depends on the extrapolated density $\mathbf{P}^{i+1, \text{ext}}$ through the extrapolated transformed potential.

3. NUMERICAL TESTS

In this section, we report some numerical results obtained with the procedure previously described in section 2. First, we discuss the significance of the few parameters that control the behavior of the algorithm and the choice of their optimal value. Then we describe the results obtained on several medium- to large-sized systems, and we compare them with the corresponding ones obtained with the traditional approaches.

All the calculations have been performed with the development version of the Gaussian suite of programs.³¹

3.1. Setup Calculations. There are three main issues involved in the set up of a simultaneous solution algorithm which will optimize the density matrix and the polarization charges, as described in Section 2. First, an initial guess for the values of the polarization charges must be formulated. Then, the optimal relative weights for the two components w_F and w_Q of the DIIS error in eq 19 must be established, and last, a suitable convergence criterion must be set on the PCM polarization charges or on the gradient of the free energy functional with respect to the charges.

We explored various possible choices for the initial guess of the polarization charges. In particular, we considered: (i) simply starting from zero charges; (ii) using a uniform distribution of charges summing up to the solute total charge; (iii) obtaining the initial charges from a diagonal approximation of the PCM equations, i.e., as

$$\tilde{q}_i^0 = -\frac{\tilde{V}_i^0}{\tilde{Y}_{ii}}$$

or (iv) from a low-accuracy solution of the PCM equations using an iterative procedure with a convergence criteria on the root-mean-square (rms) charge residue in the range of 10^{-1} – 10^{-4} . In many instances we found that starting from partially converged charges improves only slightly the overall convergence of the SCF procedure, while adding an extra computational cost that could be significant depending on the size of the system and the choice of the one-electron Hamiltonian. Our conclusion is therefore to set all the charges to zero as initial guess.

The optimal choice of the relative weights for the two components of the DIIS error, as in eq 19, has been achieved by considering the following. It should be realized immediately that the absolute value of the two components of the error may differ by orders of magnitude. In particular, at the beginning of the SCF procedure, the guess for the density is generally good as it corresponds to the diagonalization of a Fock matrix based on the Harris functional.³² Moreover, our choice of the initial guess for the polarization variables is likely to make the initial error on the charges consistently much larger than the one on the density. The big difference in the initial values of the two error components may lead to DIIS coefficients that overestimate the effect of the changes in the PCM terms of the energy functional. This, in turn, is likely to affect in a negative way the overall convergence behavior. Indeed, it has become apparent from our tests that the effect of the changes in the density is the most important one. Especially when the iterative procedure is still far from the solution, it is crucial to make the change in the density and not the change in the PCM charges, the driving force in the DIIS extrapolation. For this reason, we decided to weight the density error with a larger coefficient. After analyzing the results of a number of tests, we concluded that a good value for the ratio between w_F and w_Q is 1000:1. This ratio leads to contributions to the Pulay's DIIS matrix from the two sources of error, of roughly the same order of magnitude, while at the beginning of the SCF procedure, the density component of the error is somewhat larger. We point out that this is an arbitrary empirical choice, which is the result of a limited set of numerical tests, and it is likely to be just an approximation of the optimal weighing of the two error components. We set $w_F = 1$ and $w_Q = 0.001$, as this choice has led to good convergence behavior using both semiempirical and HF or DFT Hamiltonians.

Table 1. Systems Studied^a

	alias	N_{at}	N_{ch}	charge
diphthamide	–	48	3185	+1
cevine	–	73	3816	0
dammarane	–	84	4055	0
tuftsin	–	77	4866	0
Ca(heme)	–	73	4990	–2
substituted guanidine	sGua	84	5327	+2
Ac(Ala) ₉ NH	Ala9	99	6030	0
taxol	–	113	6460	0
(Ala) ₂₅	Ala25	259	13 767	0
(dA–dT) ₅	DNA-5	318	17 202	–8
(dA–dT) ₆	DNA-6	382	19 979	–10
(dA–dT) ₇	DNA-7	510	26 473	–12
crambin	–	642	30 271	0

^a N_{at} = number of atoms, N_{ch} = number of PCM charges, Ac = acetyl, Ala = alanine, dA = deoxyadenosyne, and dT = deoxythymine.

Finally, the convergence criterion for the polarization charges must be chosen so that the simultaneous variational minimization of the free energy functional in eq 17 is accurate enough to be used in a solute geometry optimization. In other words, the rms norm of the derivatives of eq 17 with respect to the polarization charges must be small enough that they will not reduce the accuracy of the forces, i.e., the derivatives of the free energy with respect to the atomic positions. Thus, we compared the final values of both energy and forces obtained requiring the rms norm of the PCM charge gradient to be smaller than 10^{-N} (and its maximum element to be smaller than 10^{-N+1}), with the values obtained by solving “exactly” the PCM equations by means of matrix inversion with iterative refinement (within double precision). Considering the results of a number of tests, we choose N so that the differences in energy were consistently smaller than 10^{-7} Hartrees and the rms and maximum difference in the forces were smaller than the thresholds currently used in a “very tight” geometry optimization (rms force $< 10^{-5}$ and maximum force $< 2 \times 10^{-5}$). Our results obtained using the PM6³ semiempirical method for a set of small- to medium-sized solutes show that $N = 4$ is usually enough for the requirements set forth above to be met.

3.2. Numerical Results. In this section, we report the results obtained for a set of molecules using both the CPCM and IEFPCM in their standard and variational formulation. The molecular structures are available in the Supporting Information.

To better understand the “computational size” of the systems we studied, we list in Table 1 the number of atoms (N_{at}) and the number of surface elements, i.e. the number of PCM charges N_{ch} for each solute, and we also include the value of the total charge.

We have included in our test set medium- to large or very large-sized molecules. Some of the test solutes are charged either positively or negatively in order to better explore the convergence of our algorithm in cases where the total surface charge does not add up to zero. Moreover, we performed calculations in water ($\epsilon = 78.3553$) (Tables 2 and 5), dichloromethane ($\epsilon = 8.93$) (Table 3), and cyclohexane ($\epsilon = 2.0165$) (Table 4) to test the behavior of the method with strongly polar, polar, and nonpolar solvents. All the calculations have been performed using the PM6 semiempirical method to describe the solute.

For the first set of molecules, we report in each table the number of SCF iterations and the relative total elapsed time with respect to the standard algorithm which employs matrix

Table 2. Results Obtained in Water^a

	variational		iterative		inversion
	N_{it}	time	N_{it}	time	N_{it}
PCM					
diphthamide	39	72.73	16	457.58	16
cevane	31	91.18	15	580.88	15
dammarane	40	140.32	15	812.90	15
tuftsin	50	111.51	15	595.68	15
Ca(heme)	44	65.04	22	311.19	22
sGua	40	95.54	15	286.61	15
Ala9	41	88.35	14	530.52	14
taxol	47	95.77	19	750.81	19
Ala25	49	65.76	14	139.23	14
CPCM					
diphthamide	32	75.00	16	100.00	16
cevane	28	106.67	15	213.33	15
dammarane	27	102.86	15	242.86	15
tuftsin	31	98.28	15	244.83	15
Ca(heme)	36	51.61	22	120.97	22
sGua	33	70.15	15	65.67	15
Ala9	32	85.05	14	226.17	14
taxol	38	93.38	19	321.32	19
Ala25	34	49.38	14	58.31	14

^a Timings are reported as a percentage of the time required for the matrix inversion procedure.

Table 3. Results Obtained in Dichloromethane^a

	variational		iterative		inversion
	N_{it}	time	N_{it}	time	N_{it}
PCM					
diphthamide	34	100.00	16	281.25	16
cevane	30	117.31	15	740.38	15
dammarane	33	119.05	15	806.35	15
tuftsin	43	98.56	15	558.99	15
Ca(Heme)	43	62.94	22	302.10	22
sGua	35	85.83	15	220.00	15
Ala9	34	77.29	14	534.66	14
taxol	43	88.60	19	749.51	19
Ala25	37	55.21	14	126.27	14
CPCM					
diphthamide	33	86.36	16	90.91	16
cevane	28	106.67	15	220.00	15
dammarane	28	105.71	15	248.57	15
tuftsin	29	88.14	15	238.98	15
Ca(Heme)	36	53.33	21	118.33	21
sGua	32	64.29	15	62.86	15
Ala9	32	85.05	14	230.84	14
taxol	41	101.48	19	322.96	19
Ala25	35	52.74	14	43.66	14

^a Timings are reported as a percentage of the time required for the matrix inversion procedure.

Table 4. Results Obtained in Cyclohexane^a

system	variational		iterative		inversion
	N_{it}	time	N_{it}	time	N_{it}
IEFPCM					
diphthamide	34	118.52	15	344.44	15
cevane	28	109.43	15	750.94	15
dammarane	27	104.84	15	795.16	15
tuftsin	37	118.45	15	800.00	15
Ca(Heme)	44	92.93	21	423.23	21
sGua	35	129.13	15	250.49	15
Ala9	32	101.65	14	731.32	14
taxol	40	114.10	19	1089.16	19
Ala25	35	78.56	14	191.19	14
CPCM					
diphthamide	36	84.00	15	64.00	15
cevane	32	123.33	15	216.67	15
dammarane	30	111.43	15	257.14	15
tuftsin	36	111.86	15	235.59	15
Ca(Heme)	45	67.80	21	118.64	21
sGua	40	67.06	15	74.12	15
Ala9	33	87.85	14	230.84	14
taxol	42	103.70	19	320.00	19
Ala25	37	52.66	14	42.77	14

^a Timings are reported as a percentage of the time required for the matrix inversion procedure.

inversion to solve the PCM equations. We list both the variational solution algorithm of Section 2 and the standard algorithm using an iterative approach to solve the PCM equations. We do not report the absolute timings to avoid unfair comparisons of results obtained with different computational resources.

Before discussing the results, we offer a few preliminary considerations. The relative efficiency of the standard inversion, standard iterative, and variational algorithm depends on the interplay of a number of factors, and a complete exploration of this complex landscape is beyond the scope of this contribution. The standard inversion approach is characterized by an $\mathcal{O}(N_{ch}^3)$ computational cost and an $\mathcal{O}(N_{ch}^2)$ memory requirement. In the case of CPCM the only cubic step is the matrix inversion, while the formation of the IEF matrix involves two additional matrix multiplications. On the other hand, in both the standard iterative and the variational algorithm, the matrix can be held in memory so that each iteration would have quadratic cost (matrix–vector product). However, in the case of IEF this would not change the cost of forming the matrix which would still be a $\mathcal{O}(N_{ch}^3)$ step. The alternative would be to hold in memory the **S** and **D** matrices which are required to build the IEF matrix, therefore “trading” the cost of a single cubic step for the cost six matrix–vector products per iteration and also replacing the exact diagonals of the IEF matrix with the diagonals of the **S** matrix, which are less effective as a preconditioner in the iterative solutions. Finally, a definitive assessment of the relative performance of the three algorithm will require the exploration of the linear scaling regime where the contraction of the **S** and **D** matrices with the charge vector is carried out with $\mathcal{O}(N_{ch})$ cost by means of the fast multipole method.

The standard solution by matrix inversion is the fastest method for small systems, where the $\mathcal{O}(N_{ch}^3)$ step does not

Table 5. Results for larger molecules in water.^a

	PCM				CPCM			
	variational		iterative		variational		iterative	
	N_{it}	time	N_{it}	time	N_{it}	time	N_{it}	time
DNA-5	76	53 min	34	3 h 37 min	55	17 min	28	30 min
DNA-6	96	1 h 28 min	34	4 h 46 min	49	21 min	33	45 min
DNA-7	118	3 h 15 min	29	9 h 28 min	61	46 min	36	1 h 46 min
crambin	160	5 h 20 min	57	33 h 15 min	121	1 h 53 min	57	4 h 29 min

^a All the calculations have been performed on a dual Xeon E5530/48GB RAM computer.

emerge as a bottleneck in the overall SCF calculation. On the other hand, as soon as the dimension of the system increases, the variational procedure, whose timings are nevertheless comparable with the inversion even for smaller systems, becomes a competitive alternative. It is not unexpected that an iterative approach replacing the cubic step in the number of surface elements is the only viable option for the study of large systems. For the IEF model, the variational procedure is always faster than the standard iterative solution of the equations. The comparison is more significant for CPCM, where the variational procedure competes with a very efficient standard iterative solver. For molecular sizes where the iterative procedure begins to be competitive with the matrix inversion, the standard approach is sometimes faster than the variational one, especially in the case of nonpolar solvents. Nevertheless, the overall number of matrix–vector products required by the standard iterative approach increases fairly rapidly with the dimensions of the system, and this makes the variational approach the most promising one, even for CPCM when applied to larger solute. We notice that in the case of charged systems, the variational approach shows a very good convergence behavior, resulting in the majority of cases faster than the standard iterative procedure.

In Table 5 we report some results of calculations on increasingly large systems, from 318 up to 642 atoms. We notice that the DNA fragments are heavily charged, as we did not add counterions. Indeed, we want to make sure that the variational algorithm does not show stability issues when solving for polarization charges that do not sum up to zero but rather correspond to a huge solute–solvent interaction energy. All the calculations reported in Table 5 have been performed on a dual Xeon E5530 workstation (hyper-threading was disabled) equipped with 48 GB of RAM. In this case, we report the absolute elapsed time.

As expected, the variational approach outperforms the standard iterative one for both PCM and CPCM. In particular, the comparison with CPCM proves our assumption that the variational procedure is probably the best choice for large systems, under the assumption that the one-electron model Hamiltonian being used does not represent the computational bottleneck. We also point out that, because of the challenges involved in standard iterative solution for IEF, the variational approach introduced in this work is already a viable alternative to the matrix inversion, even for fairly small systems.

4. CONCLUSIONS

In this paper we have introduced a new strategy (VPCM) to solve the Roothan equations coupled with the PCM equations. We define a free energy functional of both the electronic density and the polarization charges whose variational minimization, using alternate steps in the two variable spaces, leads to the free

energy in solution including the mutual polarization of the solute and the dielectric continuum. This approach is computationally advantageous when dealing with large systems, where the solution of the PCM problem may become the dominant step in terms of computational cost, because an approximate one-electron Hamiltonian is used.

We think that this methodology could be efficiently applied to multiscale methods where a core system is described using an accurate level of theory, while the rest of the solute is treated with a computationally cheaper approach. In this context, the introduction of solvent effects by means of PCM can easily become the computational bottleneck because of the large size of the solute. Therefore, we expect the new approach introduced in this paper to be particularly relevant in the case of ONIOM^{33–35} calculations. Indeed, the availability of an efficient method to study large to very large systems coupled to a dielectric continuum would be an important tool for the description of biological systems in solution.

■ ASSOCIATED CONTENT

S Supporting Information. Geometries of the studied molecules. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: flipparini@sns.it.

■ ACKNOWLEDGMENT

We thank Prof. Vincenzo Barone and Dr. Eric Cancès for helpful discussions. F.L. and B.M. gratefully acknowledge financial support from Gaussian, Inc.

■ REFERENCES

- (1) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.
- (2) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209–220.
- (3) Stewart, J. J. P. *J. Mol. Model.* **2007**, *13*, 1173–1213.
- (4) Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, *94*, 2027–2094.
- (5) Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999–3093.
- (6) Amovilli, C.; Barone, V.; Cammi, R.; Cancès, E.; Cossi, M.; Mennucci, B.; Pomelli, C. S.; Tomasi, J. In *Recent Advances in the Description of Solvent Effects with the Polarizable Continuum Model*; Löwdin, P.-O., Ed.; Academic Press: San Diego, CA, 1998; Vol. 32, pp 227–261.
- (7) Mennucci, B. *J. Phys. Chem. Lett.* **2010**, *1*, 1666–1674 and references therein.
- (8) Cammi, R.; Tomasi, J. *J. Comput. Chem.* **1995**, *16*, 1449–1458.

- (9) Scalmani, G.; Barone, V.; Kudin, K.; Pomelli, C.; Scuseria, G.; Frisch, M. *Theor. Chem. Acc.* **2004**, *111*, 90–100.
- (10) Cossi, M.; Scalmani, G.; Rega, N.; Barone, V. *J. Chem. Phys.* **2002**, *117*, 43–54.
- (11) York, D.; Karplus, M. *J. Phys. Chem. A* **1999**, *103*, 11060–11079.
- (12) Scalmani, G.; Frisch, M. J. *J. Chem. Phys.* **2010**, *132*, 114110.
- (13) Lipparini, F.; Scalmani, G.; Mennucci, B.; Cancès, E.; Caricato, M.; Frisch, M. J. *J. Chem. Phys.* **2010**, *133*, 014106.
- (14) Lipparini, F.; Scalmani, G.; Mennucci, B. *Phys. Chem. Chem. Phys.* **2009**, *11*, 11617–11623.
- (15) Caricato, M.; Mennucci, B.; Scalmani, G.; Trucks, G. W.; Frisch, M. J. *J. Chem. Phys.* **2010**, *132*, 084102.
- (16) Caricato, M.; Scalmani, G.; Frisch, M. J. In *Continuum Solvation Models in Chemical Physics*; Mennucci, B., Cammi, R., Eds.; Wiley: New York, 2007.
- (17) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471–2474.
- (18) Schlegel, H.; Millam, J.; Iyengar, S.; Voth, G.; Daniels, A.; Scuseria, G.; Frisch, M. J. *J. Chem. Phys.* **2001**, *114*, 9758–9763.
- (19) Iyengar, S.; Schlegel, H.; Millam, J.; Voth, G.; Scuseria, G.; Frisch, M. J. *J. Chem. Phys.* **2001**, *115*, 10291–10302.
- (20) Pulay, P. *Chem. Phys. Lett.* **1980**, *73*, 393–398.
- (21) Pulay, P. *J. Comput. Chem.* **1982**, *3*, 556–560.
- (22) Kudin, K. N.; Cancès, E.; Scuseria, G. E. *J. Chem. Phys.* **2002**, *116*, 8255–8261.
- (23) Cancès, E.; Mennucci, B.; Tomasi, J. J. *J. Chem. Phys.* **1997**, *107*, 3032–3041.
- (24) Cancès, E.; Mennucci, B. *J. Math. Chem.* **1998**, *23*, 309–326.
- (25) Mennucci, B.; Cancès, E.; Tomasi, J. J. *J. Phys. Chem. B* **1997**, *101*, 10506–10517.
- (26) Hackbusch, W. *Integral Equations - Theory and Numerical Treatment*; Birkhäuser, Basel, 1995; pp 266–286.
- (27) Klamt, A.; Schuurmann, G. *J. Chem. Soc., Perkin Trans. 2 (1972-1999)* **1993**, 799–805.
- (28) Barone, V.; Cossi, M. *J. Phys. Chem. A* **1998**, *102*, 1995–2001.
- (29) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. Comput. Chem.* **2003**, *24*, 669–681.
- (30) Greengard, L.; Rokhlin, V. *J. Comput. Phys.* **1987**, *73*, 325–348.
- (31) 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. et al. *Gaussian Development Version*, revision H.08; Gaussian Inc.: Wallingford CT, 2009.
- (32) Harris, J. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1985**, *31*, 1770–1779.
- (33) Vreven, T.; Byun, K. S.; Komaromi, I.; Dapprich, S.; Montgomery, J.; John, A.; Mo-rokuma, K.; Frisch, M. J. *J. Chem. Theory Comput.* **2006**, *2*, 815–826.
- (34) Dapprich, S.; Komaromi, I.; Byun, K.; Morokuma, K.; Frisch, M. J. *Mol. Struct.* **1999**, *461*, 1–21.
- (35) Vreven, T.; Mennucci, B.; da Silva, C.; Morokuma, K.; Tomasi, J. *J. Chem. Phys.* **2001**, *115*, 62–72.

Electron Pair Localization Function (EPLF) for Density Functional Theory and *ab Initio* Wave Function-Based Methods: A New Tool for Chemical Interpretation

Anthony Scemama,^{*,†} Michel Caffarel,[†] Robin Chaudret,[‡] and Jean-Philip Piquemal[‡]

[†]Laboratoire de Chimie et Physique Quantiques, CNRS-IRSAMC, Université de Toulouse, France

[‡]UPMC Univ Paris 06, UMR 7616, Laboratoire de Chimie Théorique, case courrier 137, 4 place Jussieu, F-75005, Paris, France, and CNRS, UMR 7616, Laboratoire de Chimie Théorique, case courrier 137, 4 place Jussieu, F-75005, Paris, France

S Supporting Information

ABSTRACT: We present a modified definition of the Electron Pair Localization Function (EPLF), initially defined within the framework of quantum Monte Carlo approaches [Scemama, A.; Caffarel, M.; Chaquin, P. *J. Chem. Phys.* **2004**, *121*, 1725] to be used in Density Functional Theories (DFT) and *ab initio* wave-function-based methods. This modified version of the EPLF—while keeping the same physical and chemical contents—is built to be analytically computable with standard wave functions or Kohn–Sham representations. It is illustrated that the EPLF defines a simple and powerful tool for chemical interpretation via selected applications including atomic and molecular closed-shell systems, σ and π bonds, radical and singlet open-shell systems, and molecules having a strong multiconfigurational character. Some applications of the EPLF are presented at various levels of theory and compared to Becke and Edgecombe’s Electron Localization Function (ELF). Our open-source parallel software implementation of the EPLF opens the possibility of its use by a large community of chemists interested in the chemical interpretation of complex electronic structures.

1. INTRODUCTION

Nowadays, when dealing with theoretical chemical interpretation, quantum chemists rely on two main strategies. The first consists of the traditional direct interpretation of the wave function through its projection onto molecular orbitals (MO) or valence bond (VB) structures (the so-called Hilbert space partitioning). The second uses a geometrical direct-space description in order to partition the electronic density into domains within the ordinary 3D space. The design of such interpretative techniques, initiated by Daudel et al.,¹ was popularized by Bader, who introduced the Quantum Theory of Atoms in Molecules (QTAIM).² Along with QTAIM, Bader introduced the concept of topological analysis, offering an atom-based partition of the molecular space grounded on the gradient dynamical system theory and using a local function, here the Laplacian of the electron density. Through the years, much effort has been devoted to the design of alternative local functions. For example, Becke and Edgecombe introduced the Electron Localization Function (ELF),³ offering access to chemically intuitive domains beyond atomic centers encompassing bonds, lone pairs, etc. Ever since, its usefulness has been demonstrated by Silvi and Savin,⁴ who extensively developed its topological analysis, although no partition of space is unique.⁵

The problem of getting an accurate description of chemical bonding gets more and more difficult as the complexity of the wave function goes beyond the single determinant approximation.⁶ Therefore, an additional natural orbital approximation was added to the ELF formalism⁷ to extend it to the correlated level, but its general applicability to *any* quantum chemical method is still subject to intense development. In that context, other methods were introduced such as the electron

localizability indicator (ELI, see ref 8 and references therein), the analysis of electronic probability distributions,^{9,10} and the Electron Pair Localization Function (EPLF).¹¹

In this work, we shall focus on this latter function, EPLF, whose main feature is giving direct access to the local (spatial) electronic correlations between spin-like and spin-unlike electronic pairs. EPLF was first introduced within the framework of quantum Monte Carlo (QMC) approaches where introducing simple and direct estimators of such local electronic correlations is particularly easy. In practice, it has been proposed to build an indicator—the electron pair localization function—based on a suitable combination of the average distances between an electron of a given spin located at point \mathbf{r} and the closest spin-like and spin-unlike electrons. EPLF has been shown to be particularly interesting to get new insights into the nature of the pairing and localization of electrons and, particularly, to understand more deeply the role of the dynamical and non-dynamical near-degeneracy correlation effects.^{12,13} From a fundamental point of view, such a result is not surprising, since the EPLF is actually related to the conditional probabilities of finding an electron at point \mathbf{r}_2 with spin σ or $\bar{\sigma}$, knowing that an electron of a given spin σ is located at some point \mathbf{r}_1 . Indeed, having such quantities at our disposal is known to be sufficient to define an exact electronic structure theory (e.g., the *exact* exchange-correlation energy of DFT can be in principle derived from such conditional probabilities; see, e.g., ref 14). The advantage of having defined the EPLF within a QMC computational scheme is that such a function can be easily calculated at various levels of approximation. Indeed, by generating QMC probability densities

Received: October 15, 2010

Published: February 02, 2011

associated with various trial wave functions, the average distances between electrons and, thus, the EPLF function can be evaluated at the Hartree–Fock, DFT, CASSCF, CI, VB, etc. levels of approximation. It is also possible to evaluate the EPLF at the fixed-node diffusion Monte Carlo level, a particularly accurate QMC approximation recovering the major part of static and dynamical correlation effects, even if in some cases the quality of the fixed-node error is not so easy to assess, see, e.g., ref 15. Besides these advantages, the main drawback of calculating the EPLF with QMC is that simulations need to be rather intensive to decrease sufficiently the statistical errors of the EPLF values at each point r of the grid employed. Indeed, a minimal resolution is needed to distinguish the subtle changes in local properties.

In this work, we propose a modified form for the EPLF allowing its *exact* computation (no statistical error) for the standard wave functions of computational chemistry written as determinantal expansions built from molecular orbitals expressed in some Gaussian basis set. The approach can also be naturally applied to DFT calculations based on a Kohn–Sham density expressed in a determinantal form. As we shall see, the proposed modification of the EPLF does not alter its chemical content. Using this modified expression, the EPLF is much more rapid to compute since its calculation requires only the evaluation of mono-electronic integrals (see below). In particular, it avoids the use of Monte Carlo sampling, which can be rather CPU-intensive for large systems, opening the possibility to perform full topological analyses in the near future. Accordingly, once introduced into standard computational chemistry packages, we believe that the EPLF will become a very useful and powerful tool for chemical interpretation accessible to a wide community of chemists.

2. EPLF: THE ORIGINAL DEFINITION

In the original definition of the EPLF, ref 11, the motivation was to define a function of \mathbb{R}^3 measuring locally the electron pairing in a molecular system. To do that, the following definition of electron pairing was first introduced: An electron i located at \mathbf{r}_i is said to be paired to an electron j located at \mathbf{r}_j if electron j is the closest electron to i . Having defined such a pairing, it has been proposed to define the amount of electron pairing at point r in terms of a quantity inversely proportional to

$$d(\mathbf{r}) = \left\langle \Psi \left| \sum_{i=1, N} \delta(\mathbf{r} - \mathbf{r}_i) \min_{j \neq i} r_{ij} \right| \Psi \right\rangle \quad (1)$$

where $d(\mathbf{r})$ can be interpreted as the average of the shortest electron–electron distance at \mathbf{r} , $\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N)$ being the N -electron wave function, and $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$.

Two different types of electron pairs are to be defined: pairs of electrons having the same spin (σ) and pairs of electrons with opposite spins ($\sigma, \bar{\sigma}$). Hence, two quantities need to be introduced:

$$d_{\sigma\sigma}(\mathbf{r}) = \left\langle \Psi \left| \sum_{i=1, N} \delta(\mathbf{r} - \mathbf{r}_i) \min_{j \neq i; \sigma_i = \sigma_j} r_{ij} \right| \Psi \right\rangle \quad (2)$$

$$d_{\sigma\bar{\sigma}}(\mathbf{r}) = \left\langle \Psi \left| \sum_{i=1, N} \delta(\mathbf{r} - \mathbf{r}_i) \min_{j; \sigma_i \neq \sigma_j} r_{ij} \right| \Psi \right\rangle \quad (3)$$

The electron pair localization function is bound in the $[-1, 1]$ interval and is defined as

$$\text{EPLF}(\mathbf{r}) = \frac{d_{\sigma\sigma}(\mathbf{r}) - d_{\sigma\bar{\sigma}}(\mathbf{r})}{d_{\sigma\sigma}(\mathbf{r}) + d_{\sigma\bar{\sigma}}(\mathbf{r})} \quad (4)$$

When the pairing of spin-unlike electrons is predominant, $d_{\sigma\sigma}(\mathbf{r}) > d_{\sigma\bar{\sigma}}(\mathbf{r})$ and $\text{EPLF}(\mathbf{r}) > 0$. When the pairing of spin-like electrons is predominant, $d_{\sigma\sigma}(\mathbf{r}) < d_{\sigma\bar{\sigma}}(\mathbf{r})$ and $\text{EPLF}(\mathbf{r}) < 0$. When the electron pairing of spin-like and spin-unlike electrons is equivalent, $\text{EPLF}(\mathbf{r}) \sim 0$.

This localization function does not depend on the type of wave function and can therefore measure electron pairing using any kind of representation: Hartree–Fock (HF), Kohn–Sham (KS), Configuration Interaction (CI), and Multi-Configurational Self-Consistent-Field (MCSCF) as well as Slater–Jastrow, Diffusion Monte Carlo (DMC), Hylleraas wave functions, etc. Due to the presence of the *min* function in the definitions of $d_{\sigma\sigma}(\mathbf{r})$ and $d_{\sigma\bar{\sigma}}(\mathbf{r})$, these quantities cannot be evaluated in an analytical way, and quantum Monte Carlo (QMC) approaches appear to be the most efficient way of computing the three-dimensional EPLF grids via a statistical sampling of $\sim \Psi^2(\mathbf{r}_1, \dots, \mathbf{r}_N)$ in the case of Variational Monte Carlo (VMC)-type calculations or $\sim \Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) \Phi_0(\mathbf{r}_1, \dots, \mathbf{r}_N)$ (Φ_0 fixed-node ground-state wave function) in the case of the more accurate Fixed-Node Diffusion Monte Carlo (FN-DMC)-type calculations^{11–13, 16} (for a detailed presentation of these various versions of QMC approaches, see, e.g., ref 17).

3. EPLF: A MODIFIED DEFINITION SUITABLE FOR DFT AND WAVE FUNCTION-BASED METHODS

Following preliminary developments,¹⁸ we propose here to introduce a modified definition of the EPLF which—in contrast with the original definition—can now be analytically computable for standard wave functions of quantum chemistry, thus avoiding the need for statistical sampling. To do that, we propose to express the *min* function appearing in the average distances in terms of Gaussian functions. More precisely, we introduce the following exact representation:

$$\min_{j \neq i} r_{ij} = \lim_{\gamma \rightarrow +\infty} \sqrt{-\frac{1}{\gamma} \ln f(\gamma; r_{ij})} \quad (5)$$

with

$$f(\gamma; r_{ij}) = \sum_{j \neq i} e^{-\gamma r_{ij}^2} \quad (6)$$

Now, our basic approximation consists in replacing, for γ large, the integrals

$$\left\langle \Psi \left| \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i) \left(\sqrt{-\frac{1}{\gamma} \ln f(\gamma; r_{ij})} \right) \right| \Psi \right\rangle \quad (7)$$

appearing in eq 1 with

$$\sqrt{-\frac{1}{\gamma} \ln \left\langle \Psi \left| \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i) f(\gamma; r_{ij}) \right| \Psi \right\rangle} \quad (8)$$

The expectation values of the minimum distances are now given by

$$d_{\sigma\sigma}(\mathbf{r}) \underset{\gamma \text{ large}}{\sim} \sqrt{-\frac{1}{\gamma} \ln \bar{f}_{\sigma\sigma}(\gamma; \mathbf{r})} \quad (9)$$

$$d_{\sigma\bar{\sigma}}(\mathbf{r}) \underset{\gamma \text{ large}}{\sim} \sqrt{-\frac{1}{\gamma} \ln \bar{f}_{\sigma\bar{\sigma}}(\gamma; \mathbf{r})} \quad (10)$$

with the two-electron integrals:

$$\bar{f}_{\sigma\sigma}(\gamma; \mathbf{r}) = \left\langle \Psi \left| \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i) \sum_{j \neq i; \sigma_i = \sigma_j}^N e^{-\gamma|\mathbf{r}_i - \mathbf{r}_j|^2} \right| \Psi \right\rangle \quad (11)$$

$$\bar{f}_{\sigma\bar{\sigma}}(\gamma; \mathbf{r}) = \left\langle \Psi \left| \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i) \sum_{j; \sigma_i \neq \sigma_j}^N e^{-\gamma|\mathbf{r}_i - \mathbf{r}_j|^2} \right| \Psi \right\rangle \quad (12)$$

When the wave function Ψ has a standard form (sum of determinants built from molecular integrals ϕ 's), such integrals can be easily obtained in terms of the following elementary contributions:

$$\phi_i(\mathbf{r}) \phi_k(\mathbf{r}) \int d\mathbf{r}' \phi_j(\mathbf{r}') \phi_l(\mathbf{r}') e^{-\gamma|\mathbf{r} - \mathbf{r}'|^2} \quad (13)$$

which in turn can be evaluated as generalized overlap integrals.

Let us now discuss our basic approximation consisting in going from eq 7 to eq 8. This approximation can be written in a more compact way as

$$\frac{\langle \sqrt{-\ln f} \rangle}{\sqrt{-\ln \langle f \rangle}^{\gamma_{\text{large}}}} \sim 1 \quad (14)$$

where the symbol $\langle Q \rangle$ denotes the integration of $Q\Psi^2$ over all particle coordinates except the i th one. For a given electronic configuration $(\mathbf{r}_1, \dots, \mathbf{r}_N)$ and γ large enough, f is dominated by a single exponential, namely, $e^{-\gamma|\mathbf{r}_i - \mathbf{r}_{j_{\min}}|^2}$, where $|\mathbf{r}_i - \mathbf{r}_{j_{\min}}|$ is the distance between the reference electron i located at \mathbf{r}_i and the closest electron labeled j_{\min} . The validity of our basic approximation is directly related to the amount of fluctuations of the quantity f when various electronic configurations are considered. Note that for a given electron i , the distance $|\mathbf{r}_i - \mathbf{r}_j|$ can vary a lot, but it is much less the case for $|\mathbf{r}_i - \mathbf{r}_{j_{\min}}|$, where the electron number j_{\min} can be different from one configuration to another. When these fluctuations are small, the ratio in eq 14 is close to 1 and the approximation is of good quality. To see what happens for larger fluctuations, let us write

$$f = f_{\min} + \delta f \quad (15)$$

A simple calculation leads to

$$\frac{\langle \sqrt{-\ln f} \rangle}{\sqrt{-\ln \langle f \rangle}} = 1 + O[(\delta f)^2] \quad (16)$$

showing that at first order in the fluctuations the ratio is still equal to 1, illustrating the validity of our approximation.

A last point to discuss is the value of γ to be chosen in practice. Because of our approximation, the limit $\gamma \rightarrow +\infty$ cannot be taken since the ratio in eq 14 goes to zero.¹⁹ Therefore, the value of γ has to be large enough to discriminate between the closest electron located at $r_{j_{\min}}$ from the other ones located at larger distances of electron i , while staying in the regime where the ratio in eq 14 stays close to one. We have found that a value of γ depending on \mathbf{r} and chosen on physical grounds allows systematic recovery of the essential features of the EPLF images calculated with QMC, that is to say, with the exact expression of the *min* function. To be effective, the discrimination of the closest electron with the other ones must be properly implemented. To do that, the value of γ is adapted to keep the leading

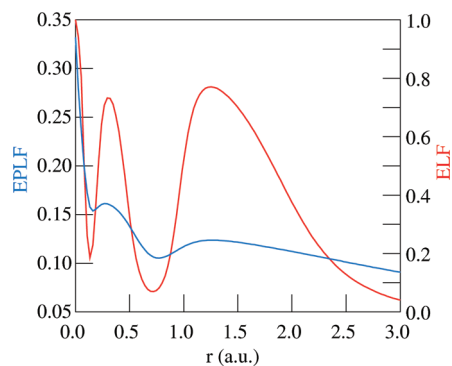


Figure 1. ELF and EPLF radial values for the argon atom as a function of the distance to the nucleus.

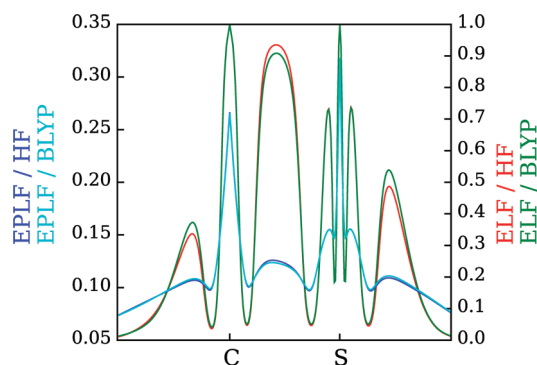


Figure 2. ELF and EPLF values in the CH_3S^- anion along the C–S axis computed using a Hartree–Fock and a BLYP determinant.

exponential $e^{-\gamma|\mathbf{r}_i - \mathbf{r}_{j_{\min}}|^2}$ significantly larger than the subleading exponential $e^{-\gamma|\mathbf{r}_i - \mathbf{r}_{j_{\text{next-min}}}|^2}$ associated with the second closest electron $j_{\text{next-min}}$. First, we define a sphere $\Omega(\mathbf{r}_i)$ centered on \mathbf{r}_i with a radius $d_{\Omega}(\mathbf{r}_i)$. Then, locally, we represent our system made of the electron located at \mathbf{r} , and its two closest neighbors by a model system of three independent particles. If one calculates the probability of finding all three particles inside the sphere, one finds

$$P_{\Omega}(\mathbf{r}_i) = \left(\frac{1}{3} \int_{\Omega(\mathbf{r}_i)} d\mathbf{r} \rho(\mathbf{r}) \right)^3 \quad (17)$$

If the density $\rho(\mathbf{r})$ is supposed as constant and equal to $\rho(\mathbf{r}_i)$, the radius $d_{\Omega}(\mathbf{r}_i)$ of the sphere can be set such that P_{Ω} is equal to a fixed value:

$$d_{\Omega}(\mathbf{r}_i) = \left(\frac{4\pi}{9} P_{\Omega}^{-1/3} \rho(\mathbf{r}_i) \right)^{-1/3} \quad (18)$$

Then, $\gamma(\mathbf{r}_i)$ is chosen in order to set a constant ratio κ between the width of $e^{-\gamma r_i^2}$ and the radius of the sphere

$$\kappa = \sqrt{2\gamma(\mathbf{r}_i) d_{\Omega}(\mathbf{r}_i)} \quad (19)$$

We obtain an expression of $\gamma(\mathbf{r}_i)$ which depends on the electron density:

$$\gamma(\mathbf{r}_i) = \frac{\kappa^2}{2} \left(\frac{4\pi}{9} P_{\Omega}^{-1/3} \rho(\mathbf{r}_i) \right)^{2/3} \quad (20)$$

In our simulations, we have found that the EPLF images obtained with QMC are properly recovered using $P_{\Omega} = 0.001$ and $\kappa = 50$.

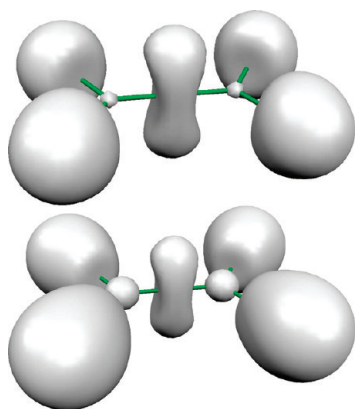


Figure 3. ELF = 0.75 (top) and EPLF = 0.12 (bottom) isosurfaces of the ethylene molecule.

4. SOME APPLICATIONS

As discussed previously, the modified form of the EPLF presented here is aimed at providing the same chemical information as the original QMC-based EPLF scheme, but without the statistical noise inherited from the QMC approach. Therefore, the interested reader can refer to existing recent publications that deal with the QMC-EPLF analysis of covalent, ionic, and multicenter bonds.^{11–13,16,20} We focus in this section on some illustrative applications highlighting the specific capabilities of the EPLF as compared to Becke and Edgecombe's ELF.

4.1. Closed-Shell Single-Determinant Systems. A first natural example to look at is the case of a closed-shell atom described at the Hartree–Fock (HF) level. Using Dunning's cc-pVDZ atomic basis set,²¹ the radial values of the EPLF and ELF for the argon atom are displayed in Figure 1. It is noted that both functions display three maximum values corresponding to the $n = 1$, $n = 2$, and $n = 3$ values of the principal quantum number. Furthermore, these maxima are essentially located at the same place. The *gross features* of the atomic shell structure are thus described in a similar way by both approaches. However, there is also a striking difference: The magnitudes of the two secondary maxima corresponding to the two most external shells are essentially identical in the ELF case but very different for the EPLF, where the outermost one is much smaller. Note that having such a difference is not surprising since EPLF is, in contrast with ELF, directly connected to electron pairing. The pairing of antiparallel electrons is likely to be the strongest in the first shell, weaker in the second shell, and the weakest in the most diffuse third shell.

EPLF and ELF were computed for the CH_3S^- methanethiolate anion, using a Hartree–Fock determinant and a Kohn–Sham determinant. The 6-31⁺⁺G** atomic basis set^{22,23} was used for both determinants, and the BLYP functional^{24,25} was used for the DFT calculation. Figure 2 compares the one-dimensional plots of the EPLF and ELF along the C–S axis of the tetrahedral CH_3S^- . As for the argon atom, the topologies of the EPLF and ELF functions are comparable for both the Hartree–Fock and the Kohn–Sham determinants. Going from the Hartree–Fock to the BLYP level, the values of the ELF are essentially the same in the core domains, become slightly smaller in the C–S bonding region, and become slightly larger on the rest of the C–S axis. As the EPLF exhibits the same trend, we conclude that for closed-shell single determinants the EPLF and ELF give qualitatively similar results.

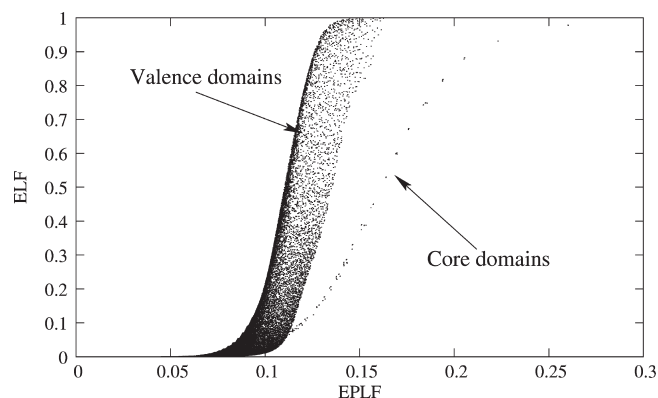


Figure 4. Correlation between the ELF and the EPLF in the ethylene molecule.

The ELF and EPLF were computed for the ethylene molecule using a HF/cc-pVDZ wave function. The isosurfaces ELF = 0.75 and EPLF = 0.12 are represented in Figure 3. These images are qualitatively similar, even if the core domains seem to be larger using the EPLF. This is due to the fact that the EPLF values are higher in the first atomic shells (as in the argon example), while the ELF has more comparable values among the shells.

To have a more quantitative visualization of the similarities and differences between the ELF and the EPLF, a correlation plot relating the values of both functions is presented in Figure 4. Three different regimes can be observed. First, a regime corresponding to the core domain where the EPLF takes its larger values. In this region, an almost perfect one-to-one correspondence is observed, thus illustrating the similarity between both localization functions. In contrast, in the valence region where the (EPLF, ELF) points are scattered, it seems to be no longer true. In fact, this is not really the case since the majority of points are found to be almost aligned along the left side of the envelope of points. To illustrate this, the median line (same number of points on each side) is represented. Finally, a last regime corresponding to the region where the ELF and EPLF values are small (say, ELF smaller than 0.05) can be defined. In such a regime, the two localization functions turn out to be fully decorrelated. However, the underlying configurations correspond to regions in space where the electronic densities are (very) small, and this case is not of great chemical interest. As a conclusion, in all chemically interesting regimes, the correlation between ELF and EPLF is high. We have found that such a conclusion is valid not only for this case but also for all molecules described by a closed-shell single determinant wave function. In this case, the qualitative information that can be obtained from an ELF and an EPLF calculation is essentially the same. This can be understood by noting that for a closed-shell monoconfigurational wave function the α electrons are independent from the β electrons, so localizing electrons is essentially equivalent to localizing antiparallel electron pairs.

4.2. Open-Shell Hartree–Fock. A wave function for the HC_2^\bullet radical was obtained at the restricted open-shell Hartree–Fock level (ROHF), using the cc-pVDZ atomic basis set. Both ELF and EPLF were computed, and the results are displayed in Figure 5. This example points out the main difference between ELF and EPLF: the localization region of the unpaired electron exhibits a maximum for ELF (high electron localization) and a minimum for EPLF (low electron pairing). EPLF can identify

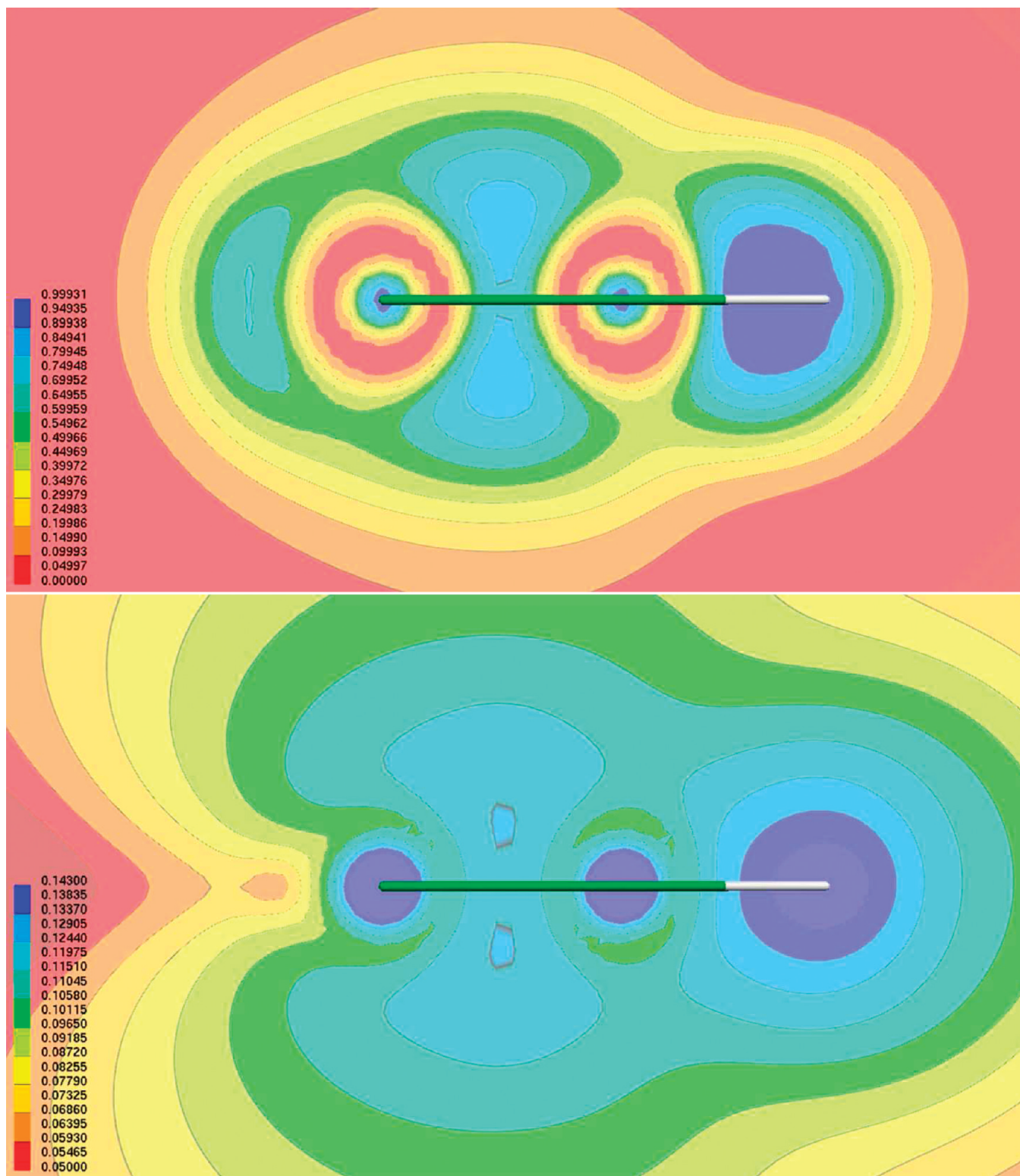


Figure 5. ELF (top) and EPLF (bottom) contour plots of the HC₂• radical in the molecular plane. Red values are the lowest and blue values are the highest.

clearly domains of electron pairing (lone pairs, core domains, and bonds), and it can additionally characterize localized unpaired electrons similarly to spin density.

4.3. Multiconfigurational Wave Function. A wave function for the singlet state of the ozone molecule was first calculated at

the HF/cc-pVDZ level. ELF and the EPLF were both calculated and give similar qualitative results (Figure 6).

Then, a complete active space wave function with eight electrons in eight orbitals (CAS(8,8)) was prepared, and EPLF was calculated (Figure 7). The EPLF obtained from the CAS wave function is

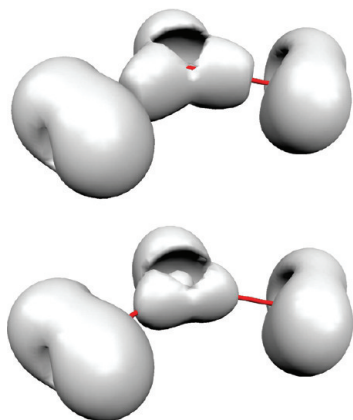


Figure 6. ELF = 0.61 (top) and ELF = 0.123 (bottom) isosurfaces of the singlet state of the ozone molecule (Hartree–Fock).

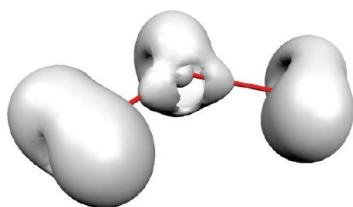


Figure 7. EPLF = 0.123 isosurface the singlet state of the ozone molecule (CAS-SCF).

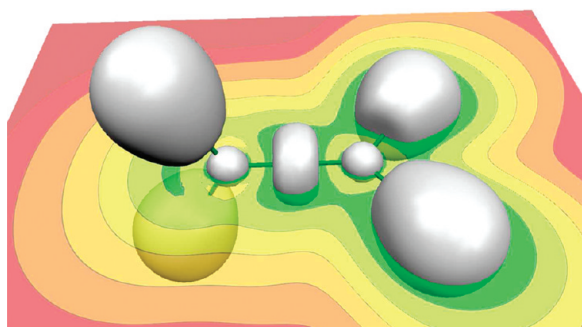


Figure 8. EPLF contour plot and isosurface of the singlet state of the twisted $\text{H}_2\text{C}^*-\text{CH}_2$ biradical. Red values of the EPLF are the lowest, and green values are the highest.

significantly different from the EPLF obtained from the HF wave function. In the HF framework, the O–O bonding domains are more connected to each other than to the lone pair domains of the central oxygen atom. With the CAS wave function, each O–O bond domain is more connected to the lone pair domains than to each other. This example shows that EPLF is an alternative to ELF in closed shell systems where a multiconfigurational method is required, as EPLF is well-defined for such cases.

4.4. Open-Shell Singlet. When the ethylene molecule is twisted with an angle of 90° along the C–C axis, the π bond breaks. Each one of the π electrons localizes on a carbon atom, giving rise to an open-shell singlet (see Figure 8), degenerate with the triplet state. In order to preserve the spin symmetry, a CAS(2,2) wave function was computed to describe the singlet state. With such a wave function, the spin density is not able to localize the unpaired electrons since the α one-electron density is

equal to the β one-electron density in every point of space. The EPLF reveals the presence of these unpaired electrons by local minimum values of the function close to the carbon atoms, in the plane perpendicular to the C–H bonds.

5. SOFTWARE

To realize the EPLF and ELF calculations presented in this paper, a code was written using the IRPF90 Fortran generator.²⁶ This code is interfaced with the Gaussian 03,²⁷ GAMESS,²⁸ and Molpro²⁹ programs. As the calculation of EPLF is more expensive than the calculation of ELF, the program has been efficiently parallelized (for both EPLF and ELF calculations) using the message passing interface (MPI) library³⁰ and exhibits a linear speedup property with the number of cores. The EPLF code is licensed under the GNU General Public License, and the source files can be downloaded from the Web at <http://eplf.sourceforge.net>.

6. CONCLUSION

In this work, we have introduced a modified version of the EPLF analytically computable for standard wave functions and DFT representations. When compared to the original EPLF defined in a QMC framework, essentially the same images are recovered. A systematic comparison of our analytical EPLF with the Electronic Localization Function (ELF) of Becke and Edgecombe has been made. For closed-shell systems, the EPLF results are shown to closely match the ELF ones. However, for other situations, the two localization functions may differ significantly (radicals, systems with strong static correlations, etc). The major advantage of the reformulated EPLF is that it can be easily computed for any kind of electronic structure method defined from single or multideterminantal wave functions. Further development will focus on the topological analysis of the EPLF, which will provide the possibility of computing various properties integrated from a partition of the three-dimensional space. As our software is available for free, it should open the possibility for any chemist to use EPLF for the understanding of complex electronic structures.

■ ASSOCIATED CONTENT

S Supporting Information. EPLF vs ELF for the xenon atom. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: scemama@irsamc.ups-tlse.fr.

■ ACKNOWLEDGMENT

Support from the French “Centre National de la Recherche Scientifique (CNRS)”, Université de Toulouse, and Université Paris 6 is gratefully acknowledged. We would like to thank IDRIS (CNRS, Orsay), CCRT (CEA/DAM, Ile-de-France), and CALMIP (Université de Toulouse) for providing us with computational resources.

■ REFERENCES

- (1) Daudel, R.; Odiot, S.; Brion, H. J. *Physique Rad.* **1954**, *15*, 804–809.
- (2) Bader, R. F. W. *Atoms in molecules: A quantum theory*; Clarendon Press: Oxford, U. K., 1990; p 438.

- (3) Becke, A. D.; Edgecombe, K. E. *J. Chem. Phys.* **1990**, *92*, 5397–5403.
- (4) Silvi, B.; Savin, A. *Nature* **1994**, *371*, 683–686.
- (5) Savin, A. *J. Chem. Sci.* **2005**, *117*, 473–475.
- (6) Silvi, B. *J. Phys. Chem. A* **2003**, *107*, 3081–3085.
- (7) Feixas, F.; Matito, E.; Duran, M.; Solà, M.; Silvi, B. *J. Chem. Theory Comput.* **2010**, *6*, 2736–2742.
- (8) Bezugly, V.; Wielgus, P.; Kohout, M.; Wagner, F. R. *J. Comput. Chem.* **2010**, *31*, 1504–1519.
- (9) Cancès, E.; Keriven, R.; Lodier, F.; Savin, A. *Theor. Chim. Acta* **2004**, *111*, 373–380.
- (10) Scemama, A.; Caffarel, M.; Savin, A. *J. Comput. Chem.* **2007**, *28* (1), 442–454.
- (11) Scemama, A.; Chaquin, P.; Caffarel, M. *J. Chem. Phys.* **2004**, *121*, 1725–1735.
- (12) Scemama, A.; Caffarel, M.; Ramírez-Solís, A. *J. Phys. Chem. A* **2009**, *113*, 9014–9021.
- (13) Caffarel, M.; Scemama, A.; Ramírez-Solís, A. *Theor. Chim. Acta* **2010**, *126*, 275–287.
- (14) Parr, R.; Yang, W. *Density-Functional Theory of Atoms and Molecules*; Oxford University Press: Oxford, U. K., 1989; p 344.
- (15) Caffarel, M.; Hernández-Lamonedá, R.; Scemama, A.; Ramírez-Solís, A. *Phys. Rev. Lett.* **2007**, *99*, 153001.
- (16) Amador-Bedolla, C.; Salomón-Ferrer, R.; Lester, W. A., Jr.; Vázquez-Martínez, J. A.; Aspuru-Guzik, A. *J. Chem. Phys.* **2007**, *126*, 204308.
- (17) Hammond, B. L.; Lester, W. A., Jr.; Reynolds, P. J. *Monte Carlo Methods in Ab Initio Quantum Chemistry. Lecture and Course Notes in Chemistry*; World Scientific: Singapore, 1994; Vol. 1.
- (18) Alary, F.; Heully, J.-L.; Scemama, A.; Garreau-de Bonneval, B.; Chane-Ching, K.; Caffarel, M. *Theor. Chim. Acta* **2010**, *126*, 243–255.
- (19) For a given electronic configuration $(\mathbf{r}_1, \dots, \mathbf{r}_N)$, f goes as $\exp(-\gamma|\mathbf{r} - \mathbf{r}_{\min}|^2)$ when γ becomes large. Accordingly, in this limit, $\langle(-\ln f)^{1/2}\rangle$ goes to a finite positive constant given by $\langle|\mathbf{r} - \mathbf{r}_{\min}|^2\rangle$. On the other hand, $\langle f \rangle$ can be written as the integral of $\exp(-\gamma|\mathbf{r} - \mathbf{r}_{\min}|^2)$ over all possible configurations, which can be rewritten as $1/(\gamma)^{1/2}(\gamma)^{1/2} \exp(-\gamma|\mathbf{r} - \mathbf{r}_{\min}|^2)$, that is, at large γ , the product of $1/(\gamma)^{1/2}$ times a Dirac distribution. Finally, in the large- γ limit, the ratio $(\langle(-\ln f)^{1/2}\rangle)/(\langle f \rangle^{1/2})$ goes to zero as $\sim 1/(\gamma)^{1/2}$.
- (20) De La Lande, A.; Salahub, D. R.; Maddaluno, J.; Scemama, A.; Pilme, J.; Parisel, O.; Gerard, H.; Caffarel, M.; Piquemal, J.-P. *J. Comput. Chem.* **2010**, [Online] DOI: 10.1002/jcc.21698.
- (21) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- (22) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213–222.
- (23) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. *J. Chem. Phys.* **1982**, *77*, 3654–3665.
- (24) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.
- (25) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.
- (26) Scemama, A. *ArXiv e-prints* 2009, [cs.SE], 0909.5012.
- (27) Frisch, M. J. et al. *Gaussian 03*; Gaussian, Inc.: Wallingford, CT, 2004.
- (28) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A., Jr. *J. Comput. Chem.* **1993**, *14* 11, 1347–1363.
- (29) Werner, H. J. et al. MOLPRO, version 2009.1. <http://www.molpro.net> (accessed Dec 7, 2009).
- (30) Gropp, W.; Lusk, E.; Doss, N.; Skjellum, A. *Parallel Comput.* **1996**, *22*, 789–828.

NCI PLOT: A Program for Plotting Noncovalent Interaction Regions

Julia Contreras-García,[†] Erin R. Johnson,[‡] Shahar Keinan,[†] Robin Chaudret,^{§,||} Jean-Philip Piquemal,^{§,||} David N. Beratan,[†] and Weitao Yang^{*,†}[†]Department of Chemistry, Duke University, Durham, North Carolina 27708[‡]School of Natural Sciences, University of California, Merced, 5200 North Lake Road, Merced, California 95343[§]UPMC Univ Paris 06, UMR 7616, Laboratoire de Chimie Théorique, case courrier 137, 4 place Jussieu, F-75005, Paris, France^{||}CNRS, UMR 7616, Laboratoire de Chimie Théorique, case courrier 137, 4 place Jussieu, F-75005, Paris, France Supporting Information

ABSTRACT: Noncovalent interactions hold the key to understanding many chemical, biological, and technological problems. Describing these noncovalent interactions accurately, including their positions in real space, constitutes a first step in the process of decoupling the complex balance of forces that define noncovalent interactions. Because of the size of macromolecules, the most common approach has been to assign van der Waals interactions (vdW), steric clashes (SC), and hydrogen bonds (HBs) based on pairwise distances between atoms according to their vdW radii. We recently developed an alternative perspective, derived from the electronic density: the non-covalent interactions (NCI) index [*J. Am. Chem. Soc.* **2010**, *132*, 6498]. This index has the dual advantages of being generally transferable to diverse chemical applications and being very fast to compute, since it can be calculated from promolecular densities. Thus, NCI analysis is applicable to large systems, including proteins and DNA, where analysis of noncovalent interactions is of great potential value. Here, we describe the NCI computational algorithms and their implementation for the analysis and visualization of weak interactions, using both self-consistent fully quantum-mechanical as well as promolecular densities. A wide range of options for tuning the range of interactions to be plotted is also presented. To demonstrate the capabilities of our approach, several examples are given from organic, inorganic, solid state, and macromolecular chemistry, including cases where NCI analysis gives insight into unconventional chemical bonding. The NCI code and its manual are available for download at <http://www.chem.duke.edu/~yang/software.htm>.

I. INTRODUCTION

Noncovalent interactions are of critical importance in many chemical, biological, and technological systems. Protein–ligand interactions,¹ the self-assembly of nanomaterials,² the folding of proteins,³ and the packing of molecular crystals⁴ are controlled by a delicate balance of numerous, weak noncovalent interactions. The large sizes of the materials and structures of interest makes understanding noncovalent forces particularly difficult.

Accurately describing noncovalent interactions, including their spatial characteristics, constitutes the first step in the process of decomposing the complex balance of chemical forces. However, even this spatial characterization of interactions has generated controversy.^{5–8} Because of the size of many systems of interest, the most common approach to examining noncovalent interactions has been to assign van der Waals interactions (vdW), steric clashes (SC), and hydrogen bonds (HBs) in terms of pairwise distances between atoms based on their vdW radii.⁹ This procedure, although lacking generality and giving rise to systematic errors (e.g., the length of HBs is generally overestimated), enables the rapid enumeration of noncovalent interactions.

More elaborate algorithms for mapping and analyzing noncovalent interactions can be derived from the electronic and kinetic-energy densities. Indeed, the theory of atoms in molecules¹⁰ has been used to understand and to quantify weak interactions on the basis of the electron density.^{10–12} This

approach relies on the fact that critical points of the density ($\nabla\rho = 0$) arise when atoms interact. If the interaction is bonding, the point is expected to be a first order saddle point. Many properties, including the density itself, its Laplacian ($\nabla^2\rho$), and the kinetic-energy density, have been found to correlate with the interaction energy in vdW and HB complexes for families of related compounds.^{11,12} A parallel approach is based on the analysis of the electron localization function (ELF).^{13,14} This is a function of the electron density and the kinetic-energy density that was developed to highlight regions of electron localization (i.e., covalent bonds, lone pairs, etc.). HBs have also been studied using this approach, since the ELF value is proportional to the strength of the HB.^{15,16}

Recently, we introduced a new approach to visualize noncovalent interactions, based on the analysis of the electron densities and their reduced gradients, $s(\vec{r})$.¹⁷ This approach was chosen for its ability to highlight interactions in the low-density regime (*vide infra*). The non-covalent interactions (NCI) index identifies interactions in a chemical system solely on the basis of the electron density and its derivatives. Indeed, a similar approach was recently introduced to analyze covalent bonds, highlighting the ability of this type of function to identify all types

Received: November 8, 2010

Published: January 25, 2011

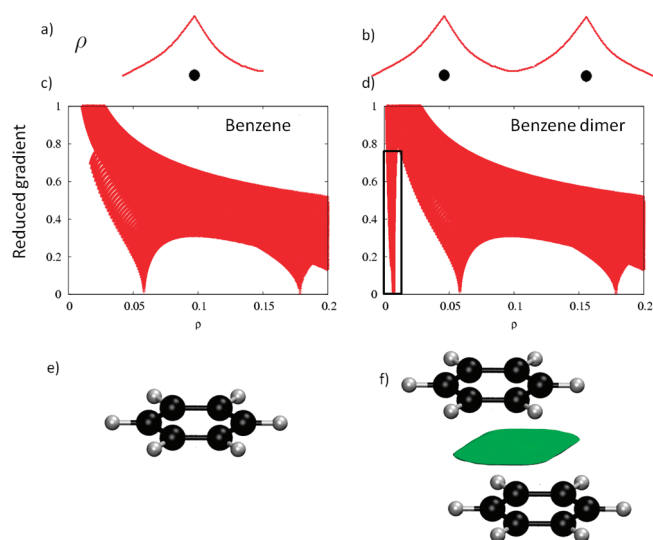


Figure 1. (a) Representative behavior of atomic density. (b) Appearance of a $s(\rho(\vec{r}))$ singularity when two atomic densities approach each other. (c, d) Comparison of the reduced density behavior for the benzene monomer and dimer; a singularity in s appears at low density values in the dimer case. (e) Benzene monomer. (f) Appearance of an intermolecular interaction surface in the benzene dimer, associated with the additional singularity in the $s(\rho)$ plot. The isosurface was generated for $s = 0.7$ au and $\rho < 0.01$ au.

of bonding situations.¹⁸ Here, we introduce a practical strategy for NCI analysis and visualization of weak interactions.

The paper first summarizes NCI theory. Next, the algorithms underlying NCI analysis are discussed. Finally, we apply the approach to chemical examples that span a diverse set of interaction types and system sizes. Emphasis is placed on the ability of the new approach to describe the bonding in systems currently under debate.

Details of the computations are found in the Supporting Information. The software and manual may be downloaded at <http://www.chem.duke.edu/~yang/software.htm>

II. THEORETICAL BACKGROUND

The NCI analysis provides an index, based on the electron density and its derivatives, that enables identification of non-covalent interactions.¹⁷ The NCI index is based on a 2D plot of the reduced density gradient, s , and the electron density, ρ , where

$$s = \frac{1}{2(3\pi^2)^{1/3}} \frac{|\nabla\rho|}{\rho^{4/3}} \quad (1)$$

When a weak inter- or intramolecular interaction is present, there is a crucial change in the reduced gradient between the interacting atoms, producing density critical points between interacting fragments (Figure 1a,b). Troughs appear in $s(\rho)$ associated with each critical point. Since the behavior of s at low densities is dominated by ρ , s tends to diverge except in the regions around a density critical point, where $\nabla\rho$ dominates, and s approaches zero. This fact is highlighted in Figure 1c,d, which shows that the main difference in the $s(\rho)$ plots between a monomer and a dimer is the steep trough at low density. When we search for the points in real space giving rise to this feature, the noncovalent region clearly appears in the (supra)molecular complex (green isosurface in Figure 1f).

Further analysis of the electron density in the troughs is required to assign the origins of these troughs (steric interactions, hydrogen bonds, etc.). The electron density values within the troughs are an indicator of the interaction strength. However, both attractive and repulsive interactions (i.e., hydrogen-bonding and steric repulsion) appear in the same region of density/reduced gradient space. To distinguish between attractive and repulsive interactions, we examine the second derivatives of the density along the main axis of variation.

On the basis of the divergence theorem,¹⁹ the sign of the Laplacian ($\nabla^2\rho$) of the density indicates whether the net gradient flux of density is entering ($\nabla^2\rho < 0$) or leaving ($\nabla^2\rho > 0$) an infinitesimal volume around a reference point. Hence, the sign of $\nabla^2\rho$ determines whether the density is concentrated or depleted at that point, relative to the surroundings. To distinguish between different types of weak interactions, one cannot use the sign of the Laplacian itself, because the sign is dominated by negative contributions from the nuclei.²⁰ Instead, contributions to the Laplacian along the axes of its maximal variation must be analyzed. These contributions are the eigenvalues λ_i of the electron-density Hessian (second derivative) matrix, such that $\nabla^2\rho = \lambda_1 + \lambda_2 + \lambda_3$, ($\lambda_1 < \lambda_2 < \lambda_3$). At the nuclei, all of the eigenvalues are negative, while away from them, $\lambda_3 > 0$. In molecules, the λ_3 values vary along the internuclear direction, while λ_1 and λ_2 report the variation of density in the plane normal to the λ_3 eigenvector. Interestingly, the second eigenvalue (λ_2) can be either positive or negative, depending on the interaction type. On the one hand, bonding interactions, such as hydrogen bonds, are characterized by an accumulation of density perpendicular to the bond, and $\lambda_2 < 0$. Nonbonded interactions, such as steric repulsion, produce density depletion, such that $\lambda_2 > 0$. Finally, vdW interactions are characterized by a negligible density overlap that gives $\lambda_2 \approx 0$. Thus, analysis of the sign of λ_2 enables us to distinguish different types of weak interactions, while the density itself enables us to assess the interaction strength.

The dependence of $s(\vec{r})$ on $\text{sign}(\lambda_2)\rho$ is shown in Figure 2a, which uses a modification of our earlier reduced gradient and density plots. The 3D plots resulting from using $\text{sign}(\lambda_2)\rho$ are shown in Figure 2, bottom. The low-density, low-reduced gradient trough in the hydrogen-bonded water dimer now lies at negative values, indicative of an attractive interaction. Conversely, the low-density, low-gradient trough for the sterically crowded bicyclo[2,2,2]octane molecule remains at positive values, indicating a repulsive interaction. Finally, the low-density, low-gradient trough for the dispersion-bound methane dimer is very near zero, with slightly negative $\text{sign}(\lambda_2) \times \rho$ negative values of -0.0025 au, indicative of weak attraction.

III. ALGORITHMIC DETAILS

Figure 3 shows the protocol for visualizing noncovalent interactions in NCIPLOT. A detailed summary of the tasks associated with each routine is shown in Table 1 (cube construction, properties, visualization, and I/O flow).

The key data-flow details are highlighted in Figure 3. The input of data is shown in red, and its processing is indicated in black. Two basic types of data constitute the input: the density information (based on wave functions or molecular geometries) and the analysis options, which determine the noncovalent interactions to be plotted. Four algorithms analyze the data: (i) the selection of interactions (through the input), (ii) the construction of the cube and the grid (in CUBE, see Table 1),

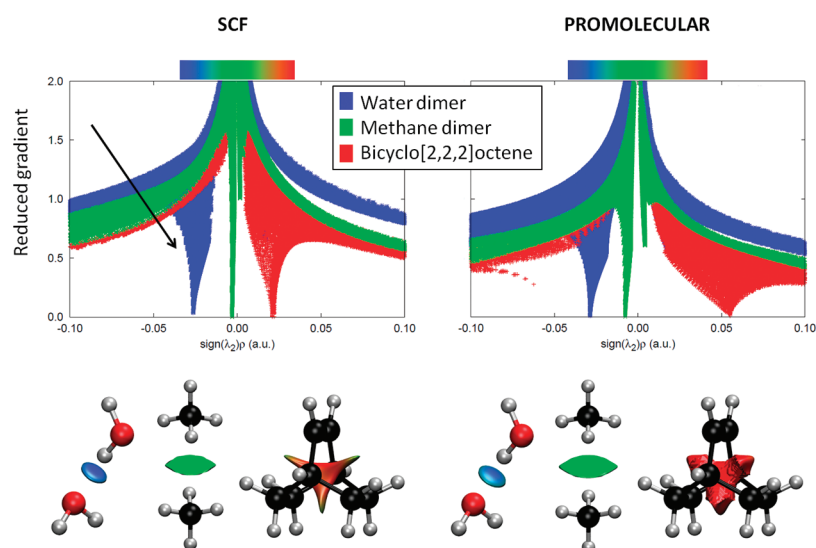


Figure 2. Top: Overlapping troughs in $s(\rho)$ plots, which can be distinguished when $\text{sign}(\lambda_2)\rho$ is used as the ordinate. Favorable interactions appear on the left, unfavorable on the right, and vdW near zero. The same $s(\rho)$ features are obtained using self-consistent (left) and promolecular (right) calculations, with a shift toward negative (stabilizing) regimes. Bottom: Taking the shift in troughs into account (i.e., changing the cutoff), the isosurface shapes remain qualitatively unaltered for selected small molecules. Figures are shown for both SCF (left) and promolecular densities (right). NCI surfaces correspond to $s = 0.6$ au and a color scale of $-0.03 < \rho < 0.03$ au for SCF densities. For promolecular densities, $s = 0.5$ au (water and methane dimers) or $s = 0.35$ au (bicyclo[2,2,2]octene), and the color scale is $-0.04 < \rho < 0.04$ au.

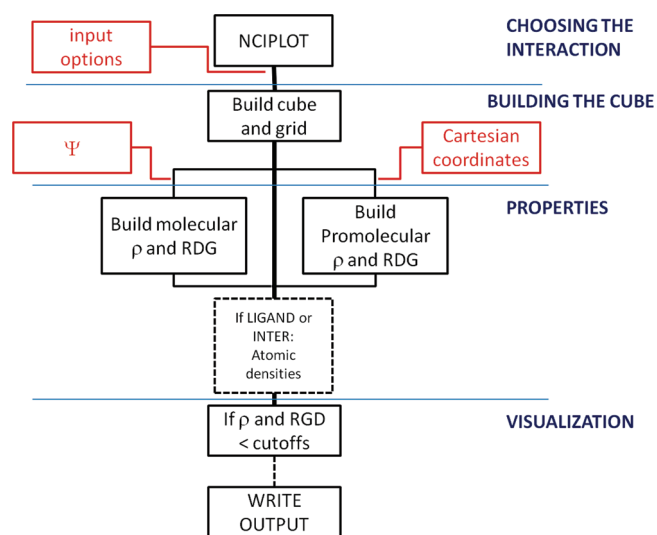


Figure 3. Flowchart for program routines for noncovalent interactions' visualization in NCIPLLOT. Red labels highlight the information that can be input by the user, whereas black labels show the internal flow of information. The flow is divided into four main algorithmic parts: input, cube construction, properties, and visualization.

(iii) the calculation of properties at each point (utilizing a number of routines), and (iv) the calculation of visualization data (carried by the main routine, NCIPLLOT). Since the input is keyword oriented, the program includes a number of parsing routines. These main features are discussed in the following sections.

III.A. Building the Cube. Interaction analysis is based on examination of local properties on a cubic grid constructed within the program. This procedure was found to be extremely efficient for computing stable properties (as is the case of NCI, see section III.B).²¹ Furthermore, this approach enables us to

discard contributions from high-density points in the construction of isosurfaces. The spatial region to be analyzed is determined, by default, in terms of the molecular geometry. Unless otherwise noted, a cube is constructed from the outermost x, y, z coordinates for all of the molecules in the input. An extra radial threshold in each direction is added to ensure that the isosurfaces are contained within the cube (no intermolecular interactions are expected in those regions, but isosurfaces can spread beyond the atoms). A practical threshold was defined as 2 Å:

$$x_i(0) = \min[x_i] - 2 \text{ \AA} \quad (2)$$

$$x_i(1) = \max[x_i] + 2 \text{ \AA} \quad (3)$$

where $x_i = x, y, z$. This step also eliminates spurious symmetry-related cancellations (e.g., if benzene were analyzed without this extra threshold, there would be no points along the C_6 axis).

It can be useful to construct a user-defined cube or to analyze the interactions only around one point or molecule (*vide infra*). All of these options are implemented in the program and are described in the Supporting Information.

III.B. Properties. Density, reduced gradient, and λ_2 values are calculated at each point on the grid. Densities can be obtained from quantum-mechanical calculations or from promolecular estimates. Topological features of the electron density in the weak interaction region are very stable with respect to the calculation method, to such an extent that these features are already contained in the sum of atomic densities (ρ_i^{at}).^{22,23} The molecular density computed from the sum of atomic contributions, also known as promolecular density (ρ^{pro}), is

$$\rho^{\text{pro}} = \sum_i \rho_i^{\text{at}} \quad (4)$$

Promolecular densities lack the relaxation introduced in a SCF Hartree–Fock or DFT calculation; however, the promolecular densities are very useful for describing large biomolecular

Table 1. NCIPLOT Tasks and Details of the Input and Output^a

MODULE	ROUTINE	TASK	INPUT	OUTPUT
MAIN	NCIPLOT	main routine		visualization
CUBE	CUBE	constructs cube and grid	geometry	cube, grid
I/O	TIMER	accumulates run times	process and resetting	elapsed times
	GETARGS	get command	arguments	argument count
	GETDATE	extracts time and date	(operatingsystem)	date and time
	ZATGUESS	provides atomic #s	atomic symbol	atomic #
	RWFN	stores wfn information	wfn file	wfn information
	RPRM	stores xyz information	xyz file	xyz information
PROPS	PROPPROM	calculates properties at \vec{r} (PROM)	\vec{r}	$\rho, s(\rho)$
	CALCHESS	calculates λ_2 at \vec{r} (PROM)	\vec{r}	λ_2
	PROPWFN	calculates properties at \vec{r} (SCF)	\vec{r}	$\rho, s(\rho), \lambda_2$
	F012	derivatives at \vec{r} (SCF)	\vec{r}	derivatives
	PHI012	summation over primitives	\vec{r}	summation
	PRI012	construction of primitives	\vec{r}	primitives
	INDEX0	angular momentum assignment	primitive type	L_i
	RS	matrix diagonalization (eispack)	matrix	eigenvalues

^aThe table is organized following the main algorithmic subdivisions of the flowchart (Figure 3): cube construction, properties, and visualization, as well as I/O flow. wfn file stands for the wavefunction file (wfn extension) commonly used for post-SCF analyses.

systems. In these systems, noncovalent interactions are crucial for describing the interplay of structure and reactivity.³ Because electron density calculations for these large systems are extremely expensive computationally, use of the promolecular density is an attractive option. When relaxed densities are compared to promolecular ones, a shift in the $s(\rho)$ troughs is observed toward bonding regimes (see Figure 2 top). Specifically, a large shift toward smaller density values is observed in the troughs corresponding to regions of nonbonded overlap, introducing less repulsion and greater stability. However, once this shift is taken into account (by changing the density cutoff), results at the self-consistent and promolecular levels are qualitatively equivalent for all of the cases considered (see Figure 2, bottom).

SCF densities are constructed from the wave function information stored in the wfn file, whereas promolecular densities are constructed from the atomic positions stored in the xyz coordinate file(s). In order to store atomic densities, fully numerical LSDA²⁴ free-atomic densities were generated for the neutral atoms H to Ar, spherically averaged over space and summed over spins. Because atomic densities are piece-wise exponentially decaying for each shell of electrons, they were then fit to one (H, He), two (Li–Ne), or three (Na–Ar) Slater-type functions of the form $\rho^{\text{at}} = \sum_j c_j e^{-r/\zeta_j}$. Once these densities are written as simple sums of exponential functions, the NCI surfaces can be calculated very efficiently for each (supra)molecule, since all of the necessary data (ρ, s, λ_2) can be obtained analytically.

III.C. Visualization: The Cutoffs. The ρ, s coordinates of the density troughs define the appropriate cutoffs for the noncovalent interactions. For example, a cutoff of $\rho < 0.05$ au is appropriate for recovering the noncovalent interactions in the benzene dimer, including the nonbonding regions at the center of each ring (Figure 1c,d). All points giving rise to ρ values above this threshold need to have their s values set to a large value. This enables the user to recover only the noncovalent interactions when $s = S$ is plotted (for some isosurface value S) because this rescaling eliminates all points in density/reduced gradient space with greater density values and $s < S$ (i.e., points with $\rho > 0.05$ au and $s < 0.5$ au would also appear in the isosurface representation

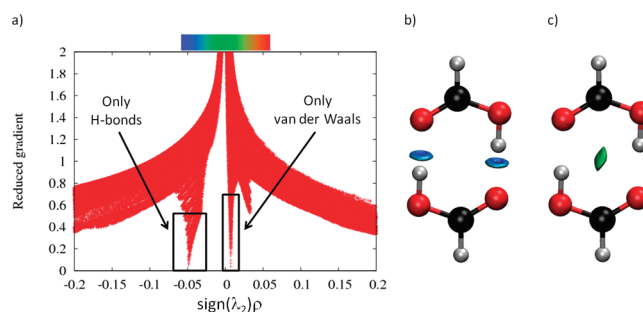


Figure 4. NCI analysis of formic acid dimer. (a) $s(\rho)$ plot for the SCF density. Peaks appear at $\rho \approx 0.01$ au for vdW and $\rho \approx 0.05$ au for hydrogen bonds. (b) If the cutoffs are set at $s = 0.7$ au and $\rho < 0.02$ au, the isosurface only recovers the vdW interactions in the system. (c) If the cutoffs are set at $s = 0.5$ au and $0.02 < \rho < 0.06$ au, only the hydrogen bonds are displayed. The NCI color scale is $-0.06 < \rho < 0.06$ au.

in Figure 1). Another example is provided in Figure 4. The formic acid dimer troughs appear at $\rho = 0.01$ au for vdW contacts, and $\rho = 0.05$ au for hydrogen bonds. If the cutoff is set to $\rho = 0.02$ au, the isosurface will only recover the vdW interactions in the system (Figure 4). Furthermore, placing a threshold for the interval $\rho = [0.02-0.06]$ au enables the user to isolate the hydrogen bonds in a similar manner.

It is convenient, therefore, to perform a preliminary run, where only $s(\rho)$ values are produced, and the user can use these data to determine optimal cutoffs. A second run can subsequently target the noncovalent interactions in a given molecule with no interference from other density regions. For this reason, the current implementation enables the user to decide which file types are to be output.

III.D. Input: Selectively Displaying the Interactions. In order to plot a certain interaction selectively, there are several constraints that can be applied to display the weak interactions of interest. Criteria for filtering the interactions include the strength of the interaction, its localization in space, and its nature (whether it is intermolecular or intramolecular). The details for each of these cases are described below.

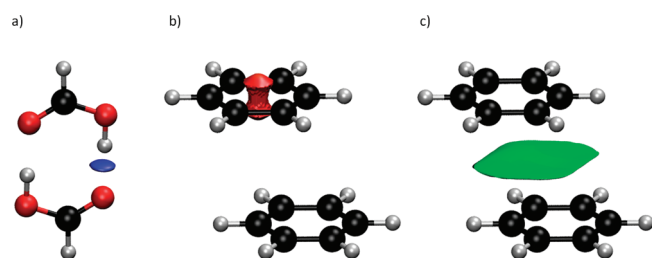


Figure 5. Results of several input options for the selective representation of a given interaction based on (a) its localization in 3D space, defined by the cube, (b) its localization in 3D space, defined by a radial threshold around a point (in this case, the center of the top benzene was chosen), (c) its inter-/intramolecular nature (in this case only intermolecular interactions are shown). NCI surfaces correspond to $s = 0.4$ au and a color scale of $-0.04 < \rho < 0.04$ au, using promolecular densities.

- Strength:** It is possible to select the interaction in terms of its strength by the choice of cutoff parameters. Figure 4 uses the information in section III.C to plot only the vdW contact region in the formic acid dimer, while avoiding the hydrogen bonds. It is also possible to define an interval range, as in the formic acid dimer case, to identify only hydrogen bonds and exclude vdW interactions. An interval of the kind $\rho = [0.02-0.06]$ au produces this result.
- Geometry:** There are two possibilities when choosing a given interaction from its location in 3D space:
 - An appropriate choice of the cube boundaries enables the selection of individual interactions (CUBE keyword in the Supporting Information). The cube in Figure 5a captures only one of the hydrogen bonds in the formic acid dimer. This option is especially suitable if there are several interactions of the same type, but only one of them is of interest. In this case, the strength criterion would not differentiate the interactions, and the geometric selection should be used.
 - An alternate implementation for the geometric criterion consists of defining the center of the cube instead of its boundaries (RADIUS keyword in the Supporting Information). This option uses the origin and length of the box sides as input, rather than the Cartesian coordinates themselves. Results for the benzene dimer are shown in Figure 5b, where the steric clashes within a single ring are highlighted by searching for noncovalent interactions near the center of the upper benzene molecule.
- Pure Intermolecular Interactions:** All of the interactions with at least a specified fraction (e.g., $f = 0.9$) of the density from a single molecule are turned off:

$$\frac{\rho_{\text{molec}}}{\rho_{\text{tot}}} = \begin{cases} \geq f & \text{intramolecular} \\ < f & \text{intermolecular} \end{cases} \quad (5)$$

This choice causes only intermolecular interactions to be plotted, screening out the intramolecular interactions. Figure 5c shows results for the benzene dimer, where the internal steric repulsions are removed, since most of their density is obtained from a single benzene contribution. In order to construct molecular densities, atoms need to be assigned to each monomer. This is readily automated if each monomer is uploaded in a different file. This proce-

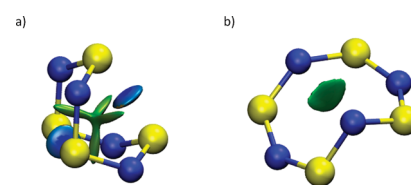


Figure 6. S_4N_4 main conformations. (a) Boat conformation. The boat conformation is more stable due to the bridging S–S bond. NCI surfaces correspond to $s = 0.4$ au and a color scale of $-0.05 < \rho < 0.05$ au.

dures enables the characterization of monomers and the construction of ρ_{molec} .

- Centered around a Given Molecule:** If the radial and the intermolecular options are used together, the intermolecular interactions around just one of the molecules will be highlighted (implemented via the LIGAND keyword, described in the Supporting Information). In this case, the desired cube is set around one entire molecule rather than around a given point. This option is particularly useful to study inclusion complexes and protein–ligand interactions, where a small molecule binds in a cavity and we wish to describe the interactions at this active site. In these cases, the interactions of interest are not only intermolecular but are localized around the smaller of the two partners (see protein–ligand interactions for HDAC8 protein in section IV.B.2).

IV. APPLICATIONS

Here, we review some examples where the applicability of NCIPLOT is illustrated for a range of systems (large and small molecules, inorganic complexes, and biological systems). Examples are chosen to highlight the capabilities of NCI and to shed light on challenging open issues as to the nature of specific interactions.

IV.A. Coordination Chemistry. IV.A.1. S–S Bond in S_4N_4 . Tetrasulfur tetranitride, S_4N_4 , is a textbook example that illustrates structure, bonding, and reactivity in main group inorganic compounds. Although discovered in 1835, S_4N_4 remains a subject of intensive study. S_4N_4 exists as an eight-membered ring. The most stable conformation is a D_{2d} “boat”, in which the ring folds back on itself to give two close S–S contacts bridging the ring.²⁵ Another possible but less-stable conformer is the C_2 “cage” structure. Both geometries were found in previous DFT calculations²⁶ and are shown in Figure 6. The greater stability of the boat conformation was explained by the presence of S–S bonding interactions.²⁷ These bonding interactions are clearly revealed by the NCI analysis and are predicted to be fairly strong, on the order of HB strengths, with $s(\rho)$ troughs at $\rho \sim 0.044$ au. There is also a nonbonding interaction through the center of the boat. Conversely, there is only a weak vdW interaction at the center of the more open cage structure. Thus, the NCI representation explains the greater stability for the boat conformation of S_4N_4 .

IV.A.2. Hg Complexes. Understanding the solvation of ions and their interactions with ligands is of prime relevance for rationalizing their bioactivity. Since metals play a decisive role in many protein active sites as cofactors, it would be useful to have simple means of describing metal–ligand interactions.

Mercury(II) is a heavy metal cation especially challenging for quantum-mechanical treatment, as both correlation and relati-

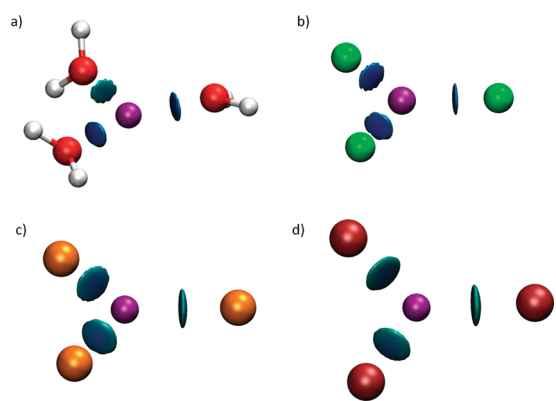


Figure 7. Complexes of Hg. (a) $[\text{Hg}(\text{H}_2\text{O})_3]^{2+}$, (b) $[\text{Hg}(\text{F})_3]^-$, (c) $[\text{Hg}(\text{Cl})_3]^-$, (d) $[\text{Hg}(\text{Br})_3]^-$. NCI surfaces correspond to $s = 0.3$ au and a color scale of $-0.1 < \rho < 0.1$ au.

vistic effects play a crucial role in its bonding and electronic structure. Figure 7a–d show how NCI is used to quickly and efficiently assess Hg complexation sites and discern the strength of binding between Hg^{2+} and its ligands. The DFT optimized geometry of the $[\text{Hg}(\text{H}_2\text{O})_3]^{2+}$ complex indicates that the three waters bound to Hg are not equivalent, with one of the waters further away than the other (2.2 Å vs 2.3 Å, see Figure 7a). In order to analyze the nature of the stabilization, we decompose the interaction energy of the complex using the RVS²⁸ (reduced variational space) procedure. Both polarization and charge transfer are significantly weaker for one water molecule than for the other two. While two water molecules show a polarization energy of -15 to -16 kcal/mol and a charge transfer energy of -10 kcal/mol, the third water shows stabilization due to polarization and charge transfer of only -12.7 kcal/mol and -6.2 kcal/mol, respectively. Specific details of the calculation are found in the Supporting Information. The weaker binding of one of the water molecules is clearly revealed by the NCI representation.

NCI also recovers the ordering of binding energies when different ligand series are analyzed. Figure 7b–d show $[\text{Hg}(\text{X})_3]$ complexes ($\text{X} = \text{F}, \text{Cl}$ or Br). F is more strongly bound to Hg than Cl than Br. This result is in agreement with the $[\text{Hg}(\text{X})_3]$ binding energies, which are -632.8 (F), -571.0 (Cl), and -562.6 kcal/mol (Br) (see calculation details in the Supporting Information).

IV.A.3. BH_3NH_3 and the Dihydrogen Bond. Recently, the term “dihydrogen bond” was coined to describe an interaction of the type $\text{D}-\text{H}\cdots\text{H}-\text{E}$, where D is a typical hydrogen donor (such as N or O).⁴⁰ The novelty of this bond is that the acceptor atom is also a hydrogen. Thus, the accepting hydrogen atom must be negatively charged, and E is an atom capable of accommodating a hydridic hydrogen. Transition metal and boron atoms are known examples of atoms serving as E atoms.

BH_3NH_3 is a widely studied example among dihydrogen-bonded complexes.^{8,35,36} Crabtree et al.³⁷ placed the $\text{NH}\cdots\text{HB}$ contacts at the upper end of the energy range for hydrogen bonds. Popelier, instead, assigned these interactions to the range of normal H-bond strengths.³⁵ Morrison and Siddick³⁶ assigned the energies to the lower end of the hydrogen-bond-strength spectrum. The most relevant point made by the later authors is that there is a large conformational difference between the native crystal structure and the gas-phase minima considered by Crabtree et al.³⁷ and Popelier.³⁵

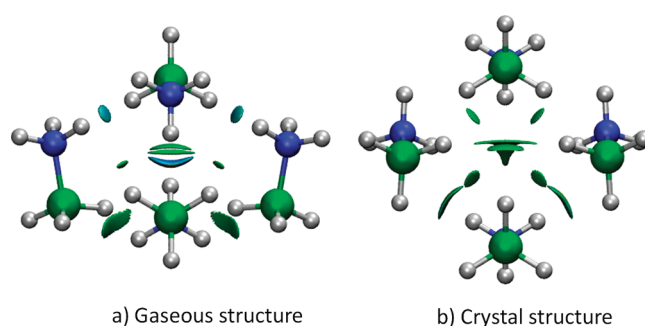


Figure 8. Dihydrogen interactions in a BH_3NH_3 tetramer in (a) the fully optimized gas-phase geometry and (b) the solid-state geometry. NCI surfaces correspond to $s = 0.4$ au and a color scale of $-0.03 < \rho < 0.03$ au.

Figure 8 shows the NCI results for two different $(\text{BH}_3\text{NH}_3)_4$ complexes. In Figure 8a, the structure was completely optimized in the gas phase, whereas in Figure 8b, the tetramer geometry was derived from the crystalline structure (with the hydrogen positions optimized to ensure correct description of the H–H contacts).³⁸ Calculations were performed using the B3LYP functional and support qualitatively different interaction types in each case. Whereas the gas-phase tetramer gives rise to a highly negative interaction energy, the crystalline structure is not stable with this density functional. This difference suggests that the solid structure is only stabilized by dispersion interactions (which are completely neglected using this functional³⁹); stronger interactions are present in the fully optimized gas-phase complex.

The NCI description is able to recover the different nature of the interactions in both tetramers, supporting Morrison and Siddick’s conclusion.³⁶ Whereas the crystalline tetramer only shows vdW contacts, the completely optimized structure gives rise to stronger dihydrogen bonds (with $\rho \approx 0.02$ au), which are closer in energy to normal hydrogen bonds.

IV.B. Large Systems. IV.B.1. High Affinity Host–Guest Complexes. Host–guest complexation, like protein–ligand binding, depends upon a balance of stabilizing and destabilizing noncovalent interactions that contribute to the net thermodynamics. These systems are well-known exemplars for understanding noncovalent interactions in more complex biomolecular systems. Recent experimental results show that synthetic host–guest systems can achieve binding affinities that rival those of the tightest-bound protein–ligand complexes. For example, the seven-unit cucurbit[7] host (CB[7], Figure 9) binds cationic adamantyl,²⁹ ferrocene derivatives,³⁰ and bicyclo[2,2,2]octane derivatives^{31–33} with binding constants of 10^9 to 10^{13} M^{-1} . Thus, a key test of our NCI method is to understand the magnitude of noncovalent interactions in host–guest complexes.

Mining minima algorithms (M2) have been used to dissect the enthalpic and entropic contributions to high binding affinities.³³ Despite the evident electrostatic complementarity of the cationic guests and the electronegative, carbonyl portions of CB[7], electrostatic interactions do not provide a significant net driving force for ligand binding. This is because the strong Coulombic attractions between the guests and CB[7] are offset by the energetic cost of removing solvating water from the cationic guests and the polar hosts upon binding. Instead, two other factors are found to be crucial for stabilizing these complexes. The entropy penalty is unusually small in relation to the enthalpy of binding. This is likely the case because of the rigidity of both

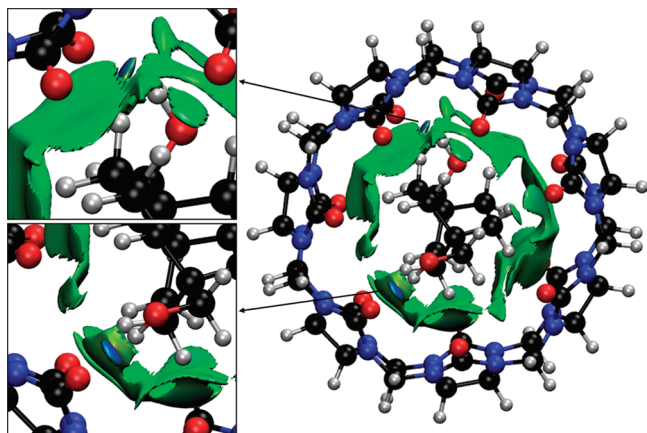


Figure 9. Cucurbit[7]uril-bicyclo[2,2,2]octane derivative inclusion complex. Anchor posts are highlighted in the insets. NCI surfaces show only intermolecular interactions. The gradient cutoff is $s = 0.5$ au, and the color scale is $-0.04 < \rho < 0.04$ au.

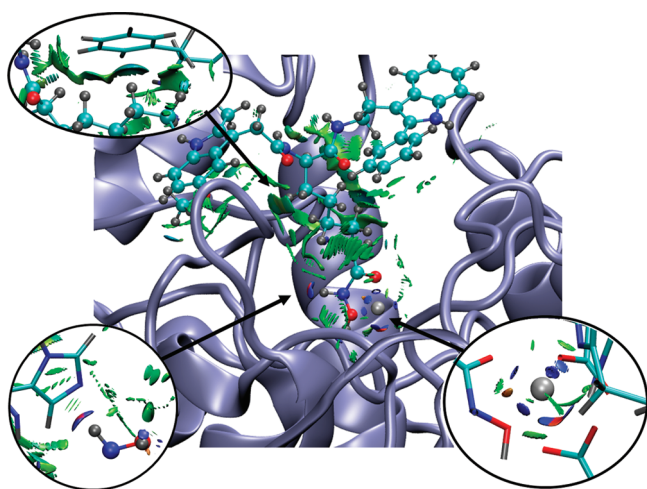


Figure 10. NCI surface around a VSX ligand in the active site of HDAC8 protein. Specific interactions are enlarged in the insets. NCI surfaces show only intermolecular interactions. The gradient cutoff is $s = 0.35$ au, and the color scale is $-0.04 < \rho < 0.02$ au.

the host and the guest structures.^{31,33} Furthermore, the host and guest molecules studied are complementary in their preferred conformations. These characteristics give rise to favorable enthalpy–entropy compensation that enhances the binding affinities.

NCI analysis allows an assessment of host–guest complementarity and the extent to which weak enthalpic interactions stabilize a complex (NCI would not identify stabilization due to purely Coulombic forces). Figure 9 shows the binding between a bicyclo[2.2.2]octane derivative and CB[7]. In agreement with M2 results, vdW interactions are established throughout the cavity. The binding of the guest is further favored by the two hydroxyl anchors: the hydroxyl substituents on the guest establish strong hydrogen bonds with two carbonyls of the CB7 host (see Figure 9 insets).

IV.B.2. Protein–Ligand Interactions. Figure 10 shows the promolecular NCI surface around a VSX ligand in the active site of the HDAC8 protein (obtained from the protein data bank, as 2v5x.pdb, ref 34), with the details of the NCI surfaces enlarged. Specific parameters for the Zn^{2+} contribution to the promolec-

ular density were calculated as described in section III.B. The NCI analysis reveals a set of complex interactions between the ligand and the protein, which arise from a combination of specific atom–atom interactions (i.e., hydrogen bonds) as well as broad surfaces indicative of stabilizing vdW interactions. On the lower right-hand side of the figure, we show the stabilization of the Zn^{2+} ion by the protein and ligand. At the lower left, we show the hydrogen bond between the ligand and His140. The top left of the figure shows the vdW surface between the ligand and the Phe139 phenyl ring. NCI analysis clearly highlights how a ligand “fits” the geometry of the active site, and the many small contributions that combine to determine the interaction energy between the ligand and protein.

V. CONCLUSIONS

Algorithms for the analysis of noncovalent interactions based on the NCI index were described. An efficient and flexible implementation was established for both SCF and promolecular densities, allowing for the analysis of noncovalent interactions in both small molecules and macromolecules. The utility of the method was illustrated with examples from organic and inorganic chemistry. Special emphasis was given to the ability of the new NCI index to provide insight into open issues in bonding.

The NCI code and a manual may be downloaded at <http://www.chem.duke.edu/~yang/software.htm>

■ ASSOCIATED CONTENT

S Supporting Information. Input files for selected examples and the programming details are available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: weitaoyang@duke.edu.

■ ACKNOWLEDGMENT

This work is supported by the National Science Foundation (CHE-09-11119 & CHE-1012357), and the National Institute of Health through the UPCMLD project (P50GM067082) and ROI-GM-061870. J.C.G. thanks the Spanish MALTA-Consolider Ingenio-2010 program under project CSD2007-00045, and E.R.J. thanks the Natural Sciences and Engineering Research Council of Canada. Support from the French National Research Agency (ANR) is acknowledged, Grant ANR-08-BLAN-0158 (R.C. and J.P.P.). Contributions from H. L. Schmider (wave function properties), A. Martín Pendás (parsing routines), E.J. Toone and Y. Wang (CB[7] data), and J. Li (S_4N_4 data) are gratefully acknowledged.

■ REFERENCES

- (1) (a) Chandler, D. *Nature* **2005**, 437, 640. (b) Panigrahi, S. K.; Desiraju, G. R. *Proteins: Struct., Funct., Bioinf.* **2007**, 67, 128.
- (2) (a) Fenniri, H.; Packiarajan, M.; Vidale, K. L.; Sherman, D. M.; Hallenga, K.; Wood, K. V.; Stowell, J. G. *J. Am. Chem. Soc.* **2001**, 123, 3854. (b) Kruse, P.; Johnson, E. R.; DiLabio, G. A.; Wolkow, R. A. *Nano Lett.* **2002**, 2, 807.
- (3) (a) Fiedler, S.; Broecker, J.; Keller, S. *Cell. Mol. Life Sci.* **2010**, 67, 1779. (b) Dill, K. A. *Biochemistry* **1990**, 29, 7133.
- (4) Silvi, B. *Phys. Rev. Lett.* **1994**, 73, 842.

- (5) Bickelhaupt, F. M.; Baerends, E. *Angew. Chem., Int. Ed.* **2003**, *42*, 4183.
- (6) Weinhold, F. *Angew. Chem., Int. Ed.* **2003**, *42*, 4188.
- (7) Cioslowski, J.; Mixon, S. T. *Can. J. Chem.* **1992**, *70*, 443.
- (8) Matta, C. F.; Hernández-Trujillo, J.; Tang, T.; Bader, R. F. W. *Chem.—Eur. J.* **2003**, *9*, 1940.
- (9) Klein, R. A. *Chem. Phys. Lett.* **2006**, *425*, 128.
- (10) Bader, R. F. W. *Atoms in Molecules: A Quantum Theory*; Oxford University Press: Oxford, U.K., 1990.
- (11) Espinosa, E.; Souhassou, M.; Lachekar, H.; Lecomte, C. *Acta Crystallogr., Sect. B* **1999**, *55*, 563.
- (12) Grabowski, S. J. *J. Phys. Chem. A* **2001**, *105*, 10739.
- (13) Becke, A. D.; Edgecombe, K. E. *J. Chem. Phys.* **1990**, *92*, 5397.
- (14) Silvi, B.; Savin, A. *Nature* **1994**, *371*, 683.
- (15) Contreras-García, J.; Recio, J. M. *Theor. Chem. Acc.* DOI: 10.1007/s00214-010-0828-1.
- (16) Alikhani, M. E.; Fuster, F.; Silvi, B. *Struct. Chem.* **2005**, *16*, 203.
- (17) Johnson, E. R.; Keinan, S.; Mori-Sánchez, P.; Contreras-García, J.; Cohen, A. J.; Yang, W. *J. Am. Chem. Soc.* **2010**, *132*, 6498.
- (18) Bohórquez, H. J.; Boyd, R. J. *Theor. Chem. Acc.* **2010**, *127*, 393.
- (19) Arfken, G. *Mathematical Methods for Physicists*; Academic Press: Orlando, FL, 1985.
- (20) Bader, R. F. W.; Essén, H. *J. Chem. Phys.* **1984**, *80*, 1943.
- (21) Noury, S.; Krokidis, X.; Fuster, F.; Silvi, B. *Comput. Chem.* **1999**, *23*, 597.
- (22) Spackman, M. A.; Maslen, E. N. *J. Phys. Chem.* **1986**, *90*, 2020.
- (23) Pendás, A. M.; Luaña, V.; Pueyo, L.; Francisco, E.; Mori-Sánchez, P. *J. Chem. Phys.* **2002**, *117*, 1017.
- (24) Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244.
- (25) DeLucia, M. L.; Coppens, P. *Inorg. Chem.* **1978**, *17*, 2336.
- (26) Pritchina, E. A.; Gritsan, N. P.; Zibarev, A. V.; Bally, T. *Inorg. Chem.* **2009**, *48*, 4075.
- (27) Gleiter, R. *J. Chem. Soc. A* **1970**, 3174. Salahub, D. R.; Messmer, R. P. *J. Chem. Phys.* **1976**, *64*, 2039.
- (28) Stevens, W. J.; Fink, W. H. *Chem. Phys. Lett.* **1987**, *139*, 15.
- (29) Liu, S.; Ruspic, C.; Mukhopadhyay, P.; Chakrabarti, S.; Zavalij, P. Y.; Isaacs, L. *J. Am. Chem. Soc.* **2005**, *127*, 1595.
- (30) Jeon, W. S.; Moon, K.; Park, S. H.; Chun, H.; Ko, Y. H.; Lee, J. Y.; Samal, S.; Selvapalam, N.; Rekharsky, M. V.; Sindelar, V.; Sobransingh, D.; Inoue, Y.; Kaifer, A. E.; Kim, K. *J. Am. Chem. Soc.* **2005**, *127*, 12984.
- (31) Rekharsky, M. V.; Mori, T.; Yang, C.; Ko, Y. H.; Selvapalam, N.; Kim, H.; Sobransingh, D.; Kaifer, A. E.; Liu, S.; Isaacs, L.; Chen, W.; Moghaddam, S.; K. Gilson, M. K.; Kim, K.; Inoue, Y. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 20737.
- (32) Liu, S.; Ruspic, C.; Mukhopadhyay, P.; Chakrabarti, S.; Zavalij, P. Y.; Isaacs, L. *J. Am. Chem. Soc.* **2005**, *127*, 15959.
- (33) Moghaddam, S.; Inoue, S.; Gilson, M. K. *J. Am. Chem. Soc.* **2009**, *131*, 4012.
- (34) Vannini, A.; Volpari, C.; Gallinari, P.; Jones, P.; Mattu, M.; Carfi, A.; Defrancesco, R.; Steinkuhler, C.; Di Marco, S. *EMBO Rep.* **2007**, *8*, 879.
- (35) Popelier, P. L. A. *J. Phys. Chem. A* **1998**, *102*, 1873.
- (36) Morrison, C. A.; Siddick, M. M. *Angew. Chem., Int. Ed. Engl.* **2004**, *116*, 4884.
- (37) Richardson, T. B.; de Gala, S.; Crabtree, R. H.; Siegbahn, P. E. M. *J. Am. Chem. Soc.* **1995**, *117*, 12875.
- (38) Klooster, W. T.; Koetzle, T. F.; Siegbahn, P. E. M.; Richardson, T. B.; Crabtree, R. H. *J. Am. Chem. Soc.* **1999**, *121*, 6337.
- (39) Johnson, E. R.; Wolkow, R. A.; DiLabio, G. A. *Chem. Phys. Lett.* **2004**, *394*, 334.
- (40) Brown, M. P.; Heseltine, R. W. *Chem. Commun. (London)* **1968**, *23*, 1551.

NOTE ADDED AFTER ASAP PUBLICATION

This article was published ASAP on January 25, 2011. The name of the first author in ref 26 has been corrected. The correct version was published on February 2, 2011.

Electron Density Based Partitioning Scheme of Interaction Energies

Marcos Mandado* and José M. Hermida-Ramón

Department of Physical Chemistry, University of Vigo, Lagoas-Marcosende s/n, ES-36310-Vigo, Galicia, Spain

ABSTRACT: In this paper, a new partitioning of the complex interaction energy is proposed. This new partitioning is based on the decomposition of the one-electron and exchange-correlation densities into unperturbed and deformation densities. Thus, the proposed energy fragmentation can be applied at the SCF level and post-SCF levels as long as the corresponding density matrices have been evaluated previously. It provides the typical description of the complex interaction as a summation of electrostatic, exchange-repulsion, and polarization terms. However, the new method allows splitting up the exchange-repulsion into exchange and Pauli-repulsion energies. A full theoretical description of the method is presented, and some examples of its application to small complexes are discussed. A comparison with results obtained using perturbation methods is also carried out, showing that the first order terms obtained from symmetry adapted perturbation theories are perfectly reproduced with the new method. A clear bridge between qualitative deformation density plots and quantitative measures of the interaction energy components can be established within the framework of this new partitioning scheme, giving rise to a graphical and very intuitive interpretation of the complex formation.

1. INTRODUCTION

The most employed method for the calculation of intermolecular energies is the supermolecular ansatz^{1–3} in which the interaction energy is calculated as the difference in the total energies of the complex and the monomers. Because of its simplicity, this approach is widely used; in fact, any ab initio or DFT program package can provide the complex and monomer energies, the latter corrected for the unwanted basis set superposition error.^{4,5} However, the information provided by the supermolecular approach about the intermolecular interaction is very poor, as it gives only the value of the intermolecular energy and does not give any physical insight into the nature of the interaction. This drawback can be solved using a perturbational treatment of intermolecular interactions. The perturbation theory has been applied to the problem of intermolecular forces since the early days of quantum chemistry, and it continues to provide the rational framework for their understanding, modeling, and computation.^{1,2,6} Thus, perturbation methods provide an interpretation of the potential energy surfaces obtained by the supermolecular approach^{1,4} and have been the keystone in the development of ab initio intermolecular force fields.^{1,7}

According to perturbation theory, all intermolecular interactions contain four fundamental physical contributions: electrostatic, exchange-repulsion, induction, and dispersion.^{1,2,6,7} The exchange-repulsion energy includes both the intermolecular electron–electron repulsion due to the Pauli exclusion principle and the exchange energy due to the exchange effects resulting from the tunneling of the electrons between interacting systems.^{1,8} The induction and dispersion energies together constitute the polarization term.^{1,2}

The conceptually simplest perturbation approach to the intermolecular interactions is the standard application of the Rayleigh–Schrödinger perturbation (RSPT) to a dimer, the so-called polarization theory.^{1,6,8,9} This approach is very successful in describing the interaction between two molecules a long distance apart, but at short range, it fails completely. Among

the perturbation theories that solve this problem, the symmetry-adapted perturbation approach of Jeziorski et al.⁶ (SAPT) has been the most used. Its popularity is mainly due to a lower computational cost in the calculation of the perturbational terms.

On the other hand, most of the perturbational approaches are based in molecular orbital theory, and they use the Heitler–London method to construct the initial wave function of the system. However, this wave function is not variationally optimized, which causes some ambiguity in the separation of the first order terms, namely, electrostatic and repulsion energies.^{10,11} This problem can be solved using a recently proposed method based on density functional theory (DFT), where the intermediate complex electron densities are obtained variationally.¹¹ Since it is based on DFT, the calculation of the energy terms requires only the knowledge of the one-electron densities. The method also provides the charge transfer contribution to the energy. However, this term is part of the short-range induction energy¹ and has been a source of controversy, giving rise to a good deal of confusion. Ayers et al.¹¹ separate the charge transfer contribution from the induction energy on the basis of a more or less arbitrary limit of the density of each fragment of the complex. Finally, the calculation of the dispersion contribution within this approach will be strongly influenced by the development of van der Waals functionals in density functional theory.

The aim of this work is to propose a new fragmentation of the intermolecular interaction energy based on the electron density that can be of general application. The energy fragmentation is performed by means of the unperturbed monomer electron densities and the intermolecular deformation densities associated with each energy term. Therefore, the method of calculation of the energy terms is given by the level of computation of the electron densities, so that the new partitioning may be considered of general application. As expected for any new

Received: November 8, 2010

Published: February 07, 2011

interaction energy partitioning, the method keeps a strong relation with other previously proposed schemes. So, it gives essentially the same first order contributions as the well-known SAPT method of Jeziorski et al.,⁶ and therefore they are directly comparable. However, the new scheme splits the short-range repulsive term into exchange and repulsion energies, giving a separate value to each of them. The second order terms are analogous to the second order SAPT contributions plus their exchange corrections, and a correspondence between them may also be expected.

If the method is considered from a practical point of view, one of its strongest points emerges: once the electron densities are obtained, the computation of all intermolecular terms takes virtually negligible time compared to the calculation of the electron densities. So the bottleneck for the calculations is the computation of the densities. Since most of quantum chemical program packages provide the electron density during the SCF or post-SCF procedures, this method will use in the most expensive step of the calculation any strategy, approach, or ansatz employed by the quantum chemical program packages to obtain the energy. This reduces significantly the computational cost, allowing the calculation of the different intermolecular energy terms in very large systems. Finally, density deformation plots associated with each intermolecular term can be obtained using this new scheme. These density plots may be a graphical and very intuitive tool to understand the nature of the different intermolecular terms and the delicate balance between them. In the next sections, we will expose the proposed method together with its application to some systems at different levels of calculation.

2. THEORETICAL DEVELOPMENTS

The total nonrelativistic energy of any molecular or atomic system can be exactly expressed in terms of the one-electron density, $\rho(\vec{r}_1)$, and the two-electron density, $\rho(\vec{r}_1, \vec{r}_2)$. Since the latter can be expressed as the sum of the product of the one-electron densities plus the exchange-correlation density, $\rho_{XC}(\vec{r}_1, \vec{r}_2)$, the expression of the energy is given by

$$E = -\frac{1}{2} \int \nabla^2 \rho(\vec{r}_1, \vec{r}_1')_{r_1' = r_1} d\vec{r}_1 + \int \hat{v}_N \rho(\vec{r}_1) d\vec{r}_1 + \frac{1}{2} \iint \frac{\rho(\vec{r}_1) \rho(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 + \frac{1}{2} \iint \frac{\rho_{XC}(\vec{r}_1, \vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 + \sum_{I=1}^N \sum_{J>I}^N \frac{Z_I Z_J}{|\vec{R}_I - \vec{R}_J|} \quad (1)$$

where \hat{v}_N represents the electrostatic potential created by the nuclei, N the number of nuclei, and Z the nuclear charge (in atomic units). In the case of two interacting molecular or atomic systems A and B , the nuclear electrostatic potential and energy, the one-electron density, and the exchange-correlation density can be expressed in terms of their values for the noninteracting systems plus their changes along the interaction. Thus, they can be written as eqs 2–345:

$$\sum_{I=1}^N \sum_{J>I}^N \frac{Z_I Z_J}{R_{IJ}} = \sum_{I=1}^{N_A} \sum_{J>I}^{N_A} \frac{Z_I Z_J}{R_{IJ}} + \sum_{I=1}^{N_B} \sum_{J>I}^{N_B} \frac{Z_I Z_J}{R_{IJ}} + \sum_{I=1}^{N_A} \sum_{J=1}^{N_B} \frac{Z_I Z_J}{|\vec{R}_I - \vec{R}_J|} \quad (2)$$

$$\hat{v}_N = \hat{v}_{N_A} + \hat{v}_{N_B} \quad (3)$$

$$\rho(\vec{r}_1) = \rho_A(\vec{r}_1) + \rho_B(\vec{r}_1) + \Delta\rho(\vec{r}_1) \quad (4)$$

$$\rho_{XC}(\vec{r}_1, \vec{r}_2) = \rho_{XC, A}(\vec{r}_1, \vec{r}_2) + \rho_{XC, B}(\vec{r}_1, \vec{r}_2) + \rho_{XC, AB}(\vec{r}_1, \vec{r}_2) + \Delta\rho_{XC}(\vec{r}_1, \vec{r}_2) \quad (5)$$

where $\Delta\rho$ is the one-electron deformation density. The third and fourth terms in eq 5 represent the intermolecular fraction of the exchange-correlation density before electron polarization and the changes experienced in the total exchange-correlation density due to the electron polarization, respectively.

The one-electron deformation density is in turn the result of two effects: on one hand, the effect over the monomers' one-electron densities of the Pauli exclusion principle, $\Delta\rho_{\text{Pauli}}$, which is given by the difference between the electron density obtained from the antisymmetrized product of the monomers' wave functions and the electron density obtained from the Hartree product; on the other hand, the effect over the one-electron density of the electron polarization due to the intermolecular interaction, $\Delta\rho_{\text{Pol}}$, which is given by the difference between the total and the Pauli deformation densities. Thus, eq 4 can be replaced by eq 6.

$$\rho(\vec{r}_1) = \rho_A(\vec{r}_1) + \rho_B(\vec{r}_1) + \Delta\rho_{\text{Pauli}}(\vec{r}_1) + \Delta\rho_{\text{Pol}}(\vec{r}_1) \quad (6)$$

By introducing eqs 2, 3, 5, and 6 in eq 1, the following expression is obtained for the energy of the interacting systems.

$$E = -\frac{1}{2} \int \nabla^2 \rho_A(\vec{r}_1, \vec{r}_1')_{r_1' = r_1} d\vec{r}_1 - \frac{1}{2} \int \nabla^2 \rho_B(\vec{r}_1, \vec{r}_1')_{r_1' = r_1} d\vec{r}_1 - \frac{1}{2} \int \nabla^2 \Delta\rho_{\text{Pauli}}(\vec{r}_1, \vec{r}_1')_{r_1' = r_1} d\vec{r}_1 + -\frac{1}{2} \int \nabla^2 \Delta\rho_{\text{Pol}}(\vec{r}_1, \vec{r}_1')_{r_1' = r_1} d\vec{r}_1 + \int \hat{v}_{N_A} \rho_A(\vec{r}_1) d\vec{r}_1 + \int \hat{v}_{N_B} \rho_B(\vec{r}_1) d\vec{r}_1 + \int \hat{v}_{N_A} \rho_B(\vec{r}_1) d\vec{r}_1 + \int \hat{v}_{N_B} \rho_A(\vec{r}_1) d\vec{r}_1 + \int \hat{v}_{N_A} \Delta\rho_{\text{Pauli}}(\vec{r}_1) d\vec{r}_1 + \int \hat{v}_{N_A} \Delta\rho_{\text{Pol}}(\vec{r}_1) d\vec{r}_1 + \int \hat{v}_{N_B} \Delta\rho_{\text{Pauli}}(\vec{r}_1) d\vec{r}_1 + \int \hat{v}_{N_B} \Delta\rho_{\text{Pol}}(\vec{r}_1) d\vec{r}_1 + \frac{1}{2} \iint \frac{\rho_A(\vec{r}_1) \rho_A(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 + \frac{1}{2} \iint \frac{\rho_B(\vec{r}_1) \rho_B(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 + \iint \frac{\rho_A(\vec{r}_1) \rho_B(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 + \frac{1}{2} \iint \frac{\Delta\rho_{\text{Pauli}}(\vec{r}_1) \Delta\rho_{\text{Pauli}}(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 + \frac{1}{2} \iint \frac{\Delta\rho_{\text{Pol}}(\vec{r}_1) \Delta\rho_{\text{Pol}}(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 + \iint \frac{\Delta\rho_{\text{Pauli}}(\vec{r}_1) \Delta\rho_{\text{Pol}}(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2$$

$$\begin{aligned}
& + \iint \frac{\Delta\rho_{\text{Pauli}}(\vec{r}_1)\rho_A(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 \\
& + \iint \frac{\Delta\rho_{\text{Pauli}}(\vec{r}_1)\rho_B(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 \\
& + \iint \frac{\Delta\rho_{\text{Pol}}(\vec{r}_1)\rho_A(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 \\
& + \iint \frac{\Delta\rho_{\text{Pol}}(\vec{r}_1)\rho_B(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 \\
& + \frac{1}{2} \iint \frac{\rho_{\text{XC},A}(\vec{r}_1, \vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 + \frac{1}{2} \iint \frac{\rho_{\text{XC},B}(\vec{r}_1, \vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 \\
& + \frac{1}{2} \iint \frac{\rho_{\text{XC},AB}(\vec{r}_1, \vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 \\
& + \frac{1}{2} \iint \frac{\Delta\rho_{\text{XC}}(\vec{r}_1, \vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 + \sum_{i=1}^{N_A} \sum_{j>i}^{N_A} \frac{Z_i Z_j}{|\vec{R}_i - \vec{R}_j|} \\
& + \sum_{i=1}^{N_B} \sum_{j>i}^{N_B} \frac{Z_i Z_j}{|\vec{R}_i - \vec{R}_j|} + \sum_{I=1}^{N_A} \sum_{j=1}^{N_B} \frac{Z_I Z_j}{|\vec{R}_I - \vec{R}_j|} \quad (7)
\end{aligned}$$

On the other hand, the summation of the energies of the isolated systems A and B is given by

$$\begin{aligned}
E_A + E_B = & -\frac{1}{2} \int \nabla^2 \rho_A(\vec{r}_1, \vec{r}_1')_{r_1'=r_1} d\vec{r}_1 \\
& -\frac{1}{2} \int \nabla^2 \rho_B(\vec{r}_1, \vec{r}_1')_{r_1'=r_1} d\vec{r}_1 \\
& + \int \hat{v}_{N_A} \rho_A(\vec{r}_1) d\vec{r}_1 + \int \hat{v}_{N_B} \rho_B(\vec{r}_1) d\vec{r}_1 \\
& + \frac{1}{2} \iint \frac{\rho_A(\vec{r}_1)\rho_A(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 \\
& + \frac{1}{2} \iint \frac{\rho_B(\vec{r}_1)\rho_B(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 \\
& + \frac{1}{2} \iint \frac{\rho_{\text{XC},A}(\vec{r}_1, \vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 \\
& + \frac{1}{2} \iint \frac{\rho_{\text{XC},B}(\vec{r}_1, \vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 \\
& + \sum_{i=1}^{N_A} \sum_{j>i}^{N_A} \frac{Z_i Z_j}{|\vec{R}_i - \vec{R}_j|} + \sum_{i=1}^{N_B} \sum_{j>i}^{N_B} \frac{Z_i Z_j}{|\vec{R}_i - \vec{R}_j|} \quad (8)
\end{aligned}$$

where all of the terms are contained explicitly within eq 7. The difference between eqs 7 and 8 is then the interaction energy, E_{int} , between A and B, expressed in terms of the deformation densities and unperturbed densities of the isolated systems. By defining the electrostatic potential created by the electrons and nuclei of molecules A and B (isolated) as

$$\hat{v}_A(\vec{r}_1) = \hat{v}_{N_A}(\vec{r}_1) + \int \frac{\rho_A(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_2 \quad (9)$$

$$\hat{v}_B(\vec{r}_1) = \hat{v}_{N_B}(\vec{r}_1) + \int \frac{\rho_B(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_2 \quad (10)$$

and grouping terms depending on the unperturbed densities and deformation densities, it is possible to separate E_{int} into different contributions with clear physical meaning (eqs 11–15).

$$E_{\text{int}} = E_{\text{elec}} + E_{\text{exch}} + E_{\text{rep}} + E_{\text{pol}} \quad (11)$$

$$\begin{aligned}
E_{\text{elec}} = & \int \hat{v}_{N_A} \rho_B(\vec{r}_1) d\vec{r}_1 + \int \hat{v}_{N_B} \rho_A(\vec{r}_1) d\vec{r}_1 \\
& + \iint \frac{\rho_A(\vec{r}_1)\rho_B(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 + \sum_{i=1}^{N_A} \sum_{j=1}^{N_B} \frac{Z_i Z_j}{|\vec{R}_i - \vec{R}_j|} \quad (12)
\end{aligned}$$

$$E_{\text{exch}} = \frac{1}{2} \iint \frac{\rho_{\text{XC},AB}(\vec{r}_1, \vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 \quad (13)$$

$$\begin{aligned}
E_{\text{rep}} = & \int \hat{v}_A \Delta\rho_{\text{Pauli}}(\vec{r}_1) d\vec{r}_1 + \int \hat{v}_B \Delta\rho_{\text{Pauli}}(\vec{r}_1) d\vec{r}_1 \\
& + \frac{1}{2} \iint \frac{\Delta\rho_{\text{Pauli}}(\vec{r}_1)\Delta\rho_{\text{Pauli}}(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 \\
& + \iint \frac{\Delta\rho_{\text{Pauli}}(\vec{r}_1)\rho_A(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 \\
& + \iint \frac{\Delta\rho_{\text{Pauli}}(\vec{r}_1)\rho_B(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 \\
& - \frac{1}{2} \int \nabla^2 \Delta\rho_{\text{Pauli}}(\vec{r}_1, \vec{r}_1')_{r_1'=r_1} d\vec{r}_1 \quad (14)
\end{aligned}$$

$$\begin{aligned}
E_{\text{pol}} = & \int \hat{v}_A \Delta\rho_{\text{Pol}}(\vec{r}_1) d\vec{r}_1 + \int \hat{v}_B \Delta\rho_{\text{Pol}}(\vec{r}_1) d\vec{r}_1 \\
& + \frac{1}{2} \iint \frac{\Delta\rho_{\text{Pol}}(\vec{r}_1)\Delta\rho_{\text{Pol}}(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 \\
& + \iint \frac{\Delta\rho_{\text{Pol}}(\vec{r}_1)\rho_A(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 \\
& + \iint \frac{\Delta\rho_{\text{Pol}}(\vec{r}_1)\rho_B(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 \\
& + \iint \frac{\Delta\rho_{\text{Pauli}}(\vec{r}_1)\Delta\rho_{\text{Pol}}(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 \\
& + -\frac{1}{2} \int \nabla^2 \Delta\rho_{\text{Pol}}(\vec{r}_1, \vec{r}_1')_{r_1'=r_1} d\vec{r}_1 \\
& + \frac{1}{2} \iint \frac{\Delta\rho_{\text{XC}}(\vec{r}_1, \vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 \quad (15)
\end{aligned}$$

The terms given above are classified as follows:

- (i) Electrostatic term, E_{elec} , which exclusively depends on the nuclear charge and one-electron density distributions of the isolated systems, accounting for the eventual presence of effective charges and multipole moments
- (ii) Exchange term, E_{exch} , depending exclusively on the unperturbed intermolecular exchange-correlation density
- (iii) Repulsion term, E_{rep} , depending on the deformation of the monomers' one-electron densities associated with the Pauli exclusion principle
- (iv) Polarization term, E_{pol} , depending on the deformation of the electron densities due to the electron polarization. In

Table 1. Comparison of Electrostatic and Exchange-Repulsion Energies for Water, Hydrogen Fluoride, and Methane Dimers Obtained at the SCF, MP2, and CCSD Levels Using the New Fragmentation Scheme and the SAPT Method (All Values in kJ mol^{-1})

	$(\text{H}_2\text{O})_2$			$(\text{HF})_2$			$(\text{CH}_4)_2$		
	SCF	MP2	CCSD	SCF	MP2	CCSD	SCF	MP2	CCSD
E_{elec}	−33.51	−33.39	−32.64	−26.40	−25.56	−25.19	−0.84	−1.00	−0.92
$E_{\text{elec}}(\text{SAPT})$	−33.51	−33.30	−32.68	−26.44	−25.44	−25.19	−0.84	−1.05	−0.96
diff (%)	0.02	0.36	0.14	0.03	0.57	0.00	3.10	2.25	3.93
$E_{\text{x-rep}}$	26.69	32.09	33.72	18.33	23.14	24.60	3.22	3.35	3.51
$E_{\text{x-rep}}(\text{SAPT})$	26.40	30.92	31.46	18.20	22.51	22.84	3.22	3.60	3.72
diff (%)	1.06	3.72	7.20	0.70	2.92	7.87	0.63	6.72	5.16

turn, the polarization energy can also be partitioned into one-electron density and exchange-correlation density terms

3. COMPUTATIONAL DETAILS

Implementation of eqs 11–15 has been performed in a program written in the Fortran language. In its current version, the program requires as input the information of the molecular orbital coefficients and the one-electron and two-electron integrals for basis functions. The Gaussian 03 program¹² provides the input information for the example calculations presented in this work. Then, the program computes the different density matrices, expressed in terms of basis functions, required for the calculation of the energy terms. The calculation of the exchange and repulsion energies is the most delicate step, and further details about their computation are given in and .

At the SCF level, the different energy terms that constitute the polarization energy are obtained from the one-electron and the exchange deformation densities. Since the total interaction energy and all its components are calculated analytically, differences between the interaction energy calculated using the supermolecular approach and that obtained as a summation of its energy components have been found to be negligible.

On the contrary, at the post-SCF level, the exchange-correlation energy term of the polarization energy is obtained as the difference between the total interaction energy and the summation of the remaining energy components. This shortcut avoids the arduous task of computing the two-electron reduced density matrices at the post-SCF level and relies on the accuracy of the calculation of the electrostatic and exchange-repulsion energies (see next section for details).

We have checked our method in three typical complexes such as water, hydrogen fluoride, and methane dimers. The 6-311++G-(2d,2p) basis set was employed in the calculations of water and hydrogen fluoride dimers. Since a good description of the methane dimer potential energy surface requires a larger basis set, we have employed for this system the aug-cc-pVQZ basis set. Energy decomposition was performed using Hartree–Fock (HF), MP2, MP4, CCD, CCSD, and QCISD electron densities.

4. EXAMPLES

Comparison with Perturbation Approaches. Although the new fragmentation can be applied on any kind of atomic or molecular interaction (covalent, ionic, van der Waals, etc.), the most interesting application is, undoubtedly, the study of weak intermolecular interactions. Perturbation and hybrid approaches

provide a solid background for the physical interpretation of the interaction energy in intermolecular complexes. Terms arising from perturbation approaches can be classified as electrostatic, exchange, and polarization, where the latter can be split up into pure induction and dispersion energies at the second-order perturbation level.

In this section, we will merge as far as possible the energy terms of eqs 11–15 with those arising from the perturbation approaches. We will start with the simplest case of SCF electron densities to further extend it to the more general case of correlated post-SCF densities.

The physical interpretation of the energy terms given in eq 11 becomes clear when the electron densities are obtained at the HF level. The electrostatic term is the classical intermolecular Coulombic interaction between electrons and nuclei of different monomers. As long as we deal with Hartree–Fock electron densities, the intramonomer correlation effects are not taken into account (in this section, we will employ the term “uncorrected” for those energy terms that do not include electron correlation corrections).

$$E_{\text{elec}}^{\text{unc}} = \int \hat{v}_{N_A} \rho_B^{\text{HF}}(\vec{r}_1) d\vec{r}_1 + \int \hat{v}_{N_B} \rho_A^{\text{HF}}(\vec{r}_1) d\vec{r}_1 + \iint \frac{\rho_A^{\text{HF}}(\vec{r}_1) \rho_B^{\text{HF}}(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 + \sum_{I=1}^{N_A} \sum_{J=1}^{N_B} \frac{Z_I Z_J}{|\vec{R}_I - \vec{R}_J|} \quad (16)$$

This is exactly the same energy as the electrostatic energy obtained with Morokuma’s method, and also the same as the electrostatic energy obtained from the hybrid methods that employ the Hartree–Fock interaction energy. The Hartree–Fock electrostatic energies calculated using the new fragmentation and the SAPT-HF method are shown in Table 1 for water, hydrogen fluoride, and methane dimers. As one can see, differences in the values obtained by both methodologies are negligible.

At the HF level, the exchange-correlation density contains exclusively the electron exchange part. Thus, $E_{\text{exch}}^{\text{unc}}$ gives exactly the exchange part of the uncorrected exchange-repulsion energy.

$$E_{\text{exch}}^{\text{unc}} = \frac{1}{2} \iint \frac{\rho_{X,AB}^{\text{HF}}(\vec{r}_1, \vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 \quad (17)$$

This is one important advantage of using electron densities to represent the interaction energy, it is possible to separate exchange energy from repulsion, which is not possible using

perturbation methods. The Hartree–Fock exchange and repulsion energy terms obtained in this work are also compared in Table 1 with the exchange-repulsion energies obtained using SAPT-HF. As in the case of electrostatic energies, the summation of the exchange plus repulsion terms gives the same results as the SAPT-HF method for the exchange-repulsion energy. As expected for interactions between closed-shell systems, the exchange part is negative, whereas the repulsion part is positive and larger in absolute value. The intermolecular exchange energy stabilizes the complex, but the Pauli repulsion exerted by electrons with the same spin in the intermolecular region overtakes the exchange energy, giving rise to a net destabilization. Details about the calculation of $\rho_{X,AB}^{\text{HF}}(\vec{r}_1, \vec{r}_2)$, and $\Delta\rho_{\text{Pauli}}^{\text{HF}}(\vec{r}_1)$ are given in .

Since the dispersion energy is not accounted for at the HF level, the polarization term (eq 15) gives directly the uncorrected induction energy. However, in this case the induction energy also reflects the exchange correction that arises in the perturbation methods.

$$\begin{aligned}
 E_{\text{ind}}^{\text{unc}} = & \int \hat{v}_A \Delta\rho_{\text{Pol}}^{\text{HF}}(\vec{r}_1) d\vec{r}_1 + \int \hat{v}_B \Delta\rho_{\text{Pol}}^{\text{HF}}(\vec{r}_1) d\vec{r}_1 \\
 & + \frac{1}{2} \iint \frac{\Delta\rho_{\text{Pol}}^{\text{HF}}(\vec{r}_1) \Delta\rho_{\text{Pol}}^{\text{HF}}(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 \\
 & + \iint \frac{\Delta\rho_{\text{Pol}}^{\text{HF}}(\vec{r}_1) \rho_A^{\text{HF}}(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 \\
 & + \iint \frac{\Delta\rho_{\text{Pol}}^{\text{HF}}(\vec{r}_1) \rho_B^{\text{HF}}(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 \\
 & + \iint \frac{\Delta\rho_{\text{Pauli}}^{\text{HF}}(\vec{r}_1) \Delta\rho_{\text{Pauli}}^{\text{HF}}(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 \\
 & + \frac{1}{2} \int \nabla^2 \Delta\rho_{\text{Pol}}^{\text{HF}}(\vec{r}_1, \vec{r}_1')_{r_1'=r_1} d\vec{r}_1 \\
 & + \frac{1}{2} \iint \frac{\Delta\rho_X^{\text{HF}}(\vec{r}_1, \vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 \quad (18)
 \end{aligned}$$

Summarizing, all energy terms obtained with the new fragmentation scheme have physical interpretation in terms of traditional electrostatic, exchange-repulsion, and induction energies derived from perturbation theory at the HF level.

The use of correlated post-SCF densities incorporates the dynamic electron correlation effect into the interaction energy. This evidently improves the results but complicates the physical interpretation of the energy fragmentation. The dispersion energy, which is intrinsically related to the dynamic electron correlation, appears to be unavoidably distributed within the polarization term. On the contrary, the electrostatic term still has a clear-cut physical meaning as the inclusion of electron correlation does not change its nature but incorporates the effect of the intramonomer dynamic electron correlation. The electrostatic term calculated at the correlated level is now the sum of the Hartree–Fock electrostatic energy plus the intramonomer correlation correction to the electrostatic energy. This term can be viewed as the “corrected” electrostatic energy.

$$\begin{aligned}
 E_{\text{elec}}^{\text{corr}} = & \int \hat{v}_{N_A} \rho_B(\vec{r}_1) d\vec{r}_1 + \int \hat{v}_{N_B} \rho_A(\vec{r}_1) d\vec{r}_1 \\
 & + \iint \frac{\rho_A(\vec{r}_1) \rho_B(\vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 + \sum_{I=1}^{N_A} \sum_{J=1}^{N_B} \frac{Z_I Z_J}{|\vec{R}_I - \vec{R}_J|} \quad (19)
 \end{aligned}$$

The electrostatic energies obtained at different post-SCF levels are compared in Table 1 with the electrostatic energies obtained with SAPT at the MP2 and CCSD levels for the reference dimers. As one can see, the MP2 and CCSD electrostatic energies obtained with the new fragmentation scheme are almost the same as those calculated with SAPT.

On the other hand, the corrected exchange energy can also be obtained straightforwardly from the unperturbed intermolecular exchange-correlation density (see for its calculation). This can be viewed as the unperturbed intermolecular exchange density corrected by the intramonomer electron correlation.

$$E_{\text{exch}}^{\text{corr}} = \frac{1}{2} \iint \frac{\rho_{X,AB}(\vec{r}_1, \vec{r}_2)}{|\vec{r}_2 - \vec{r}_1|} d\vec{r}_1 d\vec{r}_2 \quad (20)$$

The Pauli repulsion can also be obtained separately as in the case of Hartree–Fock densities. As mentioned in , the best way to compute the deformation density associated with the Pauli repulsion is through a weighted orthogonalization procedure¹³ of the natural orbitals of the monomers. One can compare in Table 1 the summation of the exchange plus Pauli repulsion energies obtained from the new fragmentation and the exchange-repulsion energies calculated with SAPT at the MP2 and CCSD levels. Results obtained by both methodologies are very similar, demonstrating the good performance of the weighted orthogonalization in the calculation of the exchange-repulsion energy. Analyzing the exchange and the Pauli repulsion energies separately, it can be observed that their absolute values go up as the degree of electron correlation increases, following the sequence HF < MP2 < CCD < CCSD \approx QCISD \approx MP4, which reflects an increase of the overlapping between the monomers’ electron densities with the electron correlation.

The disadvantage of using correlated densities in the new energy fragmentation is the apparent nonseparability of the polarization term. Thus, one can only speak “a priori” of polarization energy, including here the induction and dispersion and their exchange-correlation corrections. Separation of polarization energy into induction and dispersion seems to have only physical meaning within the second-order perturbation theory. In fact, for higher order perturbation theories, the induction and dispersion terms are not completely separable, and coupled induction-dispersion energies appear already at third-order perturbation theory.¹⁴ We are currently developing a procedure for merging the second-order RSPT with the new fragmentation scheme in order to get the induction energy from the rest of polarization energy at the post-SCF level.¹⁵

Energy Fragmentation vs Deformation Density Plots. Together with other local properties of the electron density,¹⁶ deformation density plots can be employed to rationalize at the qualitative level the stability of hydrogen bonding,¹⁷ stacking,¹⁸ or anion- π complexes.¹⁹ These deformation density plots are also frequently employed to evaluate the strength of chemical bonds.^{20,21} In the case of intermolecular interactions, deformation density plots are more properly obtained as the difference between the one-electron density of the complex and the summation of the one-electron densities of the noninteracting monomers. In the calculation, the monomers are placed at the same position that they occupy in the complex. This deformation density includes the effect that all of the interaction energy terms exert on the electron density distribution upon complexation. Thus, the effects of

Table 2. Interaction Energy Components for Water, Hydrogen Fluoride and Methane Dimers Obtained at SCF and Different Post-SCF Levels Using the New Fragmentation Scheme (All Values in kJ mol^{-1})

	SCF	MP2	CCD	CCSD	QCISD	MP4
(H_2O) ₂						
E_{elec}	-33.51	-33.39	-32.72	-32.64	-32.59	-32.68
E_{x}	-44.89	-60.08	-56.99	-57.57	-57.82	-57.91
E_{rep}	71.59	92.17	89.66	91.29	92.22	91.67
$E_{\text{x-rep}}$	26.69	32.09	32.64	33.72	34.35	33.76
E_{pol}	-8.87	-17.57	-17.91	-19.12	-19.71	-19.08
	-19.41 ^a	-39.12 ^a	-39.20 ^a	-40.96 ^a	-41.84 ^a	-41.09 ^a
	10.54 ^b	21.55 ^b	21.38 ^b	21.84 ^b	22.13 ^b	22.01 ^b
E_{int}	-15.73	-18.91	-17.82	-17.99	-17.95	-17.95
(HF) ₂						
E_{elec}	-26.40	-25.56	-25.27	-25.19	-25.15	-25.19
E_{x}	-33.39	-48.24	-44.89	-46.65	-47.15	-47.20
E_{rep}	51.71	71.38	68.20	71.30	72.68	72.22
$E_{\text{x-rep}}$	18.33	23.14	23.35	24.64	25.52	25.02
E_{pol}	-7.91	-14.10	-14.27	-15.69	-16.57	-15.94
	-8.95 ^a	-28.03 ^a	-26.61 ^a	-29.71 ^a	-31.05 ^a	-30.33 ^a
	1.05 ^b	13.93 ^b	12.34 ^b	14.02 ^b	14.56 ^b	14.39 ^b
E_{int}	-15.98	-16.48	-16.19	-16.23	-16.11	-16.07
(CH_4) ₂						
E_{elec}	-0.84	-1.00	-0.92	-0.92	-0.92	-0.96
E_{x}	-7.82	-12.76	-12.30	-12.43	-12.47	-12.51
E_{rep}	11.09	16.11	15.65	15.94	16.02	15.94
$E_{\text{x-rep}}$	3.22	3.35	3.35	3.51	3.56	3.43
E_{pol}	-0.17	-3.56	-3.18	-3.39	-3.43	-3.35
	-2.13 ^a	-7.57 ^a	-7.57 ^a	-7.91 ^a	-7.99 ^a	-7.78 ^a
	1.92 ^b	4.02 ^b	4.39 ^b	4.48 ^b	4.52 ^b	4.44 ^b
E_{int}	2.18	-1.17	-0.75	-0.84	-0.84	-0.88

^a One-electron density contribution to polarization. ^b Exchange-correlation density contribution to polarization.

exchange-repulsion and polarization cannot be separated a priori. This entails some problems for the analysis of dispersive interactions.

As an illustrative example, it is well-known that dispersion plays a key role in the stabilization of methane complexes.^{22–24} So, the Hartree–Fock interaction energy for the dimer is $0.52 \text{ kcal mol}^{-1}$, whereas the CCSD interaction energy is $-0.20 \text{ kcal mol}^{-1}$ according to the values collected in Table 2. Looking at the deformation density of Figure 1, one can see that the differences between HF and CCSD plots are negligible, and an explanation of the stabilization of the methane dimer from its deformation energy is not possible, even on a qualitative level.

On the other hand, we can confirm that the stabilization of the methane dimer at the post-SCF level mainly stems from the polarization energy, which is almost zero at the HF level (see values in Table 2). In the new energy partitioning, the deformation density can be separated into deformation density due to Pauli repulsion and deformation density due to electron polarization. Both contributions are represented separately in Figure 2 for the three reference complexes at the HF and CCSD levels. It is clear from these plots that the difference in

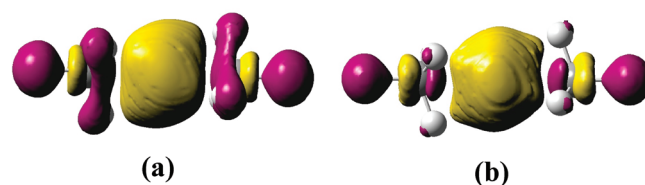


Figure 1. Total deformation density plots of the methane dimer obtained at the SCF (a) and CCSD (b) levels. Isosurface value is $4 \times 10^{-5} \text{ au}$.

the polarization energy obtained at the post-SCF level, which is associated with dispersion, is also reflected on the deformation density in a larger accumulation of electron density in the intermolecular region.

The effect of dispersion on water and hydrogen fluoride dimers is also significant. Thus, the CCSD polarization energies are around twice the Hartree–Fock energies. However, its effect on the deformation density is negligible for these complexes (see Figure 2). This is due to the large permanent electric dipoles of water and hydrogen fluoride that induce additional large dipoles, which are sufficiently well described even at the HF level. These large induced dipoles overshadow any other possible perturbations of the complex electron density.

It is also interesting to look at the different contributions to the polarization energy separately (see Table 2). Thus, the net complex stabilization due to the electron polarization comes exclusively from the change in the electron–electron and electron–nuclei Coulombic interactions (deformation of the one-electron density), whereas the change in the exchange–correlation energy always results in a destabilization of the complex. Use of correlated methods stresses this trend.

On the other hand, as mentioned before, the Pauli repulsion energy increases at the post-SCF level due to the larger overlapping between the monomers' electron densities. This can also be deduced from the deformation density plots of Figure 2 (see the Pauli repulsion contribution), which reflect a larger depletion of electron density in the intermolecular region at the post-SCF level.

5. CONCLUDING REMARKS AND FUTURE DEVELOPMENTS

An exact partitioning of the intermolecular interaction energy using one-electron and exchange–correlation densities and their deformations upon the complexation process has been presented. The methodology, which can be applied at the SCF and any post-SCF computational level, allows partitioning the interaction energy of an atomic or molecular pair into electrostatic, exchange, repulsion, and polarization terms, without significant computational cost. The slowest computational step is the previous calculation of the corresponding electron density matrices of the unperturbed systems and the complex, which can be done using any quantum chemical program package.

The particular case of weak intermolecular interactions has been treated in detail by comparison with the energy partitioning arising from perturbation methods. Uncorrelated and correlated densities have been analyzed separately, finding that the electrostatic, exchange–repulsion, and induction energy terms are identified with those obtained from the new fragmentation at the SCF and post-SCF levels. We are currently implementing a procedure to separate the induction energy from the dispersion

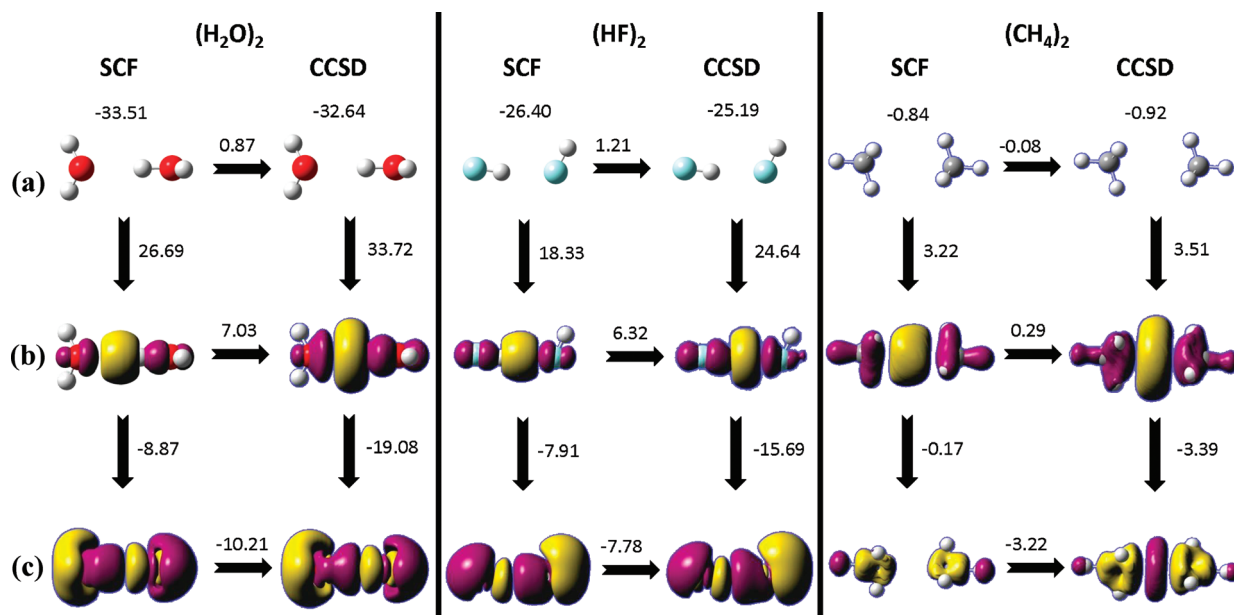


Figure 2. Deformation densities together with the energies (in kJ mol⁻¹) associated with the electrostatic (a), exchange-repulsion (b), and polarization (c) terms for water, hydrogen fluoride, and methane dimers. SCF and CCSD results are shown, and differences between them are indicated on the corresponding horizontal arrow. Isosurface values are 2×10^{-4} au for HF and H₂O and 4×10^{-5} au for CH₄. All energy values are in kJ mol⁻¹.

energy by merging the second-order RSPT with the new fragmentation scheme.

A bridge between qualitative deformation density plots and quantitative measures of the interaction energy components can be established within the framework of this new partitioning scheme, providing a graphical and very intuitive interpretation of the complex formation.

We are currently extending the method to the case of more than two molecules, which will provide insight into the nature of cooperative effects.

APPENDIX A. CALCULATION OF $\rho_{X,AB}^{\text{HF}}(\vec{r}_1, \vec{r}_2)$ AND $\Delta\rho_{\text{PAULI}}^{\text{HF}}(\vec{r})$

At the HF level, the spin-free two-electron density can be obtained by eq 21:

$$\rho^{\text{HF}}(\vec{r}_1, \vec{r}_2) = \rho^{\text{HF}}(\vec{r}_1)\rho^{\text{HF}}(\vec{r}_2) + \rho_X^{\text{HF}}(\vec{r}_1, \vec{r}_2) \quad (21)$$

where the electron exchange density, $\rho_X^{\text{HF}}(\vec{r}_1, \vec{r}_2)$, represents the difference between the nonconditioned and conditioned probability of finding an electron in r_1 when another electron of the same spin is found in r_2 . In other words, it accounts for the static correlation of the movement of an electron due to the presence of a second electron of the same spin imposed by the Pauli's exclusion principle.

The electron exchange density can be expanded in the spin orbital basis, which for the case of restricted Hartree–Fock calculations and orthogonal basis adopts the form of eq 22:

$$\rho_X^{\text{HF}}(\vec{r}_1, \vec{r}_2) = -2 \sum_{i=1}^{\text{NOO}} \sum_{j=1}^{\text{NOO}} \phi_i(\vec{r}_1) \phi_j(\vec{r}_2) \phi_j(\vec{r}_1) \phi_i(\vec{r}_2) \quad (22)$$

where NOO refers to the number of occupied orbitals. In turn, when the orbitals are constructed as a linear combination of a set

of basis functions, $\chi_{\mu\nu}$ the electron exchange density can be written as

$$\begin{aligned} & \rho_X^{\text{HF}}(\vec{r}_1, \vec{r}_2) \\ &= -2 \sum_{\mu=1}^{\text{NBF}} \sum_{\nu=1}^{\text{NBF}} \sum_{\lambda=1}^{\text{NBF}} \sum_{\sigma=1}^{\text{NBF}} D_{\mu\nu} D_{\lambda\sigma} \chi_{\mu}(\vec{r}_1) \chi_{\lambda}(\vec{r}_2) \chi_{\nu}(\vec{r}_1) \chi_{\sigma}(\vec{r}_2) \end{aligned} \quad (23)$$

where D represents the 1-electron reduced density matrix (1-RDM). The matrix elements of the 1-RDM are given by:

$$D_{\mu\nu} = \sum_{i=1}^{\text{NOO}} C_{\mu, i} C_{\nu, i} \quad (24)$$

For the particular case of an unrelaxed intermolecular complex, the electron exchange density can be expressed in terms of the monomers' molecular orbitals and partitioned into three different terms:

$$\begin{aligned} & \rho_X^{\text{HF}}(\vec{r}_1, \vec{r}_2) \\ &= -2 \sum_{i=1}^{\text{NOO}_A} \sum_{j=1}^{\text{NOO}_A} \phi_i(\vec{r}_1) \phi_j(\vec{r}_2) \phi_j(\vec{r}_1) \phi_i(\vec{r}_2) - 2 \\ & \sum_{i=1}^{\text{NOO}_B} \sum_{j=1}^{\text{NOO}_B} \phi_i(\vec{r}_1) \phi_j(\vec{r}_2) \phi_j(\vec{r}_1) \phi_i(\vec{r}_2) + \rho_{X,AB}^{\text{HF}}(\vec{r}_1, \vec{r}_2) \end{aligned} \quad (25)$$

The first two terms are the Hartree–Fock intramonomer electron exchange densities (the first two terms in eq 5), whereas the third term is the Hartree–Fock intermolecular electron exchange density. The latter results from the antisymmetrization of the product of unrelaxed wave functions of the monomers and represents the part of the two-electron density due to the electron exchange between different monomers. The calculation of this term from the unrelaxed

orbitals of the monomers has the inconvenience of these two sets of orbitals being nonorthogonal to each other. It is then necessary first to orthogonalize these orbitals in order to get an expression for the electron exchange density equivalent to eq 22. We have employed in this work the Lowdin's *symmetric orthogonalization* procedure.²⁵ Thus, denoting the new set of orthogonal orbitals with the superscript "oo", the electron exchange density can be written as

$$\begin{aligned} & \rho_X^{\text{HF}}(\vec{r}_1, \vec{r}_2) \\ &= -2 \sum_{i=1}^{\text{NOO}} \sum_{j=1}^{\text{NOO}} \phi_i^{\text{oo}}(\vec{r}_1) \phi_j^{\text{oo}}(\vec{r}_2) \phi_j^{\text{oo}}(\vec{r}_1) \phi_i^{\text{oo}}(\vec{r}_2) \end{aligned} \quad (26)$$

and the intermolecular electron exchange density can be obtained by merging eqs 25 and 26:

$$\begin{aligned} \rho_{X,AB}^{\text{HF}}(\vec{r}_1, \vec{r}_2) &= -2 \sum_{i=1}^{\text{NOO}} \sum_{j=1}^{\text{NOO}} \phi_i^{\text{oo}}(\vec{r}_1) \phi_j^{\text{oo}}(\vec{r}_2) \phi_j^{\text{oo}}(\vec{r}_1) \phi_i^{\text{oo}}(\vec{r}_2) \\ &+ 2 \sum_{i=1}^{\text{NOO}_A} \sum_{j=1}^{\text{NOO}_A} \phi_i(\vec{r}_1) \phi_j(\vec{r}_2) \phi_j(\vec{r}_1) \phi_i(\vec{r}_2) \\ &+ 2 \sum_{i=1}^{\text{NOO}_B} \sum_{j=1}^{\text{NOO}_B} \phi_i(\vec{r}_1) \phi_j(\vec{r}_2) \phi_j(\vec{r}_1) \phi_i(\vec{r}_2) \\ &+ 2 \sum_{i=1}^{\text{NOO}_A} \sum_{j=1}^{\text{NOO}_A} \phi_i(\vec{r}_1) \phi_j(\vec{r}_2) \phi_j(\vec{r}_1) \phi_i(\vec{r}_2) \\ &+ 2 \sum_{i=1}^{\text{NOO}_B} \sum_{j=1}^{\text{NOO}_B} \phi_i(\vec{r}_1) \phi_j(\vec{r}_2) \phi_j(\vec{r}_1) \phi_i(\vec{r}_2) \end{aligned} \quad (27)$$

Using basis functions, the intermolecular electron exchange density is then given by

$$\begin{aligned} \rho_X^{\text{HF}}(\vec{r}_1, \vec{r}_2) &= -2 \sum_{\mu=1}^{\text{NBF}} \sum_{\nu=1}^{\text{NBF}} \sum_{\lambda=1}^{\text{NBF}} \sum_{\sigma=1}^{\text{NBF}} (D_{\mu\nu}^{\text{oo}} D_{\lambda\sigma}^{\text{oo}} - D_{\mu\nu}^A D_{\lambda\sigma}^A \\ &- D_{\mu\nu}^B D_{\lambda\sigma}^B) \chi_{\mu}(\vec{r}_1) \chi_{\lambda}(\vec{r}_2) \chi_{\nu}(\vec{r}_1) \chi_{\sigma}(\vec{r}_2) \end{aligned} \quad (28)$$

where the superindex in the 1-RDMs denotes the set of orbitals employed for its calculation and NBF refers to the number of basis functions.

From the orbital orthogonalization proposed above, one can also obtain, in terms of the 1-RDMs, the one-electron deformation density associated to the Pauli repulsion.

$$\Delta \rho_{\text{Pauli}}^{\text{HF}}(\vec{r}) = 2 \sum_{\mu=1}^{\text{NBF}} \sum_{\nu=1}^{\text{NBF}} (D_{\mu\nu}^{\text{oo}} - D_{\mu\nu}^A - D_{\mu\nu}^B) \chi_{\mu}(\vec{r}) \chi_{\nu}(\vec{r}) \quad (29)$$

APPENDIX B. CALCULATION OF $\rho_{X,C,AB}(\vec{R}_1, \vec{R}_2)$ AND $\Delta \rho_{\text{PAULI}}(\vec{R})$ AT THE POST-SCF LEVEL

Similarly to eq 21, the exchange-correlation density for a closed-shell electron system can be written in general as

$$\rho_{XC}(\vec{r}_1, \vec{r}_2) = \rho(\vec{r}_1, \vec{r}_2) - \rho(\vec{r}_1) \rho(\vec{r}_2) \quad (30)$$

In turn, the two-electron density, $\rho(\vec{r}_1, \vec{r}_2)$, can be written in terms of the two-electron reduced density matrix (2-RDM) in the space of the basis functions:

$$\rho(\vec{r}_1, \vec{r}_2) = 2 \sum_{\mu=1}^{\text{NBF}} \sum_{\nu=1}^{\text{NBF}} \sum_{\lambda=1}^{\text{NBF}} \sum_{\sigma=1}^{\text{NBF}} D_{\mu\nu\lambda\sigma} \chi_{\mu}(\vec{r}_1) \chi_{\lambda}(\vec{r}_2) \chi_{\nu}(\vec{r}_1) \chi_{\sigma}(\vec{r}_2) \quad (31)$$

where **D** represents here the 2-RDM.

For the strict calculation of $\rho_{X,C,AB}(\vec{r}_1, \vec{r}_2)$, it is then necessary to compute the matrix elements of the 2-RDMs for the isolated monomers and for the unrelaxed complex. This is not a simple task, especially for large systems. As an alternative to the computation of the 2-RDMs, one can approximate the exchange-correlation densities in terms of natural orbitals and their occupation numbers. In this work, we have chosen the approximate form of the exchange-correlation density derived from the Pauli exclusion principle,²⁶ which was found to provide accurate results in atomic energy calculations.²⁷

$$\rho_{XC}(\vec{r}_1, \vec{r}_2) = -2 \sum_{i=1}^{\text{NNO}} \sum_{j=1}^{\text{NNO}} n_i^{1/2} n_j^{1/2} \phi_i(\vec{r}_1) \phi_j(\vec{r}_2) \phi_j(\vec{r}_1) \phi_i(\vec{r}_2) \quad (32)$$

In eq 32, n_i and n_j denote the occupation numbers of the orbitals i and j and NNO refers to the number of natural orbitals.

As in Hartree–Fock, the electron exchange-correlation density for the unrelaxed complex can be expressed in terms of the monomers' natural orbitals and partitioned into three different terms using the expression 32.

$$\begin{aligned} \rho_{XC}(\vec{r}_1, \vec{r}_2) &= -2 \sum_{i=1}^{\text{NNO}_A} \sum_{j=1}^{\text{NNO}_A} n_i^{1/2} n_j^{1/2} \phi_i(\vec{r}_1) \phi_j(\vec{r}_2) \phi_j(\vec{r}_1) \phi_i(\vec{r}_2) \\ &- 2 \sum_{i=1}^{\text{NNO}_B} \sum_{j=1}^{\text{NNO}_B} n_i^{1/2} n_j^{1/2} \phi_i(\vec{r}_1) \phi_j(\vec{r}_2) \phi_j(\vec{r}_1) \phi_i(\vec{r}_2) \\ &+ \rho_{X,C,AB}(\vec{r}_1, \vec{r}_2) \end{aligned} \quad (33)$$

In eq 33, the first two terms are the intramonomer electron exchange-correlation densities, and the last term is the intermolecular electron exchange-correlation density. As in Hartree–Fock, the intermolecular term results from the antisymmetrization of the wave function product of the unperturbed monomers and represents the part of the 2-electron density due to the electron exchange between different monomers. However, one must follow a different orthogonalization procedure to that of Hartree–Fock for the calculation of the intermolecular term from the natural orbitals of the monomers. Since the weight of each orbital is given by its occupation number (its square root in this case), one can use a weighted orthogonalization procedure¹³ to get a new set of orthogonalized orbitals. The weighted orthogonalization procedure followed in this work resembles that of Weinhold and co-workers²⁸ and also requires an additional diagonalization of the occupation matrix to get the final set of natural orbitals.²⁹ After the orthogonalization, one gets a transformation matrix that transforms both the orbitals and their occupation numbers. Thus, the exchange-correlation density can be expressed in terms of the new orthogonalized orbitals:

$$\begin{aligned} \rho_{XC}(\vec{r}_1, \vec{r}_2) &= -2 \sum_{i=1}^{\text{NNO}} \sum_{j=1}^{\text{NNO}} (n_i^{\text{oo}})^{1/2} (n_j^{\text{oo}})^{1/2} \phi_i^{\text{oo}}(\vec{r}_1) \phi_j^{\text{oo}}(\vec{r}_2) \phi_j^{\text{oo}}(\vec{r}_1) \phi_i^{\text{oo}}(\vec{r}_2) \end{aligned} \quad (34)$$

where, as in the previous appendix, they are denoted by the superscript “oo”. The intermolecular part is then given by

$$\begin{aligned} & \rho_{XC, AB}(\vec{r}_1, \vec{r}_2) \\ &= -2 \sum_{i=1}^{NNO} \sum_{j=1}^{NNO} (n_i^{oo})^{1/2} (n_j^{oo})^{1/2} \phi_i^{oo}(\vec{r}_1) \phi_j^{oo}(\vec{r}_2) \phi_j^{oo}(\vec{r}_1) \phi_i^{oo}(\vec{r}_2) \\ &+ 2 \sum_{i=1}^{NNO_A} \sum_{j=1}^{NNO_A} n_i^{1/2} n_j^{1/2} \phi_i(\vec{r}_1) \phi_j(\vec{r}_2) \phi_j(\vec{r}_1) \phi_i(\vec{r}_2) \\ &+ 2 \sum_{i=1}^{NNO_B} \sum_{j=1}^{NNO_B} n_i^{1/2} n_j^{1/2} \phi_i(\vec{r}_1) \phi_j(\vec{r}_2) \phi_j(\vec{r}_1) \phi_i(\vec{r}_2) \end{aligned} \quad (35)$$

In practice, eq 35 is calculated in terms of basis functions:

$$\begin{aligned} \rho_X^{HF}(\vec{r}_1, \vec{r}_2) &= -2 \sum_{\mu=1}^{NBF} \sum_{\nu=1}^{NBF} \sum_{\lambda=1}^{NBF} \sum_{\sigma=1}^{NBF} (1/2 D_{\mu\nu}^{oo} 1/2 D_{\lambda\sigma}^{oo} - 1/2 D_{\mu\nu}^A 1/2 D_{\lambda\sigma}^A - 1/2 D_{\mu\nu}^B 1/2 D_{\lambda\sigma}^B) \\ & \chi_{\mu}(\vec{r}_1) \chi_{\lambda}(\vec{r}_2) \chi_{\nu}(\vec{r}_1) \chi_{\sigma}(\vec{r}_2) \end{aligned} \quad (36)$$

where in general the matrix elements of $1/2\mathbf{D}$ are given by

$$1/2 D_{\mu\nu} = \sum_{i=1}^{NNO} n_i^{1/2} C_{\mu, i} C_{\nu, i} \quad (37)$$

and $C_{\mu, i}$ represents the coefficient of the basis function μ in the orbital i .

Similar to Hartree–Fock, the post-HF deformation density associated with the Pauli repulsion can be obtained from the weighted orthogonalization procedure as

$$\Delta\rho_{\text{Pauli}}(\vec{r}) = 2 \sum_{\mu=1}^{NBF} \sum_{\nu=1}^{NBF} (D_{\mu\nu}^{oo} - D_{\mu\nu}^A - D_{\mu\nu}^B) \chi_{\mu}(\vec{r}) \chi_{\nu}(\vec{r}) \quad (38)$$

where in general the density matrix elements contain also the corresponding occupation numbers.

$$D_{\mu\nu} = \sum_{i=1}^{NNO} n_i C_{\mu, i} C_{\nu, i} \quad (39)$$

AUTHOR INFORMATION

Corresponding Author

*E-mail: mandado@uvigo.es.

REFERENCES

- (1) Stone, A. J. *The theory of intermolecular forces*; Clarendon Press; Oxford University Press: Oxford, U. K., 1997.
- (2) Kaplan, I. G. *Intermolecular interactions: physical picture, computational methods, model potentials*; John Wiley & Sons: Chichester, England, 2006.
- (3) Chałasiński, G.; Szczeniński, M. M. *Chem. Rev.* **2000**, *100*, 4227.
- (4) Chałasiński, G.; Szczeniński, M. M. *Chem. Rev.* **1994**, *94*, 1723.
- (5) Van Duijneveldt, F. B.; Van Duijneveldt-Van De Rijdt, J. G. C. M.; Van Lenthe, J. H. *Chem. Rev.* **1994**, *94*, 1873.
- (6) Jeziorski, B.; Moszynski, R.; Szalewicz, K. *Chem. Rev.* **1994**, *94*, 1887.
- (7) Engkvist, O.; Åstrand, P. O.; Karlström, G. *Chem. Rev.* **2000**, *100*, 4087.
- (8) Hirschfelder, J. O. *Chem. Phys. Lett.* **1967**, *1*, 325.
- (9) Hirschfelder, J. O. *Chem. Phys. Lett.* **1967**, *1*, 363.

(10) Khaliullin, R. Z.; Cobar, E. A.; Lochan, R. C.; Bell, A. T.; Head-Gordon, M. *J. Phys. Chem. A* **2007**, *111*, 8753.

(11) Wu, Q.; Ayers, P. W.; Zhang, Y. *J. Chem. Phys.* **2009**, *131*, 164112.

(12) 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.

(13) Carlson, B. C.; Keller, J. M. *Phys. Rev.* **1957**, *105*, 102.

(14) Moszynski, R. *Theory of Intermolecular Forces: An Introductory Account*. In *Molecular Materials with Specific Interactions – Modeling and Design*; Sokalski, W. A.; Leszczynski, J., Eds.; Springer: Dordrecht, The Netherlands, 2007; Book series: Challenges and Advances in Computational Chemistry and Physics, Vol 4, pp 36–38.

(15) Mandado, M.; Hermida-Ramón, J. M. Unpublished results.

(16) Bader, R. F. W. *Atoms in Molecules: A Quantum Theory*; Oxford University Press: Oxford, U. K., 1990.

(17) Kuan-Jiuh, L.; Ming-Chu, C.; Wang, Y. *J. Phys. Chem.* **1994**, *98*, 11685.

(18) Mosquera, R. A.; González-Moa, M. J.; Estévez, L.; Mandado, M.; Graña, A. M. An electron density-based approach to the origin of stacking interactions. In *Quantum Biochemistry*; Matta, C., Ed.; Wiley-VCH: Weinheim, Germany, 2010; pp 365–387.

(19) Sanchez-Lozano, M.; Otero, N.; Hermida-Ramón, J. M.; Estévez, C. M.; Mandado, M. *J. Phys. Chem. A*. Accepted for publication.

(20) Coppens, P.; Hall, M. B. *Electron Distribution and the Chemical Bond*; Plenum Press: New York, 1982.

(21) Kremer, D.; Kraka, E. *Angew. Chem., Int. Ed.* **1984**, *23*, 627.

(22) Szczeniński, M. M.; Chałasiński, G.; Cybulski, S. M.; Scheiner, S. *J. Chem. Phys.* **1990**, *93*, 4243.

(23) Frascini, E.; Stone, A. J. *J. Comput. Chem.* **1998**, *19*, 847.

(24) Misquitta, A. J.; Stone, A. J. *Mol. Phys.* **2008**, *106*, 1631.

(25) Löwdin, P.-O. *Adv. Quantum Chem.* **1990**, *5*, 185.

(26) Müller, A. M. K. *Phys. Lett. A* **1984**, *105*, 446.

(27) Goedecker, S.; Umrigar, C. *Phys. Rev. Lett.* **1998**, *81*, 866.


(28) Reed, A. E.; Weinhold, F. *J. Chem. Phys.* **1983**, *78*, 4066.

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

Excited-State Tautomerization in the 7-Azaindole-(H₂O)_n (n = 1 and 2) Complexes in the Gas Phase and in Solution: A Theoretical Study

Hua Fang and Yongho Kim*

Department of Chemistry, Kyung Hee University, 1 Seochun-Dong, Giheung-Gu, Yongin-Si, Gyeonggi-Do, 446-701, Korea

 Supporting Information

ABSTRACT: A systematic study of the excited-state tautomerization of 7-azaindole-(H₂O)_n (n = 1 and 2) complexes in both gas and solution phases were investigated theoretically. Electronic structures and energies for the reactant, transition state (TS), and product were computed using the time-dependent density functional theory (TDDFT) and complete active space self-consistent field (CASSCF) levels with 6-31G (d,p), 6-311G(d,p), and 6-311+G(d,p) basis sets. Barrier heights and tautomerization energies were corrected by the second-order multireference perturbation theory (MRPT2) to consider the dynamic electron correlation. The solvent effect decreased the tautomerization barrier height in the 7-azaindole-H₂O complex. In the 7-azaindole-(H₂O)₂ complex, two transition states were found for two asynchronous but concerted paths: in the first, the pyrrole ring proton moved first to water; in the second, the water proton moved first to the pyridine ring. The CASSCF level with the MRPT2 correction clearly showed that the former path was much preferable to the latter. The preferable barrier height was only 1.6 kcal/mol with a zero-point energy correction, which would make the excited-state tautomerization possible. At all TDDFT levels, the TS structures and barrier heights depended on both the basis set used and the solvent effect. Most TDDFT methods failed to reproduce the CASSCF structures and MRPT2 energies. Only two methods, WB97XD/6-31G(d,p) and M062X/6-311+G(d,p), predicted two TSs for the two asynchronous paths in the 7AI-(H₂O)₂ complex but failed to reproduce the energetics. Further systematic study is necessary to test whether current TDDFT methods, including solvent effects, can be used to understand excited-state proton transfer reactions.

1. INTRODUCTION

Proton/hydrogen-atom transfer reactions have attracted much attention due to their importance for understanding many physical, chemical, and biological phenomena.¹ The characterization of the nature of the proton or hydrogen atom transfer process is of great importance not only from the point of view of spectroscopy but also from the point of view of reaction dynamics. Among many molecules, 7-azaindole (7AI) is an important model system for the study of excited state proton transfer processes since it resembles molecules of the DNA base pair.² 7AI contains one proton donor and one proton acceptor and thereby displays simple hydrogen-bonding structures upon dimerization and complexation with solvents. Therefore, the 7AI-dimer and 7AI hydrogen-bonded clusters have been studied extensively.

In an early theoretical calculation, Chaban and Gordon,^{3,4} studied the tautomerization reaction of 7AI and a 1:1 7AI/H₂O complex (7AI-H₂O) in the singlet ground (S₀) and first excited (S₁) states. They used the complete active space self-consistent field (CASSCF) method with multireference second-order perturbation theory (MCQDPT2),⁵ which included dynamic electron correlation to calculate energies and intrinsic reaction coordinates for the proton transfer process. They found that the normal 7AI form was more stable than the tautomer in the S₀ state, whereas the relative energies were reversed in the S₁ state. The activation energy for tautomerization in 7AI was dramatically reduced by complexation with one water molecule when the dynamic electron correlation was considered. Casadesús et al.⁶

investigated the tautomerization of 7AI-(H₂O)_n (n = 0–4) in the S₁ state. The geometries were optimized using the single-excitation configuration interaction (CIS) method, and the energies were calculated at the time-dependent density functional theory (TDDFT) level. The dependence of activation barriers on the number of attached water molecules was discussed.

Fernández-Ramos et al.⁷ calculated the rate constant for the excited state hydrogen transfer reaction of 7AI-water complexes in solution using the Onsager model. They concluded that concerted proton tunneling rate constants dominated the tautomerization process. However, their transition state (TS) structures, which were calculated at the CASSCF(8,8)/6-31G(d) level, were very different from those obtained from previous higher-level calculations.⁴ Deviation of the structural parameters in the TS was more than 0.2 Å (see Table 1). Otherwise, the barrier height in the 1:1 7AI/H₂O complex at the CASSCF(8,8)/6-31G(d) level was 20.64 kcal/mol,⁷ whereas the barrier height at the CASSCF(10,9)/DZP level was 18.20 kcal/mol.⁴ For the two-waters complex (the only one that can be directly compared with experimental data), the calculations indicated a concerted, though nonsynchronous, reaction path with a high-energy barrier at 16.94 kcal/mol.⁷ Recently, Kina et al.⁸ conducted *ab initio* QM/MM molecular dynamics (AIMD) simulations for the excited-state full tautomerization process in 7AI-(H₂O)_n (n = 1, 2)

Received: November 10, 2010

Published: January 19, 2011

Table 1. Geometric Parameters of Reactant, Product, and Transition States for Excited-State Proton Transfer in 7AI–H₂O Complexes^a

computational method	reactant		product	
	$r(\text{H}_{10}-\text{O}_{16})$	$r(\text{H}_{17}-\text{N}_6)$	$r(\text{N}_1-\text{H}_{10})$	$r(\text{O}_{16}-\text{H}_{17})$
CIS/6-31G(d) ^b	2.09	2.05	2.17	2.21
RICC2/TZVP ^c	1.781	1.795	1.990	1.986
CASSCF(8,8)/6-31G(d) ^d	2.166	2.131	2.182	2.196
CASSCF(10,9)/6-31G(d,p) ^e	2.136	2.126	2.190	2.196
CASSCF(10,9)/6-311G(d,p)	2.141	2.135	2.206	2.172
CASSCF(10,9)/DZP ^f	2.164	2.125	2.199	2.192
B3LYP/6-31G(d,p)	1.823	1.823	1.969	1.979
B3LYP/6-311+G(d,p)	1.890	1.831	1.963	2.052
CAM-B3LYP/6-31G(d,p)	1.794	1.808	1.958	1.950
CAM-B3LYP/6-311+G(d,p)	1.868	1.818	1.951	2.013
LC-BLYP/6-31G(d,p)	1.773	1.797	1.936	1.912
LC-BLYP/6-311+G(d,p)	1.843	1.803	1.924	1.965
M062X/6-31G(d,p)	1.821	1.836	2.006	1.970
M062X/6-311+G(d,p)	1.877	1.855	1.987	1.995
WB97XD/6-31G(d,p)	1.827	1.842	1.988	1.968
WB97XD/6-311+G(d,p)	1.889	1.843	1.973	2.022
exp. ^g	1.950	2.117		

computational method	transition state			
	$r(\text{N}_1-\text{H}_{10})$	$r(\text{H}_{10}-\text{O}_{16})$	$r(\text{O}_{16}-\text{H}_{17})$	$r(\text{H}_{17}-\text{N}_6)$
CIS/6-31G(d) ^b	1.30	1.22	1.27	1.26
CASSCF(8,8)/6-31G(d) ^d	1.101	1.484	1.269	1.237
CASSCF(10,9)/6-31G(d,p) ^e	1.258	1.220	1.086	1.444
CASSCF(10,9)/6-311G(d,p)	1.277	1.198	1.067	1.479
CASSCF(10,9)/DZP ^f	1.263	1.213	1.084	1.447
B3LYP/6-31G(d,p)	1.224	1.296	1.168	1.357
B3LYP/6-311+G(d,p)	1.151	1.407	1.264	1.253
CAM-B3LYP/6-31G(d,p)	1.221	1.289	1.151	1.369
CAM-B3LYP/6-311+G(d,p)	1.172	1.359	1.226	1.282
LC-BLYP/6-31G(d,p)	1.223	1.277	1.140	1.377
LC-BLYP/6-311+G(d,p)	1.183	1.333	1.210	1.292
M062X/6-31G(d,p)	1.233	1.272	1.123	1.408
M062X/6-311+G(d,p)	1.234	1.270	1.143	1.379
WB97XD/6-31G(d,p)	1.224	1.282	1.153	1.362
WB97XD/6-311+G(d,p)	1.176	1.349	1.220	1.283

^a Bond distances are in Å. ^b Ref 6. ^c Ref 41. ^d Ref 7. ^e Ref 42. ^f Ref 4. ^g Ref 16. The structural parameters are obtained from the fit to the experimental rotational constants.

complexes in gas and water. They found that the ESDPT takes place asynchronously in both the gas and solution phases.

The excited-state tautomerization in 7AI has been observed in the condensed phase in alcohol or water solutions.^{9–11} Although the tautomerization occurs very rapidly in aqueous solution, the excited-state multiple-proton transfer reactions in the 7AI–(H₂O)_n ($n = 1, 2$) complexes in the gas phase have been uncertain. Huang et al.¹² studied the 1:1 7AI/water complex in the first excited state in cold beams and reported a fluorescence lifetime of 8 ns, which implies that the tautomerization rate constant of ESDPT cannot be larger than approximately 10^7 s^{-1} . However, no direct observation of ESDPT in this complex has yet been made. Folmer et al.¹³ observed very fast decay profiles of the femtosecond pump–probe transients of 7AI–(H₂O)_n ($n =$

2–4) complexes and ascribed them to the excited-state multiple-proton transfer. However, the fast decay time constants did not agree with those predicted from a very sharp bandwidth of jet-cooled fluorescence excitation spectra¹⁴ and the dispersed fluorescence spectra of 7AI–(H₂O)_n ($n = 1 - 3$) complexes.¹⁵ Schmitt et al.¹⁶ measured the rotationally resolved electronic spectra of the 7AI–(H₂O)_n ($n = 1, 2$) clusters in a molecular beam. From the rotational constants, the structures in the S₀ and S₁ electronic states were determined. They also predicted the long lifetime of the excited species from the narrow bandwidth, which suggested no excited-state tautomerization in the gas phase.

However, Sakota et al.¹⁷ recently investigated the excited-state multiple-proton transfer reactions in 7AI water clusters,

Table 2. Geometric Parameters of Reactant, Product, and Transition States for Excited-State Proton Transfer in $7\text{AI}-(\text{H}_2\text{O})_2$ Complexes^a

computational method	reactant			product		
	$r(\text{H}_{10}-\text{O}_{16})$	$r(\text{O}_{20}-\text{H}_{17})$	$r(\text{H}_{19}-\text{N}_6)$	$r(\text{N}_1-\text{H}_{10})$	$r(\text{O}_{16}-\text{H}_{17})$	$r(\text{O}_{20}-\text{H}_{19})$
CIS/6-31G(d) ^b	1.79	1.75	1.84	1.77		
RICC2/TZVP ^c	1.621	1.638	1.676	1.863	1.753	1.808
CASSCF(8,8)/6-31G(d) ^d	1.978	1.894	2.013	2.064	1.915	2.021
CASSCF(10,9)/6-31G(d,p)	1.792	1.809	1.881	2.065	1.919	2.010
CASSCF(10,9)/6-311G(d,p)	1.796	1.822	1.895	2.078	1.930	2.015
B3LYP/6-31G(d,p)	1.652	1.638	1.683	1.824	1.733	1.798
B3LYP/6-311+G(d,p)	1.726	1.686	1.706	1.834	1.773	1.868
CAM-B3LYP/6-31G(d,p)	1.623	1.612	1.669	1.815	1.711	1.776
CAM-B3LYP/6-311+G(d,p)	1.698	1.662	1.693	1.822	1.749	1.840
LC-BLYP/6-31G(d,p)	1.592	1.578	1.650	1.792	1.675	1.742
LC-BLYP/6-311+G(d,p)	1.672	1.631	1.674	1.795	1.711	1.801
M062X/6-31G(d,p)	1.578	1.582	1.643	1.843	1.735	1.786
M062X/6-311+G(d,p)	1.696	1.680	1.708	1.860	1.780	1.849
WB97XD/6-31G(d,p)	1.651	1.653	1.703	1.833	1.742	1.789
WB97XD/6-311+G(d,p)	1.708	1.688	1.717	1.834	1.768	1.836
exp. ^e	1.693	1.679	1.818			

computational method	transition state					
	$r(\text{N}_1-\text{H}_{10})$	$r(\text{H}_{10}-\text{O}_{16})$	$r(\text{O}_{16}-\text{H}_{17})$	$r(\text{H}_{17}-\text{O}_{20})$	$r(\text{O}_{20}-\text{H}_{19})$	$r(\text{H}_{19}-\text{N}_6)$
CIS/6-31G(d) ^b	1.22	1.28	1.25	1.19	1.34	1.19
CASSCF(8,8)/6-31G(d) ^d	1.056	1.580	1.032	1.462	1.379	1.130
CASSCF(10,9)/6-31G(d,p) TS1	1.394	1.101	1.079	1.329	1.050	1.493
CASSCF(10,9)/6-311G(d,p) TS1	1.431	1.079	1.072	1.335	1.036	1.525
CASSCF(10,9)/6-31G(d,p) TS2	1.088	1.464	1.061	1.372	1.351	1.141
CASSCF(10,9)/6-311G(d,p) TS2	1.075	1.496	1.044	1.401	1.392	1.117
B3LYP/6-31G(d,p)	1.204	1.299	1.147	1.285	1.201	1.292
B3LYP/6-311+G(d,p)	1.123	1.428	1.122	1.314	1.312	1.191
CAM-B3LYP/6-31G(d,p)	1.209	1.281	1.138	1.283	1.178	1.309
CAM-B3LYP/6-311+G(d,p)	1.128	1.406	1.117	1.309	1.288	1.201
LC-BLYP/6-31G(d,p)	1.217	1.264	1.133	1.278	1.156	1.330
LC-BLYP/6-311+G(d,p)	1.139	1.378	1.121	1.293	1.264	1.216
M062X/6-31G(d,p)	1.248	1.236	1.126	1.293	1.126	1.375
M062X/6-311+G(d,p)	1.219	1.267	1.131	1.280	1.174	1.309
WB97XD/6-31G(d,p)	1.213	1.272	1.139	1.280	1.182	1.298
WB97XD/6-311+G(d,p)	1.129	1.400	1.117	1.306	1.301	1.188

^a Bond distances are in Å. ^b Ref 6. ^c Ref 41. ^d Ref 7. ^e Ref 16. The structural parameters are obtained from the fit to the experimental rotational constants.

$7\text{AI}-(\text{H}_2\text{O})_n$ ($n = 2, 3$), in the gas phase by combining electronic spectroscopy and quantum chemical calculations. They found that the $7\text{AI}-(\text{H}_2\text{O})_2$ and $7\text{AI}-(\text{CH}_3\text{OH})_2$ geometries were similar; the solvent molecules bridged the heteroaromatic N atom and the NH hydrogen by intermolecular hydrogen bonds, forming a cyclic structure.^{14,18} Most importantly, they successfully observed the tautomer of the $7\text{AI}-(\text{H}_2\text{O})_2$ complex in the excited state and showed that the ESTPT occurs in the gas phase.

However, there are no theoretical calculations for the excited-state proton transfer in $7\text{AI}-(\text{H}_2\text{O})_2$ complexes using the TDDFT method. Currently, the most widely used method to describe the excited states of molecules is the linear-response time-dependent density functional theory.^{19,20} Since TDDFT is computationally efficient for excited-state calculations, a large number of excited-state properties based on the TDDFT method

have been performed. TDDFT yielded surprisingly accurate absorption spectra.^{21,22} However, dramatic failures of the TDDFT method have also been recently found. For example, Ryberg states are not appropriately described due to the rapid asymptotic decay of standard exchange-correlation functionals;²³ doubly excited states are not contained owing to the linear-response formalism of TDDFT,^{24–26} and charge-transfer excited states are given lower excitation energies with the incorrect asymptotic potential energy surfaces because of spurious electron-transfer self-interaction.^{27–30}

The purpose of this work was to perform a systematic study of proton transfer in the biologically interesting $7\text{AI}-\text{H}_2\text{O}$ and $7\text{AI}-(\text{H}_2\text{O})_2$ complexes. The main interest was to determine if the TDDFT method could fit the excited state chemical reaction. In this study, five TDDFT methods, which contained hybrid

functional, long-rang correction (LC) and empirical dispersion functionals, were used to systematically investigate the tautomerization reaction in the S_1 state in both the gas and solution phases. To avoid possible TDDFT artifacts, the multiconfigurational CASSCF method was used to assess TDDFT reliability in the present study.

2. COMPUTATIONAL METHODS

Reactant, product, and TS geometries of the excited state tautomerization in the $7AI-(H_2O)_n$ ($n = 1$ and 2) complexes were fully optimized at the TDDFT and CASSCF level with 6-31G(d,p), 6-311G(d,p), and 6-311+G(d,p) basis sets using the Gaussian 09 program³¹ in the gas phase and in solution. The crucial step to conduct the CASSCF calculation was to select the proper active space. The obvious choice for an active space in 7AI complexes would include four π bonds, four corresponding antibonding orbitals, and one nitrogen lone pair, resulting in an active space of 10 electrons in nine orbitals, which was denoted as CASSCF(10,9). Calculating vibrational frequencies and establishing no imaginary frequency for the reactant and product and one imaginary frequency for the TS verified optimized structures. Single point energy calculations were also performed using the second-order multireference perturbation theory (MRPT2) for stationary points. All MRPT2 calculations were performed using the GAMESS program.³²

Analytic TDDFT gradients were calculated using the variational TDDFT formulation of Furche and Ahlrichs.³³ Several different exchange-correlation DFT potentials were used for the systems. We used Becke's three-parameter Lee–Yang–Parr hybrid functionals,³⁴ (B3LYP), Handy and co-workers' long-range corrected version of B3LYP using the Coulomb-attenuating method,³⁵ (CAM-B3LYP), a long-range-corrected version of BLYP³⁶ (LC-BLYP), the hybrid functional of Truhlar and Zhao³⁷ (M062X), and the latest functional from Head-Gordon and co-workers, which included empirical dispersion³⁶ (WB97XD).

We performed polarizable continuum model calculations using the integral equation formalism (IEFPCM)^{38–40} at the TDDFT and CASSCF levels to investigate the mechanism of tautomerization in water. The geometries of reactant, product, and TS were completely optimized in solution. Currently, the solvent effect is not implemented in the MRPT2 calculations. Therefore, the gas phase MRPT2 energies and the solvation energies at the CASSCF level were used to estimate the MRPT2 energies in water.

3. RESULTS AND DISCUSSION

Casadesús et al.⁶ examined the reaction path for the tautomerization of the $7AI-H_2O$ complex in the S_1 state using CIS with the 6-31G(d) basis set. Since CIS theory does not include higher-level electron correlation, which has an important influence on H-bonded structures, CIS does not give very accurate results. Schmitt et al.⁴¹ studied the normal and tautomer forms of $7AI-(H_2O)_n$ ($n = 1, 2$) in their ground and two lowest singlet excited states by using the RIC2 method. They found that water complexation of 7AI with one and two water molecules stabilizes the more polar L_a state, which has been observed experimentally. The CASSCF(10,9)/DZP level was the highest level reported in the literature for the $7AI-H_2O$ complex.⁴ Duong and Kim^{42,43} studied the $7AI-H_2O$ complex at the CASSCF(10,9)/6-31G(d,p) level and showed that the geometries of the reactant, product, and TS were in good agreement with

previous studies at the CASSCF(10,9)/DZP level.^{4,8} The geometry of the bare 7AI molecule⁴ is mostly unaffected by complex formation. Therefore, only the geometric parameters describing hydrogen bond (H-bond) distances are given in Tables 1 and 2. Although no excited-state tautomerization was observed for the 1:1 complex in the gas phase, Sakota et al.¹⁷ suggested that tautomerization occurs via excited-state triple proton transfer in the cyclic 1:2 complex. Therefore, it is interesting to make a systematic comparison of the structures and energies of the excited-state tautomerization between 1:1 and 1:2 complexes of 7AI with H_2O . In this study, structures of the reactant, TS, and product were obtained by TDDFT and CASSCF methods using various basis sets. To test whether the DFT functional could correctly predict stationary points of the $7AI-H_2O$ complex in the excited-state tautomerization, the TDDFT and CASSCF results were compared.

3.1. 1:1 Complex of 7AI with H_2O in the Gas Phase. Structures of the stationary points in the 1:1 7AI complex with water optimized at the CASSCF/6-31G(d,p) level are shown in Figure 1. For the $7AI-H_2O$ complex, the H-bond distances, $H_{10}-O_{16}$ and $H_{17}-N_6$ in the reactant and N_1-H_{10} and $O_{16}-H_{17}$ in the product at the B3LYP level using the 6-31G(d,p) basis set were 0.313 Å, 0.303 Å, 0.221 Å, and 0.217 Å shorter than those at the CASSCF level, respectively (Table 1). The shorter the H-bond length, the higher the H-bond energy; therefore the B3LYP level predicted strong H bonds compared to the CASSCF results. The hybrid M062X functional gave similar results to the B3LYP analysis. The H-bond distances, $H_{10}-O_{16}$ and $H_{17}-N_6$ in the reactant and N_1-H_{10} and $O_{16}-H_{17}$ in the product, were underestimated by 0.315 Å, 0.290 Å, 0.184 Å, and 0.226 Å, respectively, compared with corresponding CASSCF values. When the long-range corrected functionals (CAM-B3LYP and LC-BLYP) were used, the H-bond distances, $H_{10}-O_{16}$ and $H_{17}-N_6$ in the reactant and N_1-H_{10} and $O_{16}-H_{17}$ in the product, were on average 0.353 Å, 0.324 Å, 0.243 Å, and 0.265 Å shorter than the corresponding CASSCF values, respectively. The long-range corrected functionals predicted slightly shorter H-bond distances than in B3LYP or M062X. The WB97XD functional that included empirical dispersion was also used in the $7AI-H_2O$ complex. Compared to the CASSCF(10,9)/6-31G(d,p) results, the H-bond distances $H_{10}-O_{16}$ and $H_{17}-N_6$ in the reactant and N_1-H_{10} and $O_{16}-H_{17}$ in the product were underestimated by 0.309 Å, 0.284 Å, 0.202 Å, and 0.228 Å, respectively. Among the DFT methods used in this study, the long-range corrected functionals predicted the shortest H-bond distances (the highest H-bond energies) in the reactant and product $7AI-H_2O$ complex. When the larger 6-311+G(d,p) basis set was used, all H-bond distances except N_1-H_{10} of the product became larger at all DFT levels, although they are still shorter than the CASSCF values.

Schmitt et al.¹⁶ measured the rotationally resolved electronic spectra of the $7AI-(H_2O)_n$ ($n = 1, 2$) clusters in a molecular beam and obtained the geometry parameters from the fit to the experimental rotational constants. The comparison to calculated CASSCF(10,9) results shows that the $H_{17}-N_6$ distance in the reactant at the CASSCF(10,9)/6-31G(d,p) (or CASSCF(10,9)/6-311G(d,p)) level is consistent with the experimental value, and the $H_{10}-O_{16}$ distance at the CASSCF(10,9) level is longer than the experimental value.

TS geometries for the excited-state tautomerization in $7AI-H_2O$ were fully optimized and confirmed by frequency

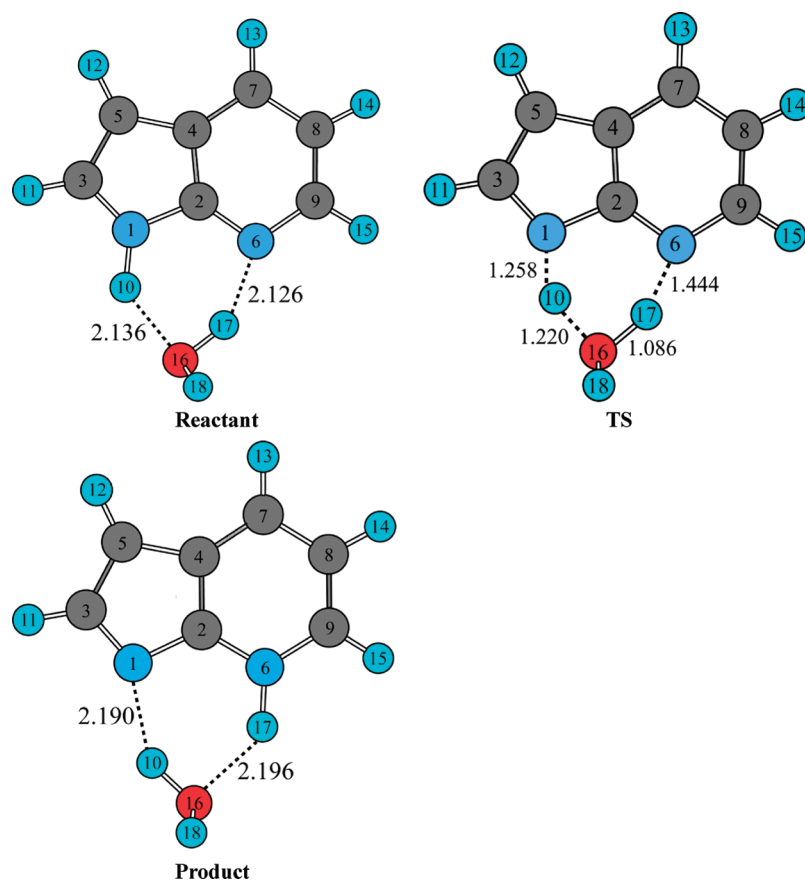


Figure 1. Excited-state structures of the reactant, TS, and product for excited-state proton transfer in the 7AI–H₂O complex.

calculations, with some of the geometric parameters listed in Table 1. The ESDPT potential energy curves in the 7AI–H₂O complex were studied at the MRPT2/CASSCF(10,9)/6-31G(d,p) level. It was shown that two protons were transferred asynchronously, but concertedly. In the 7AI–H₂O TS at the CASSCF level, the H₁₀ atom moved more than halfway along the reaction coordinate toward O₁₆ (Figure 1), whereas the H₁₇ atom rarely moved. In this asynchronous double proton transfer, the H₁₀ atom moved first, followed by the H₁₇ atom. However, in the TSs at the various TDDFT levels using the same basis sets, the H₁₀ atom moved slightly less than halfway along the reaction coordinate, and the H₁₇ atom moved slightly more than at the CASSCF level. All of the TS structures at the TDDFT/6-31G(d,p) level seemed to satisfy Hammond's postulate that the TS of the exothermic reaction resembles the reactant structure (denoted as an early TS), although the position of the two H atoms at the TS was not completely synchronized. When the 6-311+G(d,p) basis set was used, all TDDFT methods except M062X predicted completely different TS structures, in which the H₁₇ atom moved about halfway along the reaction coordinate toward N₆, whereas the H₁₀ atom moved very little. In this case, the H₁₇ atom moved first followed by the H₁₀ atom. The use of a larger 6-311+G(d,p) basis set resulted in a different asynchronous mechanism with the opposite order of proton transfer. This TS structure is similar to one reported previously at the CASSCF(8,8)/6-31G(d,p) level. It is interesting to note that the M062X method, unlike other TDDFT methods, predicted consistent TS structures irrespective of the basis set size.

A correlation plot using H-bond distances can be used to visualize hydrogen bond characteristics and proton transfer transition states. Limbach et al.^{44–46} defined the natural hydrogen bond coordinates $q_1 = 1/2(r_{\text{AH}} - r_{\text{BH}})$ and $q_2 = r_{\text{AH}} + r_{\text{BH}}$ to represent the correlation between r_{AH} and r_{BH} in many hydrogen-bonded complexes (A–H···B). For a linear H bond, q_1 represents the dislocation of H from the H-bond center, and q_2 represents the distance between the two heavy atoms. A strong H-bond results in short r_{BH} and slightly elongated r_{AH} distances. Bond distance depends on bond energy and order. Pauling suggested an exponential relationship between valence bond order and bond length. In the A–H···B complexes, the r_{AH} and r_{BH} distances depend on each other, leading to allowed r_{AH} and r_{BH} values based on the following Pauling equations under the assumption that the sum of two bond orders is conserved, $n_{\text{AH}} + n_{\text{BH}} = 1$:

$$r_{\text{AH}} - r_{\text{AH}}^0 = a_{\text{AH}} \ln n_{\text{AH}} \quad (1)$$

$$r_{\text{BH}} - r_{\text{BH}}^0 = a_{\text{BH}} \ln n_{\text{BH}} \quad (2)$$

where r_{AH}^0 and r_{BH}^0 are the equilibrium lengths of the free AH and BH bonds and a_{AH} and a_{BH} are the parameters describing the decrease of the AH and the HB unit bond valences with the corresponding distances. This type of correlation, which is called the “bond energy bond order method”, has been used for many years to study hydrogen atom transfer.^{47,48} When H is transferred from A to B in the A–H···B complex, q_1 increases from

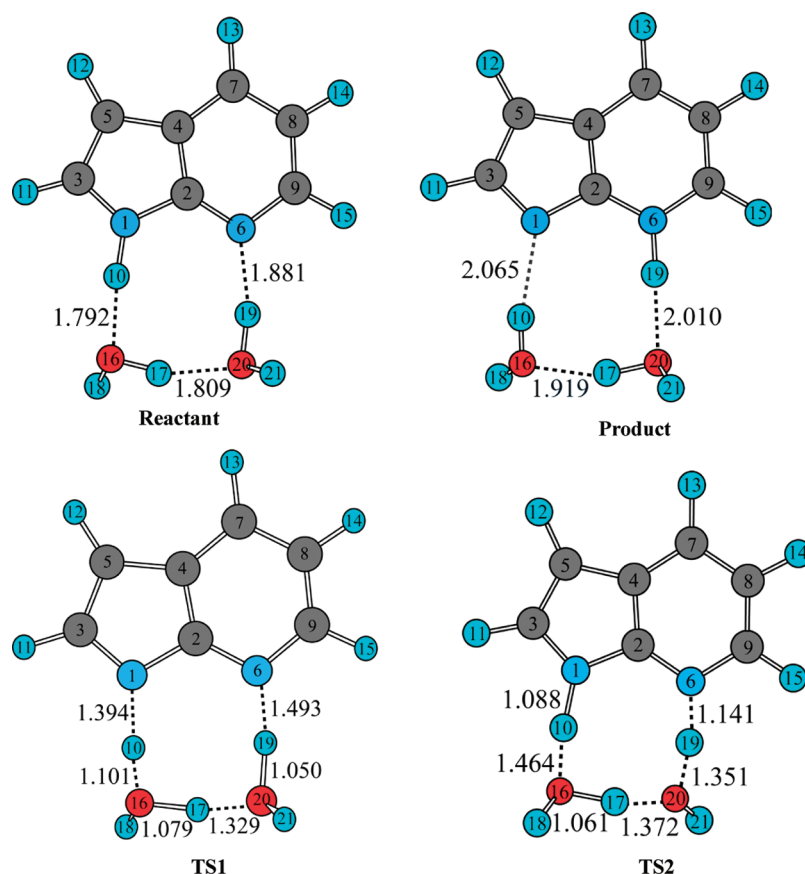


Figure 2. Excited-state CASSCF(10,9)/6-31G(d,p) structures of the reactant, product, and TS for asynchronous proton transfer in the 7AI-(H₂O)₂ complex. In the TS1 structure, H₁₀ moves first. In the TS2 structure, H₁₉ moves first.

negative to positive values, and q_2 goes through a minimum, which is located at $q_1 = 0$. Limbach et al.^{44–46} suggested that both proton transfer and hydrogen-bonding coordinates could be combined into the same correlation. This correlation can be used to study the characteristics of transition state, such as earliness or lateness, bond order, and asynchronicity. The q_1 value of TS is negative or positive when the TS is either early or late, respectively. In addition, the two TS q_1 values for the double proton transfer should be very similar and different in the synchronous and asynchronous mechanism, respectively.

The correlations between N₁-H₁₀ and H₁₀-O₁₆ distances (H₁₀ transfer) and N₆-H₁₇ and H₁₇-O₁₆ distances (H₁₇ transfer) for the 7AI-H₂O complex are depicted in Figure 3. It is interesting to note that all points for the reactant, product, and TS were very close to the black line regardless of computational level, which suggests that the sum of bond order at all stationary points was approximately conserved. All q_2 values of the reactant and product at the TDDFT level were significantly smaller than the CASSCF values; this is consistent with the fact that TDDFT overestimates H-bond strength, leading to shorter H-bond distances. In the synchronous process, two TS q_1 values for two proton transfers should be approximately the same. These q_1 values for H₁₀ and H₁₇ transfer at all TDDFT levels using the 6-31G(d,p) basis set were both negative, consistent with the early TS, but not the same. However, H₁₀ and H₁₇ transfer values at the CASSCF/6-31G(d,p) level were slightly positive and very negative, respectively, which resulted from a highly asynchronous TS (slightly late and very early TS in terms of H₁₀ and H₁₇ transfers, respectively).

When the 6-311+G(d,p) basis set was used at TDDFT levels, H₁₀ q_2 and q_1 values for the reactant became significantly larger and smaller, which moved the correlation points (q_1 and q_2) toward the upper-left side of Figure 3B, along the black line. For simplicity, the q_2 and q_1 values of H₁₀ transfer were denoted as “the H₁₀ correlation point”. The H₁₇ correlation points for the product also moved to the upper-right side along the black line. Positions of the TS on the H₁₀ and H₁₇ transfer reaction coordinates became very early and late, respectively, using larger basis sets. Only the M062X level predicted an almost identical correlation of TS for both H₁₀ and H₁₇ transfers; i.e., the locations of TSs were not dependent on the size of basis sets. The blue line represents correlations along the intrinsic reaction coordinate of the ESDPT calculated at the CASSCF(10,9)/6-31G(d,p) level. It is very interesting to note that the blue line is very different in shape from the black lines, which means that the sum of the bond order was not conserved along the intrinsic reaction coordinate. For H₁₀ transfer, the blue line was always under the black line except at the TS, which indicates that the sum of the bond order was larger than unity at both the reactant- and product-side reaction coordinates. However, for H₁₇ transfer, the blue line was above the black line near $q_1 = 0$, which means that the sum of the bond order was less than unity. It is also very interesting that the blue lines were flat at the bottom near $q_1 = 0$, which was attributed to an almost constant q_2 value that is the distance between two end atoms. These results suggest that the end atom distance rarely changed during the proton transfer near $q_1 = 0$. This finding was not surprising because it is

Table 3. Reaction Energies, Barrier Heights, and Dipole Moments for Proton Transfer in 7AI–H₂O and 7AI–(H₂O)₂ in the S₁ State at Various Levels of Theory^a

computational method	7AI–H ₂ O						7AI–(H ₂ O) ₂					
	ΔV (kcal/mol)	ΔE (kcal/mol)	μ (D)			ΔV (kcal/mol)	ΔE (kcal/mol)	μ (D)				
			R	TS	P			R	TS	P		
CIS/6-31G(d) ^b	26.94	–18.15				24.35	–15.85					
RICC2/TZVP ^c		–22.51					–20.01					
CASSCF(8,8)/6-31G(d) ^d	20.64(22.07)	–33.19	1.99	5.02	2.06	16.94	–32.45	1.51	6.88	2.21		
CASSCF(10,9)/DZP(d,p) ^e	18.20(14.7)	–31.80(–31.2)										
MCQDPT2/CASSCF(10,9)/DZP ^e	9.80(6.3)	–18.00(–17.4)										
CASSCF(10,9)/6-31G(d,p) ^f	18.08(14.39)	–28.56(–27.92)	1.78	4.28	2.05	15.84(11.04) ^g 15.26(10.72) ^h	–32.23(–32.00)	4.87	5.23 ^g 5.79 ^h	1.92		
CASSCF(10,9)/6-311G(d,p)	17.93(14.47)	–32.26(–31.52)	1.63	4.47	1.93	16.79(12.72) ^g 16.13(12.44) ^h	–31.69(–31.45)	4.95	5.57 ^g 6.05 ^h	1.92		
MRPT2/CASSCF(10,9)/6-31G(d,p)	9.72(6.03) ^e	–18.95(–18.31) ^e				6.02(1.22) ^g 8.36(3.82) ^h	–19.86(–19.63)					
MRPT2/CASSCF(10,9)/6-311G(d,p)	9.28(5.82)	–18.66(–17.92)				5.63(1.56) ^g 10.78(7.09) ^h	–19.22(–18.98)					
B3LYP/6-31G(d,p)	6.79(3.03)	–18.13(–17.64)	4.31	3.09	1.76	5.11(0.12)	–15.31(–14.59)	3.75	3.04	1.53		
CAM-B3LYP/6-31G(d,p)	6.30(2.68)	–19.68(–18.97)	4.17	3.32	1.84	4.52(–0.37)	–16.57(–15.63)	3.68	3.13	1.71		
LC-BLYP/6-31G(d,p)	5.99(2.56)	–19.23(–18.22)	3.67	3.30	2.00	3.74(–0.86)	–16.02(–14.78)	3.20	3.04	1.92		
M062X/6-31G(d,p)	5.26(2.18)	–20.22(–19.21)	4.21	3.40	1.80	2.64(–1.28)	–16.68(–15.63)	3.67	3.18	1.68		
WB97XD/6-31G(d,p)	7.88(4.04)	–19.91(–19.31)	4.24	3.32	1.86	6.32(0.96)	–16.86(–16.20)	3.74	3.17	1.70		
B3LYP/6-311+G(d,p)	9.24(5.24)	–17.62(–17.09)	4.10	3.44	1.42	6.52(2.12)	–15.27(–14.67)	3.84	4.30	1.32		
CAM-B3LYP/6-311+G(d,p)	9.09(5.21)	–18.77(–18.00)	3.87	3.39	1.59	6.26(1.83)	–16.23(–15.42)	3.71	4.12	1.60		
LC-BLYP/6-311+G(d,p)	9.09(5.41)	–17.94(–16.83)	3.23	3.20	1.85	5.84(1.38)	–15.42(–14.33)	3.20	3.64	1.91		
M062X/6-311+G(d,p)	8.28(4.90)	–19.15(–18.26)	3.88	3.33	1.51	5.23(0.76)	–16.72(–15.78)	3.74	3.16	1.60		
WB97XD/6-311+G(d,p)	10.26(6.12)	–19.00(–18.37)	4.01	3.38	1.67	7.77(3.12)	–16.46(–15.66)	3.78	4.26	1.60		

^aThe numbers in parentheses include zero-point energies. ^b Ref 6. ^c Ref 41. ^d Ref 7. ^e Ref 4. ^f Refs 42 and 43. ^g TS1. ^h TS2.

consistent with hydrogen transfer reactions of the heavy–light–heavy mass combinations, where the heavy-atom motions are well separated from the light-atom motions; the heavy atoms rarely moves while the hydrogen moves significantly.

3.2. 1:2 Complex of 7AI with H₂O in the Gas Phase. Optimized structural parameters of the reactant, product, and TS in the 7AI–(H₂O)₂ complexes are listed in Table 2, and structures at the CASSCF/6-31G(d,p) level are depicted in Figure 2. For the 7AI–(H₂O)₂ cluster, the structural parameters of the reactant obtained from the fit to the rotation constants¹⁶ are also in Table 2. Our CASSCF(10,9) H-bond distances are longer than these values. It is interesting to note that all H-bond distances in the 7AI–(H₂O)₂ complexes were smaller than those in 7AI–H₂O complexes. Particularly, all TDDFT methods using the 6-31G(d,p) basis set predicted very short H bonds for the reactant. Short H-bond distance indicates stronger H-bond strength, and linear H bonds are generally stronger than bent H bonds. The H bonds in 7AI–(H₂O)₂ complexes were more linear and shorter than those of the 7AI–H₂O complexes. H-bond distances in the reactant, H₁₀–O₁₆, H₁₇–O₂₀, and H₁₉–N₆ at the B3LYP level, were 0.140 Å, 0.171 Å, and 0.198 Å shorter than the corresponding CASSCF values, respectively (Table 2). In the product, H-bond distances, N₁–H₁₀, O₁₆–H₁₇, and O₂₀–H₁₉ at the B3LYP level, were 0.241 Å, 0.186 Å, and 0.212 Å shorter than the corresponding CASSCF values. When long-range corrected functionals (CAM-B3LYP) were used, reactant and product H-bond distances became slightly smaller than those from the uncorrected method. The M062X and WB97XD methods also predicted smaller H-bond distances compared with CASSCF results. All TDDFT methods used in this study were found to overestimate H-bond strength in the excited state, although use of the larger basis set could remedy this a little bit. Further studies are necessary to understand the

reliability of the TDDFT methods to predict H-bond strength in the excited state.

The geometry of the TS was fully optimized at the TDDFT and CASSCF(10,9) levels and was confirmed by frequency calculations. Interestingly, two TS structures at the CASSCF(10,9) level were found (Figure 2). In the first TS (denoted as TS1), the H₁₀ moved more than halfway from N₁ toward the O₁₆ atom with the H₁₇ and H₁₉ rarely moving, which generated a H₃O⁺-like moiety in a portion of the TS (at O₁₆). However, in the second TS (denoted as TS2), the H₁₉ moved more than halfway from the O₂₀ to the N₆ atom, but H₁₀ and H₁₇ rarely moved, resulting in a HO[–]-like moiety in a portion of the TS (at O₂₀). Due to the fact that only one proton moved substantially, while the other two protons moved slightly, a stepwise mechanism with a possible intermediate was predicted. However, every attempt to locate the intermediate led to either the reactant or the product. These results suggest that two concerted but asynchronous processes exist in the ESTPT, one via TS1 and the other via TS2. Very recently, Sakota et al.¹⁷ reported potential energy curves for the ESTPT reaction coordinate in the 7AI–(H₂O)₂ complex under the assumption of a synchronous process. No potential energy well for an intermediate was found, but the reaction coordinate was not intrinsic, so their argument for a concerted mechanism did not have definitive evidence. Our results are consistent with a concerted mechanism of the ESTPT; however, there are two asynchronous pathways through TS1 and TS2.

We were unable to locate two different TSs at all TDDFT levels used in this study. The bond distances listed at the TDDFT level using the 6-31G(d,p) basis set revealed that the locations of each hydrogen atom at the TS were closer to the center of the two end atoms compared with those from the CASSCF method, and their positions were close to the reactant, satisfying Hammond's postulate.

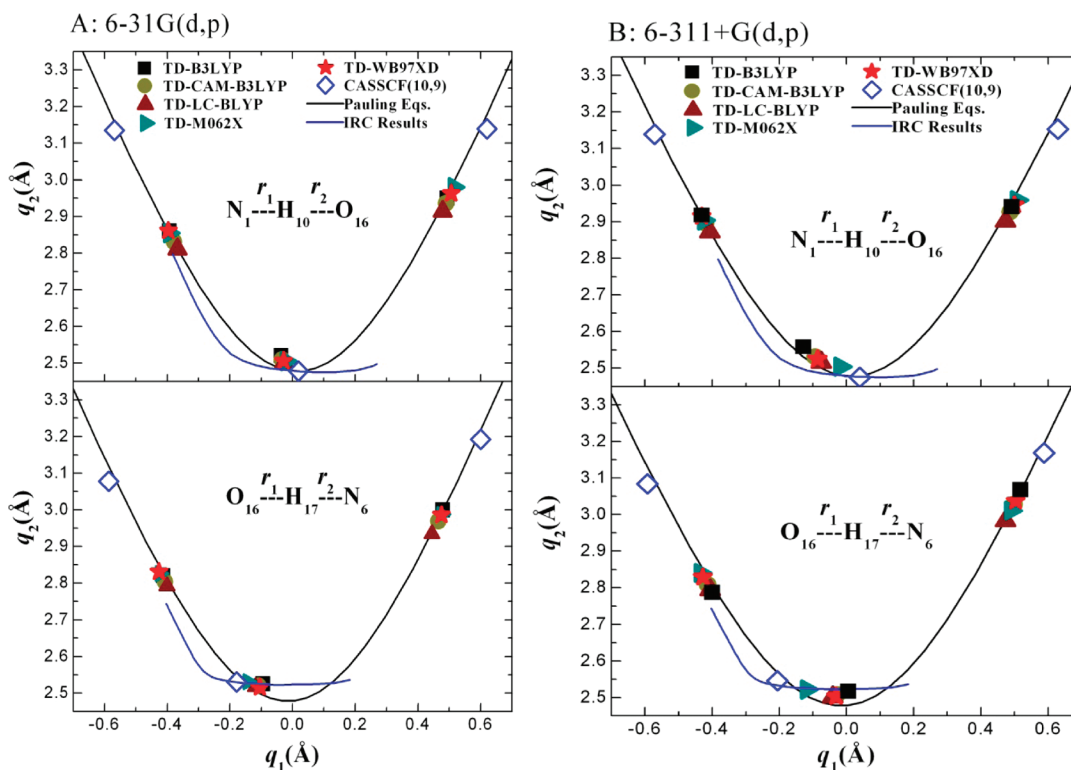


Figure 3. Correlation of the H-bond distances $q_2 = r_1 + r_2$ with the proton transfer coordinate $q_1 = 1/2(r_1 - r_2)$ for the 7AI–H₂O complex in the gas phase. Top, H₁₀ transfer; bottom, H₁₇ transfer. The black line designates the correlation that satisfies conservation of the bond order. Parameters for Pauling equations were from the literature.⁴⁵ The region above and below the black line is where the sum of bond order is smaller and larger than unity, respectively. The correlation points at the bottom near $q_1 = 0$ are for the TS, and those at the top left and right corners are for the reactant and product, respectively. The blue line represents the correlation along the intrinsic reaction coordinate (IRC) calculated at the CASSCF(10,9)/6-31G(d,p) level. The CASSCF points in B were calculated using the 6-311G(d,p) basis set.

The correlation between q_1 and q_2 for the ESTPT in 7AI–(H₂O)₂ complexes is depicted in Figure 4. All TDDFT methods predicted smaller q_2 values for the reactant and product than CASSCF, which was attributed to overestimated H-bond strength. The opposite sign of the q_1 values for H₁₀ and H₁₉ in TS1 was a clear indication of asynchronicity in the concerted proton transfer. For H₁₀ transfer, the TS1 and TS2 q_1 values at the CASSCF level were very positive and negative, respectively. The q_1 values for H₁₉ transfer were opposite those for H₁₀ transfer. These results indicate that the asynchronicity of the two processes (each via TS1 and TS2) is opposite in terms of the order of H₁₀ and H₁₉ transfers. It is interesting to note that the correlation points for TS1 and TS2 were under the solid line, which suggests that the total bond order at TS1 and TS2 is not conserved, but increased. The formation of H₃O⁺-like and HO[−]-like moieties at TS1 and TS2, respectively, might induce coulomb interactions to increase the bond order.

Recently, Limbach et al.⁴⁹ has depicted the correlated NHN hydrogen bond coordinates of various systems including the 7AI dimer in the S₁ state. They found that proton transfer is accompanied by a heavy atom motion and the hydrogen bond compression is the most important heavy atom motion. The transition state structures correspond to the strongest possible NHN hydrogen bonds. The double proton transfer in the 7AI dimer occurs stepwise via a zwitterionic intermediate. There are two possible pathways of the stepwise process depending on the order of the H atom in flight, the TSs of which having opposite signs in the q_1 values and smaller q_2 values compared

with those of the reactants. The q_1 and q_2 values of two TSs in the asynchronous and concerted reaction are consistent with those of the stepwise reaction in 7AI dimer.

At TDDFT levels using the 6-31G(d,p) basis set, all TS q_1 values for H₁₀ and H₁₉ transfers were in the middle of two points for TS1 and TS2, which were negative but closer to $q_1 = 0$ (Figure 4). All TS structures using the 6-31G(d,p) basis set resembled the reactant, satisfying Hammond's postulate, and the asynchronicity of the ESTPT was greatly reduced. When the larger 6-311+G(d,p) basis sets were used, all TS correlation points moved toward the TS2 point except for the M062X. Since the q_1 values for H₁₉ transfer became positive but those for H₁₀ transfer became more negative, the asynchronicity of the ESTPT was increased again in this case. The TS q_1 values for H₁₀ and H₁₉ at the M062X level changed very little and remained approximately in the middle of the TS1 and TS2 points, which were -0.05 and -0.135 , respectively. The negative TS q_1 values indicate early TS, satisfying Hammond's postulate.

3.3. The Energetics of Excited-State Proton Transfer in the Gas Phase. Barrier heights (ΔV), excited-state tautomerization energies (ΔE), and dipole moments (μ) for the 7AI–H₂O and 7AI–(H₂O)₂ complexes are listed in Table 3. Chaban and Gordon⁴ calculated the barrier heights in the 7AI–H₂O complex at the CASSCF(10,9)/DZP and MCQDPT2 levels. They found that the CIS and CASSCF methods overestimated the energy barriers of the excited state and suggested that dynamic electron correlation should be considered. Duong and Kim^{42,43} used the CASSCF(10,9)/6-31G(d,p) method followed by single-point

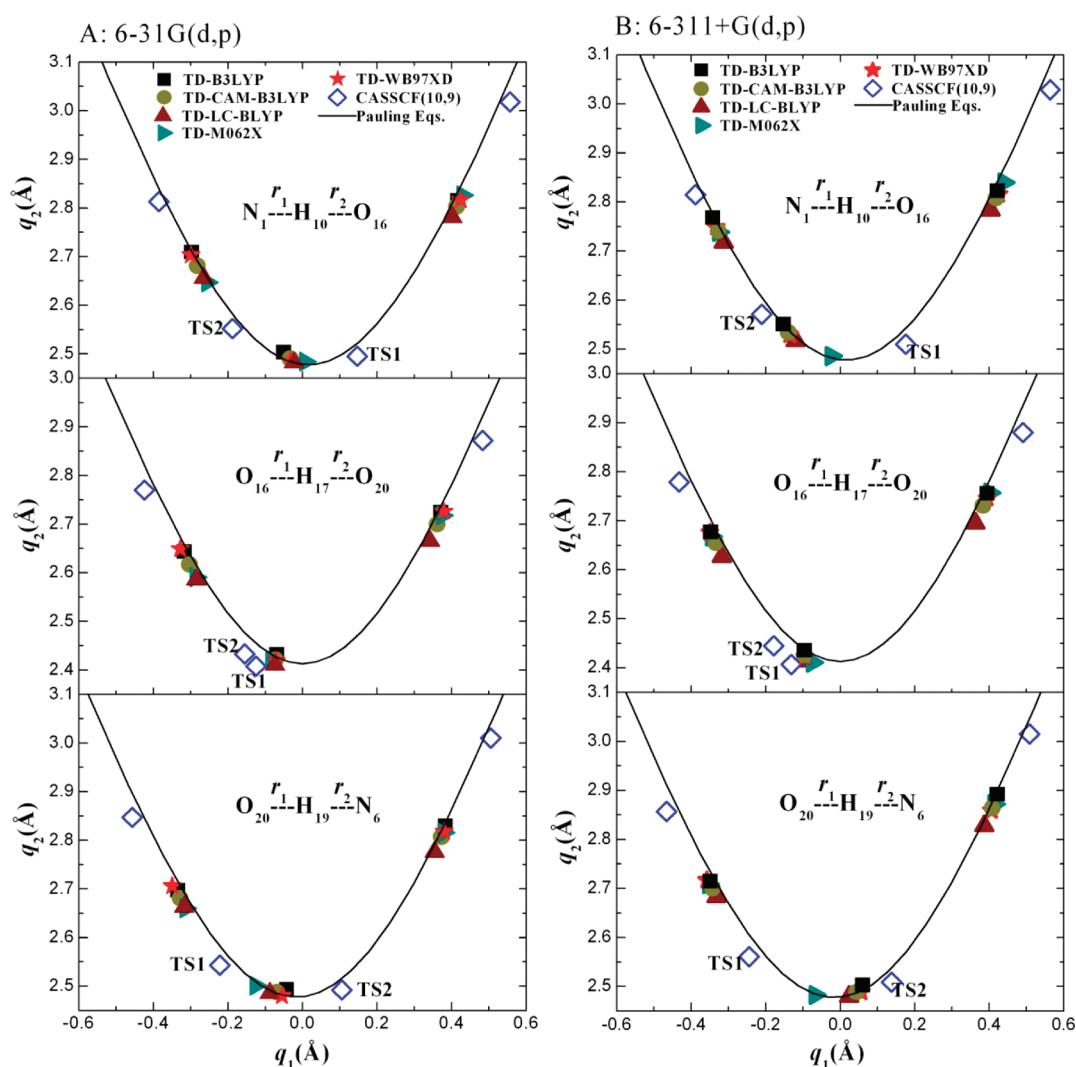


Figure 4. Correlation of the H-bond distances $q_2 = r_1 + r_2$ with the proton transfer coordinate $q_1 = 1/2(r_1 - r_2)$ for the $7\text{AI}-(\text{H}_2\text{O})_2$ complex in the gas phase. Top, H_{10} transfer; middle, H_{17} transfer; bottom, H_{19} transfer. The solid line represents the correlation for equilibrium distances calculated with Pauling equations. The CASSCF points in B were calculated using the 6-311G(d,p) basis set.

MRPT2 corrections to calculate the ESDPT energetics of the $7\text{AI}-(\text{H}_2\text{O})_2$ complex. The tautomerization energies were -18.3 and -19.0 kcal/mol with and without ZPE corrections, respectively. The ΔE values of the $7\text{AI}-(\text{H}_2\text{O})_2$ complex were -19.9 and -19.2 kcal/mol using the 6-31G(d,p) and 6-311G(d,p) basis sets, respectively, without ZPE corrections. The tautomerization energies predicted from the $7\text{AI}-(\text{H}_2\text{O})_2$ and $7\text{AI}-(\text{H}_2\text{O})_2$ complexes agreed well, within 1 kcal/mol.

The MRPT2 barrier height of the excited-state tautomerization in the $7\text{AI}-(\text{H}_2\text{O})_2$ complex was 9.7 and 9.3 kcal/mol using the 6-31G(d,p) and 6-311G(d,p) basis sets, respectively. For the triple proton transfer in the $7\text{AI}-(\text{H}_2\text{O})_2$ complex, two transition states were predicted at the CASSCF levels, and the barrier heights of TS1 and TS2 were 16.8 and 16.1 kcal/mol, respectively, using the 6-311G(d,p) basis set. TS2 was 0.7 kcal/mol lower in barrier height, and 0.3 kcal/mol with ZPE corrections. However, when we included the dynamic electron correlation at the MRPT2 level, the TS1 and TS2 barrier heights were reduced to 5.6 and 10.8 kcal/mol, respectively. It is very interesting that TS1 had an approximately 5.2 kcal/mol lower barrier height than TS2. These results suggest that the triple proton transfer in the

excited state occurred preferably via TS1. The dynamic electron correlation not only reduced barrier heights but also changed the ESTPT mechanism. When the ZPE correction was included using the frequencies calculated at the CASSCF level, the TS1 barrier height was only 1.6 kcal/mol. The difference in the ZPE-corrected barrier height between the $7\text{AI}-(\text{H}_2\text{O})_2$ and $7\text{AI}-(\text{H}_2\text{O})_2$ complexes was 4.2 kcal/mol, using the 6-311G(d,p) basis set at the MRPT2 level. This result supports the argument of Sakota et al. that the gas phase tautomerization in the excited state can occur via a concerted triple proton transfer in the $7\text{AI}-(\text{H}_2\text{O})_2$ complex,¹⁸ although that was not observed in the $7\text{AI}-(\text{H}_2\text{O})_2$ complex.

At the TDDFT level, the excited-state tautomerization energies for the $7\text{AI}-(\text{H}_2\text{O})_2$ complex agreed well with those at the MRPT2 level within 1.4 kcal/mol; however, those for the $7\text{AI}-(\text{H}_2\text{O})_2$ complex gave a maximum deviation of 4.7 kcal/mol compared with the MRPT2 values. Unlike the MRPT2 level, the TDDFT methods could not consistently reproduce the tautomerization energies of the $7\text{AI}-(\text{H}_2\text{O})_2$ and $7\text{AI}-(\text{H}_2\text{O})_2$ complexes. The ZPE-corrected barrier heights for the $7\text{AI}-(\text{H}_2\text{O})_2$ complex obtained by TDDFT methods using the 6-31G(d,p)

Table 4. Geometric Parameters of Reactant, Product, and Transition States for Excited-State Proton Transfer in 7AI–H₂O Complexes in Water^a

computational method	reactant		product	
	$r(\text{H}_{10}-\text{O}_{16})$	$r(\text{H}_{17}-\text{N}_6)$	$r(\text{N}_1-\text{H}_{10})$	$r(\text{O}_{16}-\text{H}_{17})$
CASSCF(8,8)/6-31G(d) ^b	2.189	2.118	2.168	2.200
CASSCF(10,9)/6-31G(d,p) ^c	2.255	2.098	2.179	2.254
CASSCF(10,9)/6-311G(d,p)	2.284	2.103	2.182	2.282
B3LYP/6-31G(d,p)	1.914	1.840	1.957	2.030
B3LYP/6-311+G(d,p)	2.073	1.833	1.929	2.192
CAM-B3LYP/6-31G(d,p)	1.885	1.837	1.949	1.982
CAM-B3LYP/6-311+G(d,p)	2.023	1.836	1.926	2.104
LC-BLYP/6-31G(d,p)	1.862	1.826	1.925	1.927
LC-BLYP/6-311+G(d,p)	1.982	1.823	1.903	2.025
M062X/6-31G(d,p)	1.908	1.867	2.000	1.994
M062X/6-311+G(d,p)	1.998	1.882	1.985	2.046
WB97XD/6-31G(d,p)	1.916	1.871	1.979	2.006
WB97XD/6-311+G(d,p)	2.046	1.858	1.953	2.101
computational method	transition state			
	$r(\text{N}_1-\text{H}_{10})$	$r(\text{H}_{10}-\text{O}_{16})$	$r(\text{O}_{16}-\text{H}_{17})$	$r(\text{H}_{17}-\text{N}_6)$
CASSCF(8,8)/6-31G(d) ^b	1.076	1.580	1.288	1.212
CASSCF(10,9)/6-31G(d,p) ^c	1.352	1.135	1.008	1.671
CASSCF(10,9)/6-311G(d,p)	1.364	1.124	0.996	1.727
B3LYP/6-31G(d,p)	1.242	1.276	1.177	1.348
B3LYP/6-311+G(d,p)	1.052	1.760	1.392	1.157
CAM-B3LYP/6-31G(d,p)	1.262	1.242	1.146	1.379
CAM-B3LYP/6-311+G(d,p)	1.061	1.687	1.397	1.150
LC-BLYP/6-31G(d,p)	1.275	1.221	1.135	1.387
LC-BLYP/6-311+G(d,p)	1.075	1.609	1.405	1.144
M062X/6-31G(d,p)	1.286	1.215	1.105	1.443
M062X/6-311+G(d,p)	1.304	1.198	1.111	1.433
WB97XD/6-31G(d,p)	1.271	1.230	1.140	1.384
WB97XD/6-311+G(d,p)	1.057	1.714	1.403	1.145

^a Bond distances are in Å. The solvent model is IEFPCM/UFF. Atomic radii from the UFF force field were scaled by 1.1. All hydrogens have individual spheres. ^b The Onsager solvation model was used. Ref 7. ^c Ref 42.

basis set were underestimated by between 1.99 and 3.85 kcal/mol compared with the MRPT2 value. When the 6-311+G(d,p) basis set was used, the TDDFT barriers showed better agreement with the MRPT2 values. However, as discussed previously, the TS structures were utterly dependent on the size of the basis sets. Figure 3 shows that the TS correlation points at the TDDFT level moved further away from the point of CASSCF using larger basis sets; the correlation points of H₁₀ and H₁₇ moved toward the reactant and product, respectively. Although the TDDFT barriers using the 6-311+G(d,p) basis set agreed better with the MRPT2 values, they should be used carefully because of their TS structures. As mentioned earlier, only the M062X level gave similar correlations with CASSCF, and the ZPE-corrected barrier using the larger basis sets was 4.9 kcal/mol, which is only 0.9 kcal/mol lower than the MRPT2 value.

Unlike the CASSCF level, only one TS structure was found for the 7AI–(H₂O)₂ complex at all TDDFT levels used in this study. Figure 4 shows that the TS correlation points of H₁₀ and H₁₉ were approximately in the middle of the two TS1 and TS2 points using the 6-31G(d,p) basis set but shifted toward the TS2 point using larger basis sets. Just as the TS structures depend on

the basis sets, so do barrier heights. All TDDFT barriers using the 6-31G(d,p) basis sets, except WB97XD, were underestimated compared with the corresponding MRPT2 value (Table 3). The ZPE-corrected barriers were very small, even smaller than zero in some cases. These barriers became even higher than the MRPT2 value of TS1 in most cases when the larger 6-311+G(d,p) basis set was used. As described above, the TS structures with larger basis sets are more like TS2 rather than TS1 at the CASSCF level, and the MRPT2 barrier of TS2 is much higher. Thus, the structural change of TS toward TS2 seems to increase the barrier heights at the TDDFT level. The correlation points at the M062X level varied very little and remained approximately in the same position in spite of using larger basis sets (Figure 4). The barrier height was 5.23 kcal/mol using the 6-311+G(d,p) basis set, which agreed very well with the MRPT2 value of TS1 using the 6-311G(d,p) basis set (only 0.4 kcal/mol smaller). For both 7AI–H₂O and 7AI–(H₂O)₂ complexes, the M062X level reproduced the barrier height quite well using the 6-311+G(d,p) basis set. Because the TS structures and barrier heights from the TDDFT method depend on the basis set size, one should use these methods very carefully, particularly for excited-state proton

Table 5. Geometric Parameters of Reactant, Product, and Transition States for Excited-State Proton Transfer in 7AI-(H₂O)₂ Complexes in Water^a

computational method	reactant			product		
	$r(\text{H}_{10}-\text{O}_{16})$	$r(\text{O}_{20}-\text{H}_{17})$	$r(\text{H}_{19}-\text{N}_6)$	$r(\text{N}_1-\text{H}_{10})$	$r(\text{O}_{16}-\text{H}_{17})$	$r(\text{O}_{20}-\text{H}_{19})$
CASSCF(8,8)/6-31G(d) ^b	1.992	1.894	2.006	2.060	1.914	2.040
CASSCF(10,9)/6-31G(d,p)	1.818	1.878	1.890	2.062	1.928	2.042
CASSCF(10,9)/6-311G(d,p)	1.825	1.898	1.905	2.071	1.939	2.056
B3LYP/6-31G(d,p)	1.686	1.688	1.699	1.822	1.743	1.818
B3LYP/6-311+G(d,p)	1.770	1.747	1.720	1.827	1.788	1.896
CAM-B3LYP/6-31G(d,p)	1.663	1.664	1.694	1.811	1.716	1.788
CAM-B3LYP/6-311+G(d,p)	1.746	1.723	1.718	1.813	1.758	1.857
LC-BLYP/6-31G(d,p)	1.639	1.629	1.682	1.784	1.674	1.744
LC-BLYP/6-311+G(d,p)	1.721	1.686	1.703	1.782	1.715	1.806
M062X/6-31G(d,p)	1.630	1.653	1.680	1.841	1.738	1.796
M062X/6-311+G(d,p)	1.745	1.753	1.741	1.853	1.788	1.862
WB97XD/6-31G(d,p)	1.685	1.703	1.726	1.831	1.749	1.799
WB97XD/6-311+G(d,p)	1.752	1.750	1.741	1.827	1.780	1.854

computational method	transition state					
	$r(\text{N}_1-\text{H}_{10})$	$r(\text{H}_{10}-\text{O}_{16})$	$r(\text{O}_{16}-\text{H}_{17})$	$r(\text{H}_{17}-\text{O}_{20})$	$r(\text{O}_{20}-\text{H}_{19})$	$r(\text{H}_{19}-\text{N}_6)$
CASSCF(8,8)/6-31G(d) ^b	1.043	1.668	1.002	1.563	1.345	1.159
CASSCF(10,9)/6-31G(d,p) TS1	1.642	1.005	1.080	1.330	1.014	1.602
CASSCF(10,9)/6-311G(d,p) TS1	1.681	0.995	1.090	1.308	1.007	1.619
CASSCF(10,9)/6-31G(d,p) TS2	1.035	1.693	0.994	1.580	1.371	1.132
CASSCF(10,9)/6-311G(d,p) TS2	1.030	1.719	0.984	1.620	1.351	1.142
B3LYP/6-31G(d,p)	1.148	1.388	1.130	1.317	1.313	1.195
B3LYP/6-311+G(d,p)	1.080	1.556	1.090	1.376	1.513	1.093
CAM-B3LYP/6-31G(d,p)	1.190	1.309	1.134	1.294	1.247	1.240
CAM-B3LYP/6-311+G(d,p)	1.085	1.523	1.091	1.359	1.480	1.100
LC-BLYP/6-31G(d,p)	1.297	1.185	1.139	1.273	1.134	1.362
LC-BLYP/6-311+G(d,p)	1.099	1.476	1.112	1.313	1.442	1.112
M062X/6-31G(d,p)	1.334	1.158	1.116	1.310	1.092	1.436
M062X/6-311+G(d,p) TS1	1.331	1.159	1.129	1.287	1.126	1.378
M062X/6-311+G(d,p) TS2	1.170	1.339	1.119	1.304	1.296	1.198
WB97XD/6-31G(d,p) TS1	1.314	1.171	1.134	1.288	1.134	1.365
WB97XD/6-31G(d,p) TS2	1.170	1.336	1.129	1.300	1.285	1.204
WB97XD/6-311+G(d,p)	1.092	1.498	1.110	1.323	1.525	1.083

^a Bond distances are in Å. The solvent model is IEFPCM/UFF. Atomic radii from the UFF force field were scaled by 1.1. All hydrogens have individual spheres. ^b The Onsager solvation model was used. Ref 7.

transfer reactions. No special benefit seemed to be gained from the long-range correction and the empirical dispersion in the TDDFT.

3.4. Effect of Solvation. The excited-state protropic tautomerization for 7AI in bulk solvents^{2,50–53} implies that solvation plays a key role in proton transfer dynamics. In the gas phase, Chaban and Gordon⁴ found that the excited-state proton transfer barrier was reduced to ≤ 6 kcal/mol when one water molecule was present. As such, the addition of more than one water molecule should lower the activation energy further, in agreement with the view of Siebrand and co-workers.⁵⁴ We performed IEFPCM calculations for the cyclic reactant, TS, and product at the TDDFT and CASSCF levels to understand the solvent effect. The cyclic structures of the reactant, product, and TS in solution were also confirmed by frequency calculations. Some optimized geometrical parameters for 7AI-H₂O and 7AI-(H₂O)₂

complexes in water are listed in Tables 4 and 5, respectively. The correlations between q_1 and q_2 for proton transfer in the 7AI-H₂O and 7AI-(H₂O)₂ complexes are depicted in Figures 5 and 6.

The H₁₀ correlation points moved to the upper left corner along the black line due to the solvent effect for the 7AI-H₂O reactant, irrespective of the computational level and size of the basis sets. However, those of H₁₇ were changed in the opposite way at the CASSCF level, but nearly unchanged at the TDDFT level (Figure 5). These results indicate that the H₁₀-O₁₆ and N₆-H₁₇ H bonds became longer and shorter, respectively, at the CASSCF level in water. However, at the TDDFT level, the N₆-H₁₇ H bond changed very little with a slight increase. In the 7AI-H₂O product, the H₁₀ correlation points moved slightly down to the left side due to the solvent effect; however, those of H₁₇ moved up to the right side. These results suggest that the

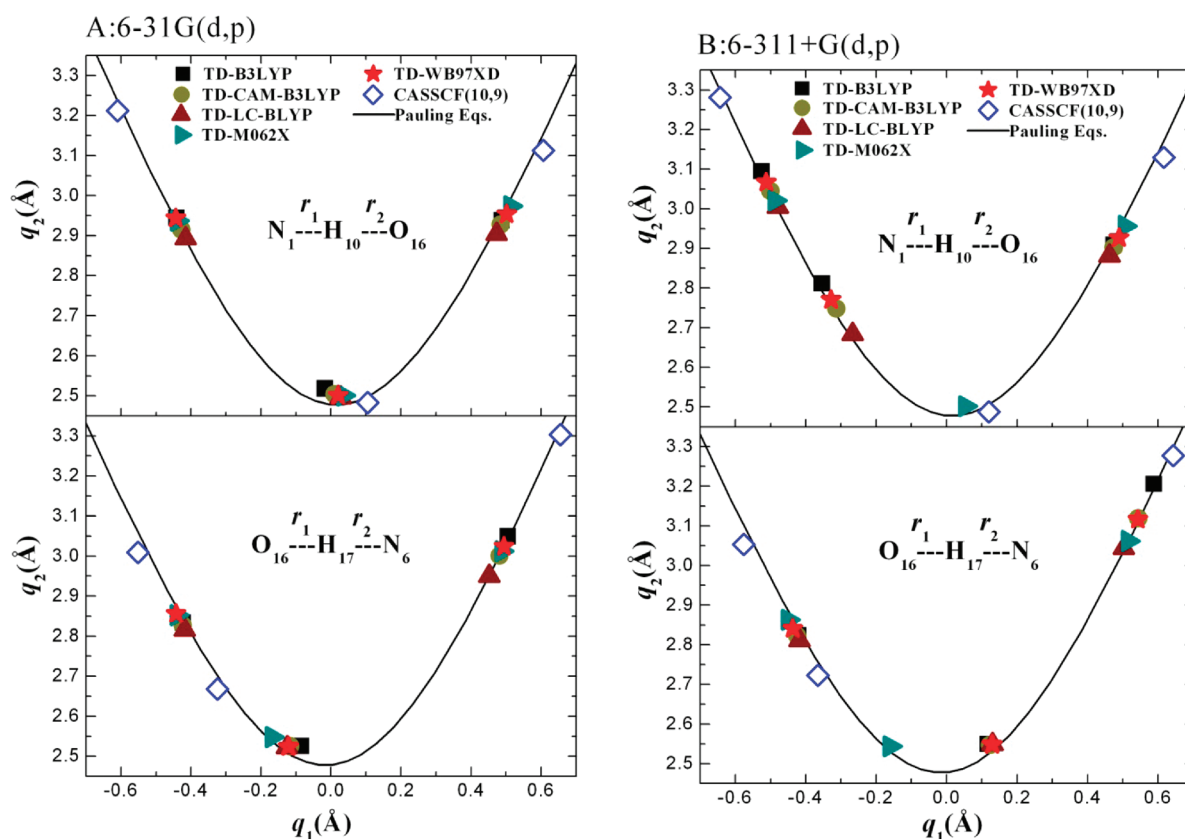


Figure 5. Correlation of the H-bond distances $q_2 = r_1 + r_2$ with the proton transfer coordinate $q_1 = 1/2(r_1 - r_2)$ for the 7AI–H₂O complex in water. Top, H₁₀ transfer; bottom, H₁₇ transfer. The solid line represents the correlation for equilibrium distances calculated with Pauling equations. The CASSCF points in B were calculated using the 6-311G(d,p) basis set.

N₁–H₁₀ and O₁₆–H₁₇ H-bond distances were slightly decreased and increased, respectively, in water. It is interesting to note that at the CASSCF level, two H bonds with a nitrogen atom as an acceptor (N₁–H₁₀ in reactant and N₆–H₁₇ in product) became shorter in water, whereas those with an oxygen acceptor became longer. These results suggest that the solvent effect could increase the excited-state basicity of the nitrogen atom in 7AI to generate slightly shorter and stronger H bonds.

At the TS, correlation of q_1 with q_2 was dependent on the solvent effect. The H₁₀ correlation points at the CASSCF level moved slightly to the right side along the Pauling equation line in water; however, those of the H₁₇ moved to the upper-left side (Figure 5). The differences in q_1 values for H₁₀ and H₁₇ transfers increased significantly at the CASSCF level; i.e., the solvent effect greatly increased the asynchronicity of the double proton transfer. At the TDDFT level using the 6-31G(d,p) basis set, the H₁₀ and H₁₇ q_1 values became slightly larger and smaller, respectively, in water. However, when the 6-311+G(d,p) basis set was used, the H₁₀ correlation points at all TDDFT levels except M062X moved greatly toward the upper-left side along the black line in water (Figure 5B), whereas those for H₁₇ transfer moved to the right side. The difference in q_1 values between H₁₀ and H₁₇ transfers increased significantly as well, but in the opposite way compared to that at the CASSCF level. The M062X level showed consistent correlation with the CASSCF level depending on the solvent effect. As a result, the TDDFT methods with the 6-311+G(d,p) basis set, except M062X, predicted completely different asynchronous double proton transfers, where the H₁₇ atom moved first

followed by the H₁₀ atom. This mechanism has the opposite order of H transfer when compared with the CASSCF level. Only the M062X level predicted the same mechanism as the CASSCF.

For the 7AI–(H₂O)₂ reactant at the CASSCF level, the H₁₀ and H₁₉ correlation points moved slightly to the upper-left side along the black line in water (Figures 4 and 6); however, those of H₁₉ rarely moved. For the product, the H₁₉ correlation points moved slightly to the upper-right side along the solid line, but those of H₁₀ and H₁₉ rarely moved. The correlation points for TS depend very much on the solvent effect. As described above, there are two TSs at the CASSCF level for two different ESTPT asynchronous mechanisms. At TS1, the H₁₀ and H₁₉ correlation points moved further away toward the right and left side in water, which indicates that the TS positions in terms of the H₁₀ and H₁₉ transfers became even more late (product-like) and early (reactant-like), respectively. At TS2, the H₁₀ and H₁₉ correlation points moved in the opposite way, which makes the H₁₀ and H₁₉ TS positions even earlier and later in water, respectively. These results suggest that the solvent effect increases asynchronicity of the ESTPT for both concerted mechanisms via TS1 and TS2. The H₁₇ correlation points for TS1 and TS2 were quite close in the gas phase but moved further apart in water; in particular, the TS2 point moved far away to the upper-left side. The position of H₁₇ at TS2 became very early in water. The increased asynchronicity in water induced the H₃O⁺-like and OH⁻-like moieties more clearly at TS1 and TS2, respectively. Even in a water solution, we were unable to find an intermediate, and every attempt led to either reactant or product.

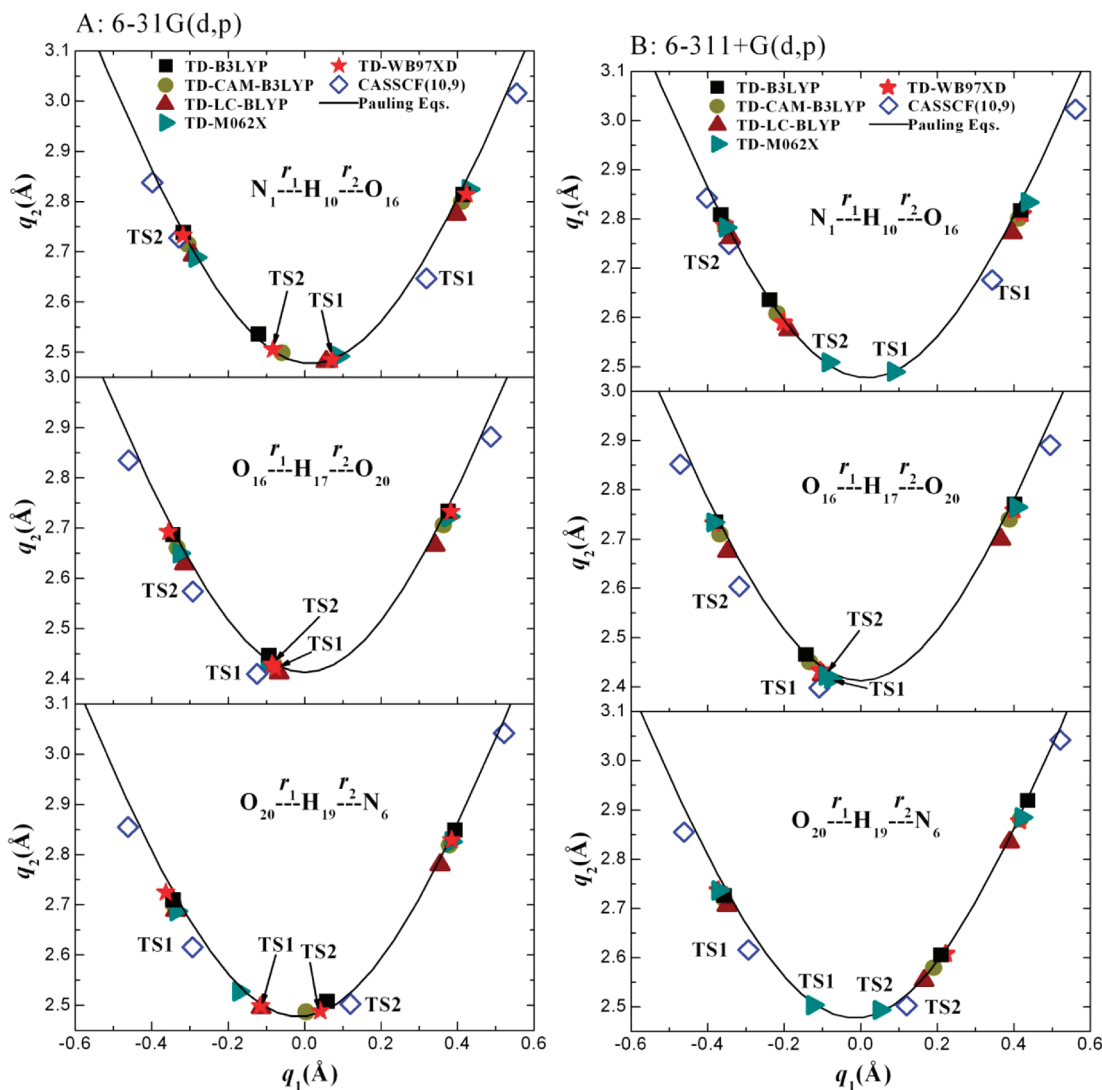


Figure 6. Correlation of the H-bond distances $q_2 = r_1 + r_2$ with the proton transfer coordinate $q_1 = 1/2(r_1 - r_2)$ for the $7\text{AI}-(\text{H}_2\text{O})_2$ complex in water. Top, H_{10} transfer; middle, H_{17} transfer; bottom, H_{19} transfer. The solid line represents the correlation for equilibrium distances calculated with Pauling equations. The CASSCF points in B were calculated using the 6-311G(d,p) basis set.

No TDDFT level predicted two TSs from two asynchronous paths in the gas phase. However, two TSs were found in water at the WD97XD/6-31G(d,p) and M062X/6-311+G(d,p) levels. As shown in Figure 6, two correlation points for TS1 and TS2 at these two levels were not separated as much as those at the CASSCF level, which implies that the asynchronicity would be predicted as being less.

Unlike the tautomerization energies in the gas phase, the MRPT2 level predicted slightly different energies in water for the $7\text{AI}-\text{H}_2\text{O}$ and $7\text{AI}-(\text{H}_2\text{O})_2$ complexes (Table 6). Tautomerization energy of the $7\text{AI}-\text{H}_2\text{O}$ complex was 4.5 kcal/mol lower (more exoergic) than that of the $7\text{AI}-(\text{H}_2\text{O})_2$ complex at the MRPT2/6-311G(d,p) level, which means that tautomerization in the $7\text{AI}-(\text{H}_2\text{O})_2$ complex is less favorable. The MRPT2 barrier height in the $7\text{AI}-\text{H}_2\text{O}$ complex with ZPE corrections was reduced by 2.8 kcal/mol in aqueous solution, which is consistent with experiments that the excited-state tautomerization was not observed in the gas phase but in solution. In the ESTPT of the $7\text{AI}-(\text{H}_2\text{O})_2$ complex, the MRPT2 barrier for TS1 was 3.9 kcal/mol lower than that for TS2; therefore

tautomerization occurred by a concerted but asynchronous mechanism via TS1. The barrier height for TS1 in water was almost unchanged compared to the value in the gas phase, which indicates that the solvent effect does not give any advantage to the ESTPT in aqueous solution. The MRPT2 barriers of the ESDPT in the $7\text{AI}-\text{H}_2\text{O}$ complex were only 1.4 and 0.8 kcal/mol higher in energy than those of the ESTPT in the $7\text{AI}-(\text{H}_2\text{O})_2$ complex with and without ZPE corrections, respectively. Considering the entropic disadvantage of forming the cyclic TS of three molecules, this energy difference seems to be too small to conclude that the ESTPT instead of the ESDPT is the preferable mechanism of the excited-state tautomerization in aqueous solution at room temperature.

Excited-state dipole moments in the gas phase and in solution are listed in Tables 3 and 6, respectively. The dipole moments of TS are larger than those of reactant at the CASSCF level. For the $7\text{AI}-\text{H}_2\text{O}$ complex, the dipole moment of the reactant and TS forms at the CASSCF(10,9)/6-311G(d,p) level in solution are 2.63 and 7.48 D, while those in the gas phase are 1.63 and 4.47 D, respectively. The changes in geometry upon solvation indicate an

Table 6. Reaction Energies, Barrier Heights, and Dipole Moments for the ESPT in 7AI–H₂O and 7AI–(H₂O)₂ Complexes in Solution^a

computational method	7AI–H ₂ O					7AI–(H ₂ O) ₂				
	ΔV (kcal/mol)	ΔE (kcal/mol)	μ (D)			ΔV (kcal/mol)	ΔE (kcal/mol)	μ (D)		
			R	TS	P			R	TS	P
CASSCF(8,8)/6-31G(d) ^b	17.94	–33.37	2.47	6.50	2.56	11.61	–32.83	2.07	9.05	3.19
CASSCF(10,9)/6-31G(d,p) ^c	14.29(12.14)	–31.81(–31.20)	2.65	6.97	2.61	14.66(11.87) ^d 13.21(10.39) ^e	–27.44(–27.78)	7.41	8.81 ^d 9.84 ^e	2.50
CASSCF(10,9)/6-311G(d,p)	13.48(11.64)	–32.17(–31.54)	2.63	7.48	2.59	15.23(12.79) ^d 13.74(10.83) ^e	–26.57(–26.95)	7.54	9.17 ^d 9.87 ^e	2.53
MRPT2/CASSCF(10,9)/6-31G(d,p)	5.93(3.78)	–22.20(–21.59)				4.84(2.05) ^d 6.31(3.49) ^e	–15.07(–15.42)			
MRPT2/CASSCF(10,9)/6-311G(d,p)	4.83(2.99)	–18.57(–17.94)				4.07(1.63) ^d 8.39(5.48) ^e	–14.10(–14.48)			
B3LYP/6-31G(d,p)	9.03(4.70)	–15.50(–15.18)	6.04	4.15	2.37	6.19(1.25)	–13.41(–12.88)	5.33	5.64	2.11
CAM-B3LYP/6-31G(d,p)	8.69(4.58)	–16.14(–15.64)	5.53	4.26	2.39	6.02(0.29)	–13.94(–13.23)	4.88	4.69	2.33
LC-BLYP/6-31G(d,p)	8.47(4.56)	–14.80(–14.13)	4.69	4.07	2.45	5.33(0.01)	–12.66(–11.80)	3.99	4.07	2.46
M062X/6-31G(d,p)	7.44(3.98)	–16.61(–16.00)	5.65	4.57	2.32	3.54(–0.38)	–14.13(–13.42)	5.01	4.94	2.24
WB97XD/6-31G(d,p)	10.38(6.17)	–16.30(–15.76)	5.65	4.31	2.40	7.94(2.18) ^d 7.89(2.38) ^e	–14.14(–13.47)	4.99	4.48 ^d 5.19 ^e	2.27
B3LYP/6-311+G(d,p)	9.65(6.90)	–14.59(–14.34)	6.09	7.17	2.49	5.40(2.76)	–12.93(–12.56)	5.49	8.51	2.10
CAM-B3LYP/6-311+G(d,p)	10.26(7.55)	–14.76(–14.26)	5.32	6.72	2.48	5.86(2.98)	–13.08(–12.57)	4.85	7.83	2.36
LC-BLYP/6-311+G(d,p)	10.96(8.14)	–13.01(–12.42)	4.41	5.78	2.51	6.32(2.83)	–11.48(–10.86)	3.96	6.52	2.50
M062X/6-311+G(d,p)	10.81(7.29)	–15.12(–14.65)	5.30	4.61	2.26	6.97(2.11) ^d 6.94(1.71) ^e	–13.53(–12.78)	5.06	4.74 ^d 5.31 ^e	2.24
WB97XD/6-311+G(d,p)	11.37(8.67)	–14.85(–14.46)	5.55	6.99	2.48	7.31(3.98)	–13.22(–12.87)	4.98	7.98	2.30

^a The numbers in parentheses include zero-point energies. The solvent model is IEFPCM/UFF. Atomic radii from the UFF force field were scaled by 1.1. All hydrogens have individual spheres. ^b Ref 7. ^c Refs 42 and 43. ^d TS1. ^e TS2.

enhancement of the ionic character of the transition states. The reactant and transition state of the 7AI–(H₂O)₂ complex, both in the gas phase and in solution at the CASSCF(10,9) level, are more polar than those of 7AI–H₂O complexes according to their dipole moments. At this level, the differences in dipole moments between the reactant and TS of the cyclic 1:1 complex are much larger than those of the 1:2 complex, which results in a stronger solvent effect to give lower barrier heights in the 1:1 complex, as listed in Tables 3 and 6. It is interesting to note that TS2 is slightly more polar than TS1. TD-DFT predicted excited-state dipole moments in a different way from the CASSCF(10,9) level of theory. The reactant and TS were more and less polar than those at the CASSCF levels, respectively. In addition to that, most TS dipole moments are slightly smaller than reactant dipole moments, except for the 7AI–(H₂O)₂ complex at the B3LYP, CAM-B3LYP, LC-BLYP, and WB97XD levels using the 6-311+G(d,p) basis sets.

All barrier heights from TDDFT depend on the solvent effect (Table 6). Because most DFT methods with a solvent effect predicted quite different TS structures from those at the CASSCF level, it may not be possible to compare barrier heights from these methods with MRPT2 values. It is interesting to note that most TDDFT barrier heights in water were larger than those in the gas phase, which is opposite of the CASSCF values. Although two asynchronous paths for the 7AI–(H₂O)₂ complex were found at the WB97XD/6-31G(d,p) and M062X/6-311+G(d,p) levels in water, there was no difference in the barrier heights between the two paths at both levels. These TDDFT

levels failed to distinguish one preferable path from the other. In addition, the barrier heights from the WB97XD/6-31G(d,p) and M062X/6-311+G(d,p) methods were 3.9 and 2.9 kcal/mol larger than the MRPT2 value of the preferable path, respectively. Further systematic study is necessary to test whether current TDDFT methods including solvent effects can be used to understand excited-state proton transfer reactions.

4. CONCLUSIONS

In the present work, systematic investigations on tautomerization processes were performed on 7AI–(H₂O)_{*n*} (*n* = 1, 2) complexes using TDDFT and CASSCF methods. Complete geometry optimization in the gas phase and in solution was performed in the S₁ state. Comparisons between the TDDFT results and CASSCF values were made carefully. The key conclusions are summarized as follows.

The dynamic electron correction is very important to the energetics of the excited-state tautomerization in 7AI–(H₂O)_{*n*} (*n* = 1, 2) complexes.

For the 7AI–(H₂O)₂ complex, CASSCF levels predicted two concerted but asynchronous paths of proton transfer in the excited-state tautomerization: one where the proton from the pyrrole ring of 7AI moved first to water and the other where the water proton moved first to the pyridine ring. Because of the asynchronous motion of protons, the H₃O⁺-like and HO[–]-like moieties were generated in the TS of the former and the latter, respectively. No difference was found between the barrier heights

of the two paths without considering the dynamic electron correlation. However, the MRPT2 correction clearly showed that the former path was much preferable to the latter.

In the gas phase, the barrier of the $7\text{AI}-(\text{H}_2\text{O})_2$ complex was 1.6 kcal/mol, which is much lower than that of the $7\text{AI}-\text{H}_2\text{O}$ complex, supporting the argument that excited-state tautomerization might occur by forming H-bonded complexes with two water molecules. The solvent effect reduced the $7\text{AI}-\text{H}_2\text{O}$ MRPT2 barrier by 2.8 kcal/mol, which is consistent with experiments in that the excited-state tautomerization was not observed in the gas phase but in solution.

All DFT methods used in this study, namely, the hybrid functional B3LYP, M062X, the functional with long-range correction CAM-B3LYP, LC-BLYP, and the WB97XD functional which includes empirical dispersion, underestimated H-bond distances in the reactant and product by about 0.1–0.4 Å. The tautomerization energies obtained by the TDDFT methods were slightly underestimated compared with the MRPT2 values, except for the $7\text{AI}-\text{H}_2\text{O}$ complex in the gas phase. No significant benefits, in terms of both structural and energetic prediction, were found from the DFT methods with long-range correction or empirical dispersion. In terms of the TS structures and the barrier height in the gas phase, the M062X method agreed best with the CASSCF with a MRPT2 correction.

At all TDDFT levels used in this study, the TS structures and barrier heights greatly depend on the basis set and the solvent effect. Only two methods, WB97XD/6-31G(d,p) and M062X/6-311+G(d,p), predicted two TSs for two asynchronous paths for the $7\text{AI}-(\text{H}_2\text{O})_2$ complex. However, the two barrier heights were almost the same in energy, and larger than the preferable MRPT2 value. Further systematic study is necessary to test whether current TDDFT methods, including solvent effects, can be used to correctly understand excited-state proton transfer reactions.

■ ASSOCIATED CONTENT

S Supporting Information. Geometric parameters of reactant, product, and transition states. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: yhkim@khu.ac.kr.

■ ACKNOWLEDGMENT

This work was supported by a grant from Kyung Hee University in 2010. We are pleased to acknowledge the support of the Center for Academic Computing at Kyung Hee University for the computing resources.

■ REFERENCES

- (1) Miyazaki, T. *Atom Tunneling Phenomena in Physics, Chemistry, and Biology*; Miyazaki, T., Ed.; Springer: Berlin, 2004; pp 1–375.
- (2) Taylor, C. A.; El-Bayoumi, M. A.; Kasha, M. *Proc. Nat. Acad. Sci.* **1969**, *63*, 253–260.
- (3) Gordon, M. S. *J. Phys. Chem.* **1996**, *100*, 3974–3979.
- (4) Chaban, M. G.; Gordon, M. S. *J. Phys. Chem. A* **1999**, *103*, 185–189.
- (5) Nakano, H. *J. Chem. Phys.* **1993**, *99*, 7983–7992.

- (6) Casadesús, R.; Moreno, M.; Lluch, J. M. *Chem. Phys.* **2003**, *290*, 319–336.
- (7) Fernandez-Ramos, A.; Smedarchina, Z.; Siebrand, W.; Zgierski, M. Z. *J. Chem. Phys.* **2001**, *114*, 7518–7526.
- (8) Kina, D.; Nakayama, A.; Noro, T.; Taketsugu, T.; Gordon, M. S. *J. Phys. Chem. A* **2008**, *112*, 9675–9683.
- (9) Chapman, C. F.; Maroncelli, M. *J. Phys. Chem.* **1992**, *96*, 8430–8441.
- (10) Chen, Y.; Gai, F.; Petrich, J. W. *J. Am. Chem. Soc.* **1993**, *115*, 10158–10166.
- (11) Chou, P.-T.; Martinez, M. L.; Cooper, W. C.; McMorrow, D.; Collins, S. T.; Kasha, M. *J. Phys. Chem.* **1992**, *96*, 5203–5205.
- (12) Huang, Y.; Arnold, S.; Sulkes, M. *J. Phys. Chem.* **1996**, *100*, 4734–4738.
- (13) Folmer, D. E.; Wisniewski, E. S.; Stairs, J. R.; Castleman, A. W., Jr. *J. Phys. Chem. A* **2000**, *104*, 10545–10549.
- (14) Nakajima, A.; Hirano, M.; Hasumi, R.; Kaya, K.; Watanabe, H.; Carter, C.; Williamson, J. M.; Miller, T. A. *J. Phys. Chem. A* **1997**, *101*, 392–398.
- (15) Hara, A.; Sakota, K.; Sekiya, H. *Chem. Phys. Lett.* **2005**, *407*, 30–34.
- (16) Vu, T. B. C.; Kalkman, I.; Meerts, W. L.; Svartsov, Y. N.; Jacoby, C.; Schmitt, M. *J. Chem. Phys.* **2008**, *128*, 214311–8.
- (17) Sakota, K.; Jouvét, C.; Dedonder, C.; Fujii, M.; Sekiya, H. *J. Phys. Chem. A* **2010**, *114*, 11161–11166.
- (18) Yokoyama, H.; Watanabe, H.; Omi, T.; Ishiuchi, S.; Fujii, M. *J. Phys. Chem. A* **2001**, *105*, 9366–9374.
- (19) Casida, M. E.; Chong, D. P. *Recent Advances In Density Functional Methods: Part I*; Chong, D. P., Ed.; World Scientific: Singapore, 1995; Chapter 5, pp 155–192.
- (20) Dreuw, A.; Head-Gordon, M. *Chem. Rev.* **2005**, *105*, 4009–4037.
- (21) Dierksen, M.; Grimme, S. *J. Chem. Phys.* **2004**, *120*, 3544–3554.
- (22) Dierksen, M.; Grimme, S. *J. Phys. Chem. A* **2004**, *108*, 10225–10237.
- (23) Van Leeuwen, R.; Baerends, E. J. *Phys. Rev. A* **1994**, *49*, 2421–2431.
- (24) Cave, R. J.; Zhang, F.; Maitra, N. T.; Burke, K. *Chem. Phys. Lett.* **2004**, *389*, 39–42.
- (25) Maitra, N. T.; Zhang, F.; Cave, R. J.; Burke, K. *J. Chem. Phys.* **2004**, *120*, 5932–5937.
- (26) Starcke, J. H.; Wormit, M.; Schirmer, J.; Dreuw, A. *Chem. Phys.* **2006**, *329*, 39–42.
- (27) Tozer, D. J.; Amos, R. D.; Handy, N. C.; Roos, B. J.; Serrano-Andres, L. *Mol. Phys.* **1999**, *97*, 859–868.
- (28) Dreuw, A.; Weisman, J. L.; Head-Gordon, M. *J. Chem. Phys.* **2003**, *119*, 2943–2946.
- (29) Sobolewski, A. L.; Domcke, W. *Chem. Phys.* **2003**, *294*, 73–83.
- (30) Dreuw, A.; Head-Gordon, M. *J. Am. Chem. Soc.* **2004**, *126*, 4007–4016.
- (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.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*; Gaussian, Inc.: Wallingford, CT, 2009.
- (32) Schmidt, M. W.; Baldrige, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.

Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347–1363.

- (33) Furche, F.; Ahlrichs, R. *J. Chem. Phys.* **2002**, *117*, 7433–7447.
- (34) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- (35) Yanai, T.; Tew, D.; Handy, N. *Chem. Phys. Lett.* **2004**, *393*, 51–57.
- (36) Iikura, H.; Tsuneda, T.; Yanai, T.; Hirao, K. *J. Chem. Phys.* **2001**, *115*, 3540–3544.
- (37) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215–241.
- (38) Cancès, E.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *107*, 3032–3041.
- (39) Cossi, M.; Barone, V.; Mennucci, B.; Tomasi, J. *Chem. Phys. Lett.* **1998**, *286*, 253–260.
- (40) Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, *106*, 5151–5158.
- (41) Svartsov, Y. N.; Schmitt, M. *J. Chem. Phys.* **2008**, *128*, 214310–9.
- (42) Duong, M. P. T.; Kim, Y. H. *J. Phys. Chem. A* **2010**, *114*, 3403–3410.
- (43) Duong, M. P. T.; Park, K. S.; Kim, Y. H. *J. Photochem. Photobiol. A: Chem.* **2010**, *214*, 100–107.
- (44) Limbach, H. H.; Pietrzak, M.; Benedict, H.; Tolstoy, P. M.; Golubev, N. S.; Denisov, G. S. *J. Mol. Struct.* **2004**, *706*, 115–119.
- (45) Limbach, H. H.; Lopez, J. M.; Kohen, A. *Phil. Trans. R. Soc. London, Ser. B* **2006**, *361*, 1399–1415.
- (46) Limbach, H. H. In *Hydrogen-Transfer Reactions*; Schowen, R. L., Klinman, J. P., Hynes, J. T., Limbach, H. H., Eds; Wiley: Weinheim, Germany, 2007; Chapter 6, pp 135–221.
- (47) Garrett, B. C.; Truhlar, D. G. *J. Am. Chem. Soc.* **1979**, *101*, 4534–4548.
- (48) Johnston, H. S. *Gas Phase Reaction Rate Theory*; Ronald Press: New York, 1966; pp 1–362.
- (49) Lopez, J. M.; Männle, F.; Wawer, I.; Buntkowsky, G.; Limbach, H. H. *Phys. Chem. Chem. Phys.* **2007**, *9*, 4498–4513.
- (50) McMorrow, D.; Aartsma, T. *Chem. Phys. Lett.* **1986**, *125*, 581–585.
- (51) Kojinberg, J.; Huizer, A. H.; Varma, C. A. G. O. *J. Chem. Soc., Faraday Trans. II* **1988**, *84*, 1163–1175.
- (52) Moog, R. S.; Maroncelli, M. *J. Phys. Chem.* **1991**, *95*, 10359–10369.
- (53) Smirnov, A. V.; English, D. S.; Rich, R. L.; Lane, J.; Teyton, L.; Schwabacher, A. W.; Luo, S.; Thornburg, R. W.; Petrich, J. W. *J. Phys. Chem. B* **1997**, *101*, 2758–2769.
- (54) Smedarchina, Z.; Siebrand, W.; Fernandez-Ramos, A.; Gorb, L.; Leszczynski, J. *J. Chem. Phys.* **2000**, *112*, 566–573.

Parameterization of a B3LYP Specific Correction for Noncovalent Interactions and Basis Set Superposition Error on a Gigantic Data Set of CCSD(T) Quality Noncovalent Interaction Energies

Severin T. Schneebeli, Arteum D. Bochevarov, and Richard A. Friesner*

Department of Chemistry, Columbia University, New York, New York 10027, United States

S Supporting Information

ABSTRACT: A vast number of noncovalent interaction energies at the counterpoise corrected CCSD(T) level have been collected from the literature to build a diverse new data set. The whole data set, which consists of 2027 CCSD(T) energies, includes most of the published data at this level. A large subset of the data was then used to train a novel, B3LYP specific, empirical correction scheme for noncovalent interactions and basis set superposition error (abbreviated as B3LYP-MM). Results obtained with our new correction scheme were directly compared to benchmark results obtained with B3LYP-D3 and M06-2X (two popular density functionals designed specifically to accurately model noncovalent interactions). For noncovalent complexes dominated by dispersion or dipole-dipole interactions, all three tested methods give accurate results with the medium-sized aug-cc-pVDZ basis set with MUEs of 0.27 (B3LYP-MM), 0.32 (B3LYP-D3), and 0.47 kcal/mol (M06-2X) (with explicit counterpoise corrections). These results validate both B3LYP-D3 and M06-2X for interactions of this type using a much larger data set than was presented in prior work. However, our new dispersion correction scheme shows some clear advantages for dispersion and dipole-dipole dominated complexes with the small LACVP* basis set, which is very popular in use due to its low associated computational cost: The MUE for B3LYP-MM with the LACVP* basis set for this subset of complexes (without explicit counterpoise corrections) is only 0.28 kcal/mol, compared to 0.65 kcal/mol for M06-2X or 1.16 kcal/mol for B3LYP-D3. Additionally, our new correction scheme also shows major improvements in accuracy for hydrogen-bonded systems and for systems involving ionic interactions, for example, cation- π interactions. Compared to B3LYP-D3 and M06-2X, we also find that our new B3LYP-MM correction scheme gives results of higher or equal accuracy for a large data set of conformer energies of di- and tripeptides, sugars, and cysteine.

INTRODUCTION

Density Functional Theory (DFT)⁷ has become an indispensable computational method to solve real world problems with quantum mechanics. The success of DFT is mainly based on its excellent balance between computational cost and accuracy, which is obtained with popular density functionals, such as B3LYP.^{8–11} B3LYP has shown good accuracy and transferability for thousands of real world systems studied during the past 15 years and therefore has become the most popular density functional for application studies in the chemical literature. However, even though B3LYP shows remarkable accuracy for many real world problems,^{12,13} it fails to accurately represent London dispersion interactions.^{14–17} Since this type of weak noncovalent interaction plays a crucial role in chemistry^{18–20} and biology,^{21–23} this represents a serious limitation for the B3LYP functional. A variety of ways to improve upon B3LYP for noncovalent interactions have been proposed in the recent literature, ranging from new density functionals^{24–26} such as M06-2X² to *a posteriori* corrections for existing density functionals,^{1,27–31} to name just a few of the most popular approaches. For a more detailed discussion of such methods, we refer the reader to the recent review literature.^{32–35}

It becomes an important task to benchmark the significant number of promising DFT methods in order to be able to

recommend the best ones to the users. For this, it is crucial to have benchmark data sets as large and diverse as possible in hand. The field of producing benchmark interaction energy data has been pioneered by Hobza and co-workers with their S22 data set,^{36,37} which has become a standard test set for noncovalent interactions. Since the publication of the S22 data set, a substantial amount of new high level benchmark noncovalent interaction energy data has appeared in the literature (for all citations, see the Supporting Information). However, since most of this data is scattered throughout the literature, assembling all of it into one database is necessary, if one wants to use it for benchmarking or training purposes. This task has been started by Hobza and co-workers with the Benchmark Energy and Geometry Database (BEGDB) database,³⁸ which contains most of the benchmark data produced in the Hobza group. In this work, we have tried to complete one of the tasks started with the BEGDB project, which is to create one large benchmark interaction energy database that contains almost all CCSD(T) data currently available in the literature. With a thorough literature search, we were able to assemble a data set of 2027 counterpoise corrected interaction energies at the CCSD(T) level. Furthermore, we also compiled a database of CCSD(T)/CBS level

Received: November 11, 2010

Published: February 11, 2011

conformational energies of di- and tripeptides, sugars, and cysteine from the literature.^{1,38,54} For future use, our databases are supplied as a whole in the Supporting Information in their current forms. However, an updated version of our noncovalent interaction energy database is also available via the internet at <http://www.noncovalent.friesnerlab.com>.

We employed our very large and diverse interaction energy database to benchmark two DFT methodologies that both have displayed good performance for noncovalent interactions, M06-2X² (a relatively new density functional) and B3LYP-D3¹ (an *a posteriori* correction scheme for the popular B3LYP density functional) when evaluated using smaller data sets. While B3LYP-D3 involves an empirical functional form as well as a few global, empirical parameters, most of its parameters are calculated using *ab initio* methods. Therefore, B3LYP-D3 is generally referred to as a *semiempirical* correction scheme.¹ Using part of our new interaction energy and conformational energy data set as the training set, we also parametrized a novel empirical correction scheme for the B3LYP functional (abbreviated as B3LYP-MM). Our correction scheme not only improves the treatment of London dispersion interactions with B3LYP but also includes a version that corrects for basis set superposition error (BSSE)³⁹ if this is not treated explicitly. It was specifically designed to yield good accuracy with any basis set with and without explicit counterpoise corrections.⁴⁰ Since B3LYP-D3 and M06-2X were not designed to correct for the large BSSE present with small basis sets, it is not surprising that we find major improvements in accuracy compared to the other benchmarked methods for the small LACVP*⁴¹ basis set (which is very popularly used because of its relatively low computational cost) without explicit counterpoise corrections. While with B3LYP-MM we get an overall mean unsigned error (MUE) of only 0.41 kcal/mol without explicit counterpoise correction and the LACVP* basis set, the corresponding MUEs with the other methods are significantly larger with 1.20 kcal/mol for M06-2X and 2.11 kcal/mol for B3LYP-D3. With our new correction scheme, special care was also taken to obtain high accuracy for hydrogen-bonded systems (for which BSSE is often especially large due to the small distances involved) and for systems with ionic interactions, including those with cation- π interactions.

Our correction is simply added to the B3LYP energy and depends only on nuclear coordinates in the same spirit as the correction schemes developed by Grimme and co-workers.^{1,28} It therefore has the big advantage of very low numerical complexity, which makes implementation of gradients and second derivatives into any existing quantum chemical code a relatively simple task. Furthermore, the time needed to calculate the correction is negligible, making the corrected DFT as fast as the original uncorrected version. The correction consists of three additive parts: A simple Lennard-Jones potential, a linear hydrogen bonding correction term, and a linear cation- π correction term. Currently, the parameters of the Lennard-Jones correction term depend only on atomic numbers and are independent of the atomic environments, while the hydrogen bonding correction parameters are identical for all types of hydrogen bonds.

Applying combination rules for the pairwise parameters of the Lennard-Jones terms allows us to keep the total number of freely adjustable parameters low. Currently, the correction contains only one parameter for each atom type (currently H, C, N, O, F, S, and Cl are implemented). For positive ions such as Li⁺ or Na⁺, Lennard-Jones terms are not needed. Lennard-Jones correction terms are also

excluded for ammonium hydrogens (any hydrogen attached to a positively charged sp³ nitrogen) since such hydrogens carry a substantial positive charge. Besides, there are four additional adjustable parameters, one global scaling factor for van der Waals radii, two parameters for the hydrogen bonding correction, and one adjustable parameter for the cation- π correction term. All 11 adjustable parameters were optimized via a least-squares fit to our very large and diverse data set of noncovalent interaction energies as well as to the data set of conformational energies.

Our correction scheme is able to yield better accuracy for small basis sets and hydrogen bonded and charged systems, mainly due to the following five major differences from similar solutions developed by others:^{1,28,32}

- (1) The functional form of our dispersion correction consists of a simple Lennard-Jones 6-12 potential, in contrast to most other correction schemes that contain attractive r^{-6} and r^{-8} terms damped at short distances.
- (2) Our correction is only applied for atom pairs more than three covalent bonds apart from each other, since we believe that special corrections should be developed for the covalent bonding region. Such corrections for covalent bonds as well as corrections for transition metals are currently being developed in our group.^{42–45} Our final goal is to merge all different correction types into a general correction scheme for B3LYP.
- (3) Hydrogen bonds are treated specially with our correction scheme. First, the Lennard-Jones dispersion correction term is not included for any hydrogen-heavy atom pair involved in a hydrogen bond. Second, a linear repulsive correction term is applied for all hydrogen-heavy atom pairs involved in a hydrogen bond, mainly in order to correct for BSSE (which we find to be especially strong for hydrogen bonds due to the short distances between the atoms involved).
- (4) Cation- π interactions obtain a special treatment with our correction scheme as well. First, the Lennard-Jones dispersion correction terms are not included for any positively charged metal ions such as Li⁺ or Na⁺ or for any ammonium hydrogens (any hydrogen attached to a positively charged sp³ nitrogens) since the positive charge of an ammonium type cation is mostly localized on the hydrogens. Second, a linear repulsive correction term is applied for cation- π interactions involving simple metal cations such as Li⁺ or Na⁺.
- (5) Since BSSE is highly dependent on the basis set used, the parameters of the correction were made dependent on the basis set and on whether counterpoise corrections are applied.

The paper is organized as follows. In the Methods section, our data set and correction scheme are discussed in detail. In the Results section, the performance of our B3LYP-MM correction scheme will be compared to the accuracy of M06-2X and B3LYP-D3. Finally, in the Conclusions, we summarize the comparisons of the various approaches.

METHODS

Interaction Energy Database. We assembled our noncovalent interaction energy database from 34 publications in the literature (the citations for all of the literature used for the database are given in the Supporting Information). Relevant data from the BEGDB project³⁸ were downloaded from <http://www.begdb.com>. Some publications did not report molecular

Table 1. Noncovalent Interaction Types Present in the Interaction Energy Database (Only Data Used for the Benchmark Studies in This Work Were Counted; for All of the Data, See the Supporting Information)

	number of data points	interaction energy range (min/max in kcal/mol)
dispersion and/or dipole-dipole dominated	1036	−11.83/3.71
hydrogen bonded, not charged	137	−22.95/0.61
with ionic interactions	542	−110.80/4.62

coordinates, but we were able to obtain most of them through correspondence with the authors.

The database contains mostly interaction energies at the counterpoise corrected CCSD(T)/CBS level, which represents the current gold standard for benchmark noncovalent interaction energies of medium-sized molecules.^{46–49} The acronym CCSD(T)/CBS represents MP2^{50,51} interaction energies (extrapolated to the complete basis set limit), which have been corrected for higher order correlation effects at the CCSD(T) level^{52,53} with a small- or medium-sized basis set.^{36,48} Recent studies have shown that for some complexes of the S22 data set, errors due to small basis sets employed to compute the CCSD(T) correction terms can be as large as 0.6 kcal/mol.^{64,65} Unfortunately, up to date, only a few highly accurate CCSD(T)/CBS interaction energies (with the CCSD(T) correction term extrapolated to the basis set limit) exist due to the extraordinary cost of such calculations, especially when applied to medium-sized molecules. Since for this study our main goal was to build a very large and diverse data set with reasonable accuracy, we assembled our interaction energy data set from almost all CCSD(T)/CBS interaction energies published in the literature without further pruning of the data set with respect to quality of the CCSD(T) corrections. The data set will however be updated in the future to reflect better quality CCSD(T) interaction energies as they become available in the literature. Furthermore, the very flexible B3LYP-MM methodology will easily allow refitting of the correction parameters to a data set of more accurate interaction energies as these become available. Besides CCSD(T)/CBS level interaction energies, several data points with interaction energies at the counterpoise corrected CCSD(T) level obtained with a medium or large basis set are also included in the database. In total, the database contains 2027 counterpoise corrected CCSD(T) interaction energy data points. For the benchmark studies reported in this work, noncovalent complexes with strongly repulsive interaction energies (>5.0 kcal/mol) were not included, mainly because the precise accuracy of the model at such geometries is generally unimportant in practical applications, as these geometries are rarely occupied due to the strongly repulsive interaction energy. Complexes containing elements other than H, C, N, O, F, S, Cl, Li, and Na were not included as well, since currently parameters for B3LYP-MM were only fit for these elements. Even though a few complexes were not used for the benchmark studies, their best estimate interaction energies as well as their coordinates are still given in the Supporting Information for potential future use.

The database contains equilibrium structures (96 data points, optimized at the CCSD(T), RI-MP2, MP2, or DFT level) and nonequilibrium structures (1931 data points). It also spans the whole range of noncovalent interactions, ranging from weakly bound dispersion dominated complexes over hydrogen bonded ones to dimers with strong electrostatic interactions. A split up of the database (only the data points that were used for the benchmark studies) into different interaction energy type categories is shown in Table 1.

Benchmark Data Set of Relative Conformational Energies.

A data set of conformer energies at the CCSD(T)/CBS level containing di- and tripeptides as well as sugars and cysteine was compiled from the literature as well. The collection of all possible energy gaps between the different conformers (a total of 700) was used as the benchmark set. The peptide benchmark energies were taken from the BEGDB project of Hobza and co-workers,^{38,54} while benchmark data for the sugars and cysteine was taken from the SCONF and the CCONF data sets of Grimme and co-workers.¹ All benchmark conformer energies as well as the corresponding structures are given together with their original citations in the Supporting Information.

Computational Details. All calculations were carried out with the Jaguar software package.⁵⁵ The pseudospectral methodology,^{56–60} which significantly speeds up the SCF iterations, was employed. Default grids and SCF convergence criteria as implemented in Jaguar were used. All structures were taken from the benchmark energy database and were not further optimized. Interaction energies were obtained by subtracting the energies of the monomers (in the geometries of the complex) from the energy of the complex. Counterpoise corrected interaction energies were obtained by employing the methodology of Boys and Bernardi.⁴⁰ With this methodology, the energies of the monomers are calculated in the basis set of the complex; i.e., basis functions are placed at the positions of all the atoms present in the complex but absent in the monomer. B3LYP-D3¹ dispersion corrections were calculated with the program provided on Professor Grimme's Web site (at <http://toc.uni-muenster.de/DFTD3>).

Near Linear Dependencies of Basis Functions. When one tries to solve the SCF equations for noncovalently bound complexes, one is often confronted with the well-known problem of numerical instabilities arising due to near linear dependencies of nonorthogonal Gaussian basis functions on atoms close in space.⁶¹ The near linear dependencies are especially pronounced for large basis sets with diffuse functions and can lead to numerical instabilities, especially if approximate numerical techniques (such as the pseudospectral method) are employed to compute integrals. Usually, the closer the noncovalently bonded atoms, the more pronounced the numerical instabilities. The near linear dependencies are manifested in very small eigenvalues of the atomic orbital overlap matrix. Therefore, the simplest solution to the problem of basis set overcompleteness is to diagonalize the overlap matrix to get a set of *canonical* orbitals as eigenvectors. Then, all *canonical* orbitals with a corresponding eigenvalue below a threshold ϵ are discarded. This simple methodology effectively removes near linear dependencies of the basis functions and has been successfully applied with Jaguar with $\epsilon = 5.0 \times 10^{-4}$.

One caveat with specifying a fixed eigenvalue cutoff ϵ is that the number and type of canonical orbitals removed can vary with the molecular geometry. Since the total number of canonical orbitals can change abruptly between two infinitesimally close geometries, the potential energy surfaces can become discontinuous. With $\epsilon = 5.0 \times 10^{-4}$, these discontinuities are virtually

nonexistent with the LACVP* basis set. With the aug-cc-pVDZ basis set^{3–6} and that ε value, the discontinuities are usually small (below 0.1 kcal/mol); they can however in a few cases reach larger values up to 0.3 kcal/mol. We believe that for most applications discontinuities of this magnitude are tolerable, especially since errors from other sources (e.g., stemming from the approximate treatment of electron correlation effects in DFT or from approximate solvation models, etc.) are often of much larger magnitude. However, if very high accuracy is required, a lower cutoff ε might have to be employed in order to minimize the discontinuities further.

Another caveat is that the number and type of canonical orbitals removed can vary between calculations that are used for energy comparisons. For example, since near linear dependencies of basis functions are often present for noncovalent dimers, but absent from the monomeric structures, the dimer is likely to contain fewer canonical orbitals than both monomers together if a fixed value of ε is employed. This will lead to an increased energy of the dimer relative to the energies of the monomers and therefore lead to a less stable complex. However, if the eigenvalue cutoff ε gets small enough, the canonical orbitals removed are already well represented by others, and their removal does not significantly increase the energy of the molecule. Furthermore, the destabilization effect on a complex due to the removal of canonical orbitals is in the opposite direction of the BSSE, which is always stabilizing the complex. Therefore, one might expect some error cancellation with the BSSE if a fixed ε is employed. In practice, we have found that, with $\varepsilon = 5.0 \times 10^{-4}$, interaction energies at the noncounterpoise corrected B3LYP/aug-cc-pVDZ level agree within a few tenths of a kilocalorie per mole with their counterpoise corrected counterparts. The same was also found for interaction energies at the B3LYP/aug-cc-pVTZ level. With the LACVP* basis set, there are usually no canonical orbital eigenvalues below $\varepsilon = 5.0 \times 10^{-4}$, such that the BSSE is not reduced with this methodology.

Corrections for London Dispersion and Basis Set Superposition Error (BSSE). As are many of the numerous correction schemes developed by others,^{1,28,32} our correction is based on an *a posteriori* energy correction, solely dependent on nuclear coordinates. Our correction, which is comprised of a London dispersion correction (E_{LDC}), a hydrogen bonding correction term (E_{HBC}), and a cation- π correction term ($E_{\pi+}$), is simply added to the DFT energy, as described in eq 1:

$$E_{\text{B3LYP-MM}} = E_{\text{B3LYP}} + E_{\text{LDC}} + E_{\text{HBC}} + E_{\pi+} \quad (1)$$

We found that the errors of B3LYP for hydrogen bonded systems and structures with cation- π interactions were very different from the errors observed for dispersion and/or dipole-dipole bound complexes. While B3LYP strongly underestimates interaction energies of most dispersion and dipole-dipole bound complexes, it often overestimates interaction energies of hydrogen bonded complexes or such with cation- π interactions, especially if no explicit counterpoise corrections are applied. Major reasons for this special behavior of hydrogen bonded complexes and systems with cation- π interactions are the very strong electrostatic interactions involved, and the weak covalent nature of the resulting interactions. To account for the special nature of hydrogen bonds and cation- π interactions, the London dispersion correction term is replaced with a special hydrogen bond or cation- π correction term (described below) for

hydrogen-bonded H \cdots acceptor pairs as well as for cation- π bonded atom pairs.

The London dispersion correction is described as a sum of Lennard-Jones (LJ) functions (eq 2), where the sum loops over all atom pairs with atoms at least four covalent bonds apart from each other.

$$E_{\text{LDC}} = \sum_{i < j} \varepsilon_{ij} \left[\left(\frac{r_{ij}^{\text{min}}}{r_{ij}} \right)^{12} - 2 \left(\frac{r_{ij}^{\text{min}}}{r_{ij}} \right)^6 \right] \quad (2)$$

The parameters r_{ij}^{min} represent the minimum distances of the LJ functions, while the parameters ε_{ij} represent the values of the correction terms at r_{ij}^{min} . We found that the simple Lennard-Jones 6-12 functional form is well suited to represent dispersion correction terms in the attractive and weakly repulsive region. However, for the strongly repulsive region (with nonbonded atoms jammed into one another), the Lennard-Jones potential is likely to greatly overestimate the van der Waals repulsion. Since most structures with strongly repulsive nonbonded contacts represent in fact highly unstable species which are difficult to accurately model even with correlated wave function methods such as CCSD(T), we have not attempted to accurately model the strongly repulsive region with the current version of B3LYP-MM. However, once a larger amount of reliable interaction energy data becomes available for strongly repulsive van der Waals contacts, it will be a relatively easy task to adjust the repulsive functional form of B3LYP-MM to allow accurate modeling of the repulsive region as well. Since Lennard-Jones dispersion correction terms are not included for hydrogen bonds as well as for cation- π interactions, the repulsive wall of the LJ 6-12 potential does not pose any problem for close hydrogen bonds and cation- π contacts.

In order to guarantee transferability of the correction, the total number of empirical parameters should be kept small compared to the number of training data points. In order to lower the total amount of empirical parameters, we introduced an empirical combination rule for the atom pairwise ε parameters (eq 3).

$$\varepsilon_{ij} = \varepsilon_i \varepsilon_j \quad (3)$$

The parameters ε_i depend only on the atomic number of atom i . This keeps the number of parameters low but at the same time limits the accuracy of the correction, since atoms in different chemical environments will be treated the same. For this work, ε parameters were only fit for the most common elements in the data set (H, C, N, O, F, S, and Cl). For metal cations such as Li^+ and Na^+ as well as for ammonium hydrogens (defined as any hydrogen attached to a positively charged sp^3 nitrogen), Lennard-Jones dispersion correction terms are not included since the contribution of dispersion to intermolecular interactions is usually very small for these positive ions. For molecules involving other elements, the correction is currently undefined. However, once even more benchmark interaction energy data become available, we are planning to determine ε parameters for more elements.

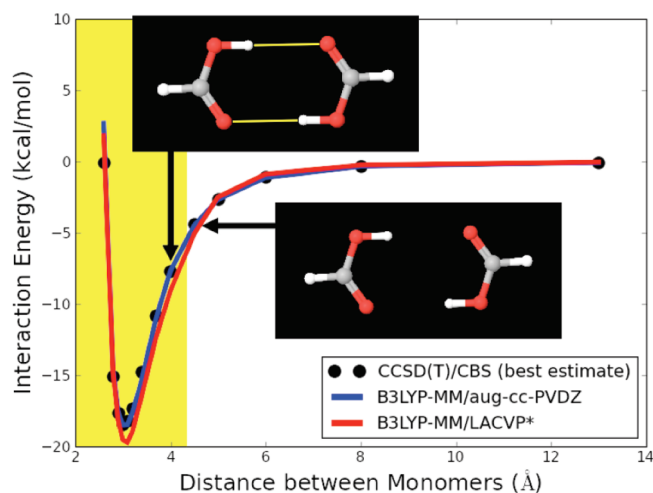
The pairwise r^{min} parameters are calculated with the additive combination rule shown in eq 4.

$$r_{ij}^{\text{min}} = q(R_i^{\text{VDW}} + R_j^{\text{VDW}}) \quad (4)$$

R_i^{VDW} represent experimental van der Waals Radii obtained from Bondi's compilation⁶² (again only dependent on the atomic number of atom i , shown in Table 2), while q is a global scaling factor.

Table 2. The Parameters R_i^{VDW} (Obtained from Bondi's Compilation⁶²)

element	R^{VDW} (Å)
H	1.20
C	1.70
N	1.55
O	1.52
F	1.47
S	1.80
Cl	1.75

**Figure 1.** Potential energy surface of the formic acid dimer. Distance between the monomers represents the distance between the two centers of mass. Hydrogen bonds are shown in yellow with the hydrogen bonded distance range (H-heavy-atom distance < 3.0 Å) highlighted in yellow.

The hydrogen bonding correction term is simply composed of a linear repulsive function with two global parameters r_0^{hb} and b^{hb} (eq 5).

$$E_{\text{HBC}} = \sum_{i < j} -b^{\text{hb}}(r_{ij} - r_0^{\text{hb}}) \quad (5)$$

The sum in eq 5 runs over all hydrogen-heavy atom pairs, which are hydrogen bonded. r_{ij} represents the distances between the hydrogen bond acceptor and the hydrogen-bonded hydrogens. Each correction term is only counted if positive, i.e., if $r_{ij} < r_0$. Hydrogen bonds were detected with the hydrogen bond detection subroutine implemented in the Schrödinger MMSHare software package.⁶³ A hydrogen bond was always assigned if two potentially hydrogen-bonding atoms (hydrogen and hydrogen-bond acceptor with the right Macromodel atom types) were closer than 3.0 Å apart from each other. In order to avoid large discontinuities in the potential energy surfaces with our correction scheme, hydrogen bond donor and acceptor angles were not considered for the assignment of hydrogen bonds. This angle independent hydrogen bond assignment ensures that the only possible discontinuity in the potential energy surface could occur at the cutoff distance of 3.0 Å. However, since all correction terms become negligible at the cutoff distance, this very small discontinuity of the potential energy surface does not pose a problem in practice. This fact is illustrated in Figure 1, which shows potential

energy surfaces of the formic acid dimer with our correction scheme. Even though the formic acid dimer contains two symmetric hydrogen bonds (which will lead to a discontinuity twice as large as with only a single hydrogen bond), Figure 1 shows that the discontinuity in the potential energy surface is virtually nonexistent.

The cation- π correction term (eq 6) is also composed of a simple linear repulsive term.

$$E_{\pi+} = \sum_{i < j} -b^{\pi+}(r_{ij} - r_0^{\pi+}) \quad (6)$$

The sum in eq 6 runs over all metal cation- sp^2/sp -carbon pairs. r_{ij} represents the distances between the atom pairs considered. All sp^2/sp carbons, except carbonyl and immine carbons, are counted for the cation- π correction term. Each correction term is only included if positive, i.e., if $r_{ij} < r_0$. The parameter $r_0^{\pi+}$ was set to 5.0 Å and not further optimized.

The parameters ϵ_i , q , b^{hb} , r_0^{hb} , and $b^{\pi+}$ were optimized via a least-squares fitting procedure to a part of our interaction energy data set as well as to a part of the data from the relative conformational energy data set. In order to avoid any discontinuity of the hydrogen bonding correction, the hydrogen bonding correction parameter b^{hb} was constrained to be ≤ 3.0 Å. The determination of optimal parameter values represents a non-linear optimization problem, to which we found a solution with a variant of the Levenberg-Marquardt algorithm as implemented in the leastsq subroutine of the scipy software package (version 0.6.0). About 75% of all interaction energy and conformational energy data was randomly selected and used as the training set for the parameters. The leftover 25% of the data was employed as a test set. In order to ensure that all types of interactions were well represented in the training and the test set, the random selection procedure for the interaction energy training set was carried out according to the following protocol:

- (1) All complexes were assigned to several ϵ -parameter groups, such that each parameter ϵ_i formed one parameter group. A complex was assigned to the parameter group of ϵ_i if its interaction energy was dependent on the value of ϵ_i . ϵ -parameter groups of common elements (e.g., C, H, or O) contained several hundred complexes while other parameter groups, such as the ones involving F or Cl only had a few tens of members.
- (2) In order to ensure that all parameter groups would be represented in the training and the test set with approximately the desired training/test ratio, all parameter groups were ordered according to increasing number of complexes. Each complex was then assigned to one single parameter group according to the following algorithm. First, the parameter group with the least amount of members was assigned all of its members. Then, the second smallest parameter group was assigned all of its members, except the ones already present in the first parameter group. Proceeding in the same manner, the n th smallest parameter group was assigned all of its members except the ones already present in the parameter groups with fewer members.
- (3) A little over half of the training set was then selected by randomly choosing 37.5% (rounded down to the nearest integer) of complexes from each parameter group.
- (4) The whole data set was also divided into the three interaction type categories shown in Table 1. The three

Table 3. Mean B3LYP-MM Parameter Values Obtained from Six Least Squares Fitting Experiments^a

parameter name	parameter units	basis set			
		LACVP*		aug-cc-pVDZ	
		un-Cp-corr ^b	Cp-corr ^c	un-Cp-corr ^b	Cp-corr ^c
ϵ_{H}	(kcal/mol) ^{0.5}	0.097 ± 0.009	0.183 ± 0.002	0.306 ± 0.003	0.313 ± 0.002
ϵ_{C}		0.589 ± 0.024	0.744 ± 0.003	0.660 ± 0.009	0.714 ± 0.013
ϵ_{N}		0.542 ± 0.019	0.744 ± 0.005	0.731 ± 0.009	0.705 ± 0.011
ϵ_{O}		0.215 ± 0.004	0.427 ± 0.004	0.595 ± 0.012	0.633 ± 0.013
ϵ_{F}		0.013 ± 0.023 ^d	0.528 ± 0.022	0.362 ± 0.010	0.540 ± 0.014
ϵ_{S}		1.117 ± 0.067	1.393 ± 0.023	1.288 ± 0.020	1.379 ± 0.027
ϵ_{Cl}		0.909 ± 0.043	1.145 ± 0.023	0.701 ± 0.054	0.974 ± 0.044
q	no units	0.895 ± 0.012	0.860 ± 0.001	0.859 ± 0.004	0.846 ± 0.005
b^{hb}	kcal/(mol × Å)	1.144 ± 0.024	1.094 ± 0.038	1.888 ± 0.064	1.816 ± 0.090
$b^{\pi+}$		0.410 ± 0.004	0.248 ± 0.006	0.130 ± 0.002	0.116 ± 0.004
r_0^{hb}	Å	3.000 ^e	2.283 ± 0.006	2.047 ± 0.002	2.035 ± 0.006
$r_0^{\pi+}$		5.000 ^f	5.000 ^f	5.000 ^f	5.000 ^f

^aError bars represent standard deviations. ^bParameters fit to interaction energies without explicit counterpoise corrections. ^cFit to counterpoise corrected interaction energies. Relative conformational energies were never counterpoise corrected. ^dParameter was constrained to be positive. During the six fitting experiments, parameter values from 0.00 to 0.06 were found. ^eParameter reached its maximum value of 3.0 Å during least-squares optimization. ^fParameter was not optimized but set to 5.0 Å.

categories are complexes with only dispersion and dipole-dipole interactions, hydrogen bonded complexes, and charged complexes.

- (5) The percentage of complexes to be selected for the training set from each *interaction type category* (defined as *f2*) was then calculated, such that the final training/test ratio would be around 3:1. *f2%* of data points from each *interaction type category* were then randomly selected for the training set, after having removed all complexes already selected for the training set in step 3.

The random selection of the training set together with the least-squares fitting was carried out six times for each set of parameters. This allowed us to obtain an estimate of the sensitivity of the parameters on the training set and is a good test to ensure there is no overfitting. Table 3 shows the mean values of all parameters (for the LACVP* and the aug-cc-pVDZ basis sets) together with their standard deviations as obtained from the six least-squares fitting experiments. Optimal parameters were determined for counterpoise and noncounterpoise corrected interaction energies. (The relative conformational energies were never counterpoise corrected, since there is no consistent way to apply counterpoise correction for total energies.)

The very small standard deviations for each parameter clearly show that there has been no overfitting and that the correction is therefore likely to be transferable to systems not included in the training set. Another fact that supports this conclusion is that the mean unsigned errors (MUEs) as well as the root mean square deviations (RMSDs) for the randomly selected training and test sets were very similar for all six fitting experiments. RMSDs and MUEs averaged over all six fitting experiments are shown in Table 4.

Since we find that BSSE is small with the aug-cc-pVDZ basis set if no halogens are involved, it is not surprising that the optimal values for the correction parameters are very similar with and without counterpoise corrected interaction energies for that basis set (Table 3). Exceptions are the halogen correction parameters

Table 4. RMSDs and MUEs Obtained with B3LYP-MM Averaged over the Six Fitting Experiments^a

basis set	cp corr ^b	MUE (kcal/mol)		RMSD (kcal/mol)	
		training set	test set	training set	test set
LACVP*	no	0.45 ± 0.01	0.45 ± 0.02	0.70 ± 0.02	0.66 ± 0.05
	yes	0.48 ± 0.01	0.49 ± 0.01	0.72 ± 0.01	0.72 ± 0.03
aug-cc-pVDZ	no	0.37 ± 0.01	0.36 ± 0.01	0.51 ± 0.01	0.50 ± 0.02
	yes	0.34 ± 0.01	0.35 ± 0.01	0.49 ± 0.01	0.49 ± 0.02

^aError bars represent standard deviations (training set: 1810 ± 6 data points; test set: 605 ± 6 data points). ^bWhether interaction energies were counterpoise corrected or not. Relative Conformational energies were never counterpoise corrected.

ϵ_{F} and ϵ_{Cl} , which vary significantly depending on whether explicit counterpoise corrections are included, due to larger counterpoise corrections for complexes involving F or Cl with the aug-cc-pVDZ basis set. For the small LACVP* basis set where BSSE is very significant (often up to several kilocalories per mole), we find clearly different correction parameters depending on whether counterpoise corrections are applied or not. Again, the halogen correction parameters are especially sensitive to the counterpoise correction issue, due to especially large counterpoise corrections for structures involving halogens.

RESULTS AND DISCUSSION

A summary of the overall statistical performance (MUEs, RMSDs, and average errors) of all of the benchmarked methods, B3LYP-D3, M06-2X, and B3LYP-MM (this work) is shown in Table 5.

In order to see how the benchmarked methods perform for complexes dominated by different kinds of interactions, results for dispersion and dipole-dipole dominated complexes are shown separately in Table 6, while Table 7 shows separate results for complexes with hydrogen bonds or ionic interactions.

Table 5. RMSDs, MUEs, and Average Errors of All Benchmarked DFT Methods (in kcal/mol)^a

basis set	counterpoise-corrected	DFT method	whole interaction energy DB (1715 data points ^b)			conformational energies DB (700 data points)		
			MUE ^c	RMSD ^d	avg ^e	MUE ^c	RMSD ^d	avg ^e
LACVP*	no	B3LYP-D3	2.11	2.87	-2.10	0.82	1.10	0.02
		B3LYP-MM	0.41	0.68	-0.04	0.55	0.71	0.05
		M06-2X	1.20	1.97	-0.94	0.78	1.04	0.10
	yes	B3LYP	2.75	3.66	1.35	1.54	2.03	0.32
		B3LYP-D3	1.21	1.91	-0.64		na ^f	
		B3LYP-MM	0.41	0.65	0.18	0.67 ^g	0.87 ^g	-0.03 ^g
		M06-2X	1.20	1.97	-0.94		na ^f	
aug-cc-pVDZ	no	B3LYP-D3	0.95	1.48	-0.72	0.39	0.49	-0.05
		B3LYP-MM	0.37	0.53	0.09	0.37	0.47	-0.07
		M06-2X	0.73	1.17	-0.30	0.61	0.79	-0.02
	yes	B3LYP	3.11	4.36	2.73	1.57	2.15	0.25
		B3LYP-D3	0.87	1.41	-0.56		na ^f	
		B3LYP-MM	0.32	0.48	0.11	0.38 ^g	0.49 ^g	-0.09 ^g
		M06-2X	0.67	1.08	-0.10		na ^f	
		B3LYP	3.25	4.57	2.89			

^a B3LYP-MM represents results obtained with the correction scheme developed in this work. ^b For the following levels of theory, a fewer number of cases were included due to SCF convergence problems: B3LYP/aug-cc-pVDZ without counterpoise correction (2 cases missing); M06-2X/LACVP* with explicit counterpoise correction (11 cases missing); M06-2X/aug-cc-pVDZ with explicit counterpoise correction (3 cases missing); M06-2X/aug-cc-pVDZ without counterpoise correction (4 cases missing). ^c Mean unsigned error. ^d Root mean square deviation. ^e Average error. ^f No consistent way to apply counterpoise corrections. ^g B3LYP energies were obtained without counterpoise corrections; however, the MM correction part was parametrized with counterpoise corrected interaction energies.

Table 6. RMSDs, MUEs, and Average Errors of All Benchmarked DFT Methods (in kcal/mol) for Dispersion and Dipole-Dipole Dominated Complexes^a

basis set	counterpoise-corrected	DFT method	dispersion and dipole-dipole dominated complexes (1036 data points ^e)		
			MUE ^b	RMSD ^c	avg ^d
LACVP*	no	B3LYP-D3	1.16	1.38	-1.15
		B3LYP-MM	0.28	0.41	0.05
		M06-2X	0.65	0.92	-0.26
	yes	B3LYP	3.15	4.09	3.14
		B3LYP-D3	0.55	0.80	0.33
		B3LYP-MM	0.33	0.46	0.21
aug-cc-pVDZ	no	M06-2X	1.06	1.36	1.03
		B3LYP	4.62	5.88	4.62
		B3LYP-D3	0.37	0.50	-0.02
	yes	B3LYP-MM	0.36	0.49	0.13
		M06-2X	0.53	0.73	0.01
		B3LYP	4.27	5.31	4.27
		B3LYP-D3	0.32	0.48	0.18
		B3LYP-MM	0.27	0.37	0.14
		M06-2X	0.47	0.63	0.31
		B3LYP	4.47	5.58	4.47

^a B3LYP-MM represents results obtained with the correction scheme developed in this work. ^b Mean unsigned error. ^c Root mean square deviation. ^d Average error. ^e For the following levels of theory, fewer number of cases were included due to SCF convergence problems: B3LYP/aug-cc-pVDZ without counterpoise correction (2 cases missing); M06-2X/aug-cc-pVDZ without counterpoise correction (4 cases missing).

Error histograms for dispersion and dipole-dipole dominated complexes, but also for hydrogen bonded and ionic ones are shown in Table 8 for B3LYP-MM, B3LYP-D3, and M06-2X. Corresponding error histograms are also shown for the conformational energy data set. Noncounterpoise corrected results are depicted for both basis sets.

In the following section, we will discuss the results obtained with both basis sets in more detail.

Results with the Medium Sized aug-cc-pVDZ Basis Set. From Table 6, it can be seen that the performance of B3LYP-D3 for dispersion and dipole-dipole dominated complexes is remarkably accurate with the aug-cc-pVDZ basis set (MUEs below

Table 7. RMSDs, MUEs, and Average Errors of All Benchmarked DFT Methods (in kcal/mol) for Complexes with Hydrogen Bonds or Ionic Interactions^a

basis set	counterpoise-corrected	DFT method	hydrogen bonded complexes (not charged; 137 data points)			complexes with ionic interactions (542 data points ^c)		
			MUE ^b	RMSD ^c	avg ^d	MUE ^b	RMSD ^c	avg ^d
LACVP*	no	B3LYP-D3	3.42	3.71	-3.42	3.58	4.36	-3.57
		B3LYP-MM	0.75	0.94	-0.31	0.56	0.95	-0.15
		M06-2X	1.85	2.16	-1.81	2.08	3.07	-2.03
		B3LYP	2.13	2.76	0.14	2.14	2.93	-1.76
	yes	B3LYP-D3	0.82	1.00	-0.51	2.57	3.18	-2.53
		B3LYP-MM	0.55	0.91	0.34	0.52	0.84	0.10
		M06-2X	0.85	1.01	0.53	1.35	2.11	-1.16
		B3LYP	3.13	4.21	3.04	1.53	2.13	-0.71
aug-cc-pVDZ	no	B3LYP-D3	0.64	0.77	-0.45	2.14	2.51	-2.12
		B3LYP-MM	0.37	0.54	0.10	0.38	0.58	0.00
		M06-2X	0.61	0.76	0.25	1.13	1.78	-1.05
		B3LYP	3.10	4.05	3.10	0.91	1.47	-0.31
	yes	B3LYP-D3	0.44	0.56	-0.25	2.05	2.41	-2.04
		B3LYP-MM	0.36	0.51	0.24	0.41	0.64	0.02
		M06-2X	0.63	0.78	0.52	1.09	1.67	-1.05
		B3LYP	3.31	4.20	3.31	0.90	1.44	-0.23

^a B3LYP-MM represents results obtained with the correction scheme developed in this work. ^b Mean unsigned error. ^c Root mean square deviation. ^d Average error. ^e For the following levels of theory, a fewer number of cases were included due to SCF convergence problems: M06-2X/LACVP* with explicit counterpoise correction (11 cases missing); M06-2X/aug-cc-pVDZ with explicit counterpoise correction (3 cases missing).

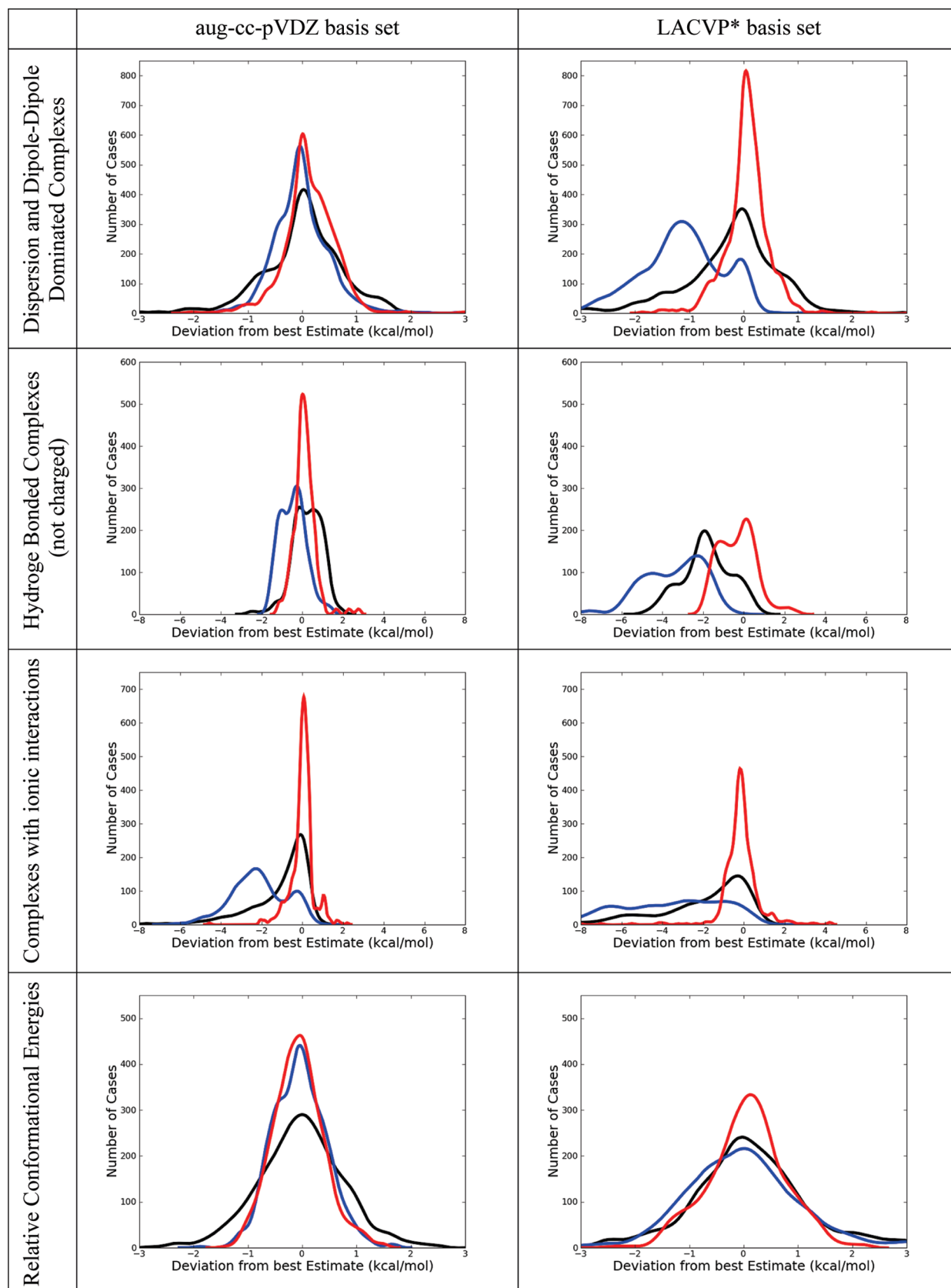
0.40 kcal/mol). B3LYP-D3 also performs very accurately for the data set of conformational energies (MUE of 0.39 kcal/mol). M06-2X gives also rather accurate results for dispersion and dipole-dipole dominated complexes as well as for the conformational energies. Its performance is however slightly worse than B3LYP-D3's. For the subset of hydrogen bonded complexes, B3LYP-D3 is slightly more accurate than M06-2X with explicit counterpoise corrections, while without counterpoise corrections both methods are of comparable accuracy. For hydrogen-bonded systems, B3LYP-D3 shows some overbinding. This overbinding is present with and without counterpoise corrections and is manifested in the negative average errors of the method. Since BSSE is always attractive, it is not surprising that the overbinding is more pronounced without explicit counterpoise corrections. For complexes with ionic interactions, both B3LYP-D3 and M06-2X show significant overbinding regardless of counterpoise corrections (average errors < -1.0 kcal/mol).

With the aug-cc-pVDZ basis set and with explicit counterpoise corrections, our B3LYP-MM method is slightly more accurate than B3LYP-D3 for dispersion and dipole-dipole dominated complexes (MUE = 0.27 kcal/mol). Without counterpoise corrections, our B3LYP-MM method still produces very accurate interaction energies. As with B3LYP-D3, we also observe a very accurate performance on the data set of conformational energies (MUE = 0.37 kcal/mol). However, in contrast to B3LYP-D3, which shows significant overbinding for hydrogen-bonded and charged complexes, the B3LYP-MM correction scheme also delivers results of high accuracy (MUEs ≤ 0.41 kcal/mol) and almost no overbinding (|average error| ≤ 0.24 kcal/mol) for hydrogen-bonded and charged systems.

Therefore, since hydrogen bonds as well as ionic interactions are of great practical importance in biology and materials science, B3LYP-MM clearly has a practical advantage over B3LYP-D3 or M06-2X in conjunction with a medium-sized basis set.

Results with the Small LACVP* Basis Set. With counterpoise corrections, B3LYP-D3 performs acceptably in the small basis set for dispersion and dipole-dipole dominated complexes, with an overall MUE of 0.55 kcal/mol, while we would not recommend M06-2X with this basis set (MUE = 1.06 kcal/mol). However, B3LYP-MM is clearly the most accurate method with counterpoise corrections in the small basis set with an overall MUE of only 0.33 kcal/mol. Even with counterpoise corrections, B3LYP-D3 seriously overbinds hydrogen-bonded complexes (average error = -0.51 kcal/mol) and complexes with ionic interactions (average error = -2.53). Without counterpoise corrections, B3LYP-D3's overbinding for hydrogen-bonded and ionic complexes clearly renders the method very inaccurate with average errors < -3.0 kcal/mol. Apart from hydrogen bonds and ionic interactions, B3LYP-D3 also shows serious overbinding for dispersion and dipole-dipole dominated complexes without counterpoise corrections in the small basis set, as can clearly be seen from the error histograms in Table 8. Without counterpoise corrections, M06-2X performs better than B3LYP-D3 in the small basis set, its overall performance is however still not convincing with a MUE of 1.20 kcal/mol.

In contrast to all other methods tested, our B3LYP-MM methodology still delivers accurate interaction and conformational energies with the small basis set without counterpoise corrections, with overall MUEs of only 0.41 and 0.55 kcal/mol that are only marginally larger than the ones obtained with the medium-sized basis set. For hydrogen-bonded complexes, B3LYP-MM shows slight overbinding (average error of -0.31 kcal/mol), which leads to a slight degradation in the overall accuracy for hydrogen-bonded complexes (MUE = 0.75 kcal/mol). However, compared to the performances of B3LYP-D3 and M06-2X for hydrogen-bonded complexes without counterpoise corrections in the small basis set (both methods show MUEs larger than 1.8 kcal/mol), the accuracy of B3LYP-MM

Table 8. Comparison of Error Distributions of B3LYP-MM (red), B3LYP-D3 (blue), and M06-2X (black)^a

^a All results are without explicit counterpoise corrections. All plots show normalized kernel density estimates (KDEs) multiplied with the number of data points. The KDEs were computed with the `density()` function in R, with Gaussian kernels and default bandwidths.

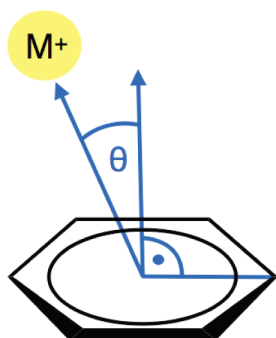


Figure 2. Definition of the angle θ .

represents a significant improvement. For complexes with ionic interactions, B3LYP-MM gives also very accurate results regardless of explicit counterpoise corrections, with MUEs ≤ 0.56 kcal/mol.

Finally, it is also noteworthy that the cation- π correction in B3LYP-MM seems to pretty uniformly improve cation- π interaction energies, regardless of whether the cation sits on top or on the side of an aromatic ring. To illustrate this, we calculated the MUEs obtained with B3LYP-MM for all benzene alkali-cation interaction energies in two different regimes of the angle θ (defined in Figure 2).

All structures were divided into conformations where the cation sits on top of the benzene ring ($\theta \leq 45^\circ$; 215 structures) and in conformations where the cation is located more toward the side of the benzene ring ($45^\circ < \theta \leq 90^\circ$; 293 structures). With B3LYP-MM, we find similar accuracy for both regimes of the angle θ regardless of basis set and explicit counterpoise corrections. The observed MUEs for the regime with $\theta > 45^\circ$ are in fact always slightly smaller (maximum difference in MUEs observed was 0.27 kcal/mol) than the corresponding MUEs for the regime with $\theta \leq 45^\circ$. This indicates that the cation- π correction contained in B3LYP-MM not only works well if the cation is directly above the benzene ring but also clearly improves interaction energies for other structures.

CONCLUSIONS

In this work, we have built a data set of highly accurate noncovalent interaction energies at the CCSD(T) level from the literature. Compared to previously published data sets, our data set has a much larger size and greater diversity. Benchmark studies of two DFT methodologies (M06-2X and B3LYP-D3) on our new interaction energy data set show high accuracy of B3LYP-D3 for dispersion and dipole-dipole dominated interactions if a medium-sized basis set such as aug-cc-pVDZ is employed with explicit counterpoise corrections (MUE = 0.32 kcal/mol). However, without counterpoise corrections, we find serious overbinding of B3LYP-D3 for hydrogen-bonded systems (average error = -0.45 kcal/mol) and for complexes with ionic interactions (average error = -2.12 kcal/mol) with the medium-sized basis set. For the small LACVP* basis sets without counterpoise corrections (a level of theory that is often employed in practice, especially for geometry optimizations), our benchmark results show that neither B3LYP-D3 nor M06-2X shows good accuracy. In contrast to B3LYP-D3 and M06-2X, we find that our new methodology (B3LYP-MM) delivers very accurate results for all types of interactions, regardless of counterpoise corrections and basis set (overall MUEs ≤ 0.41 kcal/mol). The observed improvement in accuracy with B3LYP-MM over

B3LYP-D3 and M06-2X is especially significant for hydrogen-bonded and charged systems.

The development of B3LYP-MM has demonstrated a number of interesting points:

- (1) By making parameters specific to the basis set and use of counterpoise corrections, significant improvements in accuracy can be obtained, with a particularly large improvement for small basis sets.
- (2) The addition of a small number of parameters addressing hydrogen bonding and interactions involving one or more ions can substantially improve accuracy for these systems, with a particularly large impact on ionic systems, which exhibit large errors when treated with prior alternatives in the literature.

While the data set here has quite a few data points, it does not fully cover all possible types of chemistry associated with noncovalent interactions. Unusual functional groups may in fact require the addition of new parameters to the model. The improvements to hydrogen bonded and charged systems obtained by following this path suggest that it is a fruitful one and can result in a systematically improvable, empirically corrected DFT functional; as problems are discovered, they can readily be repaired by the addition of a relatively small number of new parameters. As benchmark quantum chemical calculations become increasingly inexpensive in the future, transfer of information from such calculations to DFT-based models will become more facile and, as shown here, will be a highly effective means of achieving chemical accuracy at modest computational cost.

ASSOCIATED CONTENT

S Supporting Information. Coordinates, best estimate benchmark interaction energies, and interaction energies at all calculated DFT levels for all complexes of the noncovalent interaction energy database. All of the aforementioned information is also given for the conformational energies database. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*E-mail: rich@chem.columbia.edu.

ACKNOWLEDGMENT

R.A.F. thanks the National Institutes of Health (Grant GM-40526) for support of this work. S.T.S. is grateful for an Arun Guthikonda Memorial Fellowship. We also thank Professor Ronald Breslow and Jianing Li for helpful discussions. Finally, we would like to acknowledge Professor Seiji Tsuzuki, Professor Stacey Wetmore, Professor Ota Bludsky, and Professor Zaneta Czynnikowska for providing us with coordinates from their publications.

REFERENCES

- (1) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.
- (2) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215–241.
- (3) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- (4) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796–6806.

- (5) Woon, D. E.; Dunning, T. H. *J. Chem. Phys.* **1993**, *98*, 1358–1371.
- (6) Woon, D. E.; Dunning, T. H. *J. Chem. Phys.* **1994**, *100*, 2975–2988.
- (7) Parr, R. C.; Yang, W. *Density Functional Theory of Atoms and Molecules*; Oxford University Press: New York, 1989.
- (8) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- (9) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.
- (10) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623–11627.
- (11) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200–1211.
- (12) Singleton, D. A.; Wang, Z. H. *J. Am. Chem. Soc.* **2005**, *127*, 6679–6685.
- (13) Schneebeli, S. T.; Hall, M. L.; Breslow, R.; Friesner, R. *J. Am. Chem. Soc.* **2009**, *131*, 3965–3973.
- (14) Hobza, P.; Sponer, J.; Reschel, T. *J. Comput. Chem.* **1995**, *16*, 1315–1325.
- (15) Cerny, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2005**, *7*, 1624–1626.
- (16) Johnson, E. R.; Wolkow, R. A.; DiLabio, G. A. *Chem. Phys. Lett.* **2004**, *394*, 334–338.
- (17) Tsuzuki, S.; Luthi, H. P. *J. Chem. Phys.* **2001**, *114*, 3949–3957.
- (18) Grimme, S.; Muck-Lichtenfeld, C.; Antony, J. *J. Phys. Chem. C* **2007**, *111*, 11199–11207.
- (19) Antony, J.; Grimme, S. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2722–2729.
- (20) Stone, A. J. *The Theory of Intermolecular Forces*; Oxford University Press: Oxford, U. K., 1997.
- (21) Burley, S. K.; Petsko, G. A. *Science* **1985**, *229*, 23–28.
- (22) Hunter, C. A.; Lawson, K. R.; Perkins, J.; Urch, C. J. *J. Chem. Soc., Perkin Trans. 2* **2001**, 651–669.
- (23) Cerny, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2007**, *9*, 5291–5303.
- (24) Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2007**, *3*, 289–300.
- (25) Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157–167.
- (26) Becke, A. D.; Johnson, E. R. *J. Chem. Phys.* **2007**, *127*, 124108.
- (27) Grimme, S. *J. Chem. Phys.* **2006**, *124*, 034108.
- (28) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787–1799.
- (29) Grimme, S.; Antony, J.; Schwabe, T.; Muck-Lichtenfeld, C. *Org. Biomol. Chem.* **2007**, *5*, 741–758.
- (30) Schwabe, T.; Grimme, S. *Phys. Chem. Chem. Phys.* **2007**, *9*, 3397–3406.
- (31) Chai, J. D.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.
- (32) Riley, K. E.; Pitonak, M.; Jurecka, P.; Hobza, P. *Chem. Rev.* **2010**, *110*, 5023–63.
- (33) Grafenstein, J.; Cremer, D. *J. Chem. Phys.* **2009**, *130*, 124105.
- (34) Johnson, E. R.; Mackie, I. D.; DiLabio, G. A. *J. Phys. Org. Chem.* **2009**, *22*, 1127–1135.
- (35) Sato, T.; Nakai, H. *J. Chem. Phys.* **2009**, *131*, 224104.
- (36) Jurecka, P.; Sponer, J.; Cerny, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985–1993.
- (37) Grafova, L.; Pitonak, M.; Rezac, J.; Hobza, P. *J. Chem. Theory Comput.* **2010**, *6*, 2365–2376.
- (38) Rezac, J.; Jurecka, P.; Riley, K. E.; Cerny, J.; Valdes, H.; Pluhackova, K.; Berka, K.; Rezac, T.; Pitonak, M.; Vondrasek, J.; Hobza, P. *Collect. Czech. Chem. Commun.* **2008**, *73*, 1261–1270.
- (39) Liu, B.; McLean, A. D. *J. Chem. Phys.* **1973**, *59*, 4557–4558.
- (40) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553–566.
- (41) Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 299–310.
- (42) Knoll, E. H.; Friesner, R. A. *J. Phys. Chem. B* **2006**, *110*, 18787–18802.
- (43) Hall, M. L.; Goldfeld, D. A.; Bochevarov, A. D.; Friesner, R. A. *J. Chem. Theory Comput.* **2009**, *5*, 2996–3009.
- (44) Goldfeld, D. A.; Bochevarov, A. D.; Friesner, R. A. *J. Chem. Phys.* **2008**, *129*, 214105.
- (45) Rinaldo, D.; Tian, L.; Harvey, J. N.; Friesner, R. A. *J. Chem. Phys.* **2008**, *129*, 164108.
- (46) Sinnokrot, M. O.; Sherrill, C. D. *J. Am. Chem. Soc.* **2004**, *126*, 7690–7697.
- (47) Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M. *J. Chem. Phys.* **2005**, *122*, 144323.
- (48) Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. *J. Am. Chem. Soc.* **2000**, *122*, 3746–3753.
- (49) Pittner, J.; Hobza, P. *Chem. Phys. Lett.* **2004**, *390*, 496–499.
- (50) Moller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 0618–0622.
- (51) Headgordon, M.; Pople, J. A.; Frisch, M. J. *Chem. Phys. Lett.* **1988**, *153*, 503–506.
- (52) Pople, J. A.; Headgordon, M.; Raghavachari, K. *J. Chem. Phys.* **1987**, *87*, 5968–5975.
- (53) Scuseria, G. E.; Schaefer, H. F. *J. Chem. Phys.* **1989**, *90*, 3700–3703.
- (54) Valdes, H.; Pluhackova, K.; Pitonak, M.; Rezac, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2747–2757.
- (55) *Jaguar*, version 7.6; Schrödinger, LLC: New York, 2009.
- (56) Friesner, R. A. *Chem. Phys. Lett.* **1985**, *116*, 39–43.
- (57) Friesner, R. A. *J. Chem. Phys.* **1986**, *85*, 1462–1468.
- (58) Friesner, R. A. *J. Chem. Phys.* **1987**, *86*, 3522–3531.
- (59) Friesner, R. A. *J. Phys. Chem.* **1988**, *92*, 3091–3096.
- (60) Martinez, T. J.; Carter, E. A. Pseudospectral methods applied to the electron correlation problem. In *Modern Electronic Structure Theory*; World Scientific: Singapore, 1995; Part II, Vol 2, p 1132.
- (61) Kaminski, G. A. M. J. R.; Murphy, R. B.; Braden, D. A.; Friesner, R. A. *J. Chem. Theory Comput.* **2005**, *1*, 248–254.
- (62) Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441–451.
- (63) *MMShare*, version 18109; Schrödinger, Inc.: New York, 2009.
- (64) Pitonak, M.; Janowski, T.; Neogady, P.; Pulay, P.; Hobza, P. *J. Chem. Theory Comput.* **2009**, *5*, 1761–1766.
- (65) Takatani, T.; Hohenstein, E. G.; Malagoli, M.; Marshall, M. S.; Sherrill, C. D. *J. Chem. Phys.* **2010**, *132*, 144104.

NOTE ADDED AFTER ASAP PUBLICATION

This article was published ASAP on February 11, 2011. Spelling errors in Tables 5, 6, and 7 and the Acknowledgment section have been corrected. The correct version was published on February 16, 2011.

Density Functional Theory for Reaction Energies: Test of Meta and Hybrid Meta Functionals, Range-Separated Functionals, and Other High-Performance Functionals

Yan Zhao*

Commercial Print Engine Lab, HP Laboratories, Hewlett-Packard Co., 1501 Page Mill Road, Palo Alto, California 94304, United States

Donald G. Truhlar*

Department of Chemistry and Supercomputing Institute, University of Minnesota, Minneapolis, Minnesota 55455, United States

 Supporting Information

ABSTRACT: The present study compares the accuracy of 30 density functionals for four databases of reaction energies studied recently by Grimme and co-workers. For 20 of the density functionals, the calculations are new, and the calculations are compared to previous work for the other 10. We present the results in detail for 11 of the functionals and as mean unsigned errors for the others. The results presented in detail are for the seven most recent Minnesota functionals (M05-2X, M06-L, M06-HF, M06, M06-2X, M08-HX, and M08-SO), three range-separated functionals (HSE, LC- ω PBE, and ω B97X-D), and one dispersion-corrected global hybrid generalized gradient approximation (B97-D); the other functionals include five dispersion-corrected functionals and their uncorrected analogs, eight high-performing functionals on a recent catalytic-energies test, and the TPSSH functional because it is of special interest to compare its performance to that of M08-SO. Three of the four databases contain a total of 21 rearrangement reaction energies and 13 diverse dissociation or association energies, and the fourth contains three dissociation reaction energies of alkali metal clusters and three dissociation reaction energies of alkali-metal-cation–benzene complexes. The results are especially promising for the Minnesota hybrid meta-GGA functionals and the ω B97X-D, B2PLYP-D, and HSE functionals.

1. INTRODUCTION

Density functional theory (DFT) in the formulation of Kohn and Sham¹ is the most robust and popular electronic structure method in computational chemistry and physics. Although Kohn–Sham DFT is an exact many-body quantum mechanical theory for the ground-electronic-state properties of a given system, it depends on an unknown universal exchange–correlation (XC) functional that can only be approximated.² Many approximate functionals have been published in the literature.^{3,4} Since the exact form of the universal XC functional is unknown, approximate functionals, no matter how they were developed, need to be carefully validated before their application to real chemical problems. To this end and also in order to design new density functionals, a large number of databases have been developed and used to test density functional theory for, for example, chemical reaction barrier heights,^{5–14} thermochemistry,^{9,12–22} noncovalent interactions,^{9,13,14,23–30} transition metal chemistry,^{4,9,31–35} spectroscopy,^{9,13,36–41} and catalysis.^{42,43} Goerigk and Grimme¹⁴ recently compiled a quantum chemistry benchmark database for general main-group thermochemistry, kinetics, and noncovalent interactions, called GMTKN24, which includes 24 different subsets. Three subsets of GMTKN24 that are concerned with energies of reaction, namely SIE11 for 11 reactions that are very sensitive to self-interaction error, DC9 for nine reactions that had been shown to be very difficult for DFT, and DARC14 for Diels–Alder reactions, provide especially challenging problems for approximate density functionals. Here, we study these three

databases along with a database of alkali metal reaction energies taken from the more recent work of Grimme et al.⁴⁴ The details of these databases have been described in the original papers, and we give a brief introduction to each of the four databases in section 2.

The four databases contain 40 diverse benchmark reaction energies, and we calculate results for 20 density functionals of various types for comparison with this data. In addition, we compare to previous results for 10 other functionals for a total of 30 functionals to be compared. The first two functionals tested are range-separated functionals developed by Scuseria and co-workers, in particular the HSE^{45–47} and LC- ω PBE⁴⁸ functionals. The next seven functionals tested are the seven most recent meta and hybrid meta functionals (terms that we explain below) that we developed in Minnesota over the past half dozen years,^{9,13,20,36,50–52} in particular, M05-2X,⁵⁰ M06-L,²⁰ M06-HF,³⁶ M06,⁹ M06-2X,⁹ M08-HX,¹³ and M08-SO.¹³ The other two functionals for which we present detailed results are functionals with empirical dispersion corrections, namely, B97-D⁵³ and ω B97X-D.⁵⁴ The other functionals included in the present tests will be introduced in section 4.5.

The next two sections summarize the details of the four databases and computational methodology, and section 4 presents results and discusses them. Section 5 concludes the paper.

Received: November 16, 2010

Published: February 03, 2011

2. DATABASES

2.1. SIE11. SIE11 is a database of reaction energies for 11 systems that are especially sensitive to self-interaction error (which is discussed further below). This database includes the dissociation energies of five cationic reactions (e.g., $\text{He}_2^+ \rightarrow \text{He}^+ + \text{He}$) and six neutral reactions (e.g., $\text{LiF}_2 \rightarrow \text{Li} + \text{F}_2$). The reference data¹⁴ were obtained by coupled cluster theory. All of the reference data are given in Table 2 along with the DFT results.

2.2. DC9. DC9 is a database of nine reaction energies that have been shown to be difficult for density functionals. We have already investigated two cases in our previous studies; in particular, the energy difference between hepta-1,2,3,5,6-hexaene and hepta-1,3,5-tryne (C_7H_4) is in our πIE3 database,¹⁹ and the isomerization energy of the $(\text{CH})_{12}$ isomers was included in our paper on medium-range correlation energy.⁵⁵ The reference data¹⁴ for the DC9 database were obtained from experiments and various levels of theory, and they are listed in Table 3.

2.3. DARC14. DARC14 is a database of reaction energies for 14 typical Diels–Alder reactions; they are the reactions of butadiene, cyclopentadiene, cyclohexadiene, and furan with ethene, ethyne, maleic anhydride, and maleimide acting as dienophiles. This database was used by Johnson et al.⁵⁶ in a study of delocalization errors of density functionals. The reference data¹⁴ are based on coupled cluster calculations, and they are tabulated in Table 4.

2.4. ALK6. ALK6 is a database of six dissociation energies for alkali metal complexes and clusters. There are three alkali-metalation–benzene complexes (BzM^+), which are of special interest because Grimme et al.⁴⁴ showed that they are significantly overbound by the second generation⁵³ of density functionals with empirical dispersion corrections. The other three data are dissociation energies of nonplanar alkali metal clusters with eight atoms that have large dispersion effects. Therefore, ALK6 is a challenging database for both empirically uncorrected and dispersion-corrected density functional methods. The reference data⁴⁴ are based on estimated coupled cluster energies, and they are given in Table 5.

3. COMPUTATIONAL METHODOLOGY

The methods tested here are based on the Kohn–Sham self-consistent-field (SCF) formulation¹ of DFT in which the electronic energy is approximated as the sum of the noninteracting kinetic energy calculated from the SCF orbitals, the classical Coulomb energy calculated from the total electron density ρ , and the XC energy calculated from an XC functional (usually just called the density functional). The density functional can depend on various quantities. For all functionals considered here, the density functional depends on the spin-labeled (α and β) electron densities, ρ_α and ρ_β (whose sum is ρ), and the reduced density gradients, s_α and s_β . Most of the density functionals considered here also depend on the SCF orbitals, which are functionals of ρ_α and ρ_β . We will at first consider two ways in which the XC energy can depend on the orbitals: dependence on the orbital-dependent kinetic energy densities, τ_α and τ_β , and dependence on the Hartree–Fock exchange integrals computed from the SCF orbitals. (A third kind of dependence, namely, nonlocal second-order perturbation terms, is considered in section 4.5.) Functionals depending on ρ_α , ρ_β , s_α , s_β , τ_α , and τ_β are called local; functionals including some Hartree–Fock exchange are called nonlocal or hybrid. The hybrid functionals considered in this article can be subdivided into global hybrids, in

Table 1. Density Functionals

name	year	ref	meta?	hybrid?	X	-D?
HSE	2003	45–47	no	range-separated	100/0a	no
LC- ω PBE	2006	48, 49	no	range-separated	0/100 ^a	no
M05–2X	2006	50	yes	global	56	no
M06-L	2006	20	yes	no	0	no
M06-HF	2006	36	yes	global	100	no
M06	2008	9	yes	global	27	no
M06–2X	2008	9	yes	global	54	no
M08-HX	2008	13	yes	global	52.23	no
M08-SO	2008	13	yes	global	56.79	no
B97-D	2006	53	no	no	0	yes
ω B97X-D	2008	54	no	range-separated	22.2036/100 ^a	yes

^aThe value before the slash is X for $r_{12} \rightarrow 0$, and the value after the slash is the value of X for $r_{12} \rightarrow \infty$.

which a fixed percentage X of Hartree–Fock exchange is added to some percentage (usually $100 - X$) of local exchange, and range-separated hybrids, in which X is not fixed but depends on the interelectronic separation r_{12} (note that X is defined here only in the limit of a uniform electron gas). A final distinction we need to make is that some density functionals are defined to include an empirical dispersion correction; that is, a post-SCF molecular-mechanics dispersion term is added to the result of the SCF calculation. (In density functionals without such a correction, dispersion-like contributions are present only when there is overlap of the interacting subsystems, and such effects, like other effects of dynamical electron correlation, are present in the correlation part of the density functional.) Density functionals including empirical dispersion terms have a suffix -D.

The present study compares results from several kinds of density functionals. The properties of those considered first are specified in Table 1 by using the distinctions explained in the previous paragraph. Functionals depending only on ρ_σ and s_σ (where $\sigma = \alpha, \beta$) are called generalized gradient approximations (GGAs), and functionals depending only on ρ_σ , s_σ , and τ_σ are called meta-GGAs or meta functionals. The table shows that the initial comparisons include a meta-GGA (M06-L), six global-hybrid meta-GGAs (the other M0x functionals), three range-separated functionals, and one GGA. Two of the range-separated functionals and the GGA contain empirical dispersion corrections. Furthermore, there is a difference in philosophy for the range-separated functions. In particular, HSE is a so-called screened functional with Hartree–Fock exchange at small r_{12} and local exchange at large r_{12} , whereas LC- ω PBE and ω B97X-D are so-called long-range-corrected functionals with Hartree–Fock exchange at large r_{12} and local exchange or mainly local exchange, respectively, at small r_{12} .

The M08-SO functional may also be singled out for special attention because it satisfies the second-order gradient expansion⁵⁷ (i.e., it is correct through second order in s_σ) in both exchange and correlation; the former is a constraint avoided in most constraint-based density functionals because, prior to M08-SO, it was not possible to make a functional that satisfied this constraint and was highly accurate for bond energies. The TPSS⁵⁸ and TPSSh⁵⁹ functionals (a meta-GGA and a global-hybrid meta-GGA, respectively) are also correct through second order in both exchange and correlation. Therefore, we have included both TPSS and TPSSh in our comparison. Because M08-SO and TPSSh are built from the same ingredients (they are both

Table 2. Performance of Density Functionals for the Reaction Energies of the SIE11 Database (kcal/mol)

reaction	reference	M08-HX	M06-HF	M08-SO	LC- ω PBE	M06-2X	M05-2X	ω B97X-D	M06	HSE	M06-L	B97-D
$\text{He}_2^+ \rightarrow \text{He} + \text{He}^+$	57.4	62.3	59.5	64.5	71.3	65.6	64.9	71.2	70.6	70.0	71.6	78.4
$(\text{NH}_3)_2^+ \rightarrow \text{NH}_3 + \text{NH}_3^+$	35.3	38.5	38.9	38.1	38.3	39.1	40.1	41.5	38.7	41.6	41.2	46.4
$(\text{H}_2\text{O})_2^+ \rightarrow \text{H}_2\text{O} + \text{H}_2\text{O}^+$	37.3	43.7	39.6	44.8	44.5	44.3	43.9	47.3	46.2	48.4	51.2	55.0
$\text{C}_4\text{H}_{10}^+ \rightarrow \text{C}_2\text{H}_5 + \text{C}_2\text{H}_5^+$	35.3	37.1	39.8	36.2	35.3	37.9	39.5	38.8	38.9	38.7	39.4	41.2
$(\text{CH}_3)_2\text{CO}^+ \rightarrow \text{CH}_3 + \text{CH}_3\text{CO}^+$	22.6	24.3	25.8	25.6	27.0	25.9	27.9	28.7	24.4	29.8	26.2	25.9
$\text{ClFCl} \rightarrow \text{ClCIF}$	-1.0	0.8	0.2	5.3	-0.1	4.6	2.5	5.4	7.3	10.8	19.1	16.9
$\text{C}_2\text{H}_4 \cdots \text{F}_2 \rightarrow \text{C}_2\text{H}_4 + \text{F}_2$	1.1	1.0	0.6	0.8	0.2	0.8	0.9	0.6	0.6	1.2	1.2	2.7
$\text{C}_6\text{H}_6 \cdots \text{Li} \rightarrow \text{Li} + \text{C}_6\text{H}_6$	9.5	8.0	8.5	8.2	8.2	8.5	11.2	3.6	5.2	5.2	5.8	12.6
$\text{NH}_3 \cdots \text{ClF} \rightarrow \text{NH}_3 + \text{ClF}$	10.5	10.3	11.0	9.8	9.8	10.4	11.4	11.2	11.9	13.9	13.5	14.4
$\text{NaOMg} \rightarrow \text{MgO} + \text{Na}$	69.6	79.5	84.1	78.0	82.3	83.2	82.9	79.6	77.3	74.2	83.0	71.6
$\text{FLiF} \rightarrow \text{Li} + \text{F}_2$	94.4	106.1	84.2	108.6	103.5	107.2	112.5	109.0	120.7	115.9	128.1	128.7
MSE		3.6	1.8	4.4	4.4	5.1	6.0	5.9	6.4	7.1	9.9	11.1
MUE		3.9	3.9	4.8	4.9	5.3	6.0	7.1	7.2	7.9	10.5	11.1

Table 3. Performance of Density Functionals for the Reaction Energies of the DC9 Database (kcal/mol)

reaction	reference	ω B97X-D	M08-SO	M08-HX	M06-2X	M05-2X	M06	M06-HF	HSE	LC- ω PBE	M06-L	B97-D
2-pyridone \rightarrow 2-hydroxypyridine	-1.0	0.4	-1.5	-1.0	-1.7	-1.7	0.5	-3.4	0.2	0.3	1.8	2.1
$(\text{C}_{20})_{\text{case}} \rightarrow (\text{C}_{20})_{\text{bowl}}$	-13.3	-18.7	-25.6	-29.1	-31.4	-31.0	-19.7	-45.0	-8.5	-6.4	-22.2	-29.4
hepta-1,2,3,5,6-hexaene \rightarrow hepta-1,3,5-triyne	-14.3	-12.0	-13.1	-12.4	-13.5	-12.1	-11.7	-16.0	-6.6	-17.0	-5.8	-1.9
2 tetramethylethene \rightarrow octamethylcyclobutane	-19.2	-18.8	-18.8	-17.3	-16.6	-16.1	-14.3	-20.7	-7.6	-16.6	-10.6	-10.8
$(\text{CH})_{12} \rightarrow$ isomerization	-19.5	-23.7	-26.3	-24.9	-22.2	-20.9	-23.9	-19.4	-26.6	-51.4	-19.5	2.8
isomerization of carbo-[3]-oxocarbon	-26.9	-12.5	-7.5	-9.8	-10.5	-11.7	-6.0	-18.9	-4.9	1.3	-1.3	-16.5
$\text{N}_2\text{CH}_2 + \text{C}_3\text{H}_4 \rightarrow (\text{CH}_2)_3\text{N}_2$	-38.1	-38.3	-37.5	-37.6	-37.1	-39.0	-31.0	-43.2	-38.6	-45.8	-27.8	-24.1
$4\text{Be} \rightarrow \text{Be}_4$	-88.4	-91.4	-86.8	-86.5	-96.7	-94.0	-103.9	-71.7	-113.5	-98.2	-121.3	-98.5
$4\text{S}_2 \rightarrow \text{S}_8$	-101.0	-93.3	-98.4	-95.1	-99.9	-92.4	-105.7	-107.5	-96.2	-92.0	-95.1	-71.4
MSE		1.5	0.7	0.9	-0.9	0.3	0.7	-2.7	2.1	-0.5	2.2	8.2
MUE		4.3	5.0	5.6	5.7	6.2	7.5	8.2	9.4	11.1	11.5	14.0

global-hybrid meta-GGAs) and both are correct to second order in both exchange and correlation, the comparison of their performance will be especially interesting.

In the present study, we focus on testing against the four databases described in the previous section. None of these functionals has been tested against the four databases in previous studies except that the performance of B97-D for ALK6 has been given in a previous study.⁴⁴

The geometries for the molecules in the various databases were taken from previous work where they were optimized at different levels of theory. Details are given in previous papers.^{19,20,22}

All of the calculations in the present work employed the def2-QZVP basis set^{60,61} with the *Gaussian 09* program⁶² and a locally modified *Gaussian 03* program.⁶³

4. RESULTS AND DISCUSSION

In this section, we will present the results for the four databases, and we will gauge the quality of the tested density functionals by mean unsigned errors (MUEs), which are the mean absolute deviations of calculated values from database reference values, and by mean signed errors (MSEs, same as mean deviations), which are used to detect systematic deviations.

4.1. Performance for Self-Interaction-Error Database. Most density functionals suffer from self-interaction error to at least some extent, although it is possible (e.g., by using kinetic energy density^{18,64} or by making the Perdew–Zunger self-interaction correction)⁶⁵ to eliminate one-electron self-interaction error for

systems with an integer number of electrons. Recently, Cohen et al.⁶⁶ investigated self-interaction error by using simplified model systems, and they concluded that the error can be traced to the delocalization error (due to fractional charges, such as in the dissociated H_2^+ system) and static correlation error (due to fractional spins, such as in the dissociated H_2 system). Note, however, that approximate density functionals sometimes benefit in an indirect way from self-interaction error; for example, for transition metal systems with high multireference⁶⁷ character, local functionals having larger self-interaction errors often perform better than the Hartree–Fock theory, which has no self-interaction error, at least for physical systems, and hybrid functionals that have reduced self-interaction error. Handy and Cohen⁶⁸ concluded that local exchange functionals contain left–right correlation energies associated to some extent with the same properties that lead to self-interaction error. From a practical point of view, one can try to control these effects and hope to achieve useful accuracy for multireference systems. Tests against the SIE11 database provide a measure of whether such hopes are realized. The performance for the SIE11 database is given in Table 2.

The isomerization energy of ClFCl to ClCIF is a difficult case, and—of the 11 functionals in Table 2—only LC- ω PBE predicts the correct sign of the reaction energy in this case.

The other 10 reactions are dissociation reactions. All tested functionals give large errors for the dissociation energies of NaOMg and FLiF. Furthermore, Table 2 shows that all tested

Table 4. Performance of Density Functionals for the Reaction Energies of DARC14 Database (kcal/mol)

reactions	reference	M08-SO	M05-2X	M06-HF	M08-HX	ω B97X-D	M06-2X	HSE	M06	LC- ω PBE	M06-L	B97-D
ethene + butadiene	-43.8	-46.1	-46.7	-45.4	-46.8	-47.5	-45.8	-47.6	-44.2	-55.3	-41.7	-33.8
ethyne + butadiene	-59.3	-61.7	-63.5	-60.7	-62.6	-64.4	-61.0	-66.6	-58.5	-71.3	-58.7	-51.1
ethene + cyclopentadiene	-30.0	-29.7	-29.1	-30.4	-29.5	-29.3	-28.1	-29.1	-27.2	-36.7	-23.7	-17.5
ethyne + cyclopentadiene	-33.1	-33.4	-33.1	-33.6	-33.0	-33.4	-30.9	-35.2	-29.0	-40.3	-28.1	-22.5
ethene + cyclohexadiene	-36.5	-37.8	-38.0	-38.9	-37.8	-37.3	-36.2	-36.8	-34.4	-44.7	-30.8	-25.7
ethyne + cyclohexadiene	-48.2	-49.3	-50.5	-51.0	-49.4	-49.3	-47.1	-51.1	-43.9	-56.2	-43.0	-38.1
furan + maleic anhydride (endoproduct)	-14.4	-12.3	-12.0	-14.6	-11.1	-11.3	-10.3	-8.9	-7.4	-17.7	-2.5	0.9
furan + maleic anhydride (exoproduct)	-16.2	-13.8	-13.9	-16.3	-12.8	-13.1	-12.2	-11.1	-9.6	-19.8	-4.1	-0.7
furan + maleimide (endoproduct)	-17.2	-14.9	-14.9	-17.4	-13.9	-14.2	-13.1	-11.7	-10.1	-20.1	-5.4	-2.1
furan + maleimide (exoproduct)	-19.2	-16.6	-17.0	-19.3	-15.8	-16.1	-15.2	-14.2	-12.5	-22.4	-7.2	-4.0
cyclopentadiene + maleic anhydride (endoproduct)	-31.6	-30.9	-31.0	-35.2	-30.5	-30.6	-29.4	-27.4	-26.6	-37.1	-21.5	-17.9
cyclopentadiene + maleic anhydride (exoproduct)	-32.1	-31.3	-31.4	-35.5	-30.9	-31.1	-29.9	-28.2	-27.2	-37.8	-22.2	-18.8
cyclopentadiene + maleimide (endoproduct)	-34.1	-33.1	-33.4	-37.6	-32.9	-33.1	-31.8	-29.9	-29.0	-39.2	-24.2	-20.7
cyclopentadiene + maleimide (exoproduct)	-34.4	-33.5	-33.8	-37.8	-33.3	-33.6	-32.3	-30.6	-29.5	-39.7	-24.7	-21.5
MSE		0.4	0.1	-1.7	0.7	0.4	1.9	1.5	4.3	-6.3	8.0	12.6
MUE ^c		1.5	1.7	1.7	2.0	2.0	2.4	3.9	4.4	6.3	8.0	12.6

functionals give positive MSEs, which means that they tend to overestimate the dissociation energies in the SIE11 database. The best performers in Table 2 are M08-HX and M06-HF, whereas the worst performers are M06-L and B97-D. The latter is not surprising because both M06-L and B97-D are local functionals (they have $X = 0$ in Table 1). As shown in Table 2, the two long-range corrected density functionals (LC- ω PBE and ω B97X-D), which both have full Hartree–Fock exchange at long-range, perform better than the screened HSE functional, which does not have Hartree–Fock exchange at long-range. And yet, despite having full Hartree–Fock exchange at long-range, ω B97X-D does not perform as well as five of the hybrid-meta GGAs, and LC- ω PBE does not perform as well as three of them.

4.2. Performance for the DC9 Database. The first case in DC9 is the energy of tautomerization of 2-pyridone to 2-hydroxypyridine, which has been shown⁶⁵ to be a difficult case for many popular density functionals, which cannot predict the correct sign of the tautomeric energy. Six of the density functionals in Table 3 predict the incorrect sign of this tautomeric energy. The long-range corrections (LC- ω PBE and ω B97X-D) and empirical dispersion corrections (as in ω B97X-D and B97-D) do not solve the problem. Only when the functionals have a high percentage of Hartree–Fock exchange, as in M05-2X, M06-2X, M06-HF, M08-HX, and M08-SO, can they predict the correct sign of the tautomerization energy of 2-pyridone. This is in agreement with the finding of Piacenza and Grimme⁶⁹ that only BHandHLYP (which has 50% HF exchange) gives the correct sign, whereas B3LYP ($X = 20$), PBE ($X = 0$), and BP86 ($X = 0$) predict the wrong sign.

The second system in Table 3 is the isomerization energy of C₂₀, and it is a difficult case for our M0x functionals, and most of the M0x functionals give large errors for this case. HSE and ω B97X-D give the best results for this reaction.

The third case is one of the cases in our π IE3 database,¹⁹ and the fourth case is the reaction energy of the pericyclic addition reaction of two tetramethylethene molecules to form an octamethylcyclobutane molecule. Most of the functionals in Table 3 predict both of these reaction energies within 5 kcal/mol with the exceptions being the two local functionals (M06-L and B97-D) and the screened HSE functional, which is local at large r_{12} .

The fifth datum in Table 3 is the isomerization energy of (CH)₁₂ isomers, which have been shown by Schreiner et al.⁷⁰ and us⁵⁵ to be a difficult case for many popular density functionals. In the tests of Table 3, only LC- ω PBE and B97-D give large errors for this case.

The sixth case of DC9 is a very challenging isomerization of carbo-[3]-oxocarbon from a monocyclic isomer, which is the ring carbo-mer of the neutral oxidized form of dihydroxycyclopropenone, to a tetracyclic isomer, which is the trioxo derivative of tricyclopentadiene. This isomerization was first reported by Lepetit et al.⁷¹ Most of the tested functionals give large errors for this case, and the best performer is M06-HF, which gives an error of 8 kcal/mol.

The seventh datum of DC9 is the reaction energy for the 1,3-dipolar cycloaddition between ethylene and diazomethane, and this is a less demanding case for the tested functionals. Only the local M06-L and B97-D give large errors for this case.

The eighth and ninth cases in Table 3 are polycondensation reactions, first of four Be atoms to form Be₄ and then of four S₂ molecules to form an S₈ molecule. For Be₄, the M08-HX and M08-SO density functionals give the best agreement with the reference data, whereas M06-L overestimates the exothermicity by a large margin. For S₈, the M06-2X density functional gives the best prediction, and B97-D severely underestimates the association energy.

Overall, ω B97X-D and M08-SO give the best performance for the DC9 database, whereas M06-L and B97-D are the worst performers.

4.3. Performance for Diels–Alder Reaction Energies. Johnson et al.⁵⁶ employed the DARC14 database to illustrate the delocalization errors in DFT. They concluded that the DFT errors arise primarily from the overstabilization of the conjugated reactants relative to the unconjugated products. They also showed that M05-2X (among the 12 functionals tested in their paper) gives the best performance for Diels–Alder reactions. This is consistent with the results in Table 3, where M08-SO (which was not included in their analysis) and M05-2X give the best performance. The slight difference between our MUE for M05-2X and that of Johnson et al.⁵⁶ is due to the different basis sets employed.

Table 5. Performance of Density Functionals for the Reaction Energies of the ALK6 database (kcal/mol)

	reference	M06-2X	LC- ω PBE	M05-2X	ω B97X-D	HSE	M08-SO	M08-HX	M06	M06-HF	B97-D	M06-L
$C_6H_6 \cdots Li^+ \rightarrow C_6H_6 + Li^+$	38.4	41.2	41.3	42.7	38.5	40.3	42.1	43.5	36.1	49.1	47.7	34.6
$C_6H_6 \cdots Na^+ \rightarrow C_6H_6 + Na^+$	25.0	26.3	25.4	27.3	25.2	25.2	27.1	27.9	22.4	31.7	32.1	21.8
$C_6H_6 \cdots K^+ \rightarrow C_6H_6 + K^+$	19.2	19.7	17.6	20.0	18.7	17.8	20.1	20.5	16.7	22.7	21.3	16.6
$Li_8 \rightarrow 4 Li_2$	83.2	83.0	84.6	83.3	78.3	90.8	86.6	80.8	78.5	77.1	81.1	88.3
$Na_8 \rightarrow 4 Na_2$	54.6	54.6	52.1	51.9	51.5	53.8	54.4	51.5	57.2	49.3	45.2	70.0
$K_8 \rightarrow 4 K_2$	47.1	49.0	45.8	46.5	43.4	45.3	51.6	48.7	53.8	44.7	32.6	65.8
MSE		1.1	-0.1	0.7	-2.0	1.0	2.4	0.9	-0.5	1.2	-1.3	4.9
MUE		1.1	1.7	1.8	2.1	2.3	2.5	2.8	3.6	5.8	7.4	8.1

Table 6. Additional Density Functionals for Which Calculations Are Reported Here

name	year	ref	meta?	hybrid?	X	-D?
B3PW91	1993	72	no	global	20	no
B98	1998	73	no	global	21.98	no
VS98	1998	74	yes	no	0	no
PBE0	1999	75, 76	no	global	25	no
τ -HCTHh	2002	77	yes	global	15	no
TPSSh	2003	59	yes	global	10	no
TPSS1KCIS	2005	58, 78	yes	global	13	no
B97-3	2005	79	no	global	26.93	no
M05	2005	18	yes	global	28	no

4.4. Performance for the ALK6 Database. Grimme et al.⁴⁴ have shown that some density functionals with empirical dispersion corrections significantly overbind alkali-metal-cation–benzene complexes, as seen in the large error of B97-D for the $C_6H_6 \cdots Li^+$ complex (Table 5). Table 5 shows that M06-HF also severely overbinds these complexes, and M06-L strongly overestimates the dissociation energies of the alkali metal clusters: $M_8 \rightarrow 4 M_2$ ($M = Na, K$). Among the tested functionals, M06-2X, LC- ω PBE, and M05-2X give good performance for both kinds of data, and they have small MUEs (less than 2 kcal/mol) for the ALK6 database. The ω B97X-D density functional also gives a small MUE of 2.1 kcal/mol, which is less than the MUEs of the other dispersion-corrected methods.

4.5. Overall Performance. To keep the discussion to a reasonable length while still illustrating in detail the variation in results that can be obtained with various density functionals, we singled out 11 density functionals for detailed discussion in sections 4.1–4.4. But we have also made calculations for nine other density functionals, and in this section, we compare the mean errors for these functionals to those for the functionals already discussed and to those for 10 density functionals tested by Georik and Grimme.¹⁴ This affords a comparison of 30 density functionals for the four reaction energy databases.

The nine additional density functionals^{18,59,72–79} for which we carried out calculations are characterized in Table 6, which has the same format as Table 1. Eight of these functionals were selected because, along with M06, M05, and M06-L, they had the best performance (out of 34 density functionals tested) for the recent tests⁴³ of density functionals against a broad catalytic energies database. (Functionals with more than 28% nonlocal exchange at all or small interelectronic separations were not included in those tests.)

Finally, our comparisons include 10 of the density functionals studied by Grimme and co-workers.^{14,44,53} These consist of four

Table 7. Previously Tested Density Functionals Included in the Present Comparisons

name	year	ref	meta?	hybrid?	X	-D?
BLYP	1988	80, 81	no	no	20	no
B3LYP	1994	82	no	global	21.98	no
PBE	1996	83	no	no	0	no
TPSS	2003	58	yes	no	25	no
B2PLYP	2006	84	no	global	53(27) ^a	no
BLYP-D	1988	53	no	no	20	yes
B3LYP-D	1994	53	no	global	21.98	yes
PBE-D	1996	53	NO	no	0	yes
TPSS-D	2003	53	yes	no	25	yes
B2PLYP-D	2006	53	no	global	53(27) ^a	yes

^aThe value in parentheses is the fraction of local correlation energy that is replaced by a nonlocal second-order perturbation energy term.

standard density functionals, BLYP,^{80,81} B3LYP,⁸² PBE,⁸³ TPSS,⁵⁸ and B2PLYP,⁸⁴ plus each of these functionals with an empirical dispersion correction. The dispersion term in these functionals, like that in B97-D, is Grimme's second generation version⁵³ (sometimes⁴⁴ called D2), but an important difference from B97-D is that the parameters of B97-D were optimized in the presence of dispersion, but in the five dispersion-corrected density functionals in Table 7, the parameters were not reoptimized when dispersion was added.

B2PLYP deserves a special note because it is considered here to illustrate the effect of a third way to include a dependence on the SCF orbitals. In particular, B2PLYP includes a post-SCF term that depends on the unoccupied orbitals; it has the form of a second-order perturbation approximation to the dynamical correlation energy, although the first-order term, which does not vanish unless Hartree–Fock orbitals are used,⁸⁵ is not included (note that the Hartree–Fock SCF orbitals and the Kohn–Sham orbitals are both functionals of the same exact density). Functionals that include nonlocal correlation terms as well as nonlocal Hartree–Fock exchange are called doubly hybrid. For practical work, it is important to note that the doubly hybrid B2PLYP and B2PLYP-D methods have higher cost, steeper dependence of cost on system size, and greater basis set dependence than the other density functionals considered in Table 8.

The mean unsigned errors and average mean unsigned errors of all 30 density functionals of Tables 1, 6, and 7 are given in Table 8.

For perspective, we review the effects of empirical dispersion corrections for the performance of the approximate density functionals by first considering the comparison of the MUEs of B3LYP to those of B3LYP-D. From Table 8, we can see that

Table 8. Mean Unsigned Errors for all Four Databases (kcal/mol)

method	DARC14	SIE11	DC9	ALK6	AMUE ^d
M08-SO	1.5	4.8	5.0	2.5	3.4
M08-HX	2.0	3.9	5.6	2.8	3.6
M06-2X	2.4	5.3	2.7	1.1	3.7
ω B97X-D	2.0	7.1	4.3	2.1	3.9
M05-2X	1.7	6.0	6.2	1.8	3.9
B2PLYP-D ^b	3.7	4.6	5.5	4.4	4.5
M06-HF	1.7	3.9	8.2	5.8	4.9
M06	4.4	7.2	7.5	3.6	5.7
HSE	3.9	7.9	9.4	2.3	5.9
LC- ω PBE	6.3	4.9	11.1	1.7	6.0
PBE0	3.6	7.7	10.3	2.3	6.0
B2PLYP ^b	7.8	4.7	7.7	4.2	6.1
M05	5.4	6.9	10.4	1.8	6.1
B3PW91	6.9	7.9	9.7	3.2	6.9
B97-3	8.7	7.3	9.2	3.9	7.3
B98	8.3	8.3	8.4	4.1	7.3
τ -HCTHh	7.8	9.4	8.9	3.3	7.3
TPSSh	8.2	9.2	10.4	2.1	7.5
TPSS-D ^b	3.7	11.6	8.7	7.2	7.8
PBE-D ^b	2.8	12.6	9.9	6.8	8.0
TPSS1KCIS	9.8	8.6	11.8	2.8	8.2
PBE ^b	6.8	12.6	10.8	2.8	8.2
B3LYP-D ^b	7.6	8.1	9.9	8.2	8.5
TPSS ^b	11.1	10.7	12.3	2.3	9.1
M06-L	8.0	10.5	11.5	8.1	9.6
B97-D	12.6	11.1	14.0	7.4	11.3
B3LYP ^b	15.4	7.6	15.1	9.1	11.8
BLYP-D ^b	14.0	12.2	14.5	9.0	12.4
VS98	6.8	9.3	16.6	24.2	14.2
BLYP ^b	22.9	11.7	20.3	10.8	16.4

^a Average of MUEs for the DARC14, SIE11, DC9, and ALK6 databases. ^b Results for these functionals are taken from previous studies,^{14,44} and all DFT-D methods are the DFT-D2 versions of Grimme et al.^{44,53}

empirical dispersion corrections improve the performance of B3LYP for the DARC14 and DC9 databases by a large margin but slightly deteriorate the performance for the SIE11 and ALK6 databases. This trend can also be seen in the other four pairs of functionals involving the methods of Table 7. Although the empirical dispersion correction does not improve upon the SIE11 and ALK6 databases, and although the empirical dispersion correction can lead to unphysical results in some cases, adding a molecular mechanics term does show some improvement in the performance of density functionals that are not accurate for medium-range correlation energy.

Next, consider the results for the functionals of Table 6 for the individual databases in Table 8. The best performance for the reaction energies sensitive to self-interaction error is obtained with M05 and PBE0, with MUEs of 6.9 and 7.7 kcal/mol, as compared to 3.9 kcal/mol for M08-HX and M06-HF. However, PBE0 and M05 have MUEs of 10.3 and 10.4 kcal/mol for the DC difficult cases database. For DC9, the best performance of the Table 6 density functionals is obtained by B98 with an MUE of 8.4 kcal/mol, in comparison to 4.3 and 5.0 kcal/mol for the

ω B97X-D and M08-SO density functionals, respectively. For the Diels–Alder reactions, PBE0 and M05 do the best of the functionals in Table 6, with MUEs of 3.6 and 4.4 kcal/mol, as compared to 1.5–1.7 kcal/mol for M08-SO, M05-2X, and M06-HF, whereas B98 has an MUE of 8.3 kcal/mol; in fact, the MUE of B98 is remarkably constant at 8.3–8.4 kcal/mol for SIE11, DC9, and DARC14.

For the ALK6 reaction energy database, M05 and PBE0 are again the best of the functionals in Table 6, with MUEs of 1.8 and 2.3 kcal/mol, as compared to 1.1–1.8 kcal/mol for M06-2X, LC- ω PBE, and M05-2X.

The AMUEs in Table 8 are averages over all four reaction energy databases, and they show that M08-SO, M08-HX, and M06-2X give the best overall performance for the reaction energies considered in this article, with average mean unsigned errors of 3.7 kcal/mol or less. The comparison of the two global-hybrid meta GGAs that satisfy second-order constraints (as discussed above) shows an AMUE of 3.4 kcal/mol for M08-SO and 7.5 kcal/mol for TPSSh.

Casting a wider net and looking at the top nine entries in Table 8 (as in Tables 2–5, the density functionals are listed in Table 8 in order of increasing mean unsigned errors), we see that six Minnesota hybrid meta functionals plus ω B97X-D, B2PLYP-D, and HSE give the best performance, with average mean unsigned errors of 5.9 kcal/mol or less.

5. CONCLUSIONS

It is important to validate practical density functional approximations in order to ascertain the reliability of their predictive capabilities in many areas of chemistry. Here, we have performed benchmark calculations aimed at testing the M0x functionals against some recently developed databases of reaction energies that are challenging for popular density functionals and are outside the fitting data of the M0x functionals and most or all other functionals. We placed a special emphasis on comparing the recent Minnesota meta and hybrid meta functionals, range-separated functionals, functionals with empirical dispersion corrections, functionals that performed well on a recent test for energetic quantities relevant to catalysis, and global-hybrid meta-GGA functionals accurate to second order in the gradient expansion.

The examination of these reaction energies has allowed us to identify some especially problematic cases: (1) In the database of reaction energies sensitive to self-interaction error, the energy of isomerization of ClFCl to ClClF is a challenging case. Only LC- ω PBE predicts the correct sign for it. (2) M08-HX performs the best for the SIE11 database, but it gives large errors for the dissociation energies in NaOMg and FLiF. (3) In the DC9 database, the isomerization energy of (C₂₀)_{cage} to (C₂₀)_{bowl} and the isomerization energy of carbo-[3]-oxocarbon from monocycle to tetracycle are difficult cases for the M0x functionals. (4) In the alkali metal database, M06-HF severely overbinds alkali-metal-cation–benzene complexes, and M06-L strongly overestimates the dissociation energies of the alkali metal clusters: M₈ → M₂ (M = Na, K). (5) Empirical dispersion corrections improve the performance of approximate density functionals for the DARC14 and DC9 databases by a large margin but slightly deteriorate the performance for the SIE11 and ALK6 databases.

The final average mean unsigned error (AMUE) provides an overall assessment, averaged over the four databases. M08-SO, M08-HX, M06-2X, ω B97X-D, and M05-2X give the overall smallest AMUEs.

The generally good success of the most recent hybrid meta density functionals, especially M06-2X, M06-HX, and M08-SO, is very encouraging, as is the success of the range-separated and dispersion-corrected ω B97X-D density functional. The hybrid meta functionals were developed by a combination of satisfying constraints such as the uniform electron gas limit and the second-order gradient expansion and by fitting parameters to experimental and high-quality theoretical reference data, and they have previously been shown to perform well in many cases outside their training set, including several cases where popular older functionals fail. The ω B97X-D functional, also optimized by a parameter-fitting strategy, achieves almost as good performance in the present tests by a different route, namely, range separation and empirical dispersion rather than employing kinetic energy density.

■ ASSOCIATED CONTENT

S Supporting Information. Full results for the density functionals of Table 6. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: yan.zhao3@hp.com (Y.Z.), truhlar@umn.edu (D.G.T.).

■ ACKNOWLEDGMENT

This work was supported in part by the Air Force Office of Scientific Research (AFOSR) and the National Science Foundation (NSF).

■ REFERENCES

- (1) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, 1133. von Barth, U.; Hedin, L. *J. Phys. C* **1972**, *5*, 1629. Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 1372. Seidl, A.; Gorling, P.; Vogl, J.; Majewski, A.; Levy, M. *Phys. Rev. B* **1996**, *53*, 3764.
- (2) Kohn, W.; Becke, A. D.; Parr, R. G. *J. Phys. Chem.* **1996**, *100*, 12974.
- (3) Scuseria, G. E.; Staroverov, V. N. In *Theory and Application of Computational Chemistry: The First 40 Years*; Dykstra, C. E., Frenking, G., Kim, K. S., Scuseria, G. E., Eds.; Elsevier: Amsterdam, 2005; p 669.
- (4) Cramer, C. J.; Truhlar, D. G. *Phys. Chem. Chem. Phys.* **2009**, *11*, 10757.
- (5) Guner, V.; Khuong, K. S.; Leach, A. G.; Lee, P. S.; Bartberger, M. D.; Houk, K. N. *J. Phys. Chem. A* **2003**, *107*, 11445.
- (6) Ess, D. H.; Houk, K. N. *J. Phys. Chem. A* **2005**, *109*, 9542.
- (7) Zhao, Y.; Pu, J.; Lynch, B. J.; Truhlar, D. G. *Phys. Chem. Chem. Phys.* **2004**, *6*, 673.
- (8) Zhao, Y.; González-García, N.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 2012.
- (9) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215.
- (10) Zheng, J.; Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2009**, *5*, 808.
- (11) Grimme, S.; Mück-Lichtenfeld, C.; Würthwein, E.-U.; Ehlers, A. W.; Goumans, T. P. M.; Lammertsma, K. *J. Phys. Chem. A* **2006**, *110*, 2583.
- (12) Karton, A.; Turnopolsky, A.; Lamere, J. F.; Schatz, G. C.; Martin, J. M. L. *J. Phys. Chem. A* **2008**, *112*, 12868.
- (13) Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2008**, *4*, 1849.
- (14) Goerigk, L.; Grimme, S. *J. Chem. Theory Comput.* **2010**, *6*, 107.
- (15) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. *J. Chem. Phys.* **2000**, *112*, 7374. Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. *J. Chem. Phys.* **2005**, *123*, 1.
- (16) Izgorodina, E. I.; Coote, M. L.; Radom, L. *J. Phys. Chem. A* **2005**, *109*, 7558.
- (17) Izgorodina, E. I.; Brittain, D. R. B.; Hodgson, J. L.; Krenske, E. H.; Lin, C. J.; Namazian, M.; Coote, M. L. *J. Phys. Chem. A* **2007**, *111*, 10754.
- (18) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2*, 364.
- (19) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2006**, *110*, 10478.
- (20) Zhao, Y.; Truhlar, D. G. *J. Chem. Phys.* **2006**, *125*, 194101.
- (21) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2008**, *112*, 1095.
- (22) Johnson, E. R.; Mori-Sanchez, P.; Cohen, A. J.; Yang, W. *J. Chem. Phys.* **2008**, *129*, 204112.
- (23) Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2005**, *1*, 415.
- (24) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem.* **2005**, *109*, 5656.
- (25) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2006**, *110*, 5121.
- (26) Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2*, 1009.
- (27) Jurecka, P.; Sponer, J.; Cerny, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985.
- (28) Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2007**, *3*, 289.
- (29) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. C* **2008**, *112*, 6860.
- (30) Takatani, T.; Hohenstein, E. G.; Malagoli, M.; Marshall, M. S.; Sherill, D. C. *J. Chem. Phys.* **2010**, *132*, 144104.
- (31) Schultz, N.; Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 4388.
- (32) Schultz, N.; Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 11127.
- (33) Zhao, Y.; Truhlar, D. G. *J. Chem. Phys.* **2006**, *124*, 224105.
- (34) Furche, P.; Perdew, J. P. *J. Chem. Phys.* **2006**, *124*, 044103.
- (35) Quintal, M. M.; Karton, A.; Iron, M. A.; Boese, A. D.; Martin, J. M. L. *J. Phys. Chem. A* **2006**, *110*, 709.
- (36) Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2006**, *110*, 13126.
- (37) Jacquemin, D.; Perpète, E. A.; Scuseria, G. E.; Ciofini, I.; Adamo, C. *J. Chem. Theory Comput.* **2008**, *4*, 123.
- (38) Silva-Junior, M. R.; Schreiber, M.; Sauer, S. P. A.; Thiel, W. *J. Chem. Phys.* **2008**, *129*, 104103.
- (39) Jacquemin, D.; Perpète, E. A.; Scuseria, G.; Ciofini, I.; Adamo, C. *Chem. Phys. Lett.* **2008**, *465*, 226.
- (40) Georigk, L.; Moellmann, J.; Grimme, S. *Phys. Chem. Chem. Phys.* **2009**, *22*, 4611.
- (41) Jacquemin, D.; Perpète, E. A.; Ciofini, I.; Adamo, C.; Valero, R.; Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2010**, *6*, 2071.
- (42) Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2009**, *5*, 324.
- (43) Yang, K.; Zheng, J.; Zhao, Y.; Truhlar, D. G. *J. Chem. Phys.* **2010**, *132*, 164117.
- (44) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 164104.
- (45) Heyd, J.; Scuseria, G. E.; Ernzerhof, M. *J. Chem. Phys.* **2003**, *118*, 8207; **2006**, *124*, 219906(E).
- (46) Heyd, J.; Scuseria, G. E. *J. Chem. Phys.* **2004**, *121*, 1187.
- (47) Heyd, J.; Scuseria, G. E. *J. Chem. Phys.* **2006**, *124*, 219906.
- (48) Vydrov, O. A.; Heyd, J.; Krukav, A. V.; Scuseria, G. E. *J. Chem. Phys.* **2006**, *125*, 074106.
- (49) Vydrov, O. A.; Scuseria, G. E. *J. Chem. Phys.* **2006**, *125*, 234109.
- (50) Zhao, Y.; Schultz, N.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2*, 364.
- (51) Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157.
- (52) Zhao, Y.; Truhlar, D. G. In *Theoretical and Computational Methods in Mineral Physics: Geophysical Applications*; Wentzcovitch, R. M., Stixrude, L., Ed.; Reviews in Mineralogy and Geochemistry 71; Mineralogical Society of America: Chantilly, VA, 2010; pp 19–37.
- (53) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787.
- (54) Chai, J.-D.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615.
- (55) Zhao, Y.; Truhlar, D. G. *Org. Lett.* **2006**, *8*, 5753.

- (56) Johnson, E. R.; Mori-Sanchez, P.; Cohen, A. J.; Yang, W. *J. Chem. Phys.* **2008**, *129*, 204112.
- (57) Zhao, Y.; Truhlar, D. G. *J. Chem. Phys.* **2008**, *124*, 184109.
- (58) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys. Rev. Lett.* **2003**, *91*, 146401.
- (59) Staroverov, V. N.; Scuseria, G. E.; Tao, J.; Perdew, J. P. *J. Chem. Phys.* **2003**, *119*, 12129.
- (60) Weigend, F.; Furche, F.; Ahlrichs, R. *J. Chem. Phys.* **2003**, *119*, 12753.
- (61) Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297.
- (62) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, 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, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*; Gaussian, Inc.: Wallingford, CT, 2009.
- (63) Zhao, Y.; Truhlar, D. G. *MN-GFM 4.3*; University of Minnesota: Minneapolis, MN, 2010.
- (64) Becke, A. D. *J. Chem. Phys.* **1996**, *104*, 1040.
- (65) Ruzsinszky, A.; Perdew, J. P.; Csonka, G. L.; Vydrov, O. A.; Scuseria, G. E. *J. Chem. Phys.* **2007**, *126*, 104102.
- (66) Cohen, A. J.; Mori-Sanchez, P.; Yang, W. *Science* **2008**, *321*, 792.
- (67) Truhlar, D. G. *J. Comput. Chem.* **2006**, *28*, 73.
- (68) Handy, N. C.; Cohen, A. J. *Mol. Phys.* **2001**, *99*, 403.
- (69) Piacenza, M.; Grimme, S. *J. Comput. Chem.* **2004**, *25*, 83.
- (70) Schreiner, P. R.; Fokin, A. A.; Pascal, R. A., Jr.; de Meijere, A. *Org. Lett.* **2006**, *8*, 3635.
- (71) Lepetit, C.; Chermette, H.; Gicquel, M.; Heully, J.-L.; Chauvin, R. *J. Phys. Chem. A* **2007**, *111*, 136.
- (72) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
- (73) Schmider, H. L.; Becke, A. D. *J. Chem. Phys.* **1998**, *108*, 9624.
- (74) Van Voorhis, T.; Scuseria, G. E. *J. Chem. Phys.* **1998**, *109*, 400.
- (75) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158.
- (76) Ernzerhof, M.; Scuseria, G. E. *J. Chem. Phys.* **1999**, *110*, 5029.
- (77) Boese, A. D.; Handy, N. C. *J. Chem. Phys.* **2002**, *116*, 9559.
- (78) Kreiger, J. B.; Chen, J.; Iafrate, G. J.; Savin, A. In *Electron Correlation and Materials Properties*; Gonis, A., Kioussis, N., Eds.; Plenum: New York, 1999; p 463.
- (79) Keal, T. W.; Tozer, D. J. *J. Chem. Phys.* **2005**, *123*, 121103.
- (80) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
- (81) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
- (82) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.
- (83) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.
- (84) Grimme, S. *J. Chem. Phys.* **2006**, *124*, 034108.
- (85) Zhao, Y.; Lynch, B. J.; Truhlar, D. G. *Phys. Chem. Chem. Phys.* **2005**, *7*, 43.

All-Electron Scalar Relativistic Basis Sets for the Actinides

Dimitrios A. Pantazis^{*,†,‡} and Frank Neese^{*,†,‡}

[†]Lehrstuhl für Theoretische Chemie, Institut für Physikalische und Theoretische Chemie, Universität Bonn, Wegelerstrasse 12, D-53115 Bonn, Germany

[‡]Max-Planck-Institut für Bioanorganische Chemie, Stiftstrasse 34-36, 45470 Mülheim an der Ruhr, Germany

 Supporting Information

ABSTRACT: Increasing interest in the computational modeling of actinide compounds creates the need for alternative choices when it comes to fine tuning the computational methodology in order to best fit the problem at hand. All-electron scalar relativistic density functional theory can be a useful approach for a variety of actinide systems and would benefit from atomic basis sets geared to that level of theory. In this paper we present segmented all-electron relativistically contracted (SARC) basis sets for the complete actinide series $_{89}\text{Ac}–_{103}\text{Lr}$, optimized for use with the popular Douglas–Kroll–Hess to the second order and zeroth-order regular approximation scalar relativistic Hamiltonians. The quality of the SARC basis sets is assessed in terms of their intrinsic incompleteness and contraction errors, with respect to total energies, orbital properties, and ionization energies. Calculations on diatomic Ac and Lr molecules confirm that the valence-space construction results in negligible basis set superposition errors. The performance of the basis sets is further evaluated for molecular geometries, vibrational frequencies, and bond dissociation energies in an illustrative study of uranium fluorides UF_n ($n = 1–6$).

1. INTRODUCTION

Research into the chemistry of actinide systems¹ witnesses a steady growth owing to its wide-ranging implications for environmental safety, power production and contamination recovery. The challenges of nuclear waste disposal and the management or reprocessing of spent nuclear fuel require, among others, the understanding of actinide behavior in aqueous systems and the development of chemical procedures that will allow the successful elemental separation of mixtures containing different actinide (and lanthanide) species. Advances in these fields naturally depend upon detailed and in-depth understanding of the chemical and physical properties of molecular actinide species. Laboratory study of actinide systems is often frustrated by problems arising from the intense radioactivity of most nuclides and the high toxicity and instability of the complexes as well as the obvious sourcing limitations. This situation as well as the necessity for a comprehensive understanding of electronic structure highlights the importance of theoretical approaches to actinide research, a field that attracts ever-increasing interest in recent years.^{2–7}

Compared to experimental work, computational actinide chemistry is confronted with a different but not less daunting set of challenges. In contrast to the lanthanide 4f orbitals, the 5f orbitals of the actinides are less core-like, resulting in increased 5f orbital contribution to bonding and reactivity.^{8–13} Moreover, the energetic proximity of 5f, 6p, 6d, and 7s orbitals leads to many close-lying energy levels and a large number of accessible oxidation states, at least for the early members of the series.¹ Coupled to this electronic complexity is the significant relativistic effects¹⁴ that define the profile of actinides. Correlation and relativistic effects are both large and typically of the same magnitude,¹⁵ demanding their simultaneous and theoretically adequate treatment. Faced with these challenges, computational studies of

actinide species have so far employed with varying success a variety of methodologies from the arsenal of quantum chemistry, both in terms of approximate treatments of electron correlation and in terms of approximations to the Dirac Hamiltonian.¹⁶

With regard to the basis sets used in the description of the actinide elements, effective core potentials (ECPs)^{17–21} that are adjusted to reproduce relativistic reference data have proven successful in enabling a whole range of applications. Owing to the significant reduction in cost and the convenient inclusion of relativistic effects, the use of ECPs is now an established practice for the prediction of several properties, especially for density functional theory (DFT)^{22,23} studies of larger molecules. Inherent in this approximation, however, is the reliance on a delicate balance between the more complete empirical modeling of relativistic effects achieved through a larger ECP core and the better modeling of energetics and reactivity achieved through a larger valence space. Inevitably, this balance can be difficult to achieve in practice for all possible properties of interest.^{3,15,24–27}

Most importantly, all-electron calculations become an absolute requirement when properties of the inner shells are being probed, as for example in electron paramagnetic resonance (EPR), Mössbauer, or X-ray spectroscopy. X-ray absorption spectroscopy (XAS) is a prominent example of particular importance in actinide chemistry and speciation with applications in the study of actinide compounds as a sensitive and element-specific probe of electronic structure.^{28,29} The pre-edge and edge region (X-ray absorption near-edge spectroscopy, XANES) arises from core–electron excitations into the valence and unoccupied levels and can thus provide information about the oxidation state, the properties of the frontier orbitals, and the valence structure of

Received: December 22, 2010

Published: February 16, 2011

the actinide absorber.^{30,31} Another straightforward example that clearly requires explicit consideration of all electrons is any kind of topological analysis of electron densities,^{32–34} where the replacement of the core electrons by an effective potential may create topological inconsistencies and artifacts that extend into the bonding region.³⁵

The rich variety of available all-electron basis sets for the top part of the periodic table contrasts with the more limited choices available for the bottom part, especially for the actinide elements, while most of the existing basis sets are designed for either wave function-based correlation methods or spin-including relativistic approaches.^{36–42} In the present paper we are specifically targeting a “middle-ground” theoretical level: DFT coupled with one of the widely available scalar relativistic Hamiltonians, the Douglas–Kroll–Hess^{43–47} (DKH2, to second order), or the zeroth-order regular approximation^{48–50} (ZORA). An interesting alternative to these approaches has been implemented in the PRIRODA code,⁵¹ based on a four-component one-electron scalar relativistic approximation to the full Dirac equation where all spin–orbit terms are neglected.⁵² Generally contracted correlation-consistent basis sets developed by Laikov are used in this approach.⁵³ Two prominent examples of all-electron basis sets constructed specifically for scalar relativistic Hamiltonians are the Slater-type ZORA basis sets implemented in the Amsterdam density functional code^{54,55} and the atomic natural orbital (ANO) basis sets of Roos et al.,⁵⁶ designed on the basis of CASPT2 calculations with the DKH Hamiltonian. The applicability of the Slater basis sets is unfortunately limited within software that can make efficient use of Slater-type functions. On the other hand, whereas the generally contracted atomic natural orbital (ANO) basis sets are a reliable choice for multiconfigurational correlated ab initio calculations on small systems, they are cumbersome for DFT applications on medium to large molecules. In that case a segmented contraction is preferred to reduce the computational cost associated with the generation of two-electron integrals in most general-purpose quantum chemistry codes. A no less significant consideration is that convergence of DFT-calculated quantities in terms of polarization space is achieved much faster than in ab initio correlated calculations,^{22,57} therefore an extensive high angular momentum correlation set only adds to the cost of the calculation and not to the accuracy of the results. A recent confirmation of this can be found in the ZORA-DFT study of nuclear electric field gradient (EFG) tensors by Aquino et al.⁵⁸

In light of the above, it is desirable to have a family of standardized, compact, and efficient actinide Gaussian basis sets that are best suited to the requirements of the DFT-DKH2 and DFT-ZORA levels of theory. Building upon the same approach that was used in the development of segmented all-electron relativistically contracted (SARC) basis sets for lanthanides⁵⁹ and third-row transition metals,^{60,61} we propose a series of SARC basis sets to cover the actinide elements. Exponents of the Gaussian primitives are derived from simple rules and contraction coefficients are determined separately for the ZORA and DKH2 schemes, since these two approximations produce differing shapes for the core orbitals.⁶⁰ The quality of construction of the new basis sets is evaluated here for a range of atomic and molecular properties. All SARC basis sets are part of the freely available ORCA program package.⁶²

2. RESULTS AND DISCUSSION

2.1. Construction of Basis Sets. As with the corresponding SARC basis sets for the third-row transition metals⁶⁰ and the

Table 1. Radial Expectation Values of Innermost Orbitals (in Bohr) Determined from Spin-Averaged ROHF Calculations

	$\langle r_s \rangle$	$\langle r_p \rangle$	$\langle r_d \rangle$	$\langle r_f \rangle$
Ac	0.016994	0.060121	0.146637	0.371265
Th	0.016804	0.059413	0.144672	0.363124
Pa	0.016618	0.058720	0.142757	0.355433
U	0.016436	0.058044	0.140893	0.348091
Np	0.016258	0.057382	0.139076	0.341087
Pu	0.016084	0.056735	0.137304	0.334403
Am	0.015914	0.056103	0.135578	0.327988
Cm	0.015747	0.055485	0.133894	0.321830
Bk	0.015583	0.054880	0.132251	0.315984
Cf	0.015423	0.054288	0.130647	0.310431
Es	0.015266	0.053709	0.129082	0.305040
Fm	0.015112	0.053142	0.127553	0.299853
Md	0.014962	0.052586	0.126060	0.294860
No	0.014814	0.052042	0.124602	0.289846
Lr	0.014669	0.051510	0.123177	0.285174

lanthanides,⁵⁹ the starting point was the set of innermost radial expectation values per angular momentum, obtained from atomic ground-state restricted open-shell Hartree–Fock (ROHF) calculations. The ROHF calculations followed Zerner’s spin-averaged (SA) formalism,^{63,64} a variant of the configuration-averaged approach that averages over all states of a given spin for a given configuration. Exponents from the universal Gaussian basis set (UGBS) of de Castro and Jorge^{65,66} have served well for constructing large reference basis sets in previous studies, allowing us to obtain benchmark results that approximate the basis set limit of the employed methods. In order to ensure saturation of the one-electron reference space, we used here an extended range of UGBS primitive Gaussians that goes significantly beyond the original definition for the actinides. In the following, “UGBS” will refer to this extended uncontracted (34s25p21d17f) form that yields 333 basis functions per actinide atom.

The actinide atomic configurations and corresponding states that were employed are the following: Ac ($d^1 s^2, ^2D$), Th ($d^2 s^2, ^3F$), Pa ($f^2 d^1 s^2, ^4K$), U ($f^3 d^1 s^2, ^5L$), Np ($f^4 d^1 s^2, ^6L$), Pu ($f^6 s^2, ^7F$), Am ($f^7 s^2, ^8S$), Cm ($f^7 d^1 s^2, ^9D$), Bk ($f^9 s^2, ^6H$), Cf ($f^{10} s^2, ^5I$), Es ($f^{11} s^2, ^4I$), Fm ($f^{12} s^2, ^3H$), Md ($f^{13} s^2, ^2F$), No ($f^{14} s^2, ^1S$), and Lr ($f^{14} d^1 s^2, ^2D$). As seen from the orbital occupations, electrons in 6d orbitals are lower in energy than those in 5f orbitals early in the series. The 5f orbitals begin to be occupied at protactinium, and the 6d orbitals are not occupied again in the second half of the series. This relates to the well-known distinction between the “transition-metal-like” early actinides and the “lanthanide-like” late actinides.

The individual steps of the procedure for the construction of basis sets have been described in detail previously,^{59,60} so here we draw attention to the different parameter choices that were found to be more suitable for the actinides. Point nuclei are used throughout this work. Radial expectation values $\langle r_l \rangle$ for the first s, p, d, and f orbitals obtained from the reference SA-ROHF/UGBS calculations (Table 1) were used for deriving the exponents α_l of the tightest s, p, d, and f functions according to $\alpha_l = 2k_l f_l^2 / \pi \langle r_l \rangle^2$, where $f_l = 1, 4/3, 8/5$, and $64/35$ ($l = s, p, d, f$) and k_l is a scaling factor adjusted to 25 000, 2500, 500, and 250 for s, p, d and f functions, respectively. This leads to the generator exponents listed in Table 2. Extensive testing in atomic calculations showed

Table 2. Maximum Exponents per Angular Momentum (in Bohr⁻²) Used for the Construction of the SARC Basis Sets

	α_s	α_p	α_d	α_f
Ac	55109808.74124	782789.90010	37896.83656	3860.78816
Th	56363088.88972	801557.42322	38933.29154	4035.84133
Pa	57631858.17253	820588.66007	39984.83121	4212.38926
U	58915269.20406	839813.66697	41049.82023	4391.96026
Np	60212393.78258	859302.83100	42129.44229	4574.18427
Pu	61522220.62253	879013.38137	43223.87650	4758.86803
Am	62843653.32611	898929.09050	44331.42040	4946.84248
Cm	64183659.71229	919065.41601	45453.55550	5137.96262
Bk	65541743.54208	939440.76053	46590.64590	5327.84571
Cf	66908672.67154	960041.30864	47740.96677	5522.22151
Es	68291966.01786	980851.97769	48905.61460	5719.13577
Fm	69690927.07217	1001894.08989	50085.12219	5918.71181
Md	71095290.12593	1023192.45899	51278.52058	6120.85750
No	72522948.25117	1044695.31751	52485.58789	6334.45665
Lr	73963782.78321	1066386.17254	53706.99426	6543.71139

that these exponents are sufficiently steep to capture the essential effects of the ZORA and DKH2 scalar relativistic Hamiltonians without introducing numerical instabilities that are often encountered when extremely tight functions are employed with a point-nucleus model.⁴¹ It should be noted that the use of finite nuclei typically requires more and steeper functions to properly describe the inner core region than those employed in the present basis sets, and thus a different set of exponents would need to be generated for this model. Finite nuclear volume effects are known to affect properties that depend strongly on the inner core electronic structure, such as indirect nuclear NMR nuclear spin–spin couplings⁶⁷ or hyperfine coupling constants.⁶⁸ An evaluation of nuclear volume effects for the f-block elements will be performed separately, along with a determination of the necessary SARC basis set adaptations.

Subsequently, primitive Gaussian-type functions were obtained from the series $\alpha_i x^{-i}$ ($i = 1, 2, \dots$), where the adjustable parameter x dictates the spacing and thus the overall number of the primitives. To strike a balance between the size of the final basis set, the quality of the atomic description, the efficiency of the observed convergence behavior, and the optimal values of x were determined to be 2.20, 2.40, 2.50, and 2.60 for $l = s, p, d$ and f , respectively. Test calculations showed that it is adequate to terminate the series using cutoffs of 0.02 for s functions, 0.08 for p and d functions, and 0.2 for f functions.

The resulting uncontracted basis sets have the form (29s20p-16d12f) comprising 253 primitives in total. In the last step the innermost 9s, 8p, 7d, and 6f primitives were contracted to create the final basis sets with a [21s13p10d7f] pattern and 159 functions. In terms of size this is at most half as that of uncontracted ANO basis sets⁵⁶ and thus is deemed more appropriate for routine all-electron DFT calculations. Contraction coefficients were optimized separately for the DKH2 and ZORA scalar relativistic Hamiltonians to better fit the differential effects produced by each approximation. Finally, scaling the average of the two most diffuse f exponents by 3.75 generated a set of optional single g polarization functions. In molecular calculations employing the DKH2 or ZORA Hamiltonians these actinide basis sets would be best combined with the respective SARC basis sets for the third-row transition metals⁶⁰ and the lanthanides⁵⁹ or the relativistically recontracted variants of the Karlsruhe basis sets⁶⁹

Table 3. Estimated Incompleteness Errors (E_h) from Comparison of the UGBS and SARC Basis Sets, Obtained from Spin-Averaged ROHF Calculations with the DKH2 Hamiltonian

	UGBS (333 functions)	SARC (159 functions)	ΔE
Ac	-25678.01500	-25677.00385	1.011
Th	-26420.77386	-26419.69307	1.081
Pa	-27177.98097	-27176.83141	1.150
U	-27950.04069	-27948.81676	1.224
Np	-28737.17789	-28735.87599	1.302
Pu	-29539.63524	-29538.25221	1.383
Am	-30357.79074	-30356.32206	1.469
Cm	-31191.86262	-31190.30413	1.558
Bk	-32041.77325	-32040.12197	1.651
Cf	-32908.42477	-32906.67684	1.748
Es	-33791.95896	-33790.11137	1.848
Fm	-34692.72632	-34690.77658	1.950
Md	-35611.09371	-35609.03970	2.054
No	-36547.44572	-36545.28681	2.159
Lr	-37501.95189	-37499.68810	2.264

for lighter elements,⁶⁰ all of which are part of the ORCA basis set library.⁶² Complete listings of the presently derived SARC basis sets in input format are provided in the Supporting Information.

2.2. Evaluation of Basis Sets. Given the stated goal of constructing reasonably small basis sets, the evaluation of the SARC basis sets for the actinides proposed in this paper must first address the errors in total energies resulting from the restricted number of basis functions and from contraction of the innermost primitive Gaussians. In Table 3 we compare SAHF-DKH2 total electronic energies obtained with the [21s13p10d7f] SARC basis sets to those obtained with the uncontracted UGBS, the latter assumed to be a good approximation to the basis set limit. The energy difference, which can be taken as a numerical estimate of the incompleteness error, rises as we traverse the actinide series from 1.01 E_h for actinium to 2.26 E_h for lawrencium. These values are quite satisfying considering the significant reduction in size from the UGBS to the SARC basis sets. In fact, these incompleteness errors are even smaller than those obtained for the corresponding lanthanide SARC basis sets, which ranged from 1.07 E_h for lanthanum up to 3.98 E_h for lutetium.⁵⁹

Part of the incompleteness error results from the contraction of the SARC basis sets. To quantify the contraction error additional calculations were performed with the fully uncontracted (29s20p16d12f) SARC basis sets. The energy difference between contracted and uncontracted forms ranges from 0.042 E_h to 0.096 E_h , which again compares favorably with the contraction errors of the lanthanide basis sets, 0.033 E_h for lanthanum to 0.074 E_h for lutetium. Thus, the contraction error of the actinide SARC basis sets represents roughly a constant 4% of their total incompleteness error, suggesting that the proposed contraction pattern is a reasonable compromise between flexibility, size, and accuracy of the basis set.

As an additional evaluation metric, the energies and the radial expectation values for the orbitals of the valence region are found to either coincide or differ insignificantly between the UGBS and SARC basis sets. Specifically, the energies of occupied 6d and 7s orbitals are practically identical, while the energies of the 5f orbitals obtained by SARC tend to be slightly stabilized compared to the UGBS values, by 0.05 eV on average. As anticipated,

Table 4. SAHF-DKH2 Orbital Energies (E_h) and Radial Expectation Values (Bohr) for Plutonium: Comparison between UGBS and SARC Basis Sets

	UGBS		SARC		ΔE	$\Delta\langle r \rangle$
	E	$\langle r \rangle$	E	$\langle r \rangle$		
1s	-4488.462	0.012	-4488.374	0.012	0.087	0.000
2s	-854.015	0.054	-854.051	0.054	-0.036	0.000
2p	-711.675	0.051	-711.738	0.051	-0.063	0.000
3s	-220.340	0.142	-220.398	0.142	-0.058	0.000
3p	-180.303	0.146	-180.354	0.147	-0.051	0.000
3d	-143.568	0.135	-143.624	0.135	-0.056	0.000
4s	-58.500	0.309	-58.546	0.309	-0.046	0.000
4p	-45.153	0.331	-45.180	0.332	-0.027	0.000
4d	-31.074	0.345	-31.098	0.345	-0.024	0.000
4f	-16.615	0.340	-16.650	0.340	-0.036	0.000
5s	-13.572	0.645	-13.584	0.645	-0.012	0.000
5p	-9.283	0.721	-9.284	0.721	-0.002	0.000
5d	-4.448	0.848	-4.447	0.849	0.001	0.001
6s	-2.152	1.450	-2.155	1.449	-0.003	-0.001
6p	-1.035	1.784	-1.034	1.785	0.001	0.000
7s	-0.186	4.513	-0.186	4.509	0.000	-0.004
5f	-0.340	1.347	-0.341	1.348	-0.001	0.001

the maximum *Sf* energy difference between SARC and UGBS is observed for the heaviest actinide, lawrencium, at -0.13 eV. Larger deviations are observed for the shells with main quantum number less than five. Differences from UGBS orbital energies reach their maximum at the chemically inert 1s orbital, which is thus identified as a principal source of the incompleteness error in the SARC basis sets. As an illustrative example of the above, a detailed picture of the differences in predicted orbital energies and radial expectation values between UGBS and SARC is provided in Table 4 for the case of plutonium. In terms of practical use, it is encouraging that the radial expectation values practically coincide all the way down to the core region, since this implies that the decreased size of SARC compared to UGBS does not compromise the reproduction of scalar relativistic effects in the core region.

2.3. Ionization Energies. In contrast to the lanthanides, practically no experimental data exists for the ionization energies of the actinides beyond the first ionization energy (IE_1). From the point of view of quantum chemistry, the prediction of reliable actinide IEs is far from trivial, since it requires the simultaneous high-level treatment of electron correlation and relativity. It is not our intention here to offer a recipe for the accurate prediction of actinide ionization energies. Rather, our priority is to verify the balanced construction and contraction of the SARC basis sets, which can be inferred by the absence of bias and of systematic errors compared with the large UGBS reference.

For this purpose we have calculated the first four ionization energies of the actinides using the SAHF-DKH2 approach with both basis sets. The first electron is removed from the 6d orbital in Ac, Pa, U, Cm and Lr, and from the 7s orbital in the other elements. The second ionization results in *f*-only configurations for the cations U^{2+} to No^{2+} , with a single electron occupying the 7s orbital in Ac^{2+} , Th^{2+} , Pa^{2+} , and Lr^{2+} . Formation of the M^{3+} cations leads to a closed radon shell for Ac^{3+} and to *Sf* electronic configurations ranging from $5f^1$ (Th^{3+}) to $5f^{14}$ (Lr^{3+}). Finally, the fourth ionization involves removal of a *Sf* electron in all cases,

Table 5. Comparison of UGBS and SARC Basis Sets for the First Four Ionization Energies (eV) of the Actinides^a

	SARC				UGBS			
	IE_1	IE_2	IE_3	IE_4	IE_1	IE_2	IE_3	IE_4
Ac	4.27	10.53	16.59	46.07	4.24	10.52	16.58	46.08
Th	4.59	12.37	16.96	26.60	4.56	12.43	16.95	26.56
Pa	5.49	9.92	18.35	28.60	5.47	9.92	18.33	28.56
U	5.88	11.55	17.21	30.51	5.87	11.56	17.17	30.47
Np	3.11	13.77	18.61	32.34	3.08	13.79	18.58	32.30
Pu	2.39	13.06	19.98	34.12	2.39	13.05	19.95	34.08
Am	1.97	13.67	21.32	35.86	1.97	13.66	21.29	35.82
Cm	7.48	13.67	16.83	37.56	7.47	13.67	16.80	37.52
Bk	2.67	13.44	18.26	32.86	2.67	13.43	18.24	32.82
Cf	3.05	13.29	19.65	34.61	3.04	13.28	19.62	34.57
Es	3.44	13.11	20.99	36.31	3.44	13.10	20.96	36.27
Fm	3.86	12.91	22.30	37.98	3.86	12.89	22.27	37.94
Md	4.31	12.67	23.59	39.61	4.30	12.66	23.56	39.57
No	5.42	11.77	24.86	41.22	5.42	11.76	24.82	41.18
Lr	3.56	13.04	20.53	42.80	3.55	13.03	20.51	42.75
MAD	0.01	0.00	0.03	0.04				

^a Results of SAHF Calculations with the DKH2 Hamiltonian.

with the obvious exception of Ac^{3+} that loses a 6p electron from the radon core. The results, collected in Table 5, show negligible differences between the UGBS and SARC values. Specifically, for IE_1 and IE_2 the mean average deviation is close to zero. It is only for the third and fourth ionization energies that the deviations rise slightly, to 0.03 and 0.04 eV, respectively. Importantly, these deviations between SARC and UGBS are consistent for each step in the ionization process and entirely systematic across the whole series. No pathological cases can be identified. We note that the ZORA Hamiltonian produces identical average deviations between SARC and UGBS. Therefore, we conclude that regardless of the scalar relativistic Hamiltonian employed, the SARC basis sets are unlikely to become the accuracy-limiting factor in molecular applications involving changes in actinide oxidation states.

In terms of absolute numerical values, the uncorrelated SAHF ionization energies reported in Table 5 are obviously not to be compared with experiment or high-level calculations. They only serve as a reliable and unambiguous internal consistency check. As mentioned above, accurate prediction of ionization energies requires an adequate treatment of differential electron correlation effects, while spin-orbit coupling must additionally be taken into account especially for the third and fourth ionization processes. DFT within a scalar relativistic framework is arguably not the best-suited approach for this problem, given the compounded shortcomings of the assumed functional form, the single-determinant Kohn-Sham formulation, and the neglect of spin-dependent relativistic corrections. Nevertheless, simply to demonstrate how even an approximate DFT treatment greatly improves the predicted ionization energies, in Table 6 we compare B3LYP/SARC-DKH2 results with high-quality theoretically predicted values⁷⁰ as well as with available experimental data on IE_1 . The ab initio reference values were obtained from multi-reference averaged coupled-pair functional (CASSCF/ACPF) small-core pseudopotential (PP) calculations employing large uncontracted valence basis sets, extrapolated to the basis set limit and corrected for spin-orbit interaction and estimated PP-related

Table 6. Comparison of B3LYP/SARC-DKH2 Ionization Energies (eV) with Ab initio Reference Values⁷⁰ and Experimental Data

	B3LYP/SARC-DKH2				ACPF/PP				Expt.
	IE ₁	IE ₂	IE ₃	IE ₄	IE ₁	IE ₂	IE ₃	IE ₄	IE ₁
Ac	5.18	11.69	17.63	46.77	5.17	11.60	17.39	44.99	5.17
Th	6.21	11.98	17.91	29.00	6.25	12.11	18.30	28.45	6.31
Pa	5.63	12.16	18.53	31.23	5.81	11.96	17.73	31.24	5.90
U	5.45	11.94	19.30	33.03	6.06	11.63	19.07	33.17	6.19
Np	5.77	11.74	20.45	34.39	5.98	11.35	19.92	34.27	6.27
Pu	5.82	11.85	21.90	35.76	5.71	11.50	21.37	35.43	6.03
Am	5.88	12.02	23.45	37.34	5.71	11.71	22.34	37.26	5.97
Cm	5.60	12.50	20.67	39.23	5.68	12.17	20.36	39.06	5.99
Bk	6.13	12.31	22.33	36.49	5.90	11.96	21.93	36.52	6.20
Cf	6.24	12.45	23.60	38.50	5.96	12.03	22.84	38.12	6.28
Es	6.34	12.58	24.38	40.07	6.07	12.20	23.06	39.52	6.37
Fm	6.45	12.72	25.01	41.18	6.18	12.38	23.66	40.16	6.50
Md	6.54	12.86	26.38	42.10	6.25	12.47	24.69	40.60	6.58
No	6.64	12.99	27.69	43.68	6.33	12.58	26.05	41.96	6.65
Lr	4.56	14.53	21.96	45.28	4.78	14.25	21.52	44.12	—

errors. Thus, they represent one of the currently most reliable sets of theoretical predictions, going beyond what can be expected from the rather rudimentary B3LYP-DKH2 approach.

B3LYP values are in reasonable agreement with the ACPF reference for IE₁ and IE₂, within 0.07 and 0.25 eV on average, respectively. Larger differences are observed for the third and fourth ionization energies, where a distinction can be made between the first and second half of the actinide series: for both IE₃ and IE₄ the highest deviations, often exceeding 1 eV, are witnessed for elements in the second half of the series. In these cases B3LYP systematically overestimates the ionization energy compared to the ACPF reference. With respect to available IE₁ experimental estimates the B3LYP values yield an average deviation of 0.18 eV. In view of the inherent limitations of the present DFT scalar relativistic approach, the above results can be regarded as adequate considering the intended realm of molecular applications; the interested reader is referred to the literature for examples of more rigorous theoretical approaches to the prediction of actinide ionization energies.^{20,70–72}

2.4. Basis Set Superposition Errors. A critical factor that impacts on the accuracy of computational studies of molecular complexes is the error introduced by the incompleteness of the atomic basis sets in the valence region. This results in the well-known basis set superposition error (BSSE) and its typical manifestation is the artificial stabilization of a two-fragment system through the mutual compensation of each fragment's incomplete valence space by valence basis functions of the other. The BSSE can be magnified when a heavy and a light atom are adjacent, since the potential for overlap between a relatively big and a relatively small basis set amplifies the effect of basis set imbalance. Boys and Bernardi have proposed the counterpoise correction method (CPC) in order to obtain results that are approximately free of the superposition error,⁷³ and this is the method we use here to quantify the extent of the BSSE associated with the SARC basis sets. For this purpose we have performed calculations with the DKH2 Hamiltonian and the PBE0 hybrid functional^{74,75} on a set of diatomic actinide molecules consisting of one terminal element of the actinide series, Ac or Lr, and one

Table 7. Bond Lengths r (Å) and Dissociation Energies D_e (eV) of Actinium and Lawrencium Diatomics Computed with the PBE0 Density Functional and the DKH2 Hamiltonian, without and with BSSE Counterpoise Corrections (CPC)

	PBE0/SARC		PBE0/SARC+CPC		Δr	ΔD_e
	r	D_e	r	D_e		
AcH (¹ Σ)	2.128	2.86	2.129	2.85	0.001	−0.01
AcO (² Σ)	1.915	7.37	1.915	7.35	0.000	−0.02
AcF (¹ Σ)	2.098	6.99	2.098	6.96	0.000	−0.03
LrH (¹ Σ)	1.976	3.17	1.977	3.17	0.001	0.00
LrO (² Σ)	1.853	6.09	1.853	6.06	0.000	−0.03
LrF (¹ Σ)	1.989	7.19	1.990	7.17	0.001	−0.02

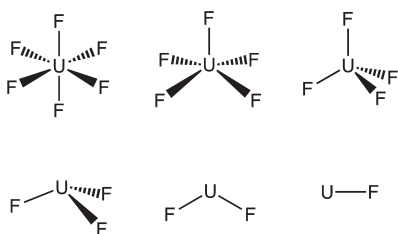
of the light atoms H, O, and F. The DKH2-recontracted versions of the def2-TZVP basis sets were used for the nonactinide elements.⁶⁰ Note that the hydrides are particularly sensitive to BSSE since they involve a combination of elements that are at diametrically opposite positions of the periodic table.

The optimized bond lengths (Table 7) display the expected contraction on passing from the lightest to the heaviest actinide. The extent of the contraction is different for each class of compound, being largest for the monohydrides and smallest for the monoxides. This has been pointed out previously for both the lanthanides⁷⁶ and the actinides,⁷⁷ and results from the fact that bonds with large force constants undergo smaller contractions along the series.⁷⁷ It is worth noting that inclusion of spin-orbit effects in addition to scalar relativity has been shown to slightly shorten the predicted bond lengths.⁷⁷ For our purposes, the most important conclusion drawn from the comparison of the computed bond lengths for each dimer with and without counterpoise corrections is that the impact of BSSE on bond lengths never exceeds 0.001 Å. Equally small is the BSSE error in predicted dissociation energies, where the uncorrected values are overstabilized by at most 0.03 eV. Thus, these results provide additional confirmation that the incompleteness errors described previously in comparison with the UGBS atomic energies indeed arise from the chemically inert core shells. In conclusion, the results on the diatomic molecules of Table 7 suggest that the SARC valence space is sufficiently extended to ensure minimal BSSE in molecular DFT calculations.

2.5. Uranium Fluorides. The series UF_{*n*} (*n* = 1–6) is quite well characterized experimentally. Especially abundant is experimental data on UF₆, which plays a central role in ²³⁵U-enrichment processes by diffusion or centrifugation. Hence, it is not surprising that the uranium fluorides, and to a lesser extent the chlorides, have seen extensive use as reference systems for the evaluation of computational methods.^{7,19,26,78–82} Here we test the SARC basis sets on the UF_{*n*} series of molecules (Scheme 1) using the ZORA Hamiltonian with the PBE0 functional, in order to make comparisons with experiment and recent theoretical studies. PBE0 has been shown to perform well for a variety of heavy-element systems⁶¹ including the actinide species under consideration.⁸⁰ ZORA-recontracted TZVP basis sets have been used for fluorine. All calculations were performed with ORCA using the one-center approximation for ZORA, tight optimization and SCF convergence criteria, increased integration grids (Grid4 in ORCA convention), and taking advantage of the chain-of-spheres (RIJCOSX) algorithm for exact exchange.⁸³

Beginning with the heaviest member first, UF₆ is predicted to have an octahedral coordination geometry (*O_h* point group) with

Scheme 1. Structures of the Uranium Fluorides



U–F bond lengths of 1.997 Å. In comparison, PBE0 calculations with small-core (60-electron) and large-core (78-electron) effective core potentials for uranium gave U–F bond lengths of 1.993 Å and 2.006 Å, respectively.^{26,80} The current PBE0/SARC value is consistent with the two estimates obtained through electron diffraction experiments, 1.999(3) and 1.996(8) Å.^{84,85} UF₅ has a doublet ground state with one unpaired electron in the 5f_{xyz} orbital of uranium. Its structure is predicted to be C_{4v} symmetric, in agreement with IR spectra obtained in Ar and Ne matrices.⁸⁶ The equatorial and axial bond lengths are almost equal, 2.019 and 2.018 Å, respectively, while the uranium atom is axially displaced (0.272 Å) above the plane defined by the four equivalent fluorine atoms, resulting in an obtuse F_{ax}–U–F_{eq} angle of 97.8°. As with UF₆, the SARC bond lengths are ca. 0.004 Å longer than small-core ECP results.^{26,80}

When T_d point group symmetry is enforced, the optimized U–F bond length in UF₄ is 2.059 Å, which is identical to the experimentally deduced value by Konings et al.⁸⁷ In comparison, PBE0 with small-core ECPs gave a value of 2.053 Å.⁸⁰ By removing all symmetry constraints the optimized structure distorts and settles into a C_{2v} symmetric geometry with two pairs of equivalent bonds, 2.054 and 2.060 Å, forming between them angles of 105.9° and 104.7°, respectively. However, the energy gain associated with this distortion is so small (less than 0.4 kcal mol⁻¹) that quantitatively distinguishing between this and other proximate structures⁸⁸ for UF₄ would probably require a more fine-grained theoretical approach and explicit consideration of competing Jahn–Teller and spin–orbit effects.⁸⁹

For UF₃ a pyramidal C_{3v} structure is predicted, with U–F bond lengths of 2.073 Å and F–U–F angles of 103.9°, as a result of the uranium atom being 0.862 Å away from the plane defined by the fluorides. UF₂ (quintet ground state) belongs to the C_{2v} point group with bond lengths of 2.067 Å and an F–U–F angle of 109.0°. Finally, the optimized bond length of UF (quartet ground state) is 2.032 Å. The small-core ECP bond lengths for the last three members of the series were reported as 2.069, 2.055, and 2.024 Å, respectively.⁸⁰ As with all UF_n species, the ECP bond lengths are consistently shorter than the SARC ones, but the predicted trend is the same with either method, the U–F bond lengths increasing in the order UF₆ < UF₅ < UF < UF₄ < UF₂ < UF₃.

Beyond geometries, an interesting metric for our testing purposes are the fluorine bond dissociation energies (BDEs) for the reactions UF_n → UF_{n-1} + F. Thermochemical studies of uranium fluorides have led to experimental BDEs for the whole series,⁹⁰ while the subject has also been the focus of recent theoretical studies.^{26,80,82} It is worth reminding that the use of large-core ECPs for BDEs leads to unacceptable errors.²⁶ In order to ensure that the consistent performance of the presently proposed basis sets extends also to this kind of

Table 8. Comparison Between Calculated PBE0 Bond Dissociation Energies (kcal mol⁻¹) and Experimental Values for the Uranium Fluoride Series

	UF ₆	UF ₅	UF ₄	UF ₃	UF ₂	UF
SARC-ZORA	70.2	94.1	133.4	143.0	144.4	159.5
60e-ECP ^a	71.4	94.6	137.2	144.7	146.6	155.9
scalar-DKH3 ^b	74.5	98.2	140.8	146.5	146.1	164.1
SO–DKH3 ^b	70.8	92.5	138.7	146.7	145.2	166.1
expt. ^c	71.0	98.0	147.0	147.9	135.0	154.9

^a Ref 73. ^b Ref 75. ^c Ref 83.

Table 9. PBE0/SARC-ZORA Vibrational Frequencies of UF₆ (cm⁻¹) Compared with PBE0/ECP Results (Ref 26) and with Experiment (Ref 91)^a

	ω ₁ (a _{1g})	ω ₂ (e _g)	ω ₃ (t _{1u})	ω ₄ (t _{1u})	ω ₅ (t _{2g})	ω ₆ (t _{2u})	MAD ν / ω
SARC-ZORA	673	536	627	184	201	142	2/3
78e-ECP	644	522	626	191	182	148	11/14
60e-ECP	677	532	631	186	198	140	4/4
expt. (ν)	667	534	626	186	200	143	
expt. (ω) ^b	672	540	634	186	200	143	

^a Mean absolute deviations (cm⁻¹) are provided with respect to both experimental fundamentals (ν) and estimated harmonic frequencies (ω). ^b Estimated harmonic frequencies.

property, we have computed the successive BDEs of the uranium fluorides at the PBE0/SARC-ZORA level. The BDE of a given species is defined as BDE(UF_n) = E(UF_{n-1}) + E(F) – E(UF_n). Zero-point energy corrections are small and were shown to account for less than 2 kcal mol⁻¹ of the total BDE,²⁶ therefore they are not included in the present calculations. The results are summarized in Table 8, where a comparison is also made with computed values from small-core ECP calculations⁸⁰ and all-electron scalar and nuclear-only spin–orbit DKH3 calculations⁸² as well as with the experimental values proposed by Hildenbrand and Lau.⁹⁰

Inspection of the computed BDEs reveals no significant discrepancy between the different theoretical approaches. The predicted values for most of the molecules agree particularly well among the different methods. Compared to the experimental estimates, UF₄ and UF₂ stand out as the two species for which experiment diverges more obviously from the theoretical predictions. Although all methods are in close agreement for the BDE of UF₂, the experimental estimate is approximately 10 kcal mol⁻¹ lower than the theoretical prediction. Exactly the opposite situation is observed for UF₄, for which the experimental estimate is consistently higher than the computed values. Overall, all theoretical methods predict the same order of monotonically increasing BDEs from UF₆ to UF, with a possible exception when passing from UF₃ to UF₂, where a small decrease might be anticipated, arising from the extra stabilization imparted to UF by the pairing of electrons in the uranium 7s orbital.

Finally, it is interesting to see how the present approach performs for the vibrational frequencies of the thoroughly characterized UF₆ molecule. In Table 9 a comparison is made between the current PBE0/SARC-ZORA harmonic vibrational frequencies, PBE0/ECP values previously obtained²⁶ with either a large-core (78-electron) or a small-core (60-electron) relativistic effective core potential for uranium, and experimental

fundamentals as well as estimated harmonic frequencies by McDowell et al.⁹¹ The PBE0/SARC-ZORA results agree remarkably well with experiment and have the lowest average deviation, 2 cm⁻¹ compared to 4 cm⁻¹ for the small-core ECP and 11 cm⁻¹ for the large-core ECP (3, 14, and 4 cm⁻¹, respectively, in comparison to the projected experimental harmonic frequencies). We consider the advantage of SARC over the small-core ECP too small in absolute terms to be interpreted as a definite improvement, particularly because the average errors of the two approaches are of the same order of magnitude as: (a) the anharmonicity effects, at least in the case of the three higher vibrations, and (b) the experimental uncertainties.⁹¹ With regard to intensities, the intensity ratio I_3/I_4 of the two IR-active vibrations, the ω_3 stretching and the ω_4 bending mode, is predicted to be 19.7 from the PBE0/SARC calculations. This value is in nearly perfect agreement with the experimental ratio of 20.⁹¹ In conclusion, the computed data on vibrational frequencies suggest that the SARC results can be expected to be at least as good as, or slightly better than results obtained with small-core pseudopotentials, and can therefore serve to either cross-validate ECPs for DFT studies or directly improve on their predictions.

3. CONCLUSIONS

Segmented all-electron relativistically contracted (SARC) basis sets are proposed for the 5f elements. These basis sets extend the existing SARC family of basis sets to encompass the whole series of the actinide elements. The basis sets are designed to be of moderate size and polarization, so that they can be used efficiently in routine all-electron DFT calculations employing one of the most popular ZORA or DKH scalar relativistic Hamiltonians. In comparison to results obtained with extended basis sets of more than double their size (UGBS), the SARC basis sets display controlled incompleteness errors that primarily originate from the inner shells. The contraction pattern that is adopted accounts for a minor (ca. 4%) and systematic percentage of the incompleteness error. Comparison of the first four ionization energies of all actinides computed with SARC and UGBS demonstrates near-coincidence for IE₁ and IE₂, while marginal deviations of 0.03 and 0.04 eV are observed for IE₃ and IE₄. Thus, despite the reduction in size, SARC valence properties are still described at a quality close to that of the much larger basis set reference. Investigation of the basis set superposition errors in six actinide diatomic molecules MX (M = Ac, Lr; X = H, O, F) revealed maximum BSSE corrections of 0.001 Å in bond lengths and 0.03 eV in bond dissociation energies. These results confirm that the SARC valence space can be considered balanced and sufficiently saturated for applications based on DFT methods, which do not require extensive polarization to obtain converged results. An illustrative application to uranium fluorides confirms that the SARC basis sets perform well for molecular properties such as geometries, vibrational frequencies and bond dissociation energies. Therefore, we expect the SARC basis sets to be a useful addition to the modern toolbox of computational actinide chemistry.

■ ASSOCIATED CONTENT

Supporting Information. Full listings of the SARC basis sets. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: pantazis@thch.uni-bonn.de (D.A.P.).

■ ACKNOWLEDGMENT

We gratefully acknowledge financial support from the Max Planck Society.

■ REFERENCES

- (1) Cotton, S. *Lanthanide and Actinide Chemistry*. 2nd ed.; John Wiley & Sons: Chichester, U.K., 2006; p 280.
- (2) Kaltsoyannis, N.; Hay, P. J.; Li, J.; Blaudeau, J. P.; Bursten, B. E. Theoretical studies of the electronic structure of compounds of the actinide elements. In *The Chemistry of the Actinide and Transactinide Elements*, 3rd ed.; Morss, L. R., Edelstein, N., Fuger, J., Eds. Springer: Dordrecht, The Netherlands, 2006; pp 1893–2012.
- (3) Schreckenbach, G.; Shamov, G. A. *Acc. Chem. Res.* **2009**, *43*, 19–29.
- (4) Kaltsoyannis, N. *Chem. Soc. Rev.* **2003**, *32*, 9–16.
- (5) Cao, X.; Dolg, M. *Coord. Chem. Rev.* **2006**, *250*, 900–910.
- (6) Vallet, V.; Macak, P.; Wahlgren, U.; Grenthe, I. *Theor. Chem. Acc.* **2006**, *115*, 145–160.
- (7) Schreckenbach, G.; Hay, P. J.; Martin, R. L. *J. Comput. Chem.* **1999**, *20*, 70–90.
- (8) Kirker, I.; Kaltsoyannis, N. *Dalton Trans.* **2011**, *40*, 124–131.
- (9) Tassell, M. J.; Kaltsoyannis, N. *Dalton Trans.* **2010**, *39*, 6719–6725.
- (10) Tsipis, A. C.; Kefalidis, C. E.; Tsipis, C. A. *J. Am. Chem. Soc.* **2008**, *130*, 9144–9155.
- (11) Vetere, V.; Maldivi, P.; Roos, B. O.; Adamo, C. *J. Phys. Chem. A* **2009**, *113*, 14760–14765.
- (12) Ingram, K. I. M.; Tassell, M. J.; Gaunt, A. J.; Kaltsoyannis, N. *Inorg. Chem.* **2008**, *47*, 7824–7833.
- (13) Hayton, T. W. *Dalton Trans.* **2010**, *39*, 1145–1158.
- (14) Dyal, K. G.; Faegri, K. *Introduction to Relativistic Quantum Chemistry*; Oxford University Press: Oxford, 2007; p 544.
- (15) de Jong, W. A.; Harrison, R. J.; Nichols, J. A.; Dixon, D. A. *Theor. Chem. Acc.* **2001**, *107*, 22–26.
- (16) Dirac, P. A. M. *Proc. R. Soc. London, Ser. A* **1928**, *117*, 610–624.
- (17) Emler, W. C.; Ross, R. B.; Christiansen, P. A. *Int. J. Quantum Chem.* **1991**, *40*, 829–846.
- (18) Küchle, W.; Dolg, M.; Stoll, H.; Preuss, H. *J. Chem. Phys.* **1994**, *100*, 7535–7542.
- (19) Hay, P. J.; Martin, R. L. *J. Chem. Phys.* **1998**, *109*, 3875–3881.
- (20) Cao, X.; Dolg, M.; Stoll, H. *J. Chem. Phys.* **2003**, *118*, 487–496.
- (21) Dolg, M.; Cao, X. *J. Phys. Chem. A* **2009**, *113*, 12573–12581.
- (22) Koch, W.; Holthausen, M. C. *A Chemist's Guide to Density Functional Theory*. 2nd ed.; Wiley-VCH: Weinheim, Germany, 2001; p 300.
- (23) Parr, R. G.; Yang, W. *Density-Functional Theory of Atoms and Molecules*; Oxford University Press: Oxford, 1989; p 352.
- (24) Odoh, S. O.; Schreckenbach, G. *J. Phys. Chem. A* **2009**, *114*, 1957–1963.
- (25) Iché-Tarrat, N.; Marsden, C. J. *J. Phys. Chem. A* **2008**, *112*, 7632–7642.
- (26) Batista, E. R.; Martin, R. L.; Hay, P. J.; Peralta, J. E.; Scuseria, G. E. *J. Chem. Phys.* **2004**, *121*, 2144–2150.
- (27) Schreckenbach, G. *Int. J. Quantum Chem.* **2005**, *101*, 372–380.
- (28) Fillaux, C.; Berthet, J. C.; Conradson, S. D.; Guillaumont, D.; Hennig, C.; Moisy, P.; Roques, J.; Simoni, E.; Shuh, D. K.; Tylliszczak, T.; Castro-Rodriguez, I.; Den Auwer, C. *C. R. Chim.* **2007**, *10*, 859–871.
- (29) Fillaux, C.; Den Auwer, C.; Guillaumont, D.; Shuh, D. K.; Tylliszczak, T. *J. Alloys Compd.* **2007**, *444–445*, 443–446.

- (30) Den Auwer, C.; Simoni, E.; Conradson, S.; Madic, C. *Eur. J. Inorg. Chem.* **2003**, 2003, 3843–3859.
- (31) Denecke, M. A. *Coord. Chem. Rev.* **2006**, 250, 730–754.
- (32) Bader, R. F. W. *Atoms in Molecules: A Quantum Theory*; Oxford University Press: Oxford, 1990; p 458.
- (33) Becke, A. D.; Edgecombe, K. E. *J. Chem. Phys.* **1990**, 92, 5397–5403.
- (34) Petit, L.; Joubert, L.; Maldivi, P.; Adamo, C. *J. Am. Chem. Soc.* **2006**, 128, 2190–2191.
- (35) Vyboishchikov, S. F.; Sierraalta, A.; Frenking, G. *J. Comput. Chem.* **1997**, 18, 416–429.
- (36) Dyal, K. *Theor. Chem. Acc.* **2007**, 117, 491–500.
- (37) Faegri, K. *Theor. Chem. Acc.* **2001**, 105, 252–258.
- (38) Faegri, K. *Chem. Phys.* **2005**, 311, 25–34.
- (39) Minami, T.; Matsuoka, O. *Theor. Chem. Acc.* **1995**, 90, 27–39.
- (40) Tsuchiya, T.; Abe, M.; Nakajima, T.; Hirao, K. *J. Chem. Phys.* **2001**, 115, 4463–4472.
- (41) Nakajima, T.; Hirao, K. *J. Chem. Phys.* **2002**, 116, 8270–8275.
- (42) Noro, T.; Sekiya, M.; Osanai, Y.; Koga, T.; Matsuyama, H. *J. Comput. Chem.* **2007**, 28, 2511–2516.
- (43) Douglas, M.; Kroll, N. M. *Ann. Phys.* **1974**, 82, 89–155.
- (44) Hess, B. A. *Phys. Rev. A: At., Mol., Opt. Phys.* **1985**, 32, 756–763.
- (45) Hess, B. A. *Phys. Rev. A: At., Mol., Opt. Phys.* **1986**, 33, 3742–3748.
- (46) Jansen, G.; Hess, B. A. *Phys. Rev. A: At., Mol., Opt. Phys.* **1989**, 39, 6016–6017.
- (47) Wolf, A.; Reiher, M.; Hess, B. A. *J. Chem. Phys.* **2002**, 117, 9215–9226.
- (48) van Lenthe, E.; Baerends, E. J.; Snijders, J. G. *J. Chem. Phys.* **1994**, 101, 9783–9792.
- (49) van Lenthe, E.; Snijders, J. G.; Baerends, E. J. *J. Chem. Phys.* **1996**, 105, 6505–6516.
- (50) van Wüllen, C. *J. Chem. Phys.* **1998**, 109, 392–399.
- (51) Laikov, D. N.; Ustynuk, Y. A. *Russ. Chem. Bull.* **2005**, 54, 820–826.
- (52) Dyal, K. G. *J. Chem. Phys.* **1994**, 100, 2118–2127.
- (53) Laikov, D. N. *Chem. Phys. Lett.* **2005**, 416, 116–120.
- (54) *Amsterdam Density Functional (ADF)*, 2007.01; SCM, Theoretical Chemistry, Vrije Universiteit: Amsterdam, The Netherlands, 2007.
- (55) te Velde, G.; Bickelhaupt, F. M.; Baerends, E. J.; Guerra, C. F.; Van Gisbergen, S. J. A.; Snijders, J. G.; Ziegler, T. *J. Comput. Chem.* **2001**, 22, 931–967.
- (56) Roos, B. O.; Lindh, R.; Malmqvist, P.-Å.; Varyazov, V.; Widmark, P.-O. *Chem. Phys. Lett.* **2005**, 409, 295–299.
- (57) Neese, F. *Coord. Chem. Rev.* **2009**, 253, 526–563.
- (58) Aquino, F.; Govind, N.; Autschbach, J. *J. Chem. Theory Comput.* **2010**, 6, 2669–2686.
- (59) Pantazis, D. A.; Neese, F. *J. Chem. Theory Comput.* **2009**, 5, 2229–2238.
- (60) Pantazis, D. A.; Chen, X. Y.; Landis, C. R.; Neese, F. *J. Chem. Theory Comput.* **2008**, 4, 908–919.
- (61) Bühl, M.; Reimann, C.; Pantazis, D. A.; Bredow, T.; Neese, F. *J. Chem. Theory Comput.* **2008**, 4, 1449–1459.
- (62) Neese, F. *ORCA – an ab initio, Density Functional and Semiempirical Program Package*, v. 2.7.0; Universität Bonn: Bonn, Germany, 2009.
- (63) Stavrev, K. K.; Zerner, M. C. *Int. J. Quantum Chem.* **1997**, 65, 877–884.
- (64) Zerner, M. C. *Int. J. Quantum Chem.* **1989**, 35, 567–575.
- (65) Jorge, F. E.; de Castro, E. V. R.; da Silva, A. B. F. *J. Comput. Chem.* **1997**, 18, 1565–1569.
- (66) de Castro, E. V. R.; Jorge, F. E. *J. Chem. Phys.* **1998**, 108, 5225–5229.
- (67) Autschbach, J. *ChemPhysChem* **2009**, 10, 2274–2283.
- (68) Malkin, E.; Malkin, I.; Malkina, O. L.; Malkin, V. G.; Kaupp, M. *Phys. Chem. Chem. Phys.* **2006**, 8, 4079–4085.
- (69) Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, 7, 3297–3305.
- (70) Cao, X.; Dolg, M. *Mol. Phys.* **2003**, 101, 961–969.
- (71) Liu, W.; Küchle, W.; Dolg, M. *Phys. Rev. A: At., Mol., Opt. Phys.* **1998**, 58, 1103.
- (72) Cao, X.; Dolg, M. *J. Mol. Struct. (THEOCHEM)* **2004**, 673, 203–209.
- (73) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, 19, 553–566.
- (74) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, 77, 3865–3868.
- (75) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, 110, 6158–6170.
- (76) Wang, S. G.; Schwarz, W. H. E. *J. Phys. Chem.* **1995**, 99, 11687–11695.
- (77) Küchle, W.; Dolg, M.; Stoll, H. *J. Phys. Chem. A* **1997**, 101, 7128–7133.
- (78) Hay, P. J.; Wadt, W. R.; Kahn, L. R.; Raffanetti, R. C.; Phillips, D. H. *J. Chem. Phys.* **1979**, 71, 1767–1779.
- (79) Han, Y. K. *J. Comput. Chem.* **2001**, 22, 2010–2017.
- (80) Batista, E. R.; Martin, R. L.; Hay, P. J. *J. Chem. Phys.* **2004**, 121, 11104–11111.
- (81) Kovács, A.; Konings, R. J. M. *J. Mol. Struct. (THEOCHEM)* **2004**, 684, 35–42.
- (82) Peralta, J. E.; Batista, E. R.; Scuseria, G. E.; Martin, R. L. *J. Chem. Theory Comput.* **2005**, 1, 612–616.
- (83) Neese, F.; Wenzmohs, F.; Hansen, A.; Becker, U. *Chem. Phys.* **2009**, 356, 98–109.
- (84) Seip, H. M. *Acta Chem. Scand.* **1965**, 19, 1955–1968.
- (85) Kimura, M.; Schomaker, V.; Smith, D. W.; Weinstock, B. *J. Chem. Phys.* **1968**, 48, 4001–4012.
- (86) Jones, L. H.; Ekberg, S. *J. Chem. Phys.* **1977**, 67, 2591–2595.
- (87) Konings, R. J. M.; Booij, A. S.; Kovács, A.; Girichev, G. V.; Giricheva, N. I.; Krasnova, O. G. *J. Mol. Struct.* **1996**, 378, 121–131.
- (88) Zhang, Y.; Li, Y.; Cao, Y. *J. Mol. Struct. (THEOCHEM)* **2008**, 864, 85–88.
- (89) David, J.; Fuentealba, P.; Restrepo, A. *Chem. Phys. Lett.* **2008**, 457, 42–44.
- (90) Hildenbrand, D. L.; Lau, K. H. *J. Chem. Phys.* **1991**, 94, 1420–1425.
- (91) McDowell, R. S.; Asprey, L. B.; Paine, R. T. *J. Chem. Phys.* **1974**, 61, 3571–3580.

Extrapolation and Scaling of the DFT-SAPT Interaction Energies toward the Basis Set Limit

Jan Řezáč^{*,†} and Pavel Hobza^{†,‡}[†]Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic and Center for Biomolecules and Complex Molecular Systems, 166 10 Prague, Czech Republic[‡]Department of Physical Chemistry, Palacky University, 771 46 Olomouc, Czech Republic Supporting Information

ABSTRACT: The dispersion energy term in the symmetry-adapted perturbation theory based on the density functional theory (DFT-SAPT) converges rather slowly with basis set size. Accurate results for large complexes, where only calculations in small basis sets are practical, can be obtained by extrapolation to the complete basis set limit (CBS). In this paper, we propose an extrapolation scheme with the variable exponent optimized specifically for the DFT-SAPT calculations in correlation-consistent basis sets with diffuse functions. Another way to improve the accuracy term at no additional cost is to scale the dispersion term by a fixed amount. We present the scaling factors averaged over a balanced set of 10 model complexes. The results of these schemes are compared to the high-quality DFT-SAPT/CBS interaction energies in small complexes obtained by fitting to a series of basis sets up to aug-cc-pV5Z and to the CCSD(T)/CBS interaction energies. It is shown that even the cheapest extrapolation scheme yields results that are limited by the accuracy of the DFT-SAPT approach rather than by the basis set convergence. Scaling the dispersion term allows accurate interaction energies as well as their components to be obtained using just the aug-cc-pVDZ basis set; such a calculation can be applied to complexes consisting of up to 50 first-row atoms.

The symmetry-adapted perturbation theory based on density functional theory^{1,2} (DFT-SAPT) is a powerful method for the decomposition of intermolecular interaction energy into physically meaningful components. The use of density functional theory as a basis for a SAPT analysis allows for an investigation of larger systems than is possible with SAPT based on the Hartree–Fock method³ or the post-HF methods. DFT-SAPT converges with the basis set size faster than the wave function theory (WFT) methods, and already the aug-cc-pVDZ basis set provides reasonable energy terms. The only exception represents the London dispersion energy, whose convergence is much slower. Furthermore, it is the most computationally intensive part of the calculation.

Problems with convergence are characteristic for the correlated WFT methods and are often addressed by an extrapolation of the energy to the complete basis set (CBS) limit from calculations in smaller basis sets. Some authors apply the extrapolation schemes developed for the WFT methods to DFT-SAPT^{4,5} without comparing the extrapolated values to a reliable estimate of the DFT-SAPT CBS limit. It has also been suggested that scaling the DFT-SAPT dispersion energy obtained in a small basis set by 10–15% could estimate the CBS limit,⁶ but this hypothesis has not been tested further. Therefore, there are no reliable data on what scaling factor should be used and how robust this approach is when applied to complexes of a different nature. The development of such a scaling is important not only as a tool for interaction energy decomposition but also as an efficient method to estimate the DFT-SAPT/CBS interaction energy from a calculation in a smaller basis set applicable to large complexes. Despite the approximations used, the DFT-SAPT/CBS interaction energies are very close (within 5%) to the benchmark CCSD(T)/CBS values.⁷

In this work, we have investigated 10 small complexes where we can use multipoint regression to obtain accurate DFT-SAPT dispersion energies at the CBS limit. From these calculations, we have obtained information that addresses both of the issues introduced above. The parameters in the fitting function can be used to design an optimal two-point extrapolation scheme that is more practical for calculations of larger systems. The ratio of the dispersion energy in a given basis set compared to the CBS limit can also be used to estimate the CBS value from just a single calculation. We present average scaling factors that can be used for complexes with a varying amount of dispersion contribution. The DFT-SAPT calculations are also compared to benchmark CCSD(T)/CBS interaction energies in order to discuss the overall accuracy of DFT-SAPT. Finally, the scaling of the dispersion term is tested on a larger set of medium-sized complexes.

For a reliable description of the dispersion, diffuse functions improve the performance of a basis set significantly. In this study, we have used the series of correlation-consistent basis sets of Dunning⁸ with diffuse functions,⁹ labeled aug-cc-pVXZ, where X = D, T, Q, and 5. The quintuple ζ aug-cc-pV5Z is the largest basis set practically usable for our model systems, and the findings indicate that it yields results close to the CBS limit.

METHODS

DFT-SAPT Calculations. All of the calculations at the DFT-SAPT level were performed using the MOLPRO 2009 package,¹⁰ employing the density fitting procedure.¹¹ We have used the PBE0AC functional recommended by the authors of the

Received: January 4, 2011

Published: February 10, 2011

Table 1. Interaction Energies (kcal/mol) from the DFT-SAPT Calculations in the aug-cc-pVXZ (X = D,T,Q,5) Basis Set, Combined into an Uncorrelated Component (E_{noD}) and the Dispersion Part (E_{D}), and Their Sum, the Total DFT-SAPT Interaction Energy^a

	E_{noD}				E_{D}				$E_{\text{noD}} + E_{\text{D}}$				CCSD(T) CBS
	D	T	Q	5	D	T	Q	5	D	T	Q	5	
water dimer	-2.276	-2.333	-2.367	-2.363	-1.988	-2.332	-2.421	-2.448	-4.264	-4.664	-4.788	-4.811	-4.996
methane dimer	0.442	0.453	0.452	0.451	-0.889	-0.976	-0.999	-1.007	-0.447	-0.523	-0.547	-0.556	-0.530
methane···water	0.113	0.041	0.042	0.042	-0.892	-1.025	-1.049	-1.055	-0.779	-0.984	-1.007	-1.012	-0.971
ammonia dimer	-0.945	-1.012	-1.018	-1.018	-1.821	-2.098	-2.163	-2.184	-2.766	-3.110	-3.181	-3.202	-3.147
ethyne dimer	-0.352	-0.368	-0.372	-0.372	-1.168	-1.302	-1.331	-1.338	-1.520	-1.671	-1.703	-1.710	-1.539
ammonia···water	-3.540	-3.646	-3.661	-3.659	-2.309	-2.700	-2.787	-2.814	-5.849	-6.346	-6.447	-6.473	-6.432
ammonia···N ₂	0.043	0.042	0.034	0.032	-0.615	-0.680	-0.697	-0.702	-0.571	-0.638	-0.663	-0.669	-0.686
formic acid···water	-4.193	-4.367	-4.425	-4.423	-4.775	-5.534	-5.752	-5.821	-8.968	-9.901	-10.177	-10.243	-10.779
methane···N ₂	0.194	0.194	0.193	0.193	-0.492	-0.540	-0.554	-0.559	-0.298	-0.346	-0.361	-0.367	-0.352
water···N ₂	-0.159	-0.174	-0.188	-0.193	-0.798	-0.899	-0.923	-0.929	-0.957	-1.073	-1.111	-1.122	-1.189
RMSE ^a									0.549	0.164	0.059	0.035	

^aThe CCSD(T)/CBS interaction energy is provided as a reference to the DFT-SAPT/CBS results.

method;¹² the calculations of the monomers have been performed in the basis set of the dimer. The shift needed to correct the asymptotic behavior of the functional¹³ was calculated as the difference between the HOMO energy of each monomer and the true ionization potential obtained from the calculation of its neutral and ionized forms. The shifts have been calculated using the same functional in the aug-cc-pVTZ basis set.

Extrapolation Scheme. Little is known on how exactly the dispersion term of DFT-SAPT changes with basis set size and therefore what function should be used for its extrapolation. A good starting point is the extrapolation schemes developed for the correlated WFT methods. We have found that both the exponential- and power-law-based schemes proposed by Helgaker et al.¹⁴ provide good results, but the best fit can be achieved with Schwenke's power-law formula¹⁵ with a variable exponent:

$$E_{\text{corr},X} = E_{\text{corr,CBS}} + AX^{-p} \quad (1)$$

where A is a constant, X is the cardinal number of the basis set, and p is the optimized parameter. This scheme is a variation of Helgaker's scheme, but the fixed exponent ($p = 3$) is replaced by an optimized parameter to achieve better fit. This additional degree of freedom can compensate for the nonideal behavior of the series of the basis set and method. Equation 1 can be easily transformed into a formula for a two-point extrapolation using the basis sets X and $X + 1$, eliminating the constant A :

$$E_{\text{corr,CBS}} = \frac{(E_{\text{corr},X+1}(X+1)^p - E_{\text{corr},X} \times X^p)}{((X+1)^p - X^p)} \quad (2)$$

CCSD(T)/CBS Calculations. The benchmark interaction energies of the complexes were calculated from the MP2/CBS value obtained with Helgaker's extrapolation scheme¹⁴ and the difference between the CCSD(T) and MP2 correlation energy (the so-called CCSD(T) correction term) calculated in a suitable basis set.¹⁶ For the 10 model complexes, the aug-cc-pVQZ and aug-cc-pVSZ basis sets were used for the MP2 calculations and the aug-cc-pVTZ for the CCSD(T) correction term; such a setup ensured results close to the CBS limit. All of those calculations have been performed using the MOLPRO 2009 package.¹⁰

Model Complexes. A set of 10 complexes of small molecules has been prepared for this study. These complexes contain up to four first-row atoms, which makes it possible to calculate the DFT-SAPT interaction energy with large basis sets. An important feature of this set is to include different types of interactions in a balanced way. Therefore, the set contains both polar and nonpolar molecules in all of the possible combinations. These complexes are water, ammonia, methane, and ethyne (T-shape) homodimers and ammonia···water, ammonia···N₂, water···N₂, water···methane, methane···N₂, and formic acid···water complexes (geometries provided in the Supporting Information). The average interaction energy in this set is -3.0 kcal/mol. The structure of these complexes was optimized at the MP2/cc-pVTZ level with the counterpoise correction. It was shown that these geometries are close to those obtained at the CCSD(T) level.¹⁷

Validation Complexes. The proposed schemes using a smaller basis set have been tested on a set of 31 larger complexes (up to nine first-row atoms). The list of these complexes and their geometries are provided in the Supporting Information. The geometry of these complexes has been optimized using the same method as used with the model complexes. The DFT-SAPT calculations in the aug-cc-pVDZ and aug-cc-pVTZ basis set have been performed for all of these complexes; aug-cc-pVQZ had been used for the first 11 of them.

RESULTS AND DISCUSSION

Multipoint Fits. For each of the model complexes, a series of DFT-SAPT interaction energies in the aug-cc-pVXZ (X = D, T, Q, 5) basis set has been calculated. All of the terms were divided into two groups and summed into the dispersion component and the remaining, nondispersion part of the energy:

$$E_{\text{D}} = E_{\text{disp}}^2 + E_{\text{exch-disp}}^2 \quad (3)$$

$$E_{\text{noD}} = E_{\text{pol}}^1 + E_{\text{exch}}^1 + E_{\text{ind}}^2 + E_{\text{exch-ind}}^2 + \delta\text{HF} \quad (4)$$

These results, along with the reference CCSD(T)/CBS calculations, are listed in Table 1. It should be noted that even

Table 2. RMSE (in kcal/mol) of the Two-Point Extrapolation Schemes As against the DFT-SAPT/CBS Values

basis sets	p	RMSE (E_D)	RMSE (E_{tot})
D→T	2.868	0.009	0.036
	3	0.015	0.065
	3.489	0.044	0.009
T→Q	2.868	0.008	0.006
	3	0.006	0.012
	3.489	0.013	0.015
Q→5	2.868	0.015	0.016
	3	0.016	0.020
	3.489	0.020	0.000

the best (in terms of basis set used) DFT-SAPT only approximates the interaction energy when compared to the CCSD(T) results. The first reason might be the difference in the underlying theory, because the DFT-SAPT treats the interaction as a perturbation only up to the second order; higher-order contributions are included only in the $\delta(\text{HF})$ term, which does not contain correlation energy. Second, the results of the DFT-SAPT are affected by the functional used. Our findings are consistent with the previously reported comparison of the DFT-SAPT and CCSD(T) interaction energies.^{4,7,18}

As mentioned above, the nondispersion part of the interaction energy converges faster with the basis set size, and the value calculated with a rather large basis set is close enough to the CBS limit. We, therefore, do not extrapolate this component as recommended in the literature.¹⁹ The DFT-SAPT/CBS energy is built from the E_{noD} calculated in the aug-cc-pVSZ basis set, and only the E_D is extrapolated to the CBS.

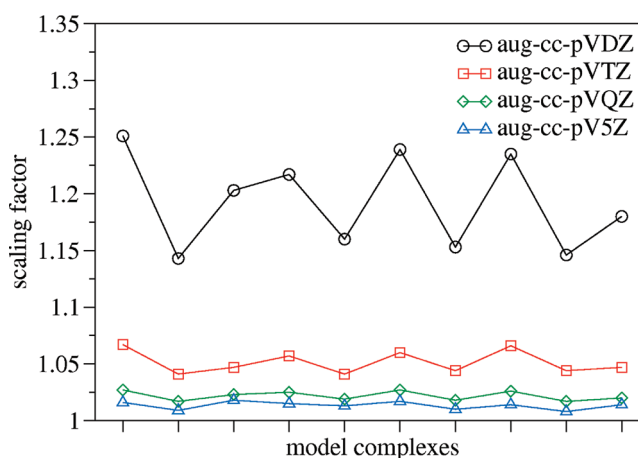
The complete series of basis sets is used to optimize the fitting function in order to find the value of the exponent p that minimizes the error between the actual E_D and the value obtained from eq 1 for each basis set. In order to build a transferable fitting function, we looked for a single value of p for the whole set of complexes. In such an optimization, the factor A and the $E_{D,\text{CBS}}$ of each system act as variable parameters. The target variable to be minimized is the sum of the squares of the residual errors.

It is known that the largest nonsystematic error is observed for the small basis sets. Therefore, a better fit can be achieved when the augmented double- ζ basis is not included in the series (the error is reduced also because the fitted set is smaller). However, for larger systems, we would like to use the smallest basis sets possible for the extrapolation, which means extrapolation from the double- and triple- ζ basis sets. We performed the fitting both on the full series (D, T, Q, 5) and on truncated series of just T, Q, and 5, but we will focus on the first one in the discussion below. For the full series, the value of $p = 2.868$ and the root-mean-square (RMS) of the residual error is 0.0039 kcal/mol. For the truncated series, we obtained $p = 3.489$ with an RMS error of 0.0003 kcal/mol. It is worth noting that both values of p are close to the theoretically predicted value of 3. The resulting $E_{D,\text{CBS}}$ values from all of these fits ($p = 2.868, 3.489, \text{ and } 3$) are very similar (the RMS difference between them being ≤ 0.01 kcal/mol). Such a difference is negligible when compared to the error with respect to the CCSD(T)/CBS reference (RMSE = 0.17 kcal/mol, which is about 6%).

Two-Point Extrapolation. Knowing the exponent p , we can use eq 2 for an extrapolation to the CBS limit from two calculations. Table 2 lists the results of the two-point

Table 3. Ratios of the Dispersion Part of the Interaction Energy (E_D) at the CBS Limit to E_D Calculated in the Finite Basis Set aug-cc-pVXZ ($X = \text{D, T, Q, 5}$)

	$f(\text{D})$	$f(\text{T})$	$f(\text{Q})$	$f(5)$
water dimer	1.251	1.067	1.027	1.016
methane dimer	1.143	1.041	1.017	1.009
methane...water	1.203	1.047	1.023	1.018
ammonia dimer	1.217	1.057	1.025	1.015
ethyne dimer	1.160	1.041	1.019	1.013
ammonia...water	1.239	1.060	1.027	1.017
ammonia...N ₂	1.153	1.044	1.018	1.010
formic acid...water	1.235	1.066	1.026	1.014
methane...N ₂	1.146	1.044	1.017	1.008
water...N ₂	1.180	1.047	1.020	1.014
average	1.193	1.051	1.022	1.013

**Figure 1.** Ratios of the dispersion part of the interaction energy E_D at the CBS limit to E_D calculated in the finite basis set aug-cc-pVXZ ($X = \text{D, T, Q, 5}$) for the model complexes considered in this work.

extrapolations using various combinations of basis sets and exponents p as the RMSE to the DFT-SAPT/CBS value obtained using $p = 2.868$. Both the dispersion part and the total interaction energy are considered.

The extrapolated values are very close to the DFT-SAPT/CBS limit, and all of them are of a quality at least comparable to DFT-SAPT/aug-cc-pVSZ. All of these errors are an order of magnitude smaller than the difference between the DFT-SAPT and CCSD(T); therefore, any extrapolation yields an estimate of the CBS limit accurate enough for practical purposes.

Scaling of the Dispersion Term. By comparing the dispersion term E_D obtained with each basis set with its CBS limit, we can obtain the scaling factor (f) needed to estimate the CBS value from just a single calculation. These factors, and their average over the set, are listed in Table 3 and plotted in Figure 1. It is clear that these factors depend on the character of the interaction, but this dependence diminishes with the increasing size of the basis set. However, even for the double- ζ basis set, the absolute deviations from the average scaling factors are always lower than $f - 1$, which means that such a scaling only improves the result when compared to the CBS limit. Our accurate estimate of the CBS limit allows us to calculate the scaling factors, which are

Table 4. Comparison of the One-Point Scaling Schemes for 10 Model Complexes^a

scaled E_D	E_{noD}	f	RMSE
aug-cc-pVDZ	aug-cc-pVDZ	1.193	0.174
aug-cc-pVDZ	aug-cc-pVTZ	1.193	0.104
aug-cc-pVTZ	aug-cc-pVTZ	1.051	0.050

^aThe dispersion part of the interaction energy is scaled by the factor f in order to get closer to the CBS limit. The errors (in kcal/mol) are relative to an accurate estimate of the CBS limit.

larger than those previously used in the literature, where a scaling of 1.1 or 1.15 was suggested for the aug-cc-pVDZ basis set⁶ without a rigorous derivation.

The most practical estimates of the E_D CBS limit obtained by such a scaling are those from smaller basis sets. If it is possible to perform the calculation in a larger basis set, another calculation in a smaller one makes it possible to use the more accurate two-point extrapolation. Here, we consider three scaling schemes: The scaling of E_D in an aug-cc-pVDZ calculation (labeled D*/D), the scaling of E_D in the aug-cc-pVTZ (labeled T*/T), and the combination of the scaled E_D from the aug-cc-pVDZ with the E_{noD} from the aug-cc-pVTZ (labeled D*/T). Since the calculation of the dispersion term is the limiting factor, the latter scheme allows us to improve the nondispersion part at a little extra expense.

The results of such calculations are summarized in Table 4 as the RMSE relative to the CBS limit. With the D*/D scheme (first line), this error is 0.17 kcal/mol, as compared to 0.55 kcal/mol in the aug-cc-pVDZ calculation without scaling. The D*/T scheme (second line) is the optimum between the computational cost and accuracy. Here, the error is 0.1 kcal/mol when compared to 0.48 kcal/mol in the same calculation without the scaling. Finally, the T*/T scheme (third line) is the most accurate one with an error of 0.05 kcal/mol. Better accuracy (0.04 kcal/mol) can be achieved with two-point extrapolation at a similar cost, but the difference is surprisingly small.

We have also tested separate scaling of the dispersion and dispersion-exchange terms of the SAPT decomposition. This approach yields worse results than scaling their sum as described before.

Validation of One-Point Scaling. In order to test the scaling schemes outside the training set, we calculated a set of 31 larger complexes in the aug-cc-pVDZ and aug-cc-pVTZ basis sets. The average interaction energy in this set is -5.5 kcal/mol. Here, the two-point extrapolation serves as a reference for the scaled calculation. We validated this reference against the aug-cc-pVTZ to aug-cc-pVQZ extrapolation for the 11 smallest complexes from this set; the RMSE is 0.027 kcal/mol. This is an order of magnitude smaller than the error of the scaling schemes we have investigated here. We therefore consider even the aug-cc-pVDZ to aug-cc-pVTZ extrapolation as a reasonable estimate of the CBS limit for this purpose. The errors, summarized in Table 5, are very similar to those observed in the validation set. The scaled double- ζ calculations are slightly more accurate than the unscaled triple- ζ ones, and the scaled triple- ζ results are very close to the CBS limit. These results confirm that the scaling factors presented here are transferable and provide a significant improvement in accuracy without any extra expense (a 4-fold improvement with the simplest scaling scheme using just the aug-cc-pVDZ calculation).

Table 5. Evaluation of the One-Point Scaling Schemes for the Validation Set of 31 Complexes As Compared to the Unscaled Results^a

scaled E_D	E_{noD}	f	RMSE (kcal/mol)
aug-cc-pVDZ	aug-cc-pVDZ	no scaling	0.684
aug-cc-pVTZ	aug-cc-pVTZ	no scaling	0.196
aug-cc-pVDZ	aug-cc-pVDZ	1.193	0.169
aug-cc-pVDZ	aug-cc-pVTZ	1.193	0.156
aug-cc-pVTZ	aug-cc-pVTZ	1.051	0.044

^aThe errors are relative to the CBS estimate extrapolated from calculations in the aug-cc-pVDZ and aug-cc-pVTZ basis sets.

CONCLUSIONS

Using accurate DFT-SAPT calculations in up to the aug-cc-pV5Z basis set, we have constructed a fitting function that can be used for the extrapolation of the DFT-SAPT dispersion energy to the CBS limit. The remaining part of the interaction energy converges faster with basis set size and is not extrapolated.

A power-law-based formula (eq 1) with the exponent $p = 2.868$ yields the best results for any pair of basis sets. The interaction energies obtained with this extrapolation are very close to the CBS limit, and the difference from the more accurate CBS estimate is an order of magnitude smaller than the accuracy of the DFT-SAPT when compared to the CCSD(T) results. Therefore, even an extrapolation using the aug-cc-pVDZ and aug-cc-pVTZ basis sets should yield results that can be safely used as a DFT-SAPT/CBS estimate.

The most affordable way of improving the DFT-SAPT calculation toward the CBS limit is a simple scaling of the dispersion. When the dispersion energy is calculated with the smallest basis set in the series, aug-cc-pVDZ, the results vary with the type of the interaction, but even this simple approach yields results better than the unscaled calculation in all of the cases we have tested. Moreover, this scheme can be improved at little extra expense by calculating the nondispersion part in a larger basis set. The resulting computational scheme offers a very efficient way of obtaining accurate total interaction energies and their components for extended complexes^{20,21} on the basis of “cheap” DFT-SAPT/aug-cc-pVDZ calculations. It should be stressed again that the DFT-SAPT/CBS interaction energies are very close to the benchmark CCSD(T)/CBS ones.

ASSOCIATED CONTENT

Supporting Information. Geometries of the model complexes, list and geometries of validation complexes, all of the components of the DFT-SAPT decomposition for each system and basis set. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*E-mail: rezac@uochb.cas.cz.

ACKNOWLEDGMENT

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.

REFERENCES

- (1) Williams, H. L.; Chabalowski, C. F. *J. Phys. Chem. A* **2001**, *105*, 646–659.
- (2) Jansen, G.; Hesselmann, A. *J. Phys. Chem. A* **2001**, *105*, 11156–11157.
- (3) Jeziorski, B.; Moszynski, R.; Szalewicz, K. *Chem. Rev.* **1994**, *94*, 1887–1930.
- (4) Hesselmann, A.; Jansen, G.; Schütz, M. *J. Am. Chem. Soc.* **2006**, *128*, 11730–11731.
- (5) Sedláč, R.; Fanfrlík, J.; Hnyk, D.; Hobza, P.; Lepšík, M. *J. Phys. Chem. A* **2010**, *114*, 11304–11311.
- (6) Berka, K.; Laskowski, R.; Riley, K. E.; Hobza, P.; Vondrášek, J. *J. Chem. Theory Comput.* **2009**, *5*, 982–992.
- (7) Riley, K. E.; Pitoňák, M.; Jurečka, P.; Hobza, P. *Chem. Rev.* **2010**, *110*, 5023–5063.
- (8) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007.
- (9) Woon, D. E.; Dunning, T. H. *J. Chem. Phys.* **1994**, *100*, 2975.
- (10) Werner, H.; Knowles, P. J.; Manby, F. R.; Schütz, M.; Celani, P.; Knizia, G.; Korona, T.; Lindh, R.; Mitrushenkov, A.; Rauhut, G.; Adler, T. B.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Goll, E.; Hampel, C.; Hesselmann, A.; Hetzer, G.; Hrenar, T.; Jansen, G.; Köppl, C.; Liu, Y.; Lloyd, A. W.; Mata, R. A.; May, A. J.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklass, A.; Palmieri, P.; Pflüger, K.; Pitzer, R.; Reiher, M.; Shiozaki, T.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T.; Wang, M.; Wolf, A. MOLPRO, version 2009.1, a package of ab initio programs; molpro, 2009.
- (11) Hesselmann, A.; Jansen, G.; Schütz, M. *J. Chem. Phys.* **2005**, *122*, 014103.
- (12) Hesselmann, A.; Jansen, G. *Chem. Phys. Lett.* **2002**, *357*, 464–470.
- (13) Gruuning, M.; Gritsenko, O. V.; van Gisbergen, S. J. A.; Baerends, E. J. *J. Chem. Phys.* **2001**, *114*, 652.
- (14) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. *Chem. Phys. Lett.* **1998**, *286*, 243–252.
- (15) Schwenke, D. W. *J. Chem. Phys.* **2005**, *122*, 014107.
- (16) Jurečka, P.; Hobza, P. *Chem. Phys. Lett.* **2002**, *365*, 89–94.
- (17) Dabkowska, I.; Jurečka, P.; Hobza, P. *J. Chem. Phys.* **2005**, *122*, 204322.
- (18) Pitoňák, M.; Riley, K. E.; Neogrády, P.; Hobza, P. *Chem. Eur. J. Chem. Phys.* **2008**, *9*, 1636–1644.
- (19) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Olsen, J. *Chem. Phys. Lett.* **1999**, *302*, 437–446.
- (20) Fiethen, A.; Jansen, G.; Hesselmann, A.; Schutz, M. *J. Am. Chem. Soc.* **2008**, *130*, 1802–+.
- (21) Podeszwa, R.; Bukowski, R.; Rice, B.; Szalewicz, K. *Phys. Chem. Chem. Phys.* **2007**, *9*, 5561–5569.

Locating Instantons in Many Degrees of Freedom

Judith B. Rommel,[†] T. P. M. Goumans,[‡] and Johannes Kästner^{*,†}[†]Computational Biochemistry Group, Institute of Theoretical Chemistry, University of Stuttgart, Stuttgart, Germany[‡]Gorlaeus Laboratories, LIC, Leiden University, Leiden, The Netherlands

ABSTRACT: We implemented and compared four algorithms to locate instantons, i.e., the most likely tunneling paths at a given temperature. These allow to calculate reaction rates, including atom tunneling, down to very low temperature. An instanton is a first-order saddle point of the Euclidean action in the space of closed Feynman paths. We compared the Newton–Raphson method to the partitioned rational function optimization (P-RFO) algorithm, the dimer method, and a newly proposed mode-following algorithm, where the unstable mode is directly estimated from the instanton path. We tested the algorithms on three chemical systems, each including a hydrogen transfer, at different temperatures. Overall, the Newton–Raphson turned out to be the most promising method, with our newly proposed mode following, being the fall-back option.

I. INTRODUCTION

Quantum tunneling of atoms plays an important role in many chemical reactions, predominantly those involving hydrogen atoms. Especially at low temperature, tunneling increases reaction rates compared to a nonquantum mechanical over-the-barrier model.

The tunneling probability depends on the mass of the particles that undergo tunneling motion and on the shape and height of the effective barrier being crossed.^{1,2} The mass dependence of the tunneling rate gives rise to large kinetic isotope effects (KIE). Atom tunneling is observed in reactions in space,^{3–5} in biological applications,^{6–9} and in other areas of chemistry. In principle, any chemical reaction will be dominated by tunneling at low enough temperature.

Many enzymes support tunneling of hydrogen atoms even at room temperature. Experimentally, this is observed by an unusually large KIE,^{6,9–13} the ratio between the reaction rates involving deuterium and protium. In computer simulations, the importance of tunneling can directly be shown by switching the effect on or off.^{7,8,14–17} A high KIE only shows that tunneling accelerates the rate-limiting step of a reaction. There is an active debate as to whether the tunneling process of atoms is crucial for the catalytic effect of some enzymes and whether it is actively promoted by the proteins. Calculations are a promising tool to provide insight into these issues.

In classical transition-state theory (TST),^{18,19} reviewed in refs 20 and 21, quantum effects can be included via the vibrations (and, thus, the zero-point energy). However, tunneling is neglected. A general transition state in a system with N degrees of freedom is a closed $N - 1$ -dimensional hypersurface encapsulating the reactant state. Commonly, however, one refers to a (classical) transition state as a first-order saddle point on the potential energy surface. The latter meaning will be used here.

In the following, we give a short overview of methods to calculate tunneling rates. Since many methods have appeared over the years, this list is necessarily incomplete.

Tunneling near the top of the barrier is accounted for by corrections to the classical reaction rate proposed by Wigner²²

and used by Bell.²³ The classical rate is multiplied by a correcting factor $\kappa(T)$, which results in exact rates for parabolic barriers. The method is applicable above a crossover temperature T_c :²⁴

$$T_c = \frac{\hbar\Omega}{2\pi k_B} \quad (1)$$

where Ω denotes the magnitude of the imaginary frequency of the unstable mode at the saddle point. To extend the range of applicability to below T_c , a third-order expansion in \hbar of $\kappa(T)$ was used.²⁵

The semiclassical approximation assumes one main tunneling path rather than taking the whole potential energy surface into account. Methods based on the transition-state theory and the semiclassical approximation can be categorized by their choice of the tunneling path. The zero-curvature tunneling approximation (ZCT)²⁶ approximates the tunneling path by the intrinsic reaction path (minimum-energy path (MEP) in mass-weighted coordinates). Even at rather high temperature, close to T_c , the most likely tunneling path will, however, deviate from the MEP. The effect is known as corner cutting.²⁷ Small-curvature tunneling (SCT)²⁸ assumes a tunneling path in the vicinity of the MEP. SCT can be expected to be a good approximation close to T_c . The other extreme for a choice of a tunneling path is the straight line path in a method known as large curvature tunneling correction (LCT).^{29,30} While minimizing the tunneling distance, LCT ignores the potential energy in the choice of the path. At low temperature, LCT can be expected to be a better approximation than SCT. It has been proposed^{31,32} to use a linear combination of the SCT and LCT paths by minimizing the tunneling action.

The optimal tunneling path within the purely statistical semiclassical approximation is found by the instanton method,^{33–37} analogous to the imaginary F method.³³ It is based on statistical Feynman path integrals.³⁸ Some of the aforementioned approximations can be regarded as an approximation to the instanton method. The instanton itself is the tunneling path with the

Received: November 15, 2010

Published: February 18, 2011

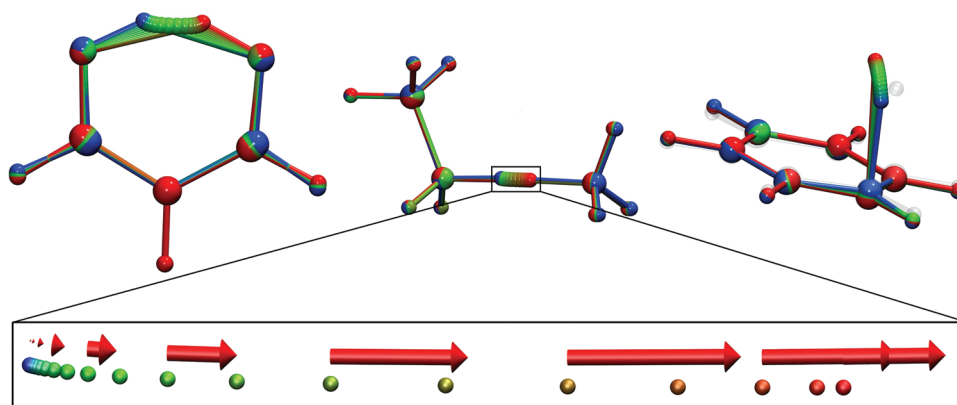


Figure 1. top: Instanton geometries of the three test cases: malonaldehyde, ammonium and methylamine as well as H + benzene. Each calculated at $T = 200$ K with $P = 20$ images. Bottom: the transition mode \mathbf{u}_{inst} indicated as arrows on the hydrogen atom transferred between ammonium and methylamine. For clarity, only the arrows corresponding to every second image are shown. The components of \mathbf{u}_{inst} on the other atoms are negligibly small on the scale shown.

highest statistical weight at a given temperature. Due to a mathematical equivalence, it can also be interpreted as a periodic orbit in the upside-down potential energy surface of the system. In the formulation of a harmonic quantum transition-state theory, many parallels between the classical TST and the instanton method were demonstrated.³⁹ It is, along with some extensions,^{40,41} increasingly used to calculate reaction rates in chemical systems.^{4,5,42–50}

Beyond TST, quantum dynamics allows to calculate tunneling rates by solving the time-dependent Schrödinger equation. The time-dependent Hartree approach^{51–54} and other methods^{55,56} were proposed. The costs of these methods grow exponentially with the number of degrees of freedom. Statistical Feynman paths can circumvent that and have also been used in simulations to estimate dynamical properties either by evaluating a centroid potential of mean force^{57–60} or by so-called ring polymer dynamics.⁶¹ Feynman path integral approaches using TST, but not the semiclassical approximation, are also in use. The centroid density method, a quantum transition-state theory (QTST), is one of these.^{24,62,63} Another QTST approach is the reversible action-space work QTST (RAW-QTST),^{39,45} which in the harmonic limit reduces to instanton theory.

This paper is organized as follows. In Section II we describe the details of instanton theory. In Section III we propose possible techniques to locate instantons. In Section IV we compare the efficiency of these algorithms for a series of test systems at different temperature. In Section V we discuss the efficiency of the algorithms, possible issues which may impede instanton optimizations, and ways to avoid them. Finally, we conclude and give an outlook at more realistic calculations as well as further problems which remain to be solved.

II. INSTANTON THEORY

The quantum statistical partition function Q is expressed via a Feynman path integral:³⁸

$$Q = \int dx \langle x | \exp^{-\beta H} | x \rangle$$

$$= \int dx \int_{\bar{x}(0)=x}^{\bar{x}(\beta\hbar)=x} \mathcal{D}\bar{x} \exp\left(-\frac{1}{\hbar} S_E[\bar{x}]\right) \quad (2)$$

with $\beta = 1/(k_B T)$, H the Hamiltonian of the system, and k_B being Boltzmann's constant. $\mathcal{D}\bar{x}$ has the "heuristic" meaning of integration over all paths \bar{x} satisfying the boundary conditions $\bar{x}(0) = \bar{x}(\beta\hbar) = x$. $\mathcal{D}\bar{x}$ can be formed into a conditional Wiener measure to give a well-defined formulation of the integral.⁶⁴ Here and in the following, points in configuration space (N dimensional for a molecule with N degrees of freedom) are denoted by italic symbols (x), while paths are denoted by upright (roman) symbols (\bar{x}). The Euclidean action functional S_E is given by

$$S_E[\bar{x}] = \int_0^{\beta\hbar} \left(\frac{m}{2} \frac{d\bar{x}(\tau)^2}{d\tau} + V(\bar{x}(\tau)) \right) d\tau \quad (3)$$

with $V(x)$ being the potential energy.

To approximate the integral over all paths in eq 2, S_E is expanded to second order around its stationary paths. The stationary condition is

$$\left[\frac{\delta S_E}{\delta \bar{x}} \right]_{x_0} = -m \frac{d^2 x_0(\tau)}{d\tau^2} + \nabla V(x_0(\tau)) = 0 \quad (4)$$

One path which fulfills eq 4 is x collapsed to the geometry of the reactant minimum ($\bar{x}(\tau) = x_{\text{RS}}$). From this, the quantum partition function of the reactant can be obtained. To use transition-state theory, the partition function of a quantum transition state (QTS), a dynamical bottleneck between the reactant and the product, is required. In contrast to the path corresponding to the reactant minimum, the instanton path x_{inst} is delocalized. Its Hessian exhibits exactly one negative eigenvalue, the eigenvector \mathbf{u}_{inst} corresponding to a movement of the whole path toward the reactant or the product. An example of \mathbf{u}_{inst} is depicted in Figure 1. The Hessian of the instanton additionally exhibits one eigenvalue which is 0. This corresponds to the arbitrary starting position of the path, i.e., to a reparametrization $\tau \rightarrow \tau + c$ of the path. For molecules, the Hessian exhibits another six 0 eigenvalues (5 for linear molecules), corresponding to the translation and rotation of the whole molecule.

To calculate the Euclidean action and, through that, the rate the Feynman path $x(\tau)$ is discretized into P' images. By introducing mass-weighted coordinates $y_i = x_i(m_i)^{1/2}$ the equation can be somewhat simplified

$$S_E = \beta\hbar \sum_{k=1}^{P'} \left(\frac{P'}{2(\beta\hbar)^2} |y_{k+1} - y_k|^2 + \frac{V(y_k)}{P'} \right) \quad (5)$$

To fulfill eq 4, a stationary point has to be searched for. Since we are interested in a rate, we specifically search for a first-order saddle point. This saddle point is the instanton.

It turns out that the instanton is delocalized along one line in the configuration space of the molecule. The Feynman path is closed by proceeding along this line forward and backward. Choosing an even number of images and starting the index $k = 1$ at an image next to one turning point, the images k and $P' - k + 1$ have identical coordinates. All images are traversed twice, the turning points lie outside of the discretized path. Thus, it is sufficient to sum over half of the images ($P = P'/2$):⁶⁵

$$S_E = \frac{2P}{\beta\hbar} \sum_{k=1}^{P-1} |y_{k+1} - y_k|^2 + \frac{\beta\hbar}{P} \sum_{k=1}^P V(y_k) \quad (6)$$

Equation 6 can be used to derive the gradient and the Hessian of S_E with respect to the coordinates of the atoms of each image.

The problem of finding an instanton has been formulated here as a saddle-point search of a discretized path. This allows to treat a high number of degrees of freedom. Instantons can, however, also be interpreted as unstable periodic orbits on the upside-down potential energy surface ($-V$). Techniques to find periodic orbits^{66,67} have previously been used to find instantons.^{68,69} In practice, these techniques are, however, only applicable to systems with a few degrees of freedom.⁶⁹

The integral over all paths in eq 2 is approximated by expanding $V(x)$ at each point, and, thus, $S_E[x]$ quadratically around the stationary path x_{inst} . Then the integral in eq 2 turns into a Gaussian integral which can be solved analytically. This results in an expression of the rate:^{34,36,70}

$$k_{\text{inst}} = \frac{1}{Q_{\text{RS}}} \sqrt{\frac{S_0}{2\pi\hbar}} \frac{P'}{\beta\hbar} \frac{1}{\sqrt{|\prod_i \lambda_i|}} \exp(-S_E[x_{\text{inst}}]/\hbar) \quad (7)$$

with Q_{RS} denoting the quantum mechanical partition function of the reactant state, λ_i are the eigenvalues of the Hessian of S_E , and the prime on the product indicates that the zero eigenvalue(s) are omitted. S_0 is twice the part of the Euclidean action depending on the length of the path:

$$S_0 = \frac{4P}{\beta\hbar} \sum_{k=1}^{P-1} |y_{k+1} - y_k|^2 \quad (8)$$

The amount of delocalization of the instanton depends on the temperature. At low temperature (high β), the effective force constant between the images, $P/(\beta\hbar)$, becomes smaller allowing the images to spread further to accommodate a lower potential energy. With increasing temperature, the path becomes more and more localized, finally collapsing to a point (the classical transition state y_{cl}) at the same T_c as given in eq 1. Instanton theory is applicable below T_c .

III. TECHNIQUES TO FIND INSTANTONS

The task is to find a first-order saddle point of S_E in the $P \times N$ -dimensional space (for a molecule with N degrees of freedom) of the coordinates of all atoms of all images of the Feynman path.⁴⁵ Generally, we use sequential cooling,^{4,5,49,50,65} i.e., we start at a

temperature below T_c and calculate an instanton and the tunneling rate. Then we successively lower the temperature, starting the search from a converged instanton. Alternative approaches and other starting paths may be used as well.

Consistently with the notation of continuous paths introduced above, N -dimensional quantities like the position of one image y_k or the classical transition mode u_{cl} are denoted by italic symbols. $P \times N$ -dimensional quantities like the complete Feynman path y or the transition mode of the instanton u_{inst} are denoted by upright (roman) symbols.

The infrastructure to calculate the Euclidean action and all the methods to optimize instantons described here were implemented in DL-FIND.^{71,72} Through the interface to ChemShell,^{73,74} they can be used with many quantum chemistry programs as well as classical force fields or quantum mechanics/molecular mechanics (QM/MM) energy expressions.

A. Starting Path. In order to facilitate the search for an instanton, an initial starting path as close as possible to the final instanton is estimated. For $T \geq T_c$ the instanton is collapsed to y_{cl} , the first-order saddle point on the potential energy surface (classical transition state). As the temperature decreases, the images spread out approximately along the classical unstable mode u_{cl} . In the quadratic region of $V(y)$, the images spread like

$$y(\tau) = y_{\text{cl}} + \Delta r u_{\text{cl}} \cos(2\pi\tau/\beta\hbar), \quad 0 \leq \tau \leq \beta\hbar \quad (9)$$

So, we use the discretized version for P images:

$$y_i = y_{\text{cl}} + \Delta r u_{\text{cl}} \cos\left(\frac{i-1/2}{P}\pi\right), \quad 0 < i \leq P \quad (10)$$

The initial spread Δr is chosen manually. It cannot be estimated from y_{cl} or its Hessian. We used $\Delta r = 0.4$ atomic units.

B. Mode-Following Methods. Minimum-mode following is an approach to search for first-order saddle points. The action is minimized in all directions but one, u_{inst} (Figure 1). Along u_{inst} the action is maximized. If u_{inst} is the eigenvector of the Hessian associated with the lowest eigenvalue, such an algorithm converges to a first-order saddle point.

The dimer method⁷⁵⁻⁷⁷ is a minimum-mode-following algorithm with the transition mode recalculated in what is called dimer rotation in each iteration. Converged rotation provides the correct u_{inst} without ever requiring the calculation of the Hessian. We implemented an improved version of the dimer rotations.⁷⁸ The limited memory version^{79,80} of the Broyden-Fletcher-Goldfarb-Shanno (BFGS)⁸¹⁻⁸⁴ optimizer was used for translations and rotations. Dimer rotations were optimized until u_{inst} changed by less than 5° in one rotation. This required dimer rotations only at the start of an optimization run. A tighter convergence criterion for the rotations was tried but led to an increase in the number of energy evaluations (i.e., a possible saving through fewer translations was lost by more rotation steps). An alternative to the dimer method, which also does not require the Hessian, is the Lanczos method,^{85,86} which has been used to locate instantons.^{4,5,49,50,65}

For the instanton search problem it turned out that an approximation of u_{inst} can easily be obtained

$$u_{\text{inst},i} \approx u_{\text{inst},i}^{\text{TM}} = \frac{1}{2}(y_{i+1} - y_{i-1}) \quad (11)$$

i.e., the transition mode is assumed to be the tangent of the instanton path. Actually, it is the tangent of only half the path and

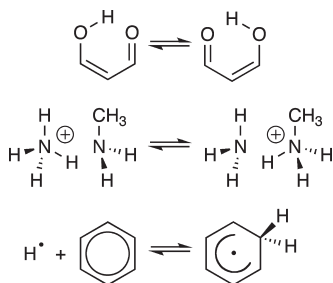


Figure 2. The test cases for which the algorithms to locate instantons were tested: malonaldehyde, ammonium and methylamine, and the addition of hydrogen to benzene.

the reversed tangent of the other half. However, as given in eq 6, we only use the first half of the instanton path as active variables in the optimization. The mode \mathbf{u}_{inst} is recalculated after each optimization step. Calculating it only at the start and keeping it constant resulted in divergence in some cases. Using a \mathbf{u}_{inst} calculated by eq 11 is called “tangent mode” (TM) in the following.

C. Hessian-Based Methods. The evaluation of the prefactor of the rate in eq 7 requires to calculate the Hessian of the potential energy surface at each image of the instanton—a computationally demanding task. In sequential cooling it can be used as a good first approximation to the Hessian of the instanton at a lower temperature, which speeds up convergence considerably.

We tested a truncated Newton–Raphson (NR) algorithm and the partitioned-rational function optimizer (P-RFO).^{87–90} NR generally converges quadratically to stationary points. However, it has the disadvantage that it converges to any stationary point, not necessarily to first-order saddle points. This is expected unproblematic since our starting structures are often already quite close to the sought-after saddle point. After convergence is achieved, the Hessian at the new instanton geometry has to be calculated. Its eigenvalue spectrum confirms whether a first-order saddle point was obtained.

The P-RFO method converges to first-order saddle points by construction. It is generally the method of choice to search for classical transition states due to its fast and reliable convergence properties.

Hessians are obtained for the individual images from previous rate calculations. With a changed temperature, they are used for the subsequent optimization. With changes in coordinates, we update the Hessians of the individual images according to the Bofill scheme.⁹¹ For small steps and a noisy gradient, the update of the Hessian may actually deteriorate it rather than improve it. So we keep the Hessian unmodified if the coordinates of the images change by less than a predefined threshold. Using the initial Hessian without updates results in less stable optimizations.

When the first instanton below T_c is calculated, a previous Hessian calculation along the full instanton path is unavailable. However, we normally evaluate the Hessian of the potential energy surface at the classical TS. We update this Hessian to the initial image positions, again using the Bofill scheme.

IV. EXAMPLES

The performance of the four optimization algorithms was tested on three chemical systems (Figure 2) at various temperature intervals. All systems were described with semiempirical

methods. These pose similar challenges to the optimization algorithms (like numeric discontinuities in the potential energy surface due to incomplete SCF convergence) while being orders of magnitude faster than density functional theory or post-Hartree–Fock methods.

The internal hydrogen transfer in malonaldehyde was simulated with the PM3 Hamiltonian.⁹² Since the reactant and product states are chemically indistinguishable, this system has a symmetric barrier. Its tunneling behavior has recently been investigated with a variety of methods.⁹³ The hydrogen transfer between ammonium and methylamine, also described with PM3, is slightly asymmetric. A very asymmetric barrier is found in the addition of hydrogen to benzene. The latter is a relevant model for the formation of H_2 in space.^{4,5} We described it with AM1.⁹⁴ In order to get a more realistic reactant-state geometry (van der Waals complex), an empirical dispersion correction originally designed for higher-order methods⁹⁵ was added (prefactor $S_6 = 1$). Energies and gradients were calculated with MNDO99⁹⁶ interfaced to ChemShell^{73,74} through DL-FIND. It should be emphasized that the aim of this study is to compare the efficiency of algorithms rather than the reproduction of experimental values. The instantons of the three test cases are depicted in Figure 1. Visualization was done using VMD 1.8.7.⁹⁷

Classical transition states for the three test systems were calculated with the superlinear converging version⁷⁸ of the dimer method⁷⁵ as implemented in DL-FIND.⁷¹ Mass-scaled coordinates (mass of hydrogen being 1) were used. Convergence was considered to be achieved for the maximum component of the gradient g_{max} being below a tolerance value (convergence criterion) of $g_{\text{tol}} = 10^{-5}$, the root-mean-square (rms) of the gradient being below 6.66×10^{-6} , the maximum component of the predicted step being below 4×10^{-5} , the rms of the predicted step being below 2.66×10^{-5} , and the last change in the energy of the dimer midpoint being below 2.22×10^{-8} , all values in atomic units. The dimer direction, which in mass-scaled coordinates converges to u_{cl} , was converged in each dimer iteration until it changed by less than 1° .

Hessians were calculated by finite differences of the gradients with two steps of 2×10^{-3} (mass-scaled atomic units) in each dimension. The crossover temperature T_c was obtained according to eq 1 as 442.3, 493.9, and 387.2 K for malonaldehyde, ammonium and methylamine, and $\text{H} + \text{benzene}$, respectively.

Instanton searches were performed in mass-weighted coordinates with masses consistent with atomic units (electron mass, m_e). That is, the mass of a hydrogen atom (^1H) is $1837.15 m_e$. This ensures that the masses in eqs 5 and 6 really drop out. On the other hand, this scales all distances up by a factor of 42.695 [= (atomic mass unit/ m_e)^{1/2}] compared to mass-scaled coordinates as defined above. Thus, g_{tol} has to be smaller by the same factor to achieve equivalent convergence. The convergence criterion for the instanton searches was $g_{\text{tol}} = 10^{-7}$ for the maximum component of the gradient of $S_E/(\beta\hbar)$ in mass-weighted coordinates (g_{max}). Suitable choices for g_{tol} will be discussed in Section V.

In each case, the first instanton was found for $T = 300$ K by starting from the classical transition state and distributing the $P = 20$ images along u_{cl} using $\Delta r = 0.4$ au as described in Section III.A. The numbers of steps needed to reach convergence are given in Table 1. In case of malonaldehyde, the instanton search using NR converged back to a state where all images are collapsed to the classical TS. This is obviously a stationary point, but it is not an instanton, i.e., it does not exhibit the correct number of eigenvalues

Table 1. Number of Steps Needed to Reach Convergence for Finding an Instanton at $T = 300$ K Starting from the Classical TS^a

system	TM	dimer	P-RFO	NR
malonal.	114 (2280)	112 (6720)	337 (6726)	64 (1280)
ammon.	103 (2060)	107 (6420)	338 (6760)	46 (920)
H + benzene	151 (3020)	355 (19 360)	137 (2740)	11 (220)

^a number of energy and gradient evaluations in parentheses, $g_{\text{tol}} = 10^{-7}$, and $P = 20$.

Table 2. Number of Steps Needed to Reach Convergence in the Different Methods at Different Temperatures^a

T (K)	P	TM	dimer	P-RFO	NR
Malonaldehyde					
275	20	81 (1620)	83 (4980)	— ^d	4 (80)
250	20	75 (1500)	79 (4740)	— ^d	6 (120)
225	20	72 (1440)	73 (4380)	— ^d	6 (120)
200	20	70 (1400)	68 (4160)	— ^d	17 (340)
100	20	75 (1500)	— ^c	293 (5860)	45 (900)
200	77	275 (21 175)	268 (61 908)	220 (16 940)	7 (539)
Ammonium and Methylamine					
275	20	53 (1060)	66 (3760)	41 (820)	5 (100)
250	20	66 (1320)	67 (4020)	670 (13 400)	8 (160)
225	20	114 (2280)	67 (4060)	— ^d	12 (240)
200	20	111 (2220)	121 (6440)	738 (14 760)	18 (360)
100	20	76 (1520)	88 (5160)	— ^d	100 (2000)
200	77	248 (19 096)	— ^d	— ^d	16 (1232)
H + Benzene					
275	20	127 (2540)	119 (6080)	263 (5260)	5 (100)
250	20	129 (2580)	123 (6160)	82 (1640)	6 (120)
225	20	142 (2840)	124 (6540)	133 (2660)	10 (200)
200	20	225 (4500)	104 (5480)	258 (5160)	17 (340)
100	20	— ^b	75 (4080)	652 (13 040)	44 (880)
200	77	528 (40 656)	463 (106 953)	— ^d	19 (1463)

^a While $g_{\text{tol}} = 10^{-7}$ and number of energy and gradient evaluations in parentheses. All calculations started out from a converged instanton at $T = 300$ K, $P = 20$. ^b Calculation converged to a wrong stationary point (all images in the reactant minimum). ^c Calculation converged to a wrong stationary point (images interchanged during the optimization). ^d Not converged.

being zero. Using $\Delta r = 0.6$, as defined in eq 10, leads to convergence to a delocalized Feynman path, the instanton.

Further instantons at lower temperature were optimized starting from the geometry and the Hessian of the instanton obtained with TM at 300 K. Here we tested the convergence for different temperature intervals rather than using sequential cooling. For $T = 200$ K we additionally performed an instanton search with more images, $P = 77$. We interpolated geometries and Hessians from the instanton with $P = 20$ and $T = 300$ K by inserting three extra images in between two consecutive ones.

The number of steps needed to achieve convergence in each case, along with the number of energy and gradient evaluations needed, are given in Table 2. The convergence behavior is exemplarily depicted in Figure 3.

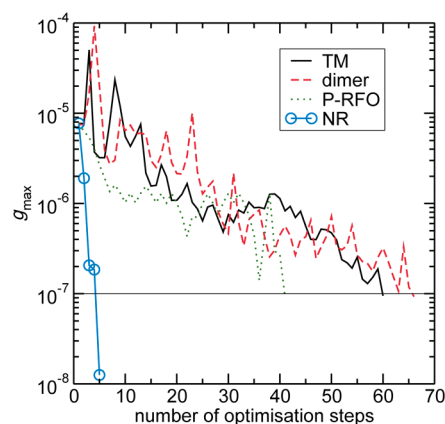


Figure 3. Convergence behavior of the four different optimization methods for ammonium and methylamine at $T = 275$ K. NR converges almost quadratically. The convergence criterion of $g_{\text{max}} = g_{\text{tol}} = 10^{-7}$ is indicated by a thin horizontal line.

In TM, P-RFO, and NR the number of energy evaluations is P times the number of optimization steps. In the dimer method, the rotations require additional energy evaluations. At least two energy and gradient calculations of the full path (dimer mid- and end-points) are required for an estimate of the rotational angle.⁷⁷ In most iterations, at least a third calculation (one dimer rotation) was required. Especially at the beginning of an optimization run, more dimer rotations may be necessary. Overall, in the dimer optimizations, the number of energy and gradient evaluations is about $3P$ times the number of optimization steps. This indicates that on average one dimer rotation step per iteration was sufficient.

Those cases for which numbers are given in Table 2 converged to the correct instanton within a maximum of 1000 optimization steps and 200 000 energy evaluations. Consistency was checked by comparing S_E , S_0 , and k_{inst} as well as the eigenvalue spectrum between the results of different optimization algorithms.

For the H + benzene case, the initial instanton search for $T = 300$ K with the TM method resulted in a somewhat problematic starting structure for the following instanton optimizations. The reason being that for this system the potential energy surface is very flat close to the reactant minimum. This results in many of the images accumulating there. A too weak convergence criterion leads to numerical noise in the image positions and, thus, to numerical noise in \mathbf{u}_{inst} in the TM method. Thus, for the TM calculations of this system, the instanton at $T = 300$ K optimized with NR was used as the starting geometry.

The influence of P and g_{tol} on the vibrational instanton rates $[\log_{10}(k_{\text{inst}})]$, ignoring changes in the rotational partition function as well as the translational partition function between the reactant and the instanton, is shown in Table 3. The error relative to the most accurate value obtained with $P = 96$ images and $g_{\text{tol}} = 10^{-9}$ is given. A difference in the logarithm of 0.1 corresponds to a rate which is off by about 25.9%.

The ammonium and methylamine case raises a particular issue in the rate calculations. The “vibrational” mode in which the two fragments rotate with respect to each other has a very low vibrational frequency of only 11.5 cm^{-1} at the classical TS. Thus, at the temperature range considered, this mode would better be described as a hindered rotator than as a harmonic oscillator. Since in this work we only investigate the effect of

Table 3. Error in the Logarithm of the Rates Compared to the Tightest Convergence Criterion g_{tol} and the Largest Number of Images for Each Case ($T = 200$ K, NR)

P	g_{tol}				
	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}
Malonaldehyde					
20	-0.05	-0.01	-0.01	0.00	0.00
39	-0.14	-0.15	0.00	0.00	0.00
58	-0.53	-0.00	0.00	0.00	0.00
77	-0.54	-0.10	-0.01	0.00	0.00
96	-0.19	-0.10	-0.01	0.00	0.00
Ammonium and Methylamine					
20	-0.45	-0.01	-0.01	-0.01	-0.01
39	-0.20	-0.01	0.00	0.00	0.00
58	-0.18	-0.27	0.00	0.00	0.00
77	-0.19	-0.41	0.00	0.00	0.00
96	-0.20	-0.40	0.00	0.00	0.00
H + Benzene					
20	0.14	0.25	0.01	0.01	0.01
39	0.18	0.31	0.00	0.00	0.00
58	0.17	-0.05	-0.01	0.00	0.00
77	0.10	0.31	0.00	0.00	0.00
96	0.21	0.29	0.01	0.00	0.00

P and g_{tol} on the rates, in Table 3 we ignored this mode just as the other six translational and rotational modes are ignored. If it were included as a vibrational mode in the rate calculations, then its low eigenvalue would cause numerical problems in automatically designating the real zero eigenvalues in the eigenvalue spectrum of the instanton Hessian, thereby compromising the rates. It is worth noting at this point that the additional low mode was handled well by the optimization algorithms.

V. DISCUSSION

We compared four different algorithms to optimize instantons. The results clearly show that the NR algorithm is the most promising one among those tested. Its near-quadratic convergence results in only few optimization steps necessary to reach tight convergence criteria. The convergence behavior is also promising for applications to systems with significantly more degrees of freedom. The methods are not restricted to sequential cooling. Using analytic potentials, we achieved convergence even with the TM method for a straight-line path as starting guess (data not shown). In these cases, a Hessian from previous calculations is not easily available. The results also point out some problems which can be expected in instanton optimizations. These will now be addressed.

It is clear from Table 2 that the P-RFO fails to converge in a number of cases. As a fixed-point iteration scheme, it sometimes reaches a periodic cycle rather than actually converging to the desired stationary point. This can be seen by plotting the maximum gradient component against the number of iterations, see Figure 4. In all cases except $T = 100$ K, a cycle is reached rather than the gradient becoming smaller and smaller. This is caused by a strong dependence of the P-RFO algorithm on an accurate Hessian. In our approach, we only update the Hessian of

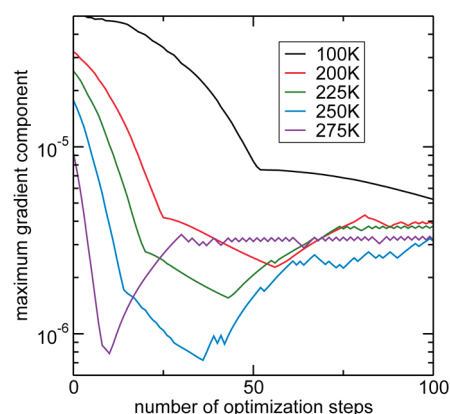


Figure 4. Convergence failures of the P-RFO method for malonaldehyde at different temperatures.

the individual images rather than recalculating it, which inevitably leads to inaccuracies. Using the Powell update⁹⁸ or no update scheme at all leads to an even worse convergence (data not shown). Recalculating the Hessian in each step would of course be prohibitively expensive. Since the other algorithms, in particular NR, generally converge faster and more reliably, instanton optimization with P-RFO is not recommended.

NR intrinsically converges to any stationary point, not necessarily first-order saddle points. While in all results shown here, it actually converged to the sought-after instanton; we observed runs in which NR converged to different states. Especially at high temperature, close to T_c , there seems to be a danger of the whole path collapsing to the classical TS. This was observed here when starting from the classical TS and searching for an instanton at $T = 300$ K in the case of malonaldehyde. A larger Δr for spreading the images in the initial path resolved that problem.

Convergence to a collapsed path can easily be detected during an optimization run by a steadily decreasing value of S_0 . According to eq 8, S_0 vanishes for any collapsed path. In cases where NR converges to a different stationary point than the instanton, the TM algorithm is recommended as a backup solution.

In two cases, H + benzene at $T = 100$ K with the TM method and the dimer method, the calculations actually converged to a path collapsed to the reactant state, see Table 2. While this is also a stationary point of S_E , it obviously is no instanton. These cases are less worrying, however, because the direct change from $T = 300$ K (starting point) to $T = 100$ K is rather extreme. Smaller temperature intervals are recommended. Even at this large temperature interval, however, NR converged well.

Convergence criteria are applied in our present implementation in the same unit system as the optimization being done, i.e., in mass-weighted coordinates with the mass of an electron as unit. To assess which g_{tol} is necessary to obtain the tunneling rate with a given accuracy, we calculated the data presented in Table 3. A criterion of $g_{\text{tol}} = 10^{-7}$ seems sufficient to ensure convergence of the logarithm (basis 10) of the rate to within 0.01. For the systems and the temperature ($T = 200$ K) studied here, $P = 20$ images are obviously sufficient, as the rate (at sufficiently small g_{tol}) is independent of the number of images. At lower temperature, however, the images become less equally spaced in configuration space. More and more images accumulate close to the minima. In these cases, more images are required to achieve converged rates as found in agreement with previous work.⁴⁹ It can also be seen from Table 3 that with more and more images,

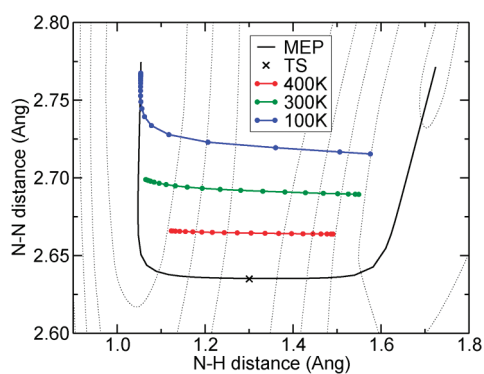


Figure 5. The potential energy surface, the MEP, and instantons at three temperatures for the case of ammonium and methylamine, projected onto a plane of two geometric coordinates. The corner cutting shortens the tunneling path by keeping the heavy nitrogen atoms fixed and focusing the movement onto the light hydrogen atom.

the rate becomes more sensitive to g_{tol} . At 20 images, the rate is already converged if $g_{\text{tol}} = 10^{-6}$ is used for malonaldehyde and ammonium and methylamine.

The remaining error at nonperfect convergence can be due to three sources, $k_{\text{inst}} = k_{S_0} \cdot k_{\text{harm}} \cdot k_{S_E}$, as apparent from eq 7 with S_0 (entering the rate as $k_{S_0} = (S_0/(2\pi\hbar))^{1/2}$), S_E (entering the rate as $k_{S_E} = \exp(-S_E/\hbar)$), and all terms covering the quadratic expansion of the potential around the stationary paths: $k_{\text{harm}} = (1/Q_{\text{RS}})(P'/\beta\hbar)(1/(\prod_i \lambda_i))^{1/2}$. The latter causes the largest effect. This is somewhat surprising as S_E enters the rate exponentially. The eigenvalues λ_i of the Hessian only enter under the square root. However, numerical noise in the Hessian affects most or all eigenvalues, which together apparently make a noticeable contribution. Thus, it seems important to calculate the Hessians of the images with high accuracy.

Finally, we discuss a property of the instanton path itself rather than the optimization of it, namely corner cutting. Tunneling is more efficient the thinner the barrier is. The relevant width of the barrier is the width in mass-weighted (or any iso-inertial) coordinates. Thus, movement of heavy atoms reduces the tunneling rate more than movement of light atoms. This results in an instanton path which, in many cases, deviates significantly from the MEP, an effect known as corner cutting.²⁷ An example is demonstrated in Figure 5. The figure shows the N–N distance plotted against one N–H distance of the ammonium and methylamine system used in this work. All numbers are of course calculated with semiempirical methods of limited accuracy. However, the qualitative conclusions also hold with more accurate methods. On the MEP, the nitrogen atoms approach each other significantly (shortening of the N–N distance by 0.14 Å) before the hydrogen atom is transferred. The movement of the heavy atoms is more and more avoided the more important tunneling becomes. At $T = 100$ K, the N–N distance decreases by only 0.06 Å. This shortening of the effective tunneling path comes at the expense of a higher energy of some images along the instanton compared to the classical TS. However, it still results in a larger tunneling rate than tunneling along the MEP.

VI. CONCLUSION

We presented a comparison of four newly implemented methods to search for instantons as saddle points on the Euclidean action surface spanned by closed Feynman paths. NR turned out to be

consistently efficient and stable. The methods were tested at various temperatures for three different small chemical systems with tunneling hydrogen atoms. We applied semiempirical methods to obtain the quantum chemical potential energies and forces. However, preliminary tests showed that the same conclusions are reached for density functional theory (DFT) calculations.

We will continue to apply these algorithms to chemically relevant systems using more accurate quantum chemical methods. Preliminary results show that the methods are also applicable to enzymatic systems with many degrees of freedom described by a QM/MM energy expression.⁹⁹

The results of this work show that instantons can be optimized rather efficiently. The geometry optimization generally requires less energy and gradient evaluations than subsequent Hessian calculations at each image, at least if the latter are done by finite-difference calculations of gradients using the two-point formula. Efficiency in the Hessian calculations can of course be gained by using analytic Hessians instead. Additionally, it may be of interest to find methods to calculate the rate using the Hessian at fewer points along the instanton. Accuracy can be improved by calculating the energies of the points along the instanton path with a higher level of theory, while performing the optimization and the Hessian calculations at the lower level. An additional challenge in instanton theory is the issue that at lower temperature, more and more images of the discretized path tend to accumulate at one end of the instanton path (the one with the smaller slope of the energy along the path). Up to some point, this can be accounted for by using more and more images. However, to apply instanton theory to really low temperature, or to calculate the limit at $T \rightarrow 0$, different schemes will have to be developed.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: kaestner@theochem.uni-stuttgart.de.

■ ACKNOWLEDGMENT

J.B.R. and J.K. thank the German Research Foundation (DFG) for financial support of the project within the Cluster of Excellence in Simulation Technology (EXC 310/1) at the University of Stuttgart. T.P.M.G. thanks The Netherlands Organization for Scientific Research (NWO) for a VENI-fellowship (700.58.404). The work was stimulated by an HPC-Europa2 project awarded to T.P.M.G. to visit J.K. Prof. H. Jónsson and Dr. A. Arnaldsson are acknowledged for helpful discussions.

■ REFERENCES

- (1) Gamow, G. *Z. Phys.* **1928**, 204.
- (2) Eckart, C. *Phys. Rev.* **1930**, 35, 1303.
- (3) Cazaux, S.; Caselli, P.; Cobut, V.; Le Bourlot, J. *Astron. Astrophys.* **2008**, 483, 495–508.
- (4) Goumans, T. P. M.; Kästner, J. *Angew. Chem., Int. Ed.* **2010**, 49, 7350–7352.
- (5) Goumans, T. P. M.; Kästner, J. *Angew. Chem.* **2010**, 122, 7508–7511.
- (6) Liang, Z.-X.; Klinman, J. P. *Curr. Opin. Struct. Biol.* **2004**, 14, 648.
- (7) Bandaria, J. N.; Cheatum, C.; Kohen, A. *J. Am. Chem. Soc.* **2009**, 131, 10151–10155.
- (8) Masgrau, L.; Roujeinikova, A.; Johannissen, L. O.; Hothi, P.; Basran, J.; Ranaghan, K. E.; Mulholland, A. J.; Sutcliffe, M. J.; Scrutton, N. S.; Leys, D. *Science* **2006**, 312, 237–241.

- (9) Yoon, M.; Song, H.; Håkansson, K.; Marsh, E. N. G. *Biochemistry* **2010**, *49*, 3168–3173.
- (10) Grant, K. L.; Klinman, J. P. *Biochemistry* **1989**, *28*, 6597.
- (11) Knapp, M. J.; Klinman, J. P. *Eur. J. Biochem.* **2002**, *269*, 3113–3121.
- (12) Antoniou, D.; Caratzoulas, S.; Kalyanaraman, C.; Mincer, J. S.; Schwartz, S. D. *Eur. J. Biochem.* **2002**, *269*, 3103–3112.
- (13) Klinman, J. P. *J. Biol. Chem.* **2006**, *281*, 3013.
- (14) Sutcliffe, M. J.; Scrutton, N. S. *Eur. J. Biochem.* **2002**, *269*, 3096.
- (15) Masgrau, L.; Roujeinikova, A.; Johannissen, L. O.; Hothi, P.; Basran, J.; Ranaghan, K. E.; Mulholland, A. J.; Sutcliffe, M. J.; Scrutton, N. S.; Leys, D. *Science* **2006**, *312*, 237.
- (16) Wang, M.; Lu, Z.; Yang, W. *J. Chem. Phys.* **2006**, *124*, 124516.
- (17) Bothma, J. P.; Gilmore, J. B.; McKenzie, R. H. *New J. Phys.* **2010**, *12*, 055002.
- (18) Arrhenius, S. *Z. Phys. Chem. (Leipzig)* **1889**, *4*, 226.
- (19) Eyring, H.; Polanyi, M. *Z. Phys. Chem., Abt. B* **1931**, 279.
- (20) Hänggi, P.; Talkner, P.; Borkovec, M. *Rev. Mod. Phys.* **1990**, *62*, 251.
- (21) Pollak, E.; Talkner, P. *Chaos* **2005**, *15*, 026116.
- (22) Wigner, E. P. *Z. Phys. Chem.* **1932**, *15*, 203.
- (23) Bell, R. The application of tunnel corrections in chemical kinetics. In *The tunnel effect in chemistry*; 1st ed.; Chapman and Hall: London, 1980; Vol. 1, pp 51–140.
- (24) Gillan, M. *J. Phys. C: Solid State Phys.* **1987**, *20*, 3621.
- (25) Wigner, E. *Trans. Faraday Soc.* **1938**, *34*, 29.
- (26) Kuppermann, A.; Truhlar, D. *J. Am. Chem. Soc.* **1971**, *93*, 1840–1851.
- (27) Marcus, R. A.; Coltrin, M. E. *J. Chem. Phys.* **1977**, *67*, 2609.
- (28) Skodje, R.; Truhlar, D.; Garrett, B. *J. Phys. Chem.* **1981**, *85*, 3019–3023.
- (29) Garrett, B. C.; Truhlar, D. G.; Wagner, A. F.; Thom H. Dunning, J. *J. Chem. Phys.* **1983**, *78*, 4400–4413.
- (30) Garrett, B. C.; Abusalbi, N.; Kouri, D. J.; Truhlar, D. G. *J. Chem. Phys.* **1985**, *83*, 2252–2258.
- (31) Liu, Y. P.; Lu, D. H.; Gonzalez-Lafont, A.; Truhlar, D. G.; Garrett, B. C. *J. Am. Chem. Soc.* **1993**, *115*, 7806–7817.
- (32) Meana-Pañeda, R.; Truhlar, D. G.; Fernández-Ramos, A. *J. Chem. Theory Comput.* **2010**, *6*, 6–17.
- (33) Langer, J. S. *Ann. Phys. (N.Y.)* **1967**, *41*, 108.
- (34) Miller, W. H. *J. Chem. Phys.* **1975**, *62*, 1899.
- (35) Coleman, S. *Phys. Rev. D: Part. Fields* **1977**, *15*, 2929.
- (36) Callan, C. G., Jr.; Coleman, S. *Phys. Rev. D: Part. Fields* **1977**, *16*, 1762.
- (37) Gildener, E.; Patrascioiu, A. *Phys. Rev. D: Part. Fields* **1977**, 423.
- (38) Feynman, R. P. *Rev. Mod. Phys.* **1948**, *20*, 367.
- (39) Mills, G.; G. K. Schenter, D. M.; Jónsson, H. RAW Quantum Transition State Theory. In *Classical and Quantum Dynamics in Condensed Phase Simulations*; 1st ed.; Berne, B. J., Ciccotti, G., Coker, D. F., Eds.; World Scientific Press: Hackensack, NJ, 1998; Vol. 1, pp 405–421.
- (40) Miller, W. H.; Zhao, Y.; Ceotto, M.; Yang, S. *J. Chem. Phys.* **2003**, *119*, 1329.
- (41) Vaníček, J.; Miller, W. H.; Castillo, J. F.; Aoiz, F. J. *J. Chem. Phys.* **2005**, *123*, 054108.
- (42) Chapman, S.; Garrett, B. C.; Miller, W. H. *J. Chem. Phys.* **1975**, *63*, 2710.
- (43) Mills, G.; Jónsson, H. *Phys. Rev. Lett.* **1994**, *72*, 1124.
- (44) Mills, G.; Jónsson, H.; Schenter, G. K. *Surf. Sci.* **1995**, *324*, 305–337.
- (45) Mills, G.; Schenter, G. K.; Makarov, D. E.; Jónsson, H. *Chem. Phys. Lett.* **1997**, *278*, 91.
- (46) Siebrand, W.; Smedarchina, Z.; Zgierski, M. Z.; Fernández-Ramos, A. *Int. Rev. Phys. Chem.* **1999**, *18*, 5.
- (47) Smedarchina, Z.; Siebrand, W.; Fernández-Ramos, A.; Cui, Q. *J. Am. Chem. Soc.* **2003**, *125*, 243–251.
- (48) Qian, T.; Ren, W.; Shi, J.; E, W.; Shen, P. *Physica A* **2007**, *379*, 491.
- (49) Andersson, S.; Nyman, G.; Arnaldsson, A.; Manthe, U.; Jónsson, H. *J. Phys. Chem. A* **2009**, *113*, 4468.
- (50) Goumans, T. P. M.; Andersson, S. *Mon. Not. R. Astron. Soc.* **2010**, *406*, 2213–2217.
- (51) Manthe, U.; Meyer, H.-D.; Cederbaum, L. S. *J. Chem. Phys.* **1992**, *97*, 3199–3213.
- (52) Meyer, H.-D.; Manthe, U.; Cederbaum, L. *Chem. Phys. Lett.* **1990**, *165*, 73–78.
- (53) Padmanaban, R.; Nest, M. *Chem. Phys. Lett.* **2008**, *463*, 263–266.
- (54) Hammer, T.; Coutinho-Neto, M. D.; Viel, A.; Manthe, U. *J. Chem. Phys.* **2009**, *131*, 224109.
- (55) Hansen, N. F.; Andersen, H. C. *J. Phys. Chem.* **1996**, 1137.
- (56) Cheney, B. G.; Andersen, H. C. *J. Chem. Phys.* **2003**, *118*, 9542–9551.
- (57) Cao, J.; Voth, G. A. *J. Chem. Phys.* **1994**, 6157.
- (58) Voth, G. A. *Adv. Chem. Phys.* **1996**, 135.
- (59) Pollak, E.; Liao, J.-L. *J. Chem. Phys.* **1998**, 2733.
- (60) Richardson, J. O.; Althorpe, S. C. *J. Chem. Phys.* **2009**, *131*, 214106.
- (61) Craig, I. R.; Manolopoulos, D. E. *J. Chem. Phys.* **2005**, *122*, 084106.
- (62) Voth, G. A.; Chandler, D.; Miller, W. H. *J. Chem. Phys.* **1989**, 7749.
- (63) Voth, G. A. *J. Phys. Chem.* **1993**, 8365.
- (64) Simon, B. *Functional integration and quantum physics*; 2nd ed.; Academic Press: New York, 1979; Vol. 1;
- (65) Arnaldsson, A. Ph.D. thesis, University of Washington: Seattle, WA, 2007.
- (66) Mestela, B.; Percival, I. *Physica D* **1987**, *24*, 172.
- (67) Baranger, M.; Davies, K. T. R.; Mahoney, J. H. *Ann. Phys. (N.Y.)* **1988**, *186*, 95.
- (68) Grobgeld, D.; Pollak, E.; Zakrzewski, J. *Physica D* **1992**, *56*, 368.
- (69) Marcinek, R.; Pollak, E. *J. Chem. Phys.* **1994**, *100*, 5894.
- (70) Messina, M.; Schenter, G. K.; Garrett, B. C. *J. Chem. Phys.* **1995**, *103*, 3430.
- (71) Kästner, J.; Carr, J. M.; Keal, T. W.; Thiel, W.; Wander, A.; Sherwood, P. *J. Phys. Chem. A* **2009**, *113*, 11856.
- (72) Kästner, J.; Keal, T. W.; Sherwood, P. DL-FIND; CCPForge: Daresbury Lab, United Kingdom; <http://ccpforge.cse.rl.ac.uk/gf/project/dl-find/>. Accessed November 15, 2010.
- (73) Sherwood, P.; de Vries, A. H.; Guest, M. F.; Schreckenbach, G.; Catlow, C. R. A.; French, S. A.; Sokol, A. A.; Bromley, S. T.; Thiel, W.; Turner, A. J.; Billeter, S.; Terstegen, F.; Thiel, S.; Kendrick, J.; Rogers, S. C.; Casci, J.; Watson, M.; King, F.; Karlsen, E.; Sjøvoll, M.; Fahmi, A.; Schäfer, A.; Lennartz, C. *J. Mol. Struct. (THEOCHEM)* **2003**, *632*, 1.
- (74) ChemShell, a Computational Chemistry Shell; Computational Science and Engineering Department: Daresbury Lab, United Kingdom; <http://www.chemshell.org>. (Accessed November 15, 2010)
- (75) Henkelman, G.; Jónsson, H. *J. Chem. Phys.* **1999**, *111*, 7010.
- (76) Olsen, R. A.; Kroes, G. J.; Henkelman, G.; Arnaldsson, A.; Jónsson, H. *J. Chem. Phys.* **2004**, *121*, 9776.
- (77) Heyden, A.; Bell, A. T.; Keil, F. J. *J. Chem. Phys.* **2005**, *123*, 224101.
- (78) Kästner, J.; Sherwood, P. *J. Chem. Phys.* **2008**, *128*, 014106.
- (79) Liu, D. C.; Nocedal, J. *Math. Program* **1989**, *45*, 503.
- (80) Nocedal, J. *Math. Comp.* **1980**, *35*, 773.
- (81) Broyden, C. G. *IMA J. Appl. Math.* **1970**, *6*, 76.
- (82) Fletcher, R. *Comp. J.* **1970**, *13*, 317.
- (83) Goldfarb, D. *Math. Comp.* **1970**, *24*, 23.
- (84) Shanno, D. F. *Math. Comp.* **1970**, *24*, 647.
- (85) Lanczos, C. *J. Res. Nat. Bur. Stand* **1951**, *45*, 255–282.
- (86) Malek, R.; Mousseau, N. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* **2000**, *62*, 7723.
- (87) Cerjan, C. J.; Miller, W. H. *J. Chem. Phys.* **1981**, *75*, 2800.
- (88) Simons, J.; Jørgensen, P.; Taylor, H.; Ozment, J. *J. Phys. Chem.* **1983**, *87*, 2745.

- (89) Banerjee, A.; Adams, N.; Simons, J.; Shepard, R. J. *Phys. Chem.* **1985**, *89*, 52–57.
- (90) Baker, J. J. *Comput. Chem.* **1986**, *7*, 385–395.
- (91) Bofill, J. M. J. *Comput. Chem.* **1994**, *15*, 1.
- (92) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209.
- (93) Wong, K. F.; Sonnenberg, J. L.; Paesani, F.; Yamamoto, T.; Vaníček, J.; Zhang, W.; Schlegel, H. B.; Case, D. A.; Cheatham, T. E., III; Miller, W. H.; Voth, G. A. *J. Chem. Theory Comput* **2010**, *6*, 2566.
- (94) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
- (95) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787.
- (96) Thiel, W. *MNDO99*, v. 6.1; Max-Planck-Institut für Kohlenforschung; Mülheim an der Ruhr, Germany, 2004.
- (97) Humphrey, W.; Dalke, A.; Schulten, K. *J. Molec. Graphics* **1996**, *14*, 33.
- (98) Powell, M. J. D. *Math. Prog.* **1971**, *26*, 1.
- (99) Rommel, J. B.; Kästner, J. to be published.

An Integrated Protocol for the Accurate Calculation of Magnetic Interactions in Organic Magnets

Vincenzo Barone,[†] Ivo Cacelli,[‡] Alessandro Ferretti,[§] Susanna Monti,[§] and Giacomo Prampolini^{*,†}

[†]Scuola Normale Superiore, piazza dei Cavalieri 7, I-56126 Pisa, Italy

[‡]Dipartimento di Chimica e Chimica Industriale, Università degli Studi di Pisa, via Risorgimento 35, I-56126 Pisa, Italy

[§]Istituto di Chimica dei Composti OrganoMetallici (ICCOM-CNR), Area della Ricerca, via G. Moruzzi 1, I-56124 Pisa, Italy

ABSTRACT: A new, fast, and efficient computational protocol for the accurate calculation of singlet–triplet magnetic splittings in organic diradicals is tested and validated. This procedure essentially consists of three steps: the adoption of modified virtual orbitals (MVO) and a mixed variational–perturbational approach (CSPA) are now combined with a third method that exploits the reduction of the configurational space dimensions achieved by fragmentation/localization criteria. This innovative approach is successfully tested on four different substituted *m*-phenylene bis(*tert*-butyl) nitroxides, which show paramagnetic behavior, by computing singlet–triplet energy gaps and comparing them with their experimental counterparts.

1. INTRODUCTION

The multidisciplinary interest in organic di- and poliradicals^{1–7} stems from the many technological applications in which they are involved, ranging from spintronics (as sensing, memory, or switching devices^{1,2,8–12}) to soft matter (as contrast agents for magnetic resonance imaging, spin labels, or mediators for controlled radical polymerization^{13–16}). Among diradicals, many substituted *m*-phenylene-bridged nitroxides have been synthesized and proposed as stable paramagnetic materials by several groups.^{12,16–23} Besides being prototypes of organic paramagnetic compounds, these species may be important as basic units of polyradical systems,^{16,17,23} thus being potentially appealing in the design of magnetic devices or biologically relevant probes. A key feature in molecular magnetism is clearly the singlet–triplet energy gap, ΔE_{ST} . This quantity is not directly accessible by experiments, but it can be derived through a numerical fit^{12,16,17,19–24} from the temperature dependence of the magnetic susceptibility (χ) obtained by superconducting quantum interference device (SQUID) magnetometry. Unfortunately, at variance with antiferromagnetic interactions, the sensitivity of this method to the values of ΔE_{ST} for ferromagnetic materials in solution is rather small^{1,18,22,24} and often allows only for a determination of a lower limit of ΔE_{ST} ,^{22,24} rather than its actual value.

From a computational point of view, the calculation of ΔE_{ST} is also challenging, due to the difficulties in defining a protocol possessing at the same time the required characteristics of accuracy and feasibility. Furthermore, in diradicals where the magnetic moieties are bridged by an unsaturated fragment, the coupling between the spins of the two unpaired electrons is strongly affected by the presence of the bridge itself, and any attempt to calculate ΔE_{ST} should take into account its presence explicitly. On the one hand, computational convenience would suggest resorting to Density Functional Theory (DFT), adopting the Noodleman broken-symmetry method^{25,26} to calculate the magnetic splitting. This method, widely applied to inorganic and organic species,^{27–33} has been recently reviewed by several authors.^{33–35} However, its theoretical foundation and robustness are rather questionable. On the

other hand, the more rigorous post-Hartree–Fock methods^{36–54} become rapidly unfeasible with the increase of the molecular dimensions. Although computational feasibility would suggest employing second order perturbation methods, the neglect of the interaction terms between the perturbers invalidates the accuracy of the results and makes variational approaches much more reliable. Among these, the Difference Dedicated Configuration Interaction methods (DDCI and DDCI2)^{37,38,42} allow for a remarkable reduction of the full CI space dimensions. Indeed, the DDCI2 scheme⁴² is certainly less expensive, and it was recently shown to yield reliable results for small organic diradicals.^{55,56} Nonetheless, some inaccuracies were reported in other cases,⁵⁷ where better performances were obtained, for paramagnetic molecules, with the adoption of the complete DDCI approach. However, the dimensions of the DDCI configurational space increase very rapidly with the molecular size, and the direct application of the DDCI scheme to large molecules remains unfeasible.

To circumvent this problem, a multilevel strategy has been recently^{56,58,59} developed in our group. The new methodology allows us to sensibly reduce the dimensions of the DDCI space with a negligible loss of accuracy. Essentially, the approach consists of three different steps. First, molecular orbitals are localized onto different moieties, and those belonging to the fragments outside the magnetic+bridge moiety are neglected.⁵⁶ In a second step, computational advantages can also be obtained by the use of modified virtual orbitals (MVOs), built adding extra charges to the magnetic sites.⁵⁸ Indeed, the size of the DDCI space can be further reduced by neglecting excitations to high energy virtual orbitals, and it has been shown⁵⁸ that faster convergence can be achieved if the virtual orbitals to be excluded are chosen among the modified rather than the canonical ones. Finally, the use of the Complementary Space Perturbative Approach (CSPA)⁵⁹ allows for a further reduction of the dimensions of the variational DDCI configurational space. In fact, only a small fraction of the

Received: October 21, 2010

Published: January 19, 2011

Table 1. Configurational Classes Considered within the DDCI Scheme^a

class	description	class ⁴²
N.2.0	the primary four-dimensional space (includes kinetic exchange)	
N.1.1	single excitations from magnetic orbitals	1p
N.0.2	double excitations from magnetic orbitals	2p
N-1.3.0	single excitations from core to magnetic orbitals (related to superexchange)	1h
N-1.2.1	single excitations from core to unoccupied orbitals (include spin polarization)	1h + 1p
N-2.4.2	double excitations from core to magnetic orbitals	2h
N-1.1.2	simultaneous excitations from core and magnetic to unoccupied orbitals	1h + 2p
N-2.3.1	double excitations from core to magnetic and unoccupied orbitals	2h + 1p

^a Each class is labeled in terms of three numbers (first column) referring to core, magnetic, and unoccupied molecular orbitals. In the last column, the corresponding notation proposed by Calzado et al.⁴² is reported. The configurational classes reported in the last two rows are not included in a DDCI2 scheme.

MVOs is considered in the variational treatment, whereas the remaining part is handled through a Möller–Plesset perturbative approach. This protocol, recently tested on a diaryl-nitroxide diradical,^{22,60} is here proposed and validated for a series of four substituted *m*-phenylene-bridged *tert*-butyl-nitroxides.^{12,16,17,21}

The paper is organized as follows: in the next section, employed methods are briefly reviewed and computational details are given. Results are presented and discussed in section 3, while main conclusions are drawn in the last section.

2. METHODS AND COMPUTATIONAL DETAILS

2.1. Methods. *2.1.1. DDCI.* The minimal description of magnetic interactions in a diradical system can be given within the four-dimensional configurational space obtained placing the two unpaired electrons in two nearly degenerate molecular orbitals (MOs) $|\phi_g\rangle$ and $|\phi_u\rangle$. To obtain a more transparent physical description, one may localize the canonical MOs through a unitary transformation into orbitals $|\phi_a\rangle$ and $|\phi_b\rangle$, which can in turn be used to generate the four basic configurations:

$$\begin{aligned}
 |^1\Psi_A\rangle &= \frac{1}{\sqrt{2}}(|\dots\phi_a\bar{\phi}_b\rangle + |\dots\phi_b\bar{\phi}_a\rangle) \\
 |^1\Psi_B\rangle &= |\dots\phi_a\bar{\phi}_a\rangle \\
 |^1\Psi_C\rangle &= |\dots\phi_b\bar{\phi}_b\rangle \\
 |^3\Psi\rangle &= \frac{1}{\sqrt{2}}(|\dots\phi_a\bar{\phi}_b\rangle - |\dots\phi_b\bar{\phi}_a\rangle)
 \end{aligned} \quad (1)$$

Among the resulting states, $|^1\Psi_A\rangle$ and $|^3\Psi\rangle$ provide the main contribution to the singlet–triplet energy gap in weakly coupled diradicals, since the ionic configurations ($|^1\Psi_B\rangle$ and $|^1\Psi_C\rangle$) lie at higher energies. Clearly, this basic configurational space alone is insufficient to provide a reliable estimate of the ΔE_{ST} ($E_S - E_T$) gap, and other classes of configurations should be considered. In the DDCI and DDCI2^{37,38,42} approaches, only those classes that are expected to give a non-negligible contribution to ΔE_{ST} are selected and employed during the calculation. Thus, besides the four configurations reported in eq 1, DDCI includes all determinants arising from single and double excitations from the magnetic to the virtual orbitals, together with double excitations from the magnetic and core orbitals to the virtual orbitals, but involving one single core-to-virtual excitation. All of these classes are reported in Table 1, where differences between the DDCI and DDCI2 approaches are underlined in the last two rows. Further details can be found in refs 56 and 58–60 and references therein.

2.1.2. Three-Step Procedure. Despite this selection, the size of DDCI configurational space increases still rapidly with the molecular dimensions, making a full variational approach with the standard DDCI scheme unfeasible for large molecules. The first step of the proposed protocol is based on fragment localization (step I): MOs are first localized onto specific moieties through the Pipek–Mezey method,⁶¹ and only those belonging to the fragments involved in the magnetic interaction (magnetic+bridge moieties) are retained for the DDCI calculation.^{55,56}

Once the virtual orbitals (VOs) localized on the “external” fragments have been removed, the Modified Virtual Orbital (MVO) scheme⁵⁸ is applied (step II) on the remaining VOs space. MVOs are determined by the following eigenvalue equation:

$$(\hat{F} + \hat{V})\phi_\mu = \epsilon_\mu\phi_\mu \quad (2)$$

where \hat{F} is the Fock operator for the restricted triplet state and \hat{V} is a supplementary nuclear potential obtained placing some extra charges q_α at the positions R_α i.e.,

$$\hat{V} = - \sum_\alpha \frac{q_\alpha}{|R_\alpha - r|}$$

where the summation is restricted to those atoms bearing the magnetic orbitals. The effective optimal supplementary charges were determined in previous work,^{58,59} on the basis of the convergence properties of the resulting MVOs in the DDCI2 or DDCI scheme, resulting in $q_\alpha = 1$ or 2, respectively. More details about the method can be found in the original paper.⁵⁸ Here, it is worth highlighting that the computed MVOs allow significant advantages, in that those having the lowest ϵ_μ 's (see eq 2) are the most involved in the spin–spin coupling, whereas decreasing effects are observed with the increase of the orbital energy.

Finally, in step III, the CSPA approach⁵⁹ is applied. In the hypothesis that the interaction terms between the perturbers play a major role in ΔE_{ST} only for those configurations which involve excitations to the lowest MVOs, the variational calculation is performed only for a fraction of MVOs, while the rest of the configurations is handled by a multireference perturbative treatment, using the so-called barycentric Möller–Plesset partition. In this way, only a small portion of the MVOs (i.e., those at low energy, thus more involved in the magnetic coupling) is active for the variational calculation, but all of them are considered for the (cheaper) perturbative calculation. A pictorial representation of steps I–III is sketched in Figure 1.

2.2. Computational Details. All molecular structures have been optimized in their triplet state at the DFT-UB3LYP/cc-pvdz

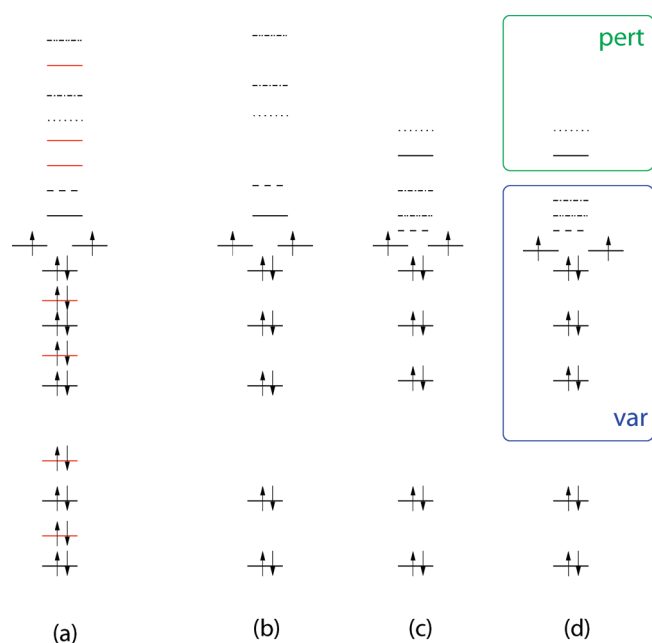


Figure 1. Scheme of the multilevel protocol. (a) All MOs are first localized into active (black) and external (red) fragments. (b) Step I: MOs on external fragments are excluded from DDCI calculations. (c) Step II: VOs are modified through MVO approach. (d) Step III: Only a fraction of orbitals is treated variationally (blue window), while other MVOs (green window) are taken into account through a perturbative approach (CSPA).

level of theory using the Gaussian package.⁶² Unless otherwise stated, all torsional potential energy surfaces have been computed by optimizing all geometries with no symmetry restrictions but the investigated torsional angles. The canonical molecular orbitals were obtained by a ROHF calculation for the triplet state, using the GAMESS code,⁶³ with the 6-31G(d) basis set. The MVOs were obtained using the Fortran program QUIOLA coded by the authors, which interfaces the GAMESS output files with the routine for the transformation of the integrals from the atomic to the molecular basis set.⁵⁵ The CI calculations were performed using the CIPSI program,^{64–66} which has been rewritten⁵⁹ in order to improve efficiency and to manage the CSPA calculations. In all CI calculations, all 1s core orbitals were always kept inactive with occupation number 2 in all configurations.

3. RESULTS AND DISCUSSION

The singlet–triplet energy gap ΔE_{ST} was computed for a group of substituted *m*-phenylene bis(*tert*-butylnitroxide) diradicals, namely, *m*-phenylene bis(*tert*-butyl) nitroxide¹⁷ (compound 1), 4,6-bis(trifluoromethyl)-*N,N*,*N*A-di-*tert*-butyl-1,3-phenylenebis(aminoxyl)²¹ (2), biphenyl-3,5-diyl bis(*tert*-butyl-nitroxide)^{12,19,20} (3), and a polyethylene glycol (PEG) functionalized (pegylated) bis(aminoxyl) diradical¹⁶ (4), whose structures and measured singlet–triplet energy gaps are reported in Figure 2 and in Table 2, respectively. Besides testing the proposed integrated strategy, our aim is also to investigate the contributions to ΔE_{ST} of the different substituents. The following discussion has been thus separated for each of the considered species.

3.1. *m*-Phenylene Bis(*tert*-butyl) Nitroxide. Diradical 1 is the smallest prototype of *m*-phenylene bis(*tert*-butyl nitroxide) diradicals, thus constituting an ideal candidate to test and validate

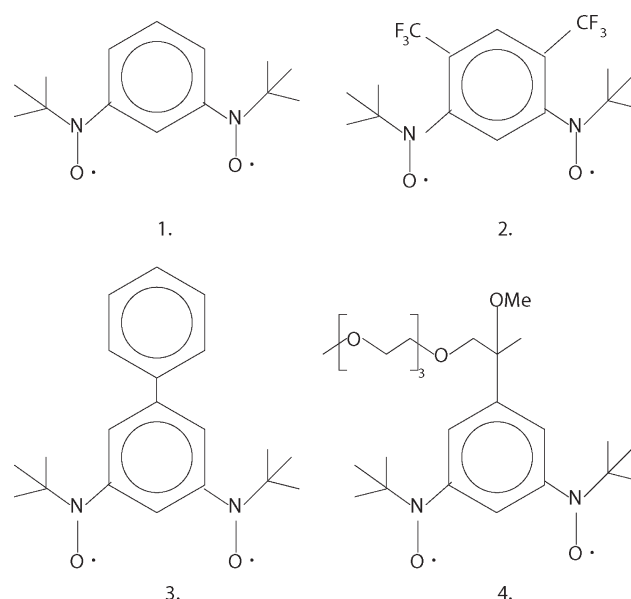


Figure 2. Substituted nitroxide diradicals studied in the present work. (1) *m*-Phenylene bis(*tert*-butyl nitroxide);¹⁷ (2) 4,6-bis(trifluoromethyl)-*N,N*,*N*A-di-*tert*-butyl-1,3-phenylene bis(aminoxyl); (3) biphenyl-3,5-diyl bis(*tert*-butyl nitroxide);^{12,19,20} (4) pegylated bis(aminoxyl) diradical.¹⁶

Table 2. Experimental Estimates of the Singlet–Triplet Energy Gap ΔE_{ST} , As Obtained from the Best Fit Theoretical Curves of SQUID Data for Compounds 1–4 of Figure 2

diradical	ΔE_{ST} (K)	ΔE_{ST} (cm ⁻¹)	ref
1	>300	>210	17
2	~80	~55	21
3	>350	>241	20
4	~650	~450	16

the computational procedure proposed in the present work. In fact, despite its dimensions not allowing a complete DDCI variational treatment, diradical 1 can be handled within the CSPA approach without resorting to fragmentation schemes. In particular, in view of the calculations to be performed on the next larger compounds, it could be computationally advantageous to substitute the *tert*-butyl moieties with methyl groups, in the hypothesis that their contribution to ΔE_{ST} can be neglected. With this aim, a *m*-phenylene bis(methyl nitroxide) diradical was optimized in two different geometries (see Figure 3). In the first case (1a), a full optimization was performed directly on the small methyl-nitroxide diradical, whereas the second geometry (1b) was obtained by substituting a hydrogen atom with each of the three methyl groups of a fully optimized geometry of the whole diradical 1. In the latter case, only the hydrogen atoms were reoptimized after substitution. Clearly, the former route is less expensive, but it does not take into account small distortions of bond lengths and angles introduced by the steric encumbrance of the *tert*-butyl groups. In Table 3, some selected coordinates are compared for the two geometries. It appears that only very small differences can be found, among which the most noticeable is the lengthening of the N2–C3 distance in diradical 1b, caused, by the repulsion between the bridge and the *tert*-butyl groups. Since ΔE_{ST} is an observable very sensitive to the chemical details of the bridge, it could be of some interest to investigate if the (very small) increase of the magnetic site–bridge distance contributes

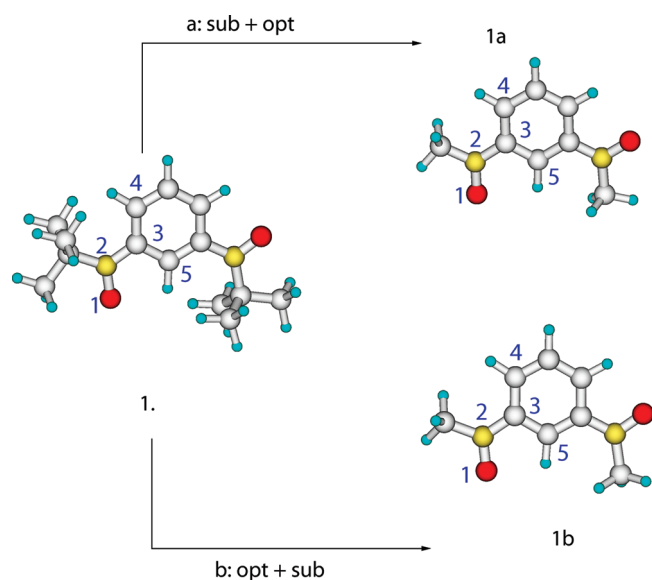


Figure 3. Two different routes to obtain a model of diradical **1** (see Figure 2 suitable for DDCI calculations). (a) *tert*-Butyl groups are first substituted with methyls (sub), and thereafter a full optimization (opt) is performed. (b) The substitution is performed on the optimized geometry of the whole diradical.

Table 3. Selected Internal Coordinates for the Two Optimized Geometries of Model Diradicals **1a** and **1b**^a

coordinate	1a	1b
O1–N2	1.276 Å	1.278 Å
N2–C3	1.407 Å	1.423 Å
C3–C4	1.409 Å	1.410 Å
O1–C5	2.721 Å	2.637 Å
O1–N2–C3	120.6°	117.6°
N2–C3–C4	121.6°	124.6°
N2–C4–C5	118.0°	116.0°
O1–N2–C3–C4	180.0°	180.0°

^a Atom numbers refer to Figure 3.

to diminishing the interaction between the magnetic sites. Furthermore, the capability of the proposed procedure to meet chemical intuition and reproduce such small effects could be considered a further proof of the protocol reliability.

Singlet–triplet energy gaps were computed combining MVO and CSPA approaches, for the whole diradical **1** and the smaller model in the two geometries **1a** and **1b**. All results are reported in Table 4, together with the variational space dimensions (VSD) and the CPU time employed, and in Figure 4, as a function of the ratio (Var %) between the MVO active to the variational treatment in the CSPA approach (VAVO), and the total number of MVOs.

The calculations performed on the whole diradical were not able to consider a DDCI variational space larger than 30% of that subtended by all VOs, due to the excessive requests of both memory and CPU time. However, it seems from the left panel of Figure 4 that the reported values are near to converge around a value of ~ 290 cm⁻¹. As far as the two smaller models are concerned, it is evident that neglecting the small differences between geometries **1a** and **1b** does have some consequence on the final estimate of ΔE_{ST} , the former being larger (326 cm⁻¹) than the

one expected for the whole original molecule. Conversely, if the geometrical effects of the presence of the *tert*-butyl groups are taken into account by adopting the b model route (see Figure 3), the results obtained are very close (286.1 cm⁻¹) to those found for the whole diradical **1**. It also appears from Table 4 that the CPU time savings amply offsets the computational burden of optimization route b. Furthermore, it is worth stressing that the adoption of the CSPA approximation allows us to handle such calculations at a reasonable computational cost, exploiting the improved convergence rate of the CSPA corrected results. Indeed, by looking at the right panel of Figure 4, where the full variational limit is reported for the model compound (in **1b** geometry), one can see that only about 30% of the MVO can be treated at a variational level, without losing much accuracy in the estimate of the final results. This allows for a savings of almost a factor 20 on CPU time. It may be worth stressing that the full variational limit, obtained with a VAVO/MVO ratio of 100%, is exactly the same value that could be obtained by employing all canonical virtual orbitals in a completely variational calculation, as reported in the original paper where MVOs were first proposed.⁵⁸

Finally, our computational results can be compared with experimental estimates reported for the whole diradical **1**. Notwithstanding, **1** is not fully persistent in solution (it isomerizes into an aminquinone imine N-oxide in a few hours). The dependence of its magnetic susceptibility χ from the temperature was measured, and an estimate of $\Delta E_{ST} > 300$ K (210 cm⁻¹) was given by the authors,¹⁷ in fair agreement with the present theoretical findings (~ 420 K).

3.2. 4,6-Trifluoromethyl, *m*-Phenylene Bis(*tert*-butyl)-nitroxide. Substitution of the of two hydrogens at positions 4 and 6 in the phenylene bridge with two trifluoromethyl groups leads to the formation of diradical **2** (see Figure 2). In this case, the modeling route based on replacing each *tert*-butyl group on the N atoms with a methyl moiety must be handled with care, since its effects on the minimum geometry are expected to be remarkably larger than those found in the previous case, due to the steric interaction between the *tert*-butyl groups and the trifluoromethyl substituents of the bridge. In fact, as can be seen from Figure 5, where the computed torsional energy surfaces are reported for *tert*-butyl and the methyl nitroxides, it appears that the presence of the larger substituent on N atoms dramatically alters the PES profile, resulting in a nonplanar minimum geometry, with the dihedral angles Φ_1 and Φ_2 placed at 65° and 125°, respectively. The dependence of the singlet–triplet gap in aromatic bridged nitroxides from such torsional dihedrals is well-known, both experimentally^{18,21,67,68} and theoretically.^{30,57} In particular it has been found that distortions from planarity in *meta* substituted aromatic bridges sensibly diminishes the stability of the triplet state, eventually leading⁶⁸ in some cases to a more stable singlet, i.e., yielding antiferromagnetic behavior. For this reason, the modeling route b, discussed in the previous section (see Figure 3), has been followed also in this case, substituting with hydrogen atoms the methyl groups of the *tert*-butyl moieties in the minimum energy nonplanar geometry, found for the whole diradical **2** (see panel a of Figure 5).

The three-step procedure has been applied on this latter configuration. First, MOs have been localized onto three fragments, namely, the *m*-phenylene bis(methyl)-nitroxide and the two trifluoromethyls. Subsequently, the MVO+CSPA scheme has been employed only on those MOs localized on the former moiety, whereas those MOs pertaining to the CF₃ fragments are discarded. The resulting ΔE_{ST} 's are reported in Table 5. By

Table 4. ΔE_{ST} Computed with the MVO/CSPA DDCI Scheme for the Whole Diradical **1** and for the Two Geometries of the Smaller Model Di-Nitroxide^a

Var (%)	target diradical (1)			model diradicals (1a and 1b)			
	VSD (10^3 det.)	CPU time	ΔE_{ST} (cm^{-1})	VSD (10^3 det.)	CPU time	ΔE_{ST}^{1a} (cm^{-1})	ΔE_{ST}^{1b} (cm^{-1})
10	573	2 days, 12 h	316.1	133	3 h	262.7	231.8
15	907	7 days, 13 h	311.9	267	12 h	332.8	295.5
25	1755	15 days, 7 h	300.3	439	23 h	322.6	287.4
30	2844	20 days, 12 h	296.2	648	1 day, 12 h	319.8	286.3
35				894	2 days, 1 h	318.2	284.7
40				1500	3 days, 13 h	318.5	284.6
50				1726	4 days, 13 h	318.7	284.9
55				2254	5 days, 15 h	323.0	285.2
100				5417	17 days, 12 h	326.5	286.1

^a The ratio VAVO/MVO is reported in the first column (Var %), while the dimensions of the DDCI space treated variationally (VSD, expressed in number of configurations) appears in the second column.

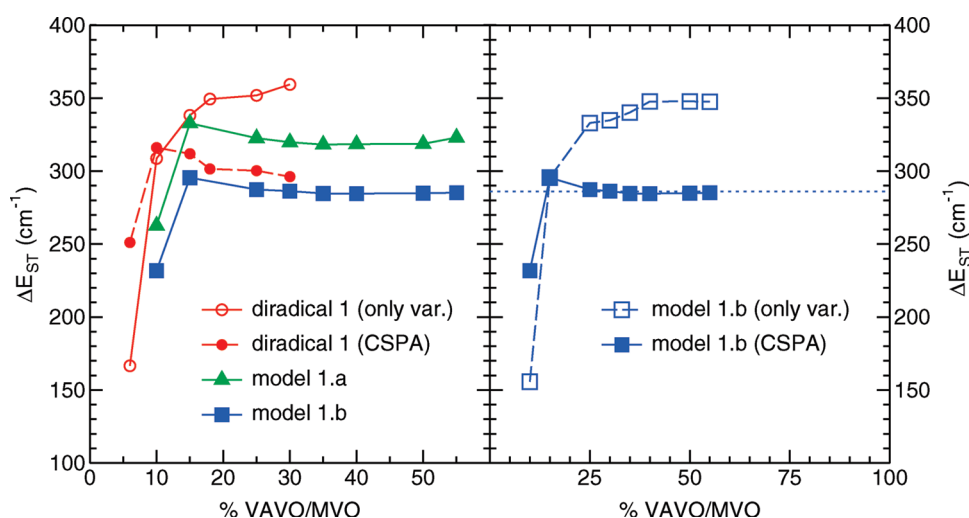


Figure 4. ΔE_{ST} as a function of the VAVO/MVO ratio. Left panel: Singlet–triplet energy gaps are compared for the whole (**1**, red symbols) and the model diradicals (**1a** and **1b**). For the larger compound, results obtained without the CSPA correction are also reported with empty red circles. Right panel: Convergence of the CSPA uncorrected (empty triangles) and corrected (full triangles) approach is compared for geometry **1b** of the model compound. The dotted line is the result of a complete variational treatment.

looking at these results, it appears that the singlet–triplet interaction is reduced by an order of magnitude with respect to diradical **1**, due to the substitution of the trifluoromethyl moieties, whose major effect is to displace outside the plane containing the aromatic ring the oxygen atoms bearing the unpaired electrons. This remarkable reduction is also in agreement with the experimental findings. In fact, despite an unexpected antiferromagnetic behavior being experimentally registered²¹ at room temperature, the triplet state was found to be the most stable at very low temperatures, where the minimum energy conformer is expected to be more populated. In these conditions, a best fit estimate of 80 K ($\sim 55 \text{ cm}^{-1}$) was reported²¹ for ΔE_{ST} , i.e., well below the lower limit value of 300 K given¹⁷ for its parent homologue **1**. It is also worth pointing out that, thanks to CSPA, the computed value is nearly converged by treating variationally only less than 20% of all MVOs, thus allowing a remarkable savings of computational time.

3.3. Biphenyl-3,5-diyl Bis(*tert*-butyl Nitroxide). Diradical **3** is obtained from diradical **1** by substitution of a hydrogen atom, placed in position 5 on the phenylene bridge, with a phenyl ring (see Figure 2). As in previous cases, modeling route b has been

followed; i.e., the whole diradical has been first optimized, and thereafter the methyl groups of the *tert*-butyl-nitroxides were substituted with hydrogen atoms. Differently from compound **2**, this diradical retains a planar structure of the *m*-phenylene bis-nitroxide moiety; that is, the oxygen atoms are coplanar to the bridge aromatic ring. Conversely, the phenyl substituent in position 5 is found at $\sim 40^\circ$ with respect to the bridge, so that a negligible conjugation between the two rings is expected. The latter observation allows us to apply the fragmentation procedure, projecting out from the DDCI configurational space all MOs localized onto the substituent phenyl ring. Computed ΔE_{ST} as a function of the VAVO/MVO ratio is reported in Figure 6. As in previous cases, convergence around a value of $\sim 300 \text{ cm}^{-1}$ is reached rather quickly, and a reliable estimate of the magnetic energy gap can be obtained by applying a variational approach to only 20–25% of the total number of MVOs, leaving the remaining VOs for the perturbative correction of the CSPA.

The value obtained is again in agreement with experimental estimates. Indeed, notwithstanding the first measures¹⁹ performed on the α phase of diradical **3** crystal unexpectedly revealing a

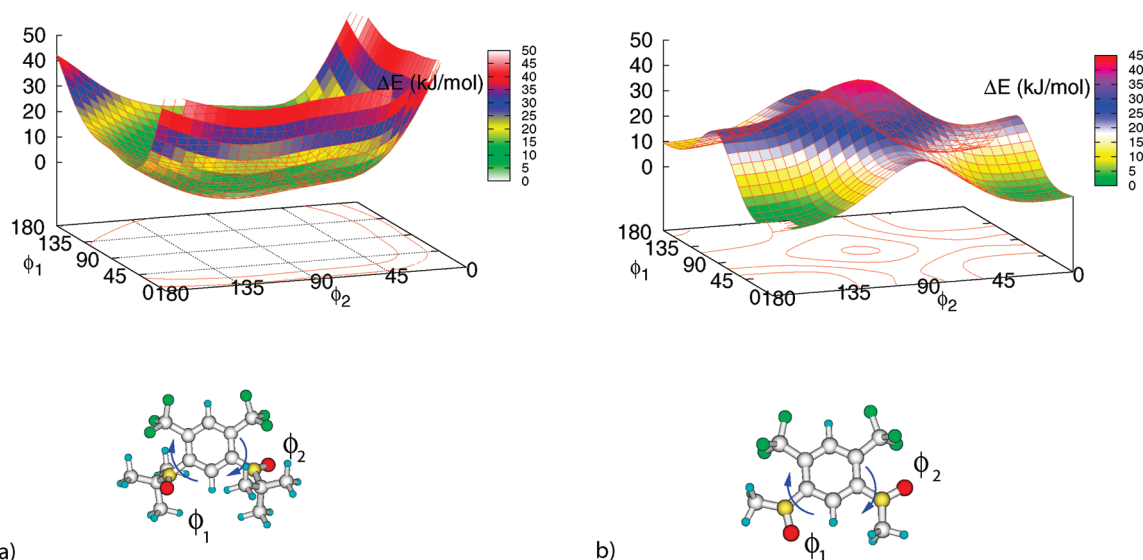


Figure 5. Torsional potential energy profiles for 4,6-trifluoromethyl, *m*-phenylene bis(*tert*-butyl)-nitroxide diradical **2** (panel a) and the 4,6-trifluoromethyl, *m*-phenylene bis(methyl)-nitroxide diradical (panel b). Dihedrals Φ_1 and Φ_2 are indicated with blue arrows.

Table 5. ΔE_{ST} Computed with the Three-Step DDCI Scheme for Diradical **2**

VAVO/MVO (%)	detors	CPU time	ΔE_{ST} (cm^{-1})
10	114676	5 h	21.9
15	239296	17 h	26.7
25	400116	1 days, 6 h	26.6
30	597136	1 days, 21 h	26.5
35	830356	2 days, 9 h	26.4
45	1099776	3 days, 3 h	26.3
50	1405396	3 days, 19 h	26.5
55	1747216	4 days, 22 h	26.5
65	1747216	5 days, 14 h	26.9

diamagnetic nature, further investigations²⁰ proved the existence of a second (β) crystal phase, in which the triplet state was shown to be the most stable. In particular, the antiferromagnetic response of the crystal in the α phase was attributed to intermolecular interactions between nitroxides of neighboring molecules, while an intramolecular origin was inferred for phase β . In the latter case, both crystallographic data (bridge-nitroxides planarity and 40.2° intraring torsion²⁰) and the best fit estimate of the singlet–triplet gap ($\Delta E_{ST} > 350$ K) are in agreement with our computed data (39.5° and 431 K).

3.4. Pegylated Bis(aminoxyl) Diradical. The dimensions of diradical **4** compelled us to choose a simpler model to represent the polymeric substituent. Despite the effect on ΔE_{ST} of both the reciprocal orientation between the polymeric and magnetic bearing moieties and the conformational mobility of the polyethylene glycol substituent needing to be taken into account, the aim of the present work is to test the sensitivity of the proposed computational protocol to different substituents. For this reason, only two models, in their fully optimized geometry, will be taken into account for the calculation of the singlet–triplet energy splitting of diradical **4**. Besides the substitution of the *tert*-butyl groups (which was again performed according to the aforementioned modeling route b), the polymeric chain was represented by two alkoxy substituents, namely $(\text{CH}_3\text{O})(\text{CH}_3)_2\text{C}-$ (diradical **4a**)

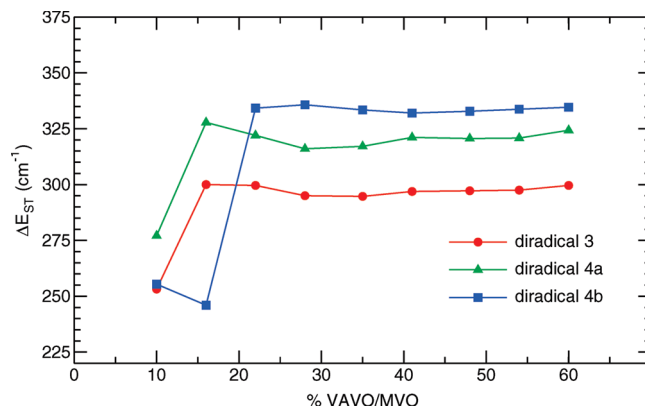


Figure 6. Singlet–triplet splittings ΔE_{ST} as a function of the VAVO/MVO ratio, computed for diradical **3** (circles) and two different models of diradical **4** (triangles and squares for a and b, respectively).

and $(\text{CH}_3\text{O})(\text{CH}_3)(\text{CH}_3\text{CH}_2\text{OCH}_2)\text{C}-$ (diradical **4b**), whose minimum energy conformations are shown respectively in panels a and b of Figure 7. The two geometries are identical in their *m*-phenylene moiety, whereas some differences can be spotted in the relative orientation of the aliphatic substituents with respect to the bridge. Indeed, in compound **4a**, the oxygen of the methoxyl group lies in the same plane as the aromatic ring, causing the two methyls to be in symmetric position with respect to the bridge. Conversely, in model **4b**, the methoxyl group is out-of-plane by $\sim 12^\circ$, and the polymeric fragment is nearly perpendicular to the ring plane.

For both models, the same fragmentation scheme has been applied, retaining for the DDCI calculations only those MOs localized in the nitroxides+bridge moiety. Computed singlet–triplet magnetic splittings are reported in Figure 6 as a function of the VAVO/MVO ratio. ΔE_{ST} 's are in both cases greater than the one computed for diradical **3**, with model **4b** showing a slightly larger value (≈ 490 K) than model **4a** (≈ 470 K). The agreement with the experimental data¹⁶ is less satisfactory than for the first three diradicals studied, even if the best fit estimate of ~ 650 K

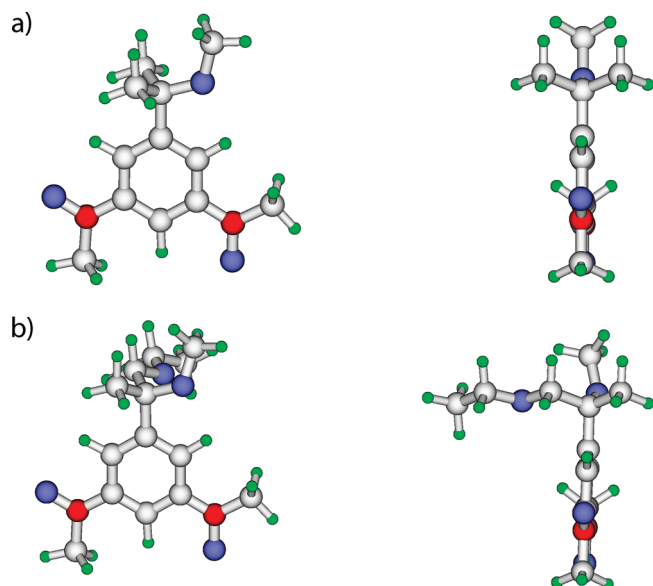


Figure 7. Front and side views of model diradicals **4a** and **4b** in their optimized geometries.

given by Rajca and co-workers could be somewhat overestimated (see note 17 in ref 16). However, it should be pointed out that, besides not considering the whole polymer chain, the present calculations do not include any dynamical effect, due to conformational mobility and/or interaction with the solvent, that could affect the ΔE_{ST} value.

4. CONCLUSIONS

A new integrated strategy for an accurate calculation of singlet–triplet magnetic splittings has been tested and validated. The procedure essentially consists of three steps, where the MVO and CSPA approaches, recently proposed by our group,^{58,59} are now combined on a DDCI configurational space previously reduced by fragmentation/localization criteria. This approach has been tested successfully on four different substituted *m*-phenylene bis(*tert*-butyl) nitroxides, which are experimentally known to possess a stable triplet state.

The fragmentation schemes sensibly reduce the DDCI space dimensions but still allow evaluation of the effect of different substituents, as they are able to distinguish between very similar models (e.g., model **4a** and **4b**). From a computational point of view, the application of the CSPA scheme allowed us to remarkably reduce the dimensions of the variationally treated space without a sensible loss in the accuracy of the calculations. More important, the optimal VAVO/MVO ratio (25–30%) seems not to depend on the diradical under study and could be confidently applied to larger systems. On the other hand, if *a priori*, “chemical” models (e.g., the substitution of the *tert*-butyl groups with smaller methyl moieties) are to be adopted, attention must be paid to the effect that such changes can introduce to the target diradical geometry, as ΔE_{ST} was shown to be rather sensitive to even small geometrical changes involving the bridge connecting the two magnetic moieties.

Finally, considering the difficulties in the experimental estimate of such a quantity, and the uncertainty that it is known^{1,18,22,24} to affect the best fit procedure, the proposed computational route could be considered as a powerful auxiliary technique for the accurate determination of magnetic interactions.

AUTHOR INFORMATION

Corresponding Author

*E-mail: giacomo.prampolini@sns.it.

REFERENCES

- (1) Rajca, A. *Chem. Rev.* **1994**, *94*, 871.
- (2) Rajca, A.; Wongsriratanakul, J.; Rajca, S. *Science* **2001**, *294*, 1503.
- (3) Kahn, O. *Molecular Magnetism*; VCH Publishers Inc.: New York, 1993; p 393.
- (4) Coronado, E.; Delhaès, P.; Gatteschi, D.; Milles, J. S. *Molecular Magnetism: From Molecular Assemblies to the Devices*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1996; NATO ASI Series Vol. E321.
- (5) Gatteschi, D.; Sessoli, R.; Villain, J. *Molecular nanomagnets*; Oxford University Press: Oxford, U. K., 2006.
- (6) Rajca, A.; Wongsriratanakul, J.; Rajca, S. *J. Am. Chem. Soc.* **2004**, *126*, 6608.
- (7) Fukuzaki, E.; Nishide, H. *J. Am. Chem. Soc.* **2006**, *128*, 996.
- (8) Matsuda, K.; Iwamura, H. *Curr. Opin. Solid State Mater. Sci.* **1997**, *2* (4), 446–450.
- (9) Koivisto, B. D.; Hicks, R. G. *Coord. Chem. Rev.* **2005**, *249*, 2612–2630.
- (10) Murata, H.; Miyajima, D.; Nishide, H. *Macromolecules* **2006**, *39*, 6331–6335.
- (11) Itoh, T.; Hirai, K.; Tornioka, H. *Bull. Chem. Soc. Jpn.* **2007**, *80*, 138–157.
- (12) Nishimaki, H.; Ishida, T. *J. Am. Chem. Soc.* **2006**, *132*, 9598.
- (13) Zhdanov, R. *Bioactive Spin Labels*; Springer-Verlag: Berlin, 1992.
- (14) Marx, L.; Rassat, A. *Chem. Commun.* **2002**, 632.
- (15) Francese, G.; Dunand, F.; Loosli, C.; Merbach, A.; Decurtins, S. *Magn. Reson. Chem.* **2003**, *43*, 81.
- (16) Spagnol, G.; Shiraishi, K.; Rajca, S.; Rajca, A. *Chem. Commun.* **2005**, 5047.
- (17) Ishida, T.; Iwamura, H. *J. Am. Chem. Soc.* **1991**, *113*, 4238.
- (18) Dvolutzky, M.; Chiarelli, R.; Rassat, A. *Angew. Chem., Int. Ed.* **1992**, *31*, 180.
- (19) Kurokawa, G.; Ishida, T.; Nogami, T. *Chem. Phys. Lett.* **2004**, *392*, 74.
- (20) Nishimaki, H.; Mashiyama, S.; Yasui, M.; Nogami, T.; Ishida, T. *Chem. Mater.* **2006**, *18*, 3602.
- (21) Rajca, A.; Lu, K.; Rajca, S.; Ross, C., II. *Chem. Commun.* **1999**, 1249.
- (22) Rajca, A.; Shiraishi, K.; Rajca, S. *Chem. Commun.* **2009**, 4372.
- (23) Rajca, A.; Vale, M.; Rajca, S. *J. Am. Chem. Soc.* **2008**, *130*, 9099.
- (24) Rajca, A.; Shiramishi, K.; Pink, M.; Rajca, S. *J. Am. Chem. Soc.* **2007**, *129*, 7232.
- (25) Noodleman, L.; Norman, J. G. *J. Chem. Phys.* **1979**, *70*, 4903.
- (26) Noodleman, L. *J. Chem. Phys.* **1981**, *74*, 5737.
- (27) Barone, V.; Bencini, A.; Ciofini, I.; Daul, C. A. *J. Phys. Chem. A* **1999**, *103*, 4275.
- (28) Adamo, C.; Barone, V.; Bencini, A.; Totti, F.; Ciofini, I. *Inorg. Chem.* **1999**, *38*, 1996.
- (29) Ali, M. E.; Datta, S. N. *J. Phys. Chem. A* **2006**, *110*, 2776.
- (30) Ali, M. E.; Roy, A. S.; Datta, S. N. *J. Phys. Chem. A* **2007**, *111*, 5523.
- (31) Illas, F.; Moreira, I. de P. R.; Bofill, J. M.; Filatov, M. *Theor. Chem. Acc.* **2006**, *116*, 587.
- (32) Rivero, P.; Moreira, I. de P. R.; Illas, F.; Scuseria, G. E. *J. Chem. Phys.* **2008**, *129*, 184110.
- (33) Bencini, A. *Inorg. Chim. Acta* **2008**, *361*, 3820.
- (34) Moreira, I. de P. R.; Illas, F. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1645.
- (35) Bencini, A.; Totti, F. *J. Chem. Theory Comput.* **2009**, *5*, 144.
- (36) de Loth, P.; Cassoux, P.; Daudey, J. P.; Malrieu, J. P. *J. Am. Chem. Soc.* **1981**, *103*, 4007.

- (37) Miralles, J.; Daudey, J. P.; Caballol, R. *Chem. Phys. Lett.* **1992**, *198*, 555.
- (38) Miralles, J.; Castell, O.; Caballol, R.; Malrieu, J. P. *Chem. Phys.* **1993**, *172*, 33.
- (39) Castell, O.; Caballol, R.; Subra, R.; Grand, A. *J. Phys. Chem.* **1995**, *99*, 154.
- (40) de Graaf, C.; Sousa, C.; Moreira, I. de P. R.; Illas, F. *J. Phys. Chem. A* **2001**, *105*, 11371.
- (41) Angeli, C.; Calzado, C. J.; Cimiraglia, R.; Evangelisti, S.; Guihéry, N.; Leininger, T.; Malrieu, J.-P.; Maynau, D.; Ruitz, J. V. P.; Sparta, M. *Mol. Phys.* **2003**, *101*, 1389.
- (42) Calzado, C. J.; Cabrero, J.; Malrieu, J. P.; Caballol, R. *J. Chem. Phys.* **2002**, *116*, 2728.
- (43) Calzado, C. J.; Cabrero, J.; Malrieu, J. P.; Caballol, R. *J. Chem. Phys.* **2002**, *116*, 3985.
- (44) Cabrero, J.; Ben Amor, N.; de Graaf, C.; Illas, F.; Caballol, R. *J. Phys. Chem. A* **2000**, *104*, 9983.
- (45) Neese, F. *J. Chem. Phys.* **2003**, *119*, 9428.
- (46) Calzado, C. J.; Angeli, C.; Taratiel, D.; Caballol, R.; Malrieu, J.-P. *J. Chem. Phys.* **2009**, *131*, 044327.
- (47) de Graaf, C.; Caballol, R.; Romo, S.; Poblet, J. M. *Theor. Chem. Acc.* **2009**, *123*, 3.
- (48) Calzado, C. J.; Angeli, C.; Caballol, R.; Malrieu, J.-P. *Theor. Chem. Acc.* **2010**, *126*, 185.
- (49) Monari, A.; Maynau, D.; Malrieu, J.-P. *J. Chem. Phys.* **2010**, *133*, 44106.
- (50) Queral, N.; Taratiel, D.; de Graaf, C.; Caballol, R.; Cimiraglia, R.; Angeli, C. *J. Comput. Chem.* **2008**, *29*, 994.
- (51) Rota, J.-B.; Norel, L.; Train, C.; Ben Amor, N.; Maynau, D.; Robert, V. *J. Am. Chem. Soc.* **2008**, *130*, 10380.
- (52) Bastardis, R.; Guihéry, N.; de Graaf, C. *J. Chem. Phys.* **2008**, *129*, 104102.
- (53) Maurice, R.; Guihéry, N.; Bastardis, R.; de Graaf, C. *J. Chem. Theory Comput.* **2010**, *6*, 55.
- (54) Calzado, C. J.; Angeli, C.; de Graaf, C.; Caballol, R. *Theor. Chim. Acc.* **2010**, DOI:10.1007/s00214-010-0831-6.
- (55) Barone, V.; Cacelli, I.; Ferretti, A.; Girlanda, M. *J. Chem. Phys.* **2008**, *128*, 174303.
- (56) Barone, V.; Cacelli, I.; Ferretti, A. *J. Chem. Phys.* **2009**, *130*, 94306.
- (57) Barone, V.; Cacelli, I.; Cimino, P.; Ferretti, A.; Monti, S.; Prampolini, G. *J. Phys. Chem. A* **2009**, *113*, 15150.
- (58) Barone, V.; Cacelli, I.; Ferretti, A.; Prampolini, G. *Phys. Chem. Chem. Phys.* **2009**, *11*, 3854.
- (59) Barone, V.; Cacelli, I.; Ferretti, A.; Prampolini, G. *J. Chem. Phys.* **2009**, *131*, 224103.
- (60) Barone, V.; Cacelli, I.; Ferretti, A.; Prampolini, G. *Phys. Chem. Chem. Phys.* **2010** under revision.
- (61) Pipek, J.; Mezey, P. G. *J. Chem. Phys.* **1989**, *90*, 4916.
- (62) 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.
- (63) Schmidt, M. W.; Baldridge, K. K.; Boats, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347.
- (64) Huron, B.; Malrieu, J. P.; Rancurel, P. *J. Chem. Phys.* **1973**, *58*, 5745.
- (65) Evangelisti, S.; Daudey, J. P.; Malrieu, J. P. *Chem. Phys.* **1983**, *75*, 91.
- (66) Cimiraglia, R.; Persico, M. *J. Comput. Chem.* **1987**, *8*, 39.
- (67) Shultz, D.; Boal, A.; Lee, H.; Farmer, G. *J. Org. Chem.* **1999**, *64*, 4386.
- (68) Shultz, D.; Fico, R.; Kampf, H. L. J.; Kirschbaum, K.; Pinkerton, A.; Boyle, P. *J. Am. Chem. Soc.* **2003**, *125*, 15426.

Electron Transport Suppression from Tip– π State Interaction on Si(100)-2 \times 1 Surfaces

K. P. Dou,^{†,‡} W. Fan,[‡] T. A. Niehaus,^{§,||} T. Frauenheim,^{||} C. L. Wang,[†] X. H. Zhang,[‡] and R. Q. Zhang^{*,†,⊥}

[†]School of Physics, Shandong University, Jinan 250100, P.R. China

[‡]Nano-organic Photoelectronic Laboratory, Technical Institute of Physics and Chemistry, Chinese Academy of Sciences, Beijing 100190, China

[§]Institute I—Theoretical Physics, University of Regensburg, D-93040 Regensburg, Germany

^{||}Bremen Center for Computational Material Science, University Bremen, 28334 Bremen, Germany

[⊥]Center of Super-Diamond and Advanced Films and Department of Physics and Materials Science, City University of Hong Kong, Hong Kong Special Administrative Region, China

S Supporting Information

ABSTRACT: We investigate the electron transport between a scanning tunneling microscope tip and Si(100)-2 \times 1 surfaces with four distinct configurations by performing calculations using density functional theory and the nonequilibrium Green's function method. Interestingly, we find that the conducting mechanism is altered when the tip–surface distance varies from large to small. At a distance larger than the critical value of 4.06 Å, the conductance is increased with a reduction in distance owing to the π state arising from the silicon dimers immediately under the tip; this in turn plays a key role in facilitating a large transmission probability. In contrast, when the tip is closer to the substrate, the conductance is substantially decreased because the π state is suppressed by the interaction with the tip, and its contribution in the tunneling channels is considerably reduced.

INTRODUCTION

Surface states originating from dangling bonds (DBs) play a pivotal role in providing active reaction sites^{1–5} with covalent attachments and assisting conductance^{6–11} in electron transport. For instance, the π/π^* states induced from silicon surface dimerization are of great importance for understanding the conducting mechanisms.^{9,10} Moreover, such π -like states are important in organic molecular electronics and have been intimately associated with the formation of some particular phenomena such as negative differential resistance (NDR) in recent studies.^{12,13} In the theoretical research of Fan et al.,¹² the mismatch of the energy alignment between the nonlocal π^* state of the conducting molecule and the metallic lead states results in NDR, while the NDR mechanism in the theoretical investigation by Bevan et al.¹³ of a styrene adsorbed on a silicon substrate is attributed to the bias-induced pulling of the π state of the molecule out of the resonance window and into the band gap of the silicon substrate.

The DBs of an unpassivated silicon surface can serve as localized charge trap sites, and their corresponding energy levels often appear within the band gap close to the Fermi level and hence determine much of the electronic behavior of the surface. Intense research has been devoted to exploring the functionality of the DB in forming or characterizing reconstructed surfaces both experimentally^{6–8} and theoretically.^{9–11} One area of heavy interest is electron transport through crystallized surfaces with DBs. Hata et al.⁶ presented a refined interpretation of images of Si(100) produced by a scanning tunneling microscope (STM) and attributed the typical high surface bias used to observe the

π/π^* state to the influence of the adsorbed defects. Brown et al.⁷ found that the π/π^* state shows different signatures in band bending in the presence of positively charged defects when imaging both filled and empty states. Recently, D'angelo et al.⁸ experimentally determined the metallic DB state of Si(111)-7 \times 7 through a metal-to-insulator transition process induced by Na adsorption.

In general, theoretical transport studies of DBs are performed in two directions: along⁹ and perpendicular^{10,11} to the surface. Much work has been done in these regards. Doumergue et al.⁹ reported theoretical results on the conductance along a DB line stabilized on a Si(001)-2 \times 1:H surface in two different configurations. They found that the DB line with a Peierls distortion creates a small energy gap straddling the Fermi level compared with that without the Peierls distortion. Theoretical investigations on vertical transport through silicon surface states refer mainly to the simulations of an STM tip–sample interaction image. Ono et al.¹⁰ found that the surface π states of the bare silicon dimer dramatically influence the STM image on a hydrogen-terminated Si(001) surface. Jelínek et al.¹¹ observed a substantial decrease of conductance during approach of the tip to the Si(111)-7 \times 7 surface and revealed that this unusual feature results from the formation of the strong covalent bond between the tip apex atom and the adatom on the surface. We conjectured that such an abnormal conductance drop could also be observed on a Si(001)-2 \times 1 surface on which DB states π/π^* were

Received: September 1, 2010

Published: January 10, 2011

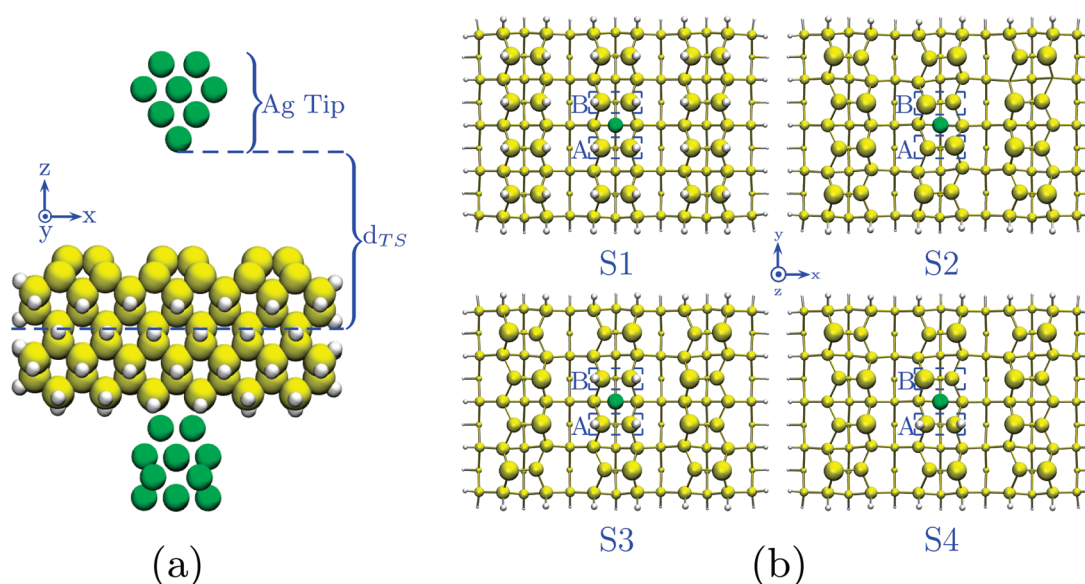


Figure 1. (a) Schematic diagram of the tip–substrate model. The slab consists of nine silicon layers. (b) Top view of upper four layers of different silicon slabs under the tip after relaxation: the fully hydrogen-terminated surface containing three monohydride chains (S1), the clean surface with buckled dimers (S2), the two central silicon dimers, A and B, are hydrogen terminated (S3), and only dimer A is hydrogen terminated (S4). The yellow and white balls are silicon and hydrogen atoms, respectively. The green ball represents the hollow site where the tip locates above. The upper layer atoms are bigger balls, while lower layer atoms are smaller ones.

considered to be localized mainly on the silicon dimer atoms.^{6,14,15} To explore such a possibility, in the present work, we investigated the electron transport between an STM tip and Si(100)- 2×1 surfaces with four distinct configurations by performing calculations using density functional theory and the nonequilibrium Green's function (NEGF) method.

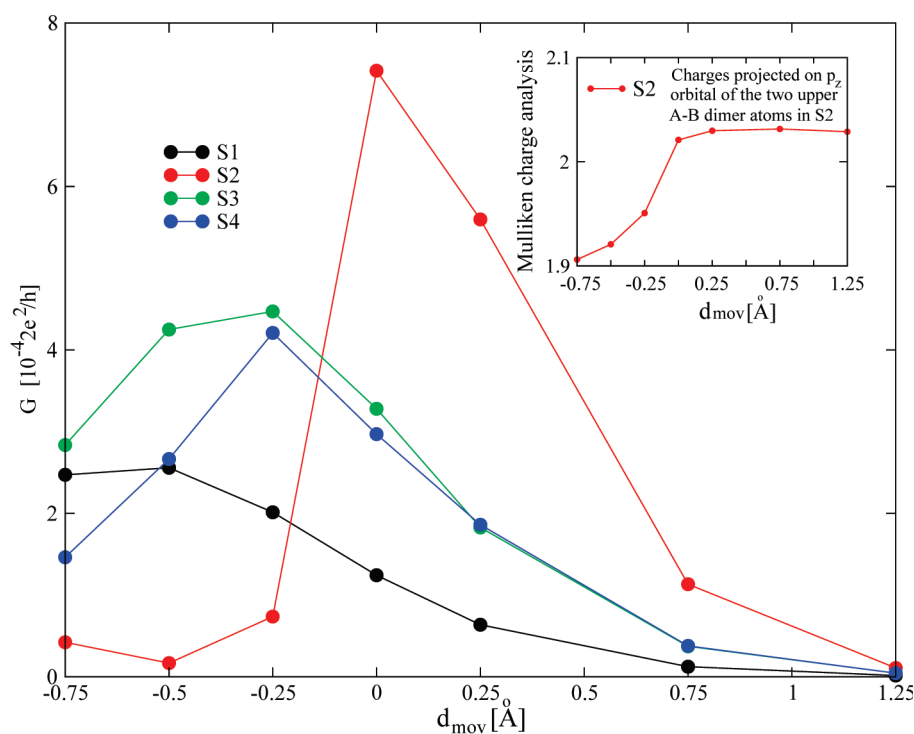
MODELS AND METHOD

To model the silicon surface, we adopted a slab structure with the top layer containing three adjacent silicon dimer rows. Each row, consisting of four silicon dimers together with eight underlying silicon layers, represented the reconstructed Si(100)- 2×1 surface. We considered four models for the Si(100)- 2×1 surface, which Figure 1b depicts: a fully hydrogen-terminated surface containing three monohydride chains (S1), a purely clean surface with buckled dimers (S2), both central silicon dimers A and B, which are hydrogen terminated (S3), and a hydrogen-terminated dimer A only (S4).

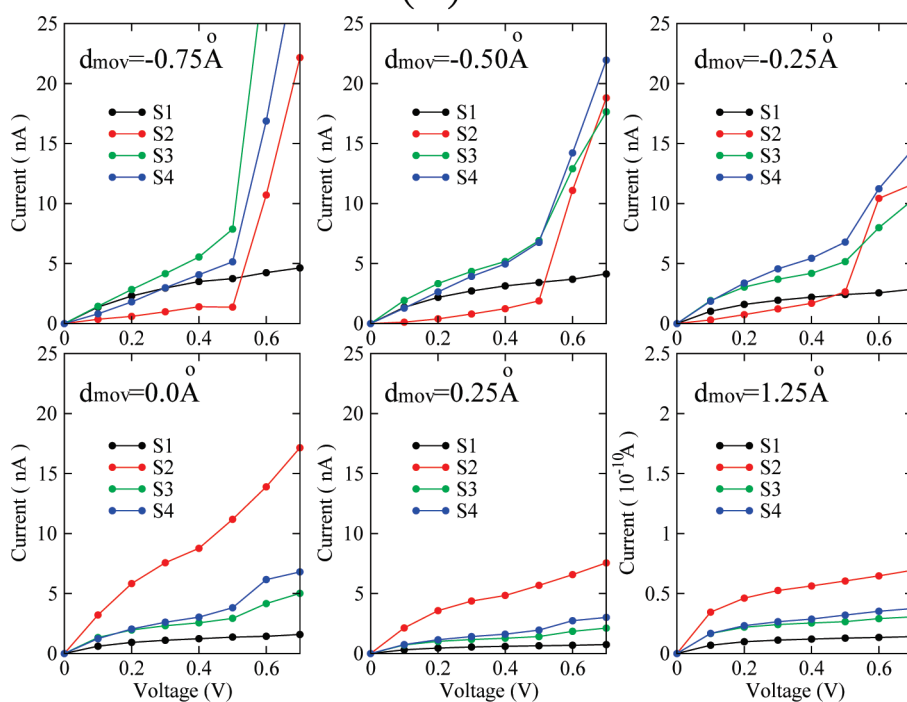
To reduce computational cost, we adopted two ultrathin semi-infinite Ag nanowires as the source and drain electrodes, which we connected to the STM tip and the silicon substrate, respectively. The tip was modeled by a 14 silver-atom cone-shaped structure, which was obtained by sharpening the terminal unit cell of the source electrode neighboring the silicon substrate. The lateral and bottom sides of the slab were terminated with hydrogen atoms except for the four center silicon atoms on the bottom side, which were coupled to the drain electrode. These terminating hydrogen atoms were used to eliminate the unnecessary DBs of the silicon slab so as to create a chemical environment similar to that of an actual Si(100)- 2×1 surface. Although partial surface oxidation cannot be ruled out under experimental conditions, this surface model allows for the study of the transport characteristics of dimer states in a well-defined environment.¹¹ We have used our models to successfully mimic the flipping dynamics of bare dimers next to the H-passivated

dimers revealed by DFT calculation²² and experimentation.²³ The details are shown in the Supporting Information. The tip–substrate model is schematically shown in Figure 1a. We performed a structural optimization with the first unit cell of the drain electrode and the bottom four silicon layers of the slab, with the other five layers of atoms of the slab fixed (in the absence of the tip). We then located the tip always above the hollow site of the two central silicon surface dimers (refer to A and B in Figure 1b) in all four models and gradually moved it toward the substrate. At each instant state after moving a distance, the topmost four layers of the substrate were fully relaxed from their presetting symmetrically reconstructed 2×1 structure, with the other parts fixed. From Figure 1b, we can see that for the S2, S3, and S4 models, the initial presetting symmetrical silicon dimers relaxed to favorable buckling structures owing to the approach of the tip. During the geometry relaxations, the fifth layer of the silicon slab was always kept fixed in its bulk configuration, which acted as a reference to define the distance between the tip and the substrate.

Our theoretical study of the Si(100) surface was performed with the gDFTB¹⁶ code, which is an extension of the NEGF method of electron transport via density-functional-based tight binding (DFTB).¹⁷ This method allows an efficient treatment of systems composed of a large number of atoms to high precision. During the nonequilibrium process, the chemical potential of the tip, μ_T , was kept at the equilibrium Fermi level of the source electrode, while the chemical potential of the substrate was shifted following the applied bias by $\mu_s = \mu_T + eV_b$. We performed transport calculations by sampling only the Γ point of the surface Brillouin zone. We included one unit layer of the drain electrode neighboring the silicon substrate, as well as the 14 silver-atom tip and the nine atomic layers of the Si(100) surface as the scattering region. The tunneling current of the system was calculated on the basis of Landauer theory.¹⁸ Details on the gDFTB implementation are available in a recent review.¹⁹



(a)



(b)

Figure 2. (a) Equilibrium conductance of four models shown in Figure 1b as a function of the tip movement at $V_b = 0.0$ V. The inset plots the Mulliken charges projected on the p_z orbital of the two upper silicon dimer atoms on A and B also as a function of the tip movement at $V_b = 0.0$ V in S2. (b) I – V curves of four silicon slabs shown in Figure 1b at different tunneling distances. The case at $d_{\text{mov}} = 0$ Å is considered to be the switch point for alternative tunneling mechanisms dominating the electron transport.

RESULTS AND DISCUSSION

During the geometric optimization, the upper layers of the substrate were relaxed and reconstructed, which caused the distance from the tip to the upper layers to be ill-defined. We

therefore fixed the middle layer of the slab during the relaxation and defined a distance, d_{TS} , between this layer and the tip as the tip–substrate distance variable for the following presentation. As shown in Figure 1a, the so-defined tip–substrate distance

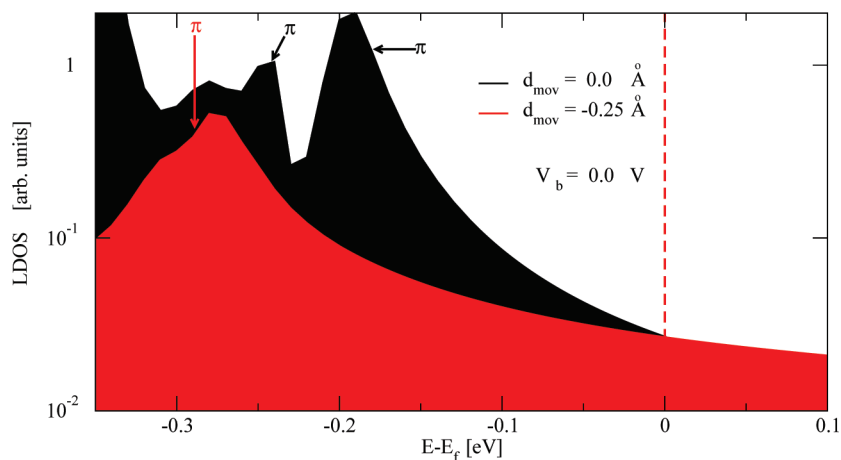


Figure 3. LDOS of silicon atoms on two central dimers A and B in S2 at $d_{\text{mov}} = 0 \text{ \AA}$ and $d_{\text{mov}} = -0.25 \text{ \AA}$ for zero bias. The red dashed line located at 0 indicates the Fermi level of the source electrode.

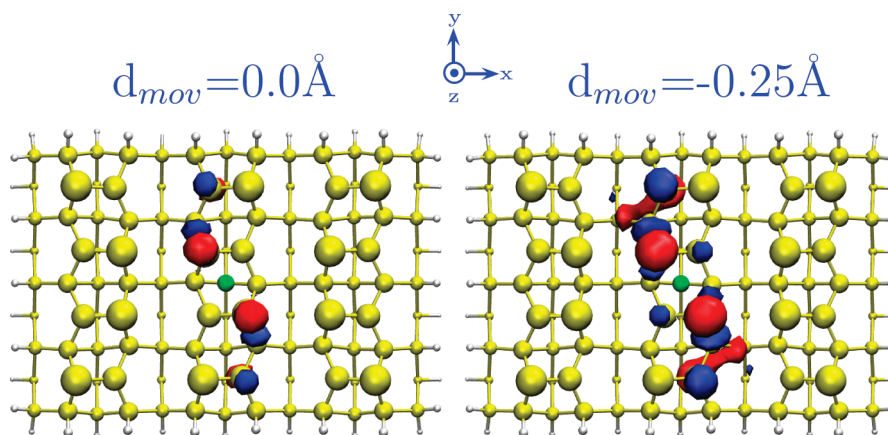


Figure 4. Evolution of the HOMO state on the silicon surface in S2 at two typical tip movement representatives $d_{\text{mov}} = 0 \text{ \AA}$ and $d_{\text{mov}} = -0.25 \text{ \AA}$ under zero bias.

transformed the ill-defined quantity into one that was well-defined. With the d_{TS} , we could perform tunneling calculations with varied tip–substrate distances. So that the tip-moving process is well explained, we defined the $d_{\text{TS}} = 9.04 \text{ \AA}$ model as the static location $d_{\text{mov}} = 0 \text{ \AA}$. At this location, the average distance between the tip apex atom and the upper dimer atoms on A and B is 4.06 \AA in the clean surface model S2. With this reference position, a positive d_{mov} indicated that the tip was lifted up, whereas a negative one meant it was dipping down. In this paper, we consider five movements of the tip starting from its original position ($d_{\text{mov}} = 0 \text{ \AA}$).

Figure 2a represents a characteristic set of equilibrium conductances as a function of the tip movement. It shows that at a large tip–substrate distance, the conductance increases quickly with the decrease of the tip–substrate distance and then undergoes a sudden drop after $d_{\text{mov}} = 0 \text{ \AA}$ in the clean surface model S2. It appears that the closer tip–substrate distance does not increase the capability of tunneling for carriers but instead strengthens the suppression of electron transport, although the interatomic coupling should be enhanced between the tip apex atom and the silicon surface atoms, especially the atoms of the two central dimers A and B.

In the following, we uncover the origin of the electron transport suppression by virtue of the local density of states

(LDOS) on the silicon atoms of dimers A and B in S2 at two typical tip movement representatives, $d_{\text{mov}} = 0 \text{ \AA}$ and -0.25 \AA , as shown in Figure 3. Note that in both cases there are dimer-derived surface states in the band gap known as π states.^{20,21} It is clearly seen that the π state located at the atoms of the two central silicon dimers A and B is greatly suppressed and is pushed away from the Fermi level for the d_{mov} varying from 0 \AA to -0.25 \AA . The change of the localized DB π state on the dimer atoms is induced by the tip proximity. To obtain an intuitive image of such changes resulting from the tip– π interaction, it is necessary to investigate the evolution of the frontier orbitals of the silicon surface in the scattering region at different d_{mov} values. In our transport calculations, the Fermi level of the source electrode is close to the occupied surface π states rather than being in the middle of the gap. So the highest occupied molecular orbitals (HOMOs) are the key states in determining the transport behavior of the model at equilibrium and at low biases. As presented in Figure 4, at the large tip–substrate separation $d_{\text{mov}} = 0$, the HOMO is localized mainly on the atoms of the two central dimers A and B, showing a localized DB π character at equilibrium in S2. When the tip further approaches the substrate, the HOMO becomes delocalized, and this delocalization carries further along the silicon surface and down to the lower layers of the substrate rather than upward to the tip. Consequently, the

projection of the charge on the DB π state to the transport direction is reduced. It is known that the DB π -like state on such a dimerized surface originates mainly from the p_z bond orbital of the silicon dimer atoms pointing normally to the surface plane.¹⁵ We also analyzed the variation of the Mulliken charges projected on the p_z orbitals of the two upper silicon dimer atoms on A and B, which were closest to the tip apex atom in S2. As shown in the inset of Figure 2a, there exists a sudden drop after $d_{\text{mov}} = 0 \text{ \AA}$ and a monotonic loss tendency on the p_z orbitals of the two upper silicon atoms immediately under the tip with the approach of the tip, hence limiting the tunneling through them.

Similar conductance behavior also appeared for models S3 and S4 but not at all for S1. This is because the DBs were terminated by hydrogen atoms with different coverage in the three models, and the orbital hybridization around the hydrogen atoms and the Si–H bonds was stronger and less delocalized than in the DB π state. Thus, the transport suppression in S3 and S4 was not as remarkable as that in S2. In S1, on the other hand, all of the DB states were eliminated by hydrogen chemisorption, and there was no longer a tip– π interaction. Consequently, the tip proximity did not suppress transport.

For better clarity on how the DB π state suppression induced transport discrepancies in the four models, we further performed calculations under bias to obtain the current–voltage (I – V) characteristics of all models. The I – V curves are presented in Figure 2b. The plots are divided into two rows, which show quite different characteristics. In the lower row of Figure 2b, the current of the S2 model is always larger than that of the other models. The order of the current magnitude is always $I(\text{S2}) > I(\text{S4}) > I(\text{S3}) > I(\text{S1})$. At a positive d_{mov} , the tip– π interaction is still not strong enough to suppress the transport; meanwhile, the DB π state is more widely expanded than the states around the hydrogen-adsorbed dimer.¹⁰ As a result, the current of S2 is larger than that of the other three models. In the first row at the negative d_{mov} , the current of S2 is obviously suppressed as a result of the tip– π interaction increase, as mentioned in the equilibrium discussion.

On the basis of the above discussion, when $d_{\text{mov}} \geq 0$, the dominant tunneling should be from the widely expanded π state of the surface dimers, whereas when $d_{\text{mov}} < 0$, the strong hybridization of the atomic orbitals around the Si–H bonds should be the main contribution in the conducting channels.

SUMMARY

We have studied vertical electron transport through an STM tip to a Si(100)- 2×1 reconstructed substrate. The π states from the silicon dimers immediately below the tip assist tunneling primarily at large tip–substrate distances. However, the π state is delocalized when the tip approaches the substrate and leads to transport suppression. Such suppression can be reduced or even eliminated by hydrogen chemisorptions.

ASSOCIATED CONTENT

S Supporting Information. Simulation for the flipping dynamics of bare dimers next to the H-passivated dimers on silicon surface by using our models. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*E-mail: aprqz@cityu.edu.hk.

ACKNOWLEDGMENT

The work described in this paper was supported by grants from National Basic Research Program of China (Grant No. 2006CB933000) and the Research Grants Council of the Hong Kong Special Administrative Region, China (CityUS/CRF/08, CityU 103609).

REFERENCES

- (1) Lopinski, G. P.; Wayner, D. D. M.; Wolkow, R. A. Self-directed growth of molecular nanostructures on silicon. *Nature (London)* **2000**, *406*, 48.
- (2) Teague, L. C.; Boland, J. J. STM Study of Multiple Bonding Configurations and Mechanism of 1,3-Cyclohexadiene Attachment on Si(100)- 2×1 . *J. Phys. Chem. B* **2003**, *107*, 3820.
- (3) Moriarty, P.; Upward, M. D.; Dunn, A. W.; Ma, Y. R.; Beton, P. H. C_{60} -terminated Si surfaces: Charge transfer, bonding, and chemical passivation. *Phys. Rev. B* **1998**, *57*, 362.
- (4) Hossain, M. Z.; Yamashita, Y.; Mukai, K.; Yoshinobu, J. Microscopic observation of precursor-mediated adsorption process of NH_3 on Si(100)c(4×2) using STM. *Phys. Rev. B* **2003**, *68*, 235322.
- (5) Naydenov, B.; Teague, L. C.; Ryan, P.; Boland, J. J. Contact Formation Dynamics: Mapping Chemical Bond Formation between a Molecule and a Metallic Probe. *Nano Lett.* **2006**, *6*, 1752.
- (6) Hata, K.; Yasuda, S.; Shigeokawa, H. Reinterpretation of the scanning tunneling microscopy images of Si(100)-(2×1) dimers. *Phys. Rev. B* **1999**, *60*, 8164.
- (7) Brown, G. W.; Grube, H.; Hawley, M. E.; Schofield, S. R.; Curson, N. J.; Simmons, M. Y.; Clark, R. G. Imaging charged defects on clean Si(100)-(2×1) with scanning tunneling microscopy. *J. Appl. Phys.* **2002**, *92*, 820.
- (8) D'angelo, M.; Takase, K.; Miyata, N.; Hirahara, T.; Hasegawa, S. Conductivity of the Si(111) 7×7 dangling-bond state. *Phys. Rev. B* **2009**, *79*, 035318.
- (9) Doumergue, P.; Pizzagalli, L.; Joachim, C.; Altibelli, A.; Baratoff, A. Conductance of a finite missing hydrogen atomic line on Si(001)-(2×1)-H. *Phys. Rev. B* **1999**, *59*, 15910.
- (10) Ono, T.; Horie, S.; Endo, K.; Hirose, K. First-principles study of the tunnel current between a scanning tunneling microscopy tip and a hydrogen-adsorbed Si(001) surface. *Phys. Rev. B* **2006**, *73*, 245314.
- (11) Jelínek, P.; Švec, M.; Pou, P.; Perez, R.; Cháb, V. Tip-Induced Reduction of the Resonant Tunneling Current on Semiconductor Surfaces. *Phys. Rev. Lett.* **2008**, *101*, 176101.
- (12) Fan, W.; Zhang, R. Q.; Rocha, A. R.; Sanvito, S. Energy alignment induced negative differential resistance: The role of hybrid states in aromatic molecular devices. *J. Chem. Phys.* **2008**, *129*, 074710.
- (13) Bevan, K. H.; Kienle, D.; Guo, H.; Datta, S. First-principles nonequilibrium analysis of STM-induced molecular negative-differential resistance on Si(100). *Phys. Rev. B* **2008**, *78*, 035303.
- (14) Raza, H.; Raza, T. Z.; Kan, E. C. Electrical transport in a two-dimensional electron and hole gas on a Si(001)-(2×1) surface. *Phys. Rev. B* **2008**, *78*, 193401.
- (15) Srivastava, G. P. Clean Semiconductor Surfaces. In *Theoretical Modelling of Semiconductor Surfaces*, 1st ed.; World Scientific: River Edge, NJ, 1999; Vol 8, p 176.
- (16) Carlo, A. D.; Pecchia, A.; Latessa, L.; Frauenheim, Th.; Seifert, G. Tight-Binding DFT for Molecular Electronics (gDFTB). *Lect. Notes Phys.* **2005**, *680*, 153.
- (17) Frauenheim, Th.; Seifert, G.; Elstner, M.; Niehaus, T. A.; Köhler, C.; Amkreutz, M.; Sernberg, M.; Hajnal, Z.; Carlo, A. D.; Suhai, S. Atomistic simulations of complex materials: ground-state and excited-state properties. *J. Phys.: Condens. Matter* **2002**, *14*, 3015.
- (18) Büttiker, M.; Imry, Y.; Landauer, R.; Pinhas, S. Generalized many-channel conductance formula with application to small rings. *Phys. Rev. B* **1985**, *31*, 6207.
- (19) Pecchia, A.; Salvucci, L.; Penazzi, G.; Carlo, A. D. Non-equilibrium Green's functions in density functional tight binding: method and application. *New J. Phys.* **2008**, *10*, 065022.

(20) Nakayama, K. S.; Alemany, M. M. G.; Sugano, T.; Ohmori, K.; Kwak, H.; Chelikowsky, J. R.; Weaver, J. H. Electronic structure of Si(001)-c(4 × 2) analyzed by scanning tunneling spectroscopy and *ab initio* simulations. *Phys. Rev. B* **2006**, *73*, 035330.

(21) Krüger, P.; Pollmann, J. Dimer Reconstruction of Diamond, Si, and Ge (001) Surfaces. *Phys. Rev. Lett.* **1995**, *74*, 1155.

(22) Chen, D. X.; Boland, J. J. Chemisorption-induced disruption of surface electronic structure: Hydrogen adsorption on the Si(100)-2 × 1 surface. *Phys. Rev. B* **2002**, *65*, 165336.

(23) Buehler, J. E.; Boland, J. J. Dimer Preparation That Mimics the Transition State for the Adsorption of H₂ on the Si(100)-2 × 1 Surface. *Science* **2000**, *290*, 506.

■ NOTE ADDED AFTER ASAP PUBLICATION

This article was published ASAP on January 10, 2011. The spelling of the third author's name has been corrected. The correct version was published on February 9, 2011.

Silicon Nanocrystal Functionalization: Analytic Fitting of DFTB Parameters

Fabio Trani^{*,†,‡} and Vincenzo Barone^{†,‡}[†]Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126, Pisa, Italy[‡]Infn Sezione di Pisa

ABSTRACT: A density functional tight binding (DFTB) scheme has been applied to functionalized silicon nanocrystals. Using an analytic functional representation of DFTB parameters, the scheme has been used to compute the adsorption energies in the organic functionalization of reconstructed Si(100) and H-terminated Si(111) surfaces of hundreds-of-atoms nanocrystals. We adopt an ONIOM(QM:QM') approach that corrects the overbinding of DFTB, obtaining nice agreement with high-level reaction energies and structural configurations.

1. INTRODUCTION

The synthesis of bright, stable, water-soluble silicon nanocrystals has opened the route to realistic biomedical applications of functionalized silicon nanoparticles as cellular probes.^{1–3} The recent chemical synthesis techniques allow for an excellent control of nanoparticle size-dispersion, shape, and passivation that allows a suitable design of their fluorescence properties.¹ As is well-known, the chemical environment plays the role of a key agent that can deeply modify the nanocrystal electronic structure and optical spectra. Oxygen-free silicon nanostructures, for instance, alkyl-terminated nanocrystals, are characterized by a strong photoluminescence, with high quantum yield, in the blue spectral region.⁴ This new generation of blue-emitting, water stable silicon nanocrystals makes them ideal candidates for applications in medical biosensing, in particular as fluorescent tumoral markers.^{3,5,6} At variance with other compounds, such as CdSe nanorods, silicon nanocrystals are nontoxic and biocompatible, and this justifies the enormous interest in such materials.

On the theoretical side, several papers have been published on the functionalization of silicon nanocrystals,^{7–9} but comprehensive studies are still lacking. New integrated computational approaches are currently being developed to simulate the interaction of silicon nanocrystals with biological systems. In particular, many body and density functional approaches allow for the description of few-nanometer structures, while huge, realistic nanocrystals, functionalized by large organic molecules (such as DNA fragments or organic dyes), have hardly been studied because of the demanding resources required. A detailed description of the mutual interplay between the silicon nanoparticle excited states and the energy levels of biological molecules can lead to very promising results, with high impact technological implications.

Density functional tight binding is a powerful approach that has been receiving wide interest from the computational chemistry community, because of its reliable description of organic molecules, as well as that of inorganic structures.^{10–15} The method gives an accurate description of the core Hamiltonian, better than the standard quantum chemistry semiempirical

approaches. Furthermore, it takes into account the charge transfer in a self-consistent way (self-consistent charge density functional tight binding), thus allowing for an accurate prediction of covalent and ionic bonds. [The self-consistent charge density functional tight binding is usually named SCC-DFTB. For the sake of simplicity, we use in the whole paper the notation DFTB for the SCC-DFTB approach.] The method has been used for the description of organic molecules as a high level approach in a QM/MM multilevel scheme,¹² or as a low-level method in QM/QM'.¹⁶ A recent extension of the method to include third-order expansion of the total energy has been shown to improve the description of hydrogen-bonding interactions and proton affinities of biological systems.¹⁷ Moreover, the time dependent extension of the method (TD-DFTB) is becoming a very powerful and reliable tool to calculate the spectroscopic properties of biological systems and nanomaterials.^{18,19} Another important field of application is in solvation effects and simulations of molecules in solution.²⁰

In all of these issues, the calculation of the forces is very challenging, and the availability of explicit, analytic gradients is in most cases an invaluable tool in speeding up the procedure. This motivated us to do an analytic fitting of DFTB parameters for applications to nanomaterials.

In this paper, an analytic formulation of DFTB parameters for silicon-based materials is provided. The first part of the paper is devoted to the description of the fitting procedure of a tabulated set of DFTB parameters involving silicon, with a presentation of the results of the fitting. In the second part, the method is applied to the description of functionalized silicon nanocrystals and is checked against more accurate results. Among the many possible applications, we show how DFTB can be used together with density functional theory in a multilevel approach, based on an ONIOM scheme,^{21,22} for a high-level description of hundreds-of-atoms functionalized silicon nanocrystals. It will be shown that the results obtained using ONIOM are very accurate, both in terms of the geometry and in terms of the energetics.

Received: October 25, 2010

Published: January 10, 2011

2. THE METHOD

Semiempirical tight binding approaches have been used for several years in the field of materials science.²³ The transferability of the parameters from the bulk to nanostructures made it a simple and transparent framework for describing the electronic properties of hundreds- to thousands-of-atoms systems.^{24,25} However, the lack of reliable interaction parameters to represent the organic molecules was the main drawback of this approach, which was unfit to model hybrid organic–inorganic structures. With the progress of synthesis techniques and the use of nanostructures in realistic biotechnological applications, the theoretical problem has moved from the description of free-standing semiconductor nanocrystals to their interaction with an external organic environment, and the use of different, more reliable computational tools was required.

Quantum chemistry semiempirical approaches, such as MNDO, AM1, and PM6,²⁶ offer an invaluable tool for structural optimizations and reaction path descriptions of thousands-of-atoms structures. More accurate than classical force field models in all situations involving the breaking and formation of chemical bonds, semiempirical methods are very useful in characterizing organic molecules and small clusters. But the parametrization of the core Hamiltonian is often oversimplified. At variance with semiempirical approaches, DFTB furnishes a more complex parametrization of the core Hamiltonian that can lead to a more accurate treatment of the excited states, especially in complex environments like semiconductor nanocrystals.

DFTB is a method derived by a second-order expansion of the exact DFT Hamiltonian with respect to the electronic charge density variations.²⁷ The method is based on a minimal parametrization of the core Hamiltonian, where the onsite parameters (onsite energies, Hubbard terms, spin–spin interaction terms) are calculated from isolated atoms.²⁸ For the two center parameters, contracted Slater-type atomic orbitals are obtained from the solution of the Kohn–Sham equation for free atoms, where a confining potential is added to mimic the presence of a molecular or solid state environment. From the Slater-type atomic orbitals, the overlap integrals are calculated as a function of the atom–atom distance. Instead, the contracted atomic densities are used to form the interatomic potential of a diatomic molecule,²⁹ whose matrix elements with the contracted orbitals give the two-center interaction terms of the core Hamiltonian.³⁰ Repulsive interatomic energy E_{rep} takes into account the exchange-correlation contributions that are not included in the core Hamiltonian, in the self-consistent charge contribution, or in the spin–spin interaction terms.¹⁵ It is written as a sum of two-body potentials V^{rep} and is obtained by comparison to DFT total energies. The parametrization of the repulsive term is a delicate point of the scheme. In particular, it strongly depends on the exchange-correlation functional chosen as a reference (GGA, B3LYP...) and determines the goodness of a DFTB parametrization. Automatic fitting schemes have been proposed in the literature, with the hope of providing a unique potential able to describe most of the features of interest (vibrational spectra, geometries, reaction path energetics, proton affinities).³¹

The method is competitive against semiempirical models for biological systems, where it has been used with a remarkable success.^{16,28,32} For semiconductor nanocrystals, DFTB keeps the advantages of semiempirical tight binding for huge structures. It has been applied for time dependent calculations of excited state properties of silicon nanocrystals.^{33,34} Recent efforts propose

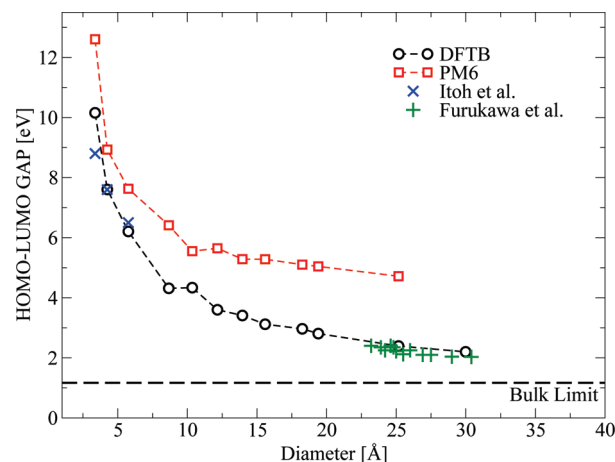


Figure 1. The energy gap of silicon nanocrystals as a function of their size. DFTB and PM6 results are respectively indicated in black circles and red squares. Experimental data for small-³⁷ and medium-size³⁸ nanocrystals are shown as blue crosses and green plus symbols. The lines are guides for the eyes.

DFTB parametrizations that can be applied as well to periodic systems, showing that the method is able to fairly reproduce the band structures.^{35,36} Thus, if well parametrized, the method is virtually able to cover a wide range of systems, from molecular systems to thousands-of-atoms structures, from nanowires to slabs to bulk systems. But, as a matter of fact, the quality of the results depends on the parametrization.

A common benchmark used to check the validity of a parametrization to describe semiconductor nanostructures is the prediction of the optical gap as a function of the size.²⁴ In Figure 1, the highest occupied molecular orbital (HOMO) to lowest unoccupied molecular orbital (LUMO) energy gap is reported for a set of silicon nanocrystals, upon increasing their size. The energy gap was calculated using DFTB and PM6. As a comparison, experimental data of optical gaps for small- and medium-size nanocrystals are shown, together with the indirect gap of bulk silicon. The calculations have been done using the Gaussian package,³⁹ with the *pb0-3* set of parameters for DFTB.⁴⁰ The nanocrystals were built cutting silicon clusters of spherical shape, with a silicon atom in the center, and terminating all the surface atoms with hydrogen.²⁵ To have an idea of the number of atoms, $\text{Si}_{705}\text{H}_{300}$ is the largest nanocrystal here considered (3 nm diameter), while $\text{Si}_{191}\text{H}_{148}$ corresponds to a 2 nm diameter nanocrystal. Figure 1 shows that this set of parameters nicely reproduces the gap energies for silicon nanocrystals. The DFTB results cross the experimental data, and they tend to the correct limit of the indirect band gap of bulk silicon. It is known that for small molecules the correlation effects can be significant, and the experimental gap, measured from the absorption threshold or photoluminescence spectra, usually differs from the HOMO–LUMO gap. But such a difference becomes negligible upon increasing the size, and for huge nanocrystals, the HOMO–LUMO gap represents quite an accurate estimation of the absorption threshold. The main result is that DFTB reproduces the right trend of the energy gap versus the size. On the contrary, the accuracy of PM6 is significantly reduced upon increasing the nanocrystal size, and for hundreds-of-atoms structures, the results are several electronvolts far from the bulk limit.

It is known from the literature that a minimal basis set can hardly reproduce the bulk silicon conduction bands, and either

an extended basis set⁴¹ or a three-center parametrization²⁴ is needed for an accurate reproduction of the band structures. Figure 1 demonstrates that, while PM6 is unable to reproduce the electronic spectra, going to a large energy gap for huge nanocrystals, DFTB can be used for UV/vis spectra calculations, since it yields the correct trend of the energy gap with the size. Nevertheless, a quantitative treatment of the excited states can be performed only after an improvement of the DFTB parametrization, which has to furnish accurate band structures for the bulk limit, as has been recently shown with TiO₂ and ZnO parametrizations.^{35,36} Work along these lines is in progress.

3. DFTB ANALYTIC PARAMETRIZATION

A version of the DFTB method was recently implemented in the Gaussian package,³⁹ based on the analytic fitting of the Hamiltonian and overlap matrix elements, as well as that of the interatomic repulsive term.^{13,15} The great advantage of using an analytic formulation of DFTB parameters consists in having a functional form for the energy gradients and Hessians, which do not need to be calculated numerically.^{32,42,43} Starting from an existing parametrization, we performed an analytic fitting of the tabulated DFTB parameters (Hamiltonian, overlap, and repulsive potential) adopting a functional representation implemented in Gaussian.

We used the *pbc-0-3* set of parameters.⁴⁰ The Si–Si parameters were initially developed by Frauenheim et al.⁴⁴ They were later modified by Sieck to include a self-consistent charge contribution.³⁰ As we show in Figure 1, this parametrization describes free-standing silicon nanocrystals well. But the method only gives qualitative results for organic molecules. We thus integrate the parameters of Si–Si and Si–X interactions, to the set already present in Gaussian,³⁹ which is more accurate in describing biological systems.¹⁶ In the original DFTB set of parameters, the overlap and Hamiltonian matrix elements are tabulated for a dense mesh of interatomic distances. Instead, the repulsive potential is represented using a collection of cubic splines. Cubic splines are not a good approach for the calculation of Hessians, because of piecewise linear functions in the second-order derivatives. For this reason, automatic procedures to calculate the repulsive potentials were recently proposed, with the use of higher-order splines, to get sufficiently smooth functions to adequately reproduce the Hessian.³¹ The use of an analytic functional form overcomes this problem, with smooth second-order derivatives.

We used the following functional form, implemented in Gaussian, for the fitting:

$$F(R) = \sum_{i=1}^{10} C_i \exp(-\alpha\beta^i R) \quad (1)$$

The fitting scheme was based on a trust region algorithm to obtain a starting estimation of the solution, followed by a Levenberg–Marquard algorithm to reach high precision results. The procedure was repeated starting from several initial configurations, in order to minimize the dependence on the starting guess. As an example of the accuracy of our results, we report in Figure 2 the overlap and Hamiltonian Si–Si *sp* matrix elements, and the repulsive potential, according to the tabulated set of parameters, and calculated using the fitted values. We also report in the insets the error, defined as the difference between the curves. The redundant functional form reported in eq 1 produces

artificial oscillations at large interatomic separation. As much as we could, we reduced the oscillating behavior near the cutoff radius (see Figure 2), but the oscillations cannot be fully deleted. A major problem related to them is the presence of a discontinuity at the cutoff radius that leads to divergences in the calculation of energy gradients. In order to solve this problem, beyond a given value (chosen at $0.85R_{\text{cutoff}}$), in the Gaussian package,³⁹ the analytic expression in eq 1 is multiplied by a polynomial function that makes the curve smoothly tend toward zero.

In Table 1, the root-mean-square errors are reported for the Si–X (X = Si, C, H, O, N) DFTB parameters. It can be observed that the fitting scheme works better for overlap and Hamiltonian matrix elements, while for the repulsive potential, the errors are slightly larger, albeit still satisfactory. We checked the analytic parameters against the tabulated ones for several molecules and found accurate geometries and total energies. We underline that the use of a functional form for the parameters does not modify the quality of the results, which are uniquely determined by the tabulated parametrization described in the literature.³⁰

4. RESULTS

In this section, a few applications of the method for functionalized silicon nanocrystals are reported. First, we studied the adsorption of organic molecules on the Si(100) surface of silicon nanocrystals with increasing size. We take as a case study the 1-amino-3-cyclopentene (ACP) adsorption, which is a significant example of [2 + 2] cycloaddition of a cyclic organic molecule with an amino group, on the Si(100) reconstructed surface.⁴⁵ The results have been much debated in the literature, especially in regard to their unexpected difference, even within density functional theory, upon changing the basis set, the exchange-correlation functional, the model used to simulate the system. Non-negligible variations of the energetics were reported when moving from pseudopotential to all-electron basis sets, from GGA to B3LYP functionals, and from cluster to periodic surface models.^{45,46} Nevertheless, the PBE0 exchange-correlation functional has been shown to give reaction path energetics comparable to MP2 results.⁴⁶ We thus performed DFT calculations using Gaussian,³⁹ with the PBE0 functional and the recently developed N07D basis set.⁴⁷

We first considered the ACP adsorption on the small nanocrystals previously studied in the literature, Si₉H₁₂ and Si₂₉H₃₂;⁴⁵ then we increased the nanocrystal size. But, instead of simply choosing a spherical shape, with lots of single-bonded silicon atoms at the surface, we chose a more compact shape and a smaller number of H terminations. The nanocrystals are fully H-passivated except for a Si–Si dimer on the (100) surface. In Figure 3, the ACP-functionalized silicon nanocrystals are shown for several sizes.

For small nanocrystals, we performed PBE0/N07D calculations, and a direct comparison with DFTB calculations was possible. The results are reported in Table 2. We found that, while DFTB correctly reproduces the geometry, the adsorption energies are more than 10 kcal mol⁻¹ too large in absolute value. The tendency of DFTB to overbind is known in the literature, as well as its accuracy in reproducing molecular structures.^{48,49} In the present case, the systematic error could be due to the pure exchange-correlation functional used as a reference in the parametrization. In order to determine the origin of the overbinding, we performed local spin-density calculations (LSDA) for small nanocrystals, using the N07D basis set, and report the results in

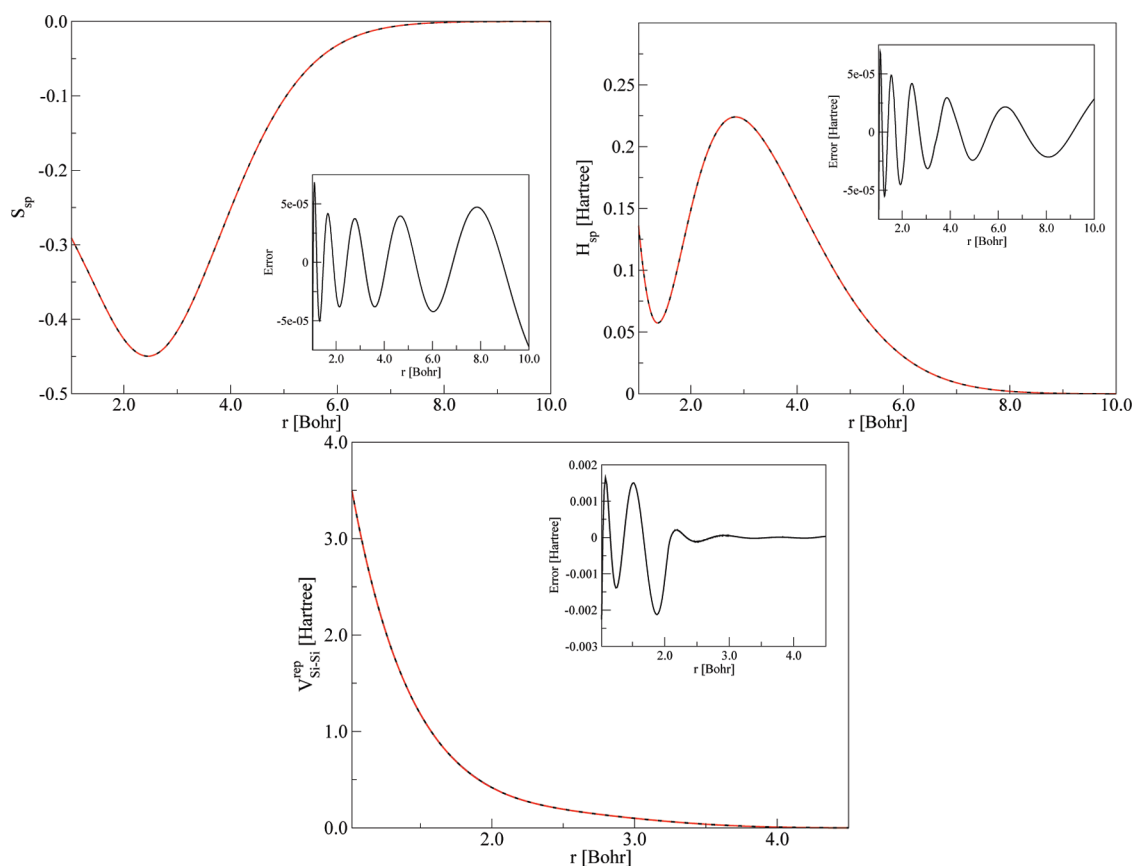


Figure 2. Some results obtained applying the fitting scheme to Si–Si parameters. The overlap and Hamiltonian matrix elements for the *sp* interaction, and the repulsive potential, are shown as a function of the interatomic distance. Black solid lines refer to the *pbc-0-3* tabulated parameters.^{30,44} Red dashed lines are the curves obtained as a result of the fitting. The errors, defined as difference between reference data and fitted values, are reported in the insets.

Table 1. Root Mean Square of the Difference between the Analytic Fitting and the Original Set of Parameters, for Overlap, Hamiltonian Matrix and Repulsive Potential (Energies in Hartree)

	Si–Si	Si–C	Si–H	Si–O	Si–N
S	2.0×10^{-5}	9.0×10^{-6}	5.1×10^{-6}	3.1×10^{-6}	3.2×10^{-6}
H^0	2.1×10^{-5}	1.0×10^{-5}	8.9×10^{-6}	9.5×10^{-6}	4.3×10^{-6}
V^{rep}	1.4×10^{-4}	5.6×10^{-4}	6.8×10^{-5}	1.3×10^{-4}	1.6×10^{-4}

the last column of Table 2. There is nice agreement with DFTB; the adsorption energies correspond within 2 kcal mol^{-1} . Thus, the DFTB error could have been inherited by the well-known LDA overbinding.^{46,50}

An interesting improvement consists in performing ONIOM-(QM:QM') calculations, with DFTB as a low-level approach and density functional theory as a high-level method, as recently proposed for the study of enzymes.¹⁶ Previous ONIOM calculations have shown that it can be a powerful approach to the description of functionalized silicon surfaces.⁵¹ We used DFT PBE0/N07D for high-level calculations on the model system and DFTB in its analytic formulation for calculations on the real system. According to the ONIOM scheme, the DFT total energy of a huge functionalized silicon nanocrystal is estimated as

$$\Delta E^{\text{ONIOM}} = \Delta E^{\text{DFTB}}(R) + \Delta E^{\text{DFT}}(M) - \Delta E^{\text{DFTB}}(M) \quad (2)$$

In the previous equation, R is the real system (ACP on the whole nanocrystal), while M is the model system, constituted by ACP and the nanocrystal active site.

A delicate point in using ONIOM approaches consists in the choice of the model system. In the present work, it is constituted by the organic molecule, ACP, and the smallest cluster here considered, Si_9H_{12} . In Figure 4, we graphically report the ONIOM scheme for ACP on $\text{Si}_{158}\text{H}_{96}$. The part used for the model system is highlighted; the remaining part of the nanocrystal is only calculated within DFTB. We report in Table 2 the results obtained for several nanocrystals. It can be noted that the ONIOM approach represents a large improvement over DFTB. For both the nanocrystals that we used for the comparison, $\text{Si}_{29}\text{H}_{32}$ and $\text{Si}_{33}\text{H}_{32}$, the relative error in the energetics decreases from about 21–24% to 1–2%, with an absolute error smaller than 1 kcal mol^{-1} . In particular, the ONIOM approach works better for the more compact nanocrystal $\text{Si}_{33}\text{H}_{32}$. It is interesting that, while DFTB adsorption energies are almost independent of the size (small to big) and the shape (decrease of H terminations), the ONIOM approach allows for the recovery of a dependence on the size and shape that is in good agreement with PBE0. Moreover, as expected, the adsorption energy changes very little for the largest nanocrystals. The adsorption energy limit has a value that is in fair agreement with PBE0 results.

In Table 3, the main structural data are reported for the optimized configurations of ACP on $\text{Si}_{29}\text{H}_{32}$. The geometries have been obtained using DFTB, PBE0/N07D, and the ONIOM

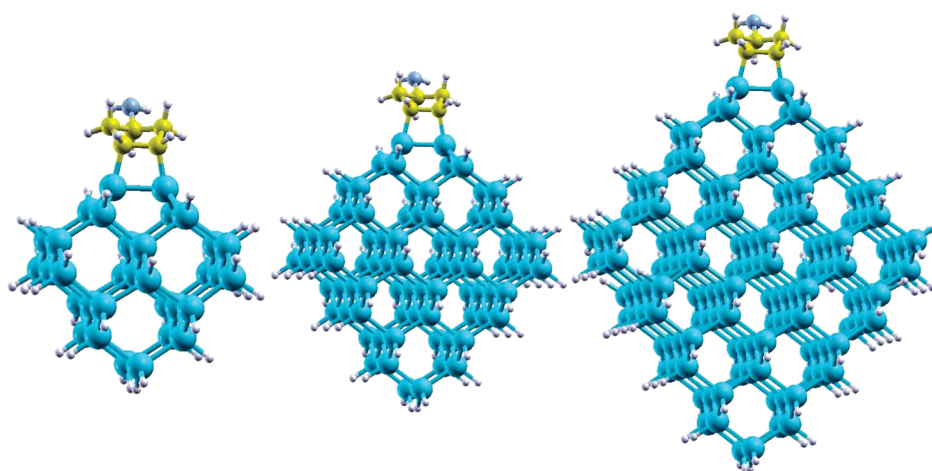


Figure 3. ACP adsorption on $\text{Si}_{33}\text{H}_{32}$, $\text{Si}_{82}\text{H}_{60}$, and $\text{Si}_{158}\text{H}_{96}$.

Table 2. Adsorption Energies for a [2 + 2] Cycloaddition Reaction of ACP on the Si(100) Surface, for Several Nanocrystals of Increasing Size [Energies in kcal mol^{-1}]

	DFTB	ONIOM	PBE0	LSDA
Si_9H_{12}	-60.99		-47.29	-59.51
$\text{Si}_{29}\text{H}_{32}$	-60.45	-47.80	-48.88	-61.12
$\text{Si}_{33}\text{H}_{32}$	-60.52	-49.33	-49.98	-62.22
$\text{Si}_{82}\text{H}_{60}$	-60.51	-50.06		
$\text{Si}_{158}\text{H}_{96}$	-60.46	-50.13		

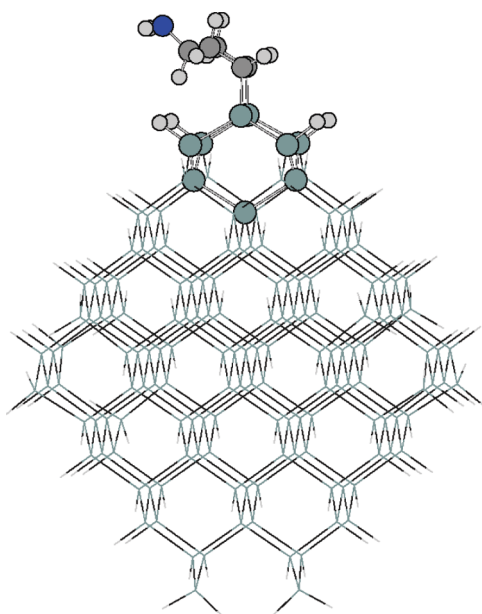


Figure 4. ACP functionalization of $\text{Si}_{158}\text{H}_{96}$. The part highlighted in the picture constitutes the model system used for ONIOM calculations.

approach previously described. In agreement with the literature, we find that DFTB is a valid tool to perform fast structural optimizations, with a fair agreement with a high-level approach. Indeed, the errors are smaller than 0.02 \AA for the lengths, and 1° on the angles, with respect to the PBE0 calculation. The slight underestimation of the Si–C distances is an artifact of the DFTB

Table 3. Optimized Geometry for the Adsorption of an ACP Molecule on $\text{Si}_{29}\text{H}_{32}$ ^a

	DFTB	ONIOM	PBE0
Si(1)–Si(2)	2.348	2.348	2.345
Si(2)–Si(3)	2.344	2.362	2.358
Si(1)–C(1)	1.947	1.956	1.959
C(1)–C(2)	1.56	1.583	1.582
Si(3)–Si(2)–Si(4)	111.6	110.6	111.1
Si(1)–C(1)–C(2)	101.7	101.3	101.2

^a The lengths are in \AA , the angles in degrees. Labeling of atoms follows Festa et al.⁴⁵

overbinding, and it could be overcome by an improved parametrization. Using the ONIOM approach as described above, the errors decrease dramatically, becoming smaller than 0.004 \AA for the distances.

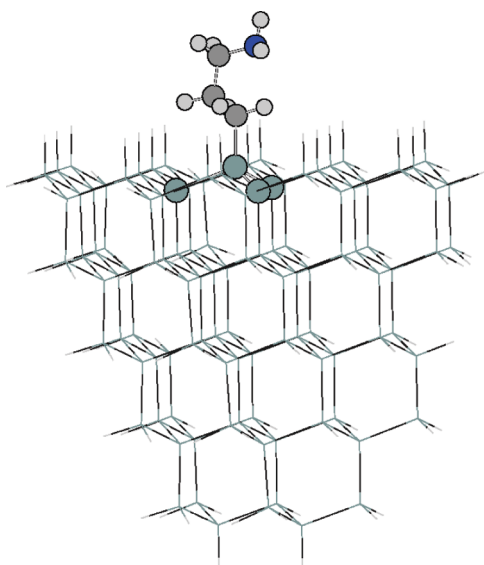
To conclude, it is worth noting that, using just a few silicon atoms for the model system, the ONIOM approach yields results comparable to full DFT/PBE0 calculations, proving geometries and energies for hundreds-of-atoms nanocrystals, with a much lighter computational effort.

As a second case, we studied the adsorption of an allylamine molecule on the Si(111) surface, whose interest for DNA sensing is huge.^{52,53} In particular, the functionalization with a propylamine chain was studied upon increasing the H-terminated Si(111) surface area and depth. At variance with the first case (ACP on silicon nanocrystals), the clusters have been modeled to be almost conically shaped, in order to represent a piece of a surface. Just like before, we studied the smallest cluster, Si_4H_{10} , then $\text{Si}_{16}\text{H}_{28}$, and, finally, a hundreds-of-atoms cluster, $\text{Si}_{109}\text{H}_{88}$. The results are reported in Table 4.

Also in this system, there is an overbinding of DFTB with respect to PBE0/N07D, with a large relative error. We performed LSDA calculations, to see whether in this case there is agreement with DFTB, too. As reported in Table 4, the trend is confirmed, but in this case, DFTB results show a larger overbinding than LSDA. In order to understand the role of the different interaction terms in DFTB, we performed some calculations on small systems, from which it emerges that the main source of error could rely on the parametrization of the Si–C repulsive term.

Table 4. Reaction Energies for the Functionalization of a Si(111) Surface by Propylamine [Energies in kcal mol⁻¹]

	DFTB	ONIOM	PBE0	LSDA
Si ₄ H ₁₀	-54.48		-32.95	-43.47
Si ₁₆ H ₂₈	-54.47	-32.88	-32.29	-47.61
Si ₁₀₉ H ₈₈	-54.78	-33.54		

**Figure 5.** Functionalization of Si₁₀₉H₈₈ with a propylamine chain. Illustration of the model system in ONIOM.

Just like before, we applied an ONIOM(DFT:DFTB) approach, where both the chain and the minimal cluster (Si₄H₁₀) form the model system. Figure 5 illustrates the use of ONIOM in this situation, and the conical shape that we used to model the surface. In this case, the number of atoms in the model system is so small that the ONIOM calculations are extremely fast. Even in this system, the ONIOM approach largely corrects the DFTB overbinding, with an absolute error smaller than 1 kcal mol⁻¹.

5. CONCLUSIONS

An analytic parametrization of DFTB has been proposed for Si-based materials, on the grounds of the DFTB method implemented in the Gaussian computational suite. The method allows the description of thousands-of-atoms structures, in particular, the organic functionalization of silicon surfaces and nanocrystals. As an illustration of the method, we applied the method to the adsorption of ACP on the (100) surface of Si nanocrystals, and the propylamine functionalization of Si(111) surfaces. We found an overbinding in the energetics of about 10 kcal mol⁻¹ and more than 20 kcal mol⁻¹, respectively. Nevertheless, using an ONIOM(DFT:DFTB) approach with only a few atoms close to the nanocrystal active site, good agreement with DFT/PBE0 results is gained, with a drastic decrease of the relative error, and an accuracy better than 1 kcal mol⁻¹ on the adsorption energies. Work along an improvement of the parametrization, and the use of the ONIOM approach for the study of huge functionalized silicon nanostructures, is in progress.

AUTHOR INFORMATION

Corresponding Author

*E-mail: fabio.trani@sns.it

ACKNOWLEDGMENT

The authors thank Giacomo Prampolini, for support on the fitting scheme. Part of the calculations have been performed using the CASPUR advanced computing facilities.

REFERENCES

- (1) Warner, J.-H.; Hoshino, A.; Yamamoto, K.; Tilley, R. D. *Angew. Chem., Int. Ed.* **2005**, *44*, 4550–4554.
- (2) He, Y.; Kang, Z.-H.; Li, Q.-S.; Tsang, C. H. A.; Fan, C.-H.; Lee, S.-T. *Angew. Chem., Int. Ed.* **2009**, *48*, 128.
- (3) Erogbogbo, F.; Yong, K.-T.; Roy, I.; Xu, G.; Prasad, P. N.; Swihart, M. T. *ACS Nano* **2008**, *2*, 873–878.
- (4) Rosso-Vasic, M.; Spruijt, E.; van Lagen, B.; De Cola, L.; Zuilhof, H. *Small* **2008**, *4*, 1835–1841.
- (5) Lin, V. S.-Y. *Nat. Mater.* **2009**, *8*, 252–253.
- (6) Park, J.-H.; Gu, L.; von Maltzahn, G.; Ruoslahti, E.; Bhatia, S. N.; Sailor, M. J. *Nat. Mater.* **2009**, *8*, 331–336.
- (7) Reboredo, F. A.; Galli, G. *J. Phys. Chem. B* **2005**, *109*, 1072–1078.
- (8) Puzder, A.; Williamson, A. J.; Grossman, J. C.; Galli, G. *Phys. Rev. Lett.* **2002**, *88*, 097401.
- (9) Li, Q. S.; Zhang, R. Q.; Lee, S. T.; Niehaus, T. A.; Frauenheim, T. *J. Chem. Phys.* **2008**, *128*, 244714.
- (10) Frauenheim, T.; Seifert, G.; Elstner, M.; Niehaus, T.; Kohler, C.; Amkreutz, M.; Sternberg, M.; Hajnal, Z.; Di Carlo, A.; Suhai, S. *J. Phys.: Condens. Matter* **2002**, *14*, 3015.
- (11) Aradi, B.; Hourahine, B.; Frauenheim, T. *J. Phys. Chem. A* **2007**, *111*, 5678–5684.
- (12) de M. Seabra, G.; Walker, R. C.; Elstner, M.; Case, D. A.; Roitberg, A. E. *J. Phys. Chem. A* **2007**, *111*, 5655–5664.
- (13) Zheng, G.; Witek, H. A.; Bobadova-Parvanova, P.; Irlle, S.; Musaev, D. G.; Prabhakar, R.; Morokuma, K.; Lundberg, M.; Elstner, M.; Köhler, C.; Frauenheim, T. *J. Chem. Theory Comput.* **2007**, *3*, 1349–1367.
- (14) de M. Seabra, G.; Walker, R. C.; Roitberg, A. E. *J. Phys. Chem. A* **2009**, *113*, 11938–11948.
- (15) Zheng, G.; Lundberg, M.; Jakowski, J.; Vreven, T.; Frisch, M. J.; Morokuma, K. *Int. J. Quantum Chem.* **2009**, *109*, 1841–1854.
- (16) Lundberg, M.; Sasakura, Y.; Zheng, G.; Morokuma, K. *J. Chem. Theory Comput.* **2010**, *6*, 1413–1427.
- (17) Yang, Y. H.; York, D.; Cui, Q.; Elstner, M. *J. Phys. Chem. A* **2007**, *111*, 10861–10873.
- (18) Niehaus, T. A.; Suhai, S.; Della Sala, F.; Lugli, P.; Elstner, M.; Seifert, G.; Frauenheim, T. *Phys. Rev. B* **2001**, *63*, 085108.
- (19) Cui, G.; Fang, W.; Yang, W. *Phys. Chem. Chem. Phys.* **2010**, *12*, 416–421.
- (20) Hou, G.; Zhu, X.; Cui, Q. *J. Chem. Theory Comput.* **2010**, *6*, 2303–2314.
- (21) Dapprich, S.; Komaromi, I.; Byun, K.; Morokuma, K.; Frisch, M. *THEOCHEM* **1999**, *461–462*, 1–21.
- (22) Vreven, T.; Byun, K. S.; Komaromi, I.; Dapprich, S.; Montgomery, J. A.; Morokuma, K.; Frisch, M. J. *J. Chem. Theory Comput.* **2006**, *2*, 815–826.
- (23) Delerue, C.; Lannoo, M. *Nanostructures, Theory and Modelling; NanoScience and Technology*; Springer: Berlin, 2004; pp 31–38; 60–70; 112–128; 197–200.
- (24) Niquet, Y. M.; Delerue, C.; Allan, G.; Lannoo, M. *Phys. Rev. B* **2000**, *62*, 5109–5116.
- (25) Trani, F.; Cantele, G.; Ninno, D.; Iadonisi, G. *Phys. Rev. B* **2005**, *72*, 075423.
- (26) Stewart, J. J. *Mol. Model.* **2007**, *13*, 1173–1213.

- (27) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. *Phys. Rev. B* **1998**, *58*, 7260–7268.
- (28) Zheng, G.; Irle, S.; Morokuma, K. *Chem. Phys. Lett.* **2005**, *412*, 210–216.
- (29) Porezag, D.; Frauenheim, T.; Köhler, T.; Seifert, G.; Kaschner, R. *Phys. Rev. B* **1995**, *51*, 12947–12957.
- (30) Sieck, A. Ph.D. thesis, University of Paderborn, 2000. <http://ubdata.uni-paderborn.de/ediss/06/2000/sieck> (accessed Dec 06, 2010), pp 26–30;109–120.
- (31) Gaus, M.; Chou, C.-P.; Witek, H.; Elstner, M. *J. Phys. Chem. A* **2009**, *113*, 11866–11881.
- (32) Witek, H. A.; Irle, S.; Morokuma, K. *J. Chem. Phys.* **2004**, *121*, 5163.
- (33) Li, Q. S.; Zhang, R. Q.; Lee, S. T.; Niehaus, T. A.; Frauenheim, T. *Appl. Phys. Lett.* **2008**, *92*, 053107.
- (34) Wang, Y.; Zhang, R.; Frauenheim, T.; Niehaus, T. A. *J. Phys. Chem. C* **2009**, *113*, 12935.
- (35) Moreira, N.; Dolgonos, G.; Aradi, B.; da Rosa, A.; Frauenheim, T. *J. Chem. Theory Comput.* **2009**, *5*, 605–614.
- (36) Dolgonos, G.; Aradi, B.; Moreira, N. H.; Frauenheim, T. *J. Chem. Theory Comput.* **2010**, *6*, 266–278.
- (37) Itoh, U.; Toyoshima, Y.; Onuki, H.; Washida, N.; Ibuki, T. *J. Chem. Phys.* **1986**, *85*, 4867–4872.
- (38) Furukawa, S.; Miyasato, T. *Phys. Rev. B* **1988**, *38*, 5726–5729.
- (39) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian Development Version*; Gaussian Inc.: Wallingford, CT, 2010.
- (40) The parameters were downloaded from dftb.org (accessed Dec 06, 2010).
- (41) Jancu, J.-M.; Scholz, R.; Beltram, F.; Bassani, F. *Phys. Rev. B* **1998**, *57*, 6493–6507.
- (42) Heringer, D.; Niehaus, T. A.; Wanko, M.; Frauenheim, T. *J. Comput. Chem.* **2007**, *28*, 2589.
- (43) Hratchian, H. P.; Parandekar, P. V.; Raghavachari, K.; Frisch, M. J.; Vreven, T. *J. Chem. Phys.* **2008**, *128*, 034107.
- (44) Frauenheim, T.; Weich, F.; Köhler, T.; Uhlmann, S.; Porezag, D.; Seifert, G. *Phys. Rev. B* **1995**, *52*, 11492.
- (45) Festa, G.; Cossi, M.; Barone, V.; Cantele, G.; Ninno, D.; Iadonisi, G. *J. Chem. Phys.* **2005**, *122*, 184714.
- (46) Cantele, G.; Trani, F.; Ninno, D.; Cossi, M.; Barone, V. *J. Phys.: Condens. Matter* **2006**, *18*, 2349–2365.
- (47) Barone, V.; Cimino, P. *Chem. Phys. Lett.* **2008**, *454*, 139–143.
- (48) Elstner, M. *J. Phys. Chem. A* **2007**, *111*, 5614–5621.
- (49) Otte, N.; Scholten, M.; Thiel, W. *J. Phys. Chem. A* **2007**, *111*, 5751–5755.
- (50) Pan, W.; Zhu, T.; Yang, W. *J. Chem. Phys.* **1997**, *107*, 3981–3985.
- (51) Santos, H.; Ujaque, G.; Ramos, M.; Gomes, J. *J. Comput. Chem.* **2006**, *27*, 1892–1897.
- (52) Barone, V.; Cacelli, I.; Ferretti, A.; Monti, S.; Prampolini, G. *Phys. Chem. Chem. Phys.* **2009**, *11*, 10644–10656.
- (53) Barone, V.; Cacelli, I.; Ferretti, A.; Monti, S.; Prampolini, G. *Phys. Chem. Chem. Phys.* **2010**, *12*, 4201–9.

Edge Stabilities of Hexagonal Boron Nitride Nanoribbons: A First-Principles Study

Rajdip Mukherjee^{*,†} and Somnath Bhowmick^{*,‡}

[†]Department of Materials Engineering, Indian Institute of Science, Bangalore, India

[‡]Materials Research Center, Indian Institute of Science, Bangalore, India

ABSTRACT: We investigate the comparative stability of sp^2 bonded planar hexagonal boron nitride (h-BN) nanoribbon (BNNR) edges, using first principles calculations. We find that the pristine armchair edges have the highest degree of stability. Pristine zigzag edges are metastable, favoring planar reconstructions [in the form of 5–7 rings] that minimizes the energy. Our investigation further reveals that the pristine zigzag edges can be stabilized against 5–7 reconstructions by passivating the dangling bonds at the edges by other elements, such as hydrogen (H) atoms. Electronic and magnetic properties of nanoribbons depend on the edge shapes and are strongly affected by edge reconstructions.

1. INTRODUCTION

The area of research on two-dimensional material has seen tremendous growth in the past few years, driven by fundamental physics as well as potential next generation device applications. Although graphene has been the frontrunner until now,^{1,2} single layer hexagonal boron nitride (h-BN) and its hybrids with graphene have drawn a great deal of attention of late.^{3–6}

Similar to graphene,^{7,8} single (or a few) layers of h-BN sheets have been prepared using micromechanical cleaving and chemically derived routes.^{9,10} One-dimensional nanoribbons of graphene¹¹ and h-BN¹² have also been successfully synthesized. Due to novel electronic and magnetic properties,^{1,2,13–16} nanoribbons are of great interest. On the basis of the shape, edges of nanoribbons can be classified as zigzag and armchair, as shown in Figure 1a and b, respectively. Electronic and magnetic properties of the nanoribbons depend on edge shapes.^{1,4,5} Thus, determining the stability of different edges is crucial for the purpose of technological applications of nanoribbons in future generation devices.

Structural instabilities at the edges and resulting reconstructions are well-known in graphene nanoribbons.^{17–22} Certain reconstructions are found to induce compressive stress along the edge, which is released by warping of the nanoribbon.²² The stability of various different edges of graphene nanoribbons can be controlled by experimental conditions and depend on the nature of edge passivation.¹⁹ As reported by Koskinen et al.,¹⁸ edge reconstructions are also self-passivating for metastable unpassivated zigzag graphene nanoribbons and lower the edge energy. Surprisingly, the stability of BNNR edges has hardly been discussed in the literature. Although Ding et al.²³ have studied the pristine BNNR edges as a function of H-passivation, to the best of our knowledge, the possibility of self-passivating edge reconstructions in h-BN nanoribbons remains unexplored.

In this work, we report a comparative study of the stability of different edges of h-BN nanoribbons. Other than the regular pristine armchair and zigzag edge, we also consider a reconstructed edge, made of 5–7 rings. Such reconstructions create homoelemental B–B and N–N bonds [see Figure 1c], which are higher

in energy than the B–N bonds and thus unfavorable.²⁴ Hence, unlike graphene nanoribbons (see previous paragraph), 5–7 reconstructions are not guaranteed to lower the edge formation energy of unpassivated zigzag BNNR, and we investigate whether they can be self-passivating in the latter. Other than the self-passivating reconstructions, nanoribbon edges can also be stabilized by saturating the dangling bonds with other elements, such as H atoms.¹⁹ Our study also reveals how H-passivation affects the stability of various different BNNR edges: pristine and 5–7 reconstructed.

This paper is organized as follows. We describe computational details in section 2. The main findings of our work are reported in section 3. First, we discuss the edge stability of nanoribbons: unpassivated BNNRs in section 3.1 and H-passivated BNNRs in section 3.2. We present a comparison between graphene and h-BN nanoribbons in section 3.3. Keeping in mind the importance of the electronic and magnetic properties of nanoribbons, we briefly describe how they are affected by the edge reconstructions in section 3.4. The paper is concluded in section 4.

2. METHOD

We use first-principles calculations as implemented in the PWSCF code,²⁵ with a plane-wave basis set and ultrasoft pseudo-potential, and the electron exchange-correlation is treated within a generalized gradient approximation (GGA), as described by Perdew et al.²⁶ We use an energy cutoff for the plane-wave basis for wave functions (charge density) of 40 (400) Ry. Nanoribbons are simulated using a supercell geometry, with a vacuum layer of ~ 15 Å between any two periodic images of the BNNR. A k -point grid of $1 \times 12 \times 1$ k points (periodic direction of the ribbon along the y axis) is used for sampling Brillouin zone integrations. The length of periodicity L [$3a$, $(3a)^{1/2}$, and $2 \times (3a)^{1/2}$ for pristine armchair, pristine zigzag, and edge reconstructed zigzag, respectively] is fixed according to the equilibrium lattice parameter of h-BN ($a = 1.45$ Å). We allow the structure to fully relax

Received: November 6, 2010

Published: January 10, 2011

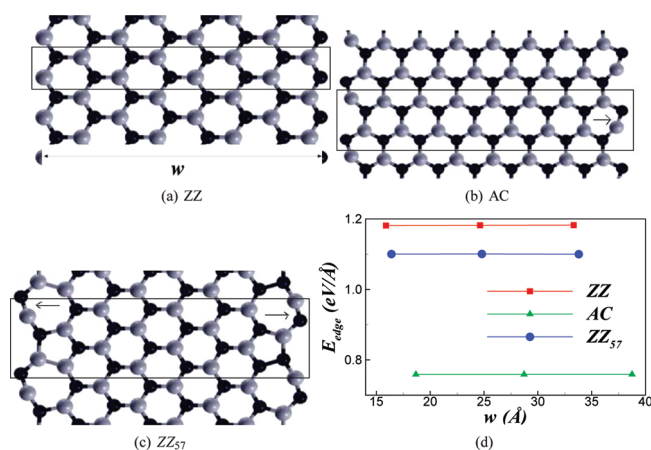


Figure 1. Unpassivated (a) pristine ZZ and (b) pristine AC h-BN nanoribbon of width w . The ribbons are periodic in the direction perpendicular to w . The rectangular box represents the repeat unit of each nanoribbon. N (B) atoms are shown in black (gray). Arrows mark the triple bonds (length ~ 1.3 Å). (c) Edge reconstructed zigzag h-BN nanoribbon— ZZ_{57} ; 5–7 defects create homoelemental B–B and N–N bonds. (d) Edge formation energy [see eq 1 and Table 1] as a function of w ; E_{edge} does not depend on ribbon width. Pristine AC nanoribbon has the lowest edge formation energy, followed by ZZ_{57} and pristine ZZ.

until the force on every atom is less than 10^{-3} Ry/au and total energy changes are smaller than 10^{-4} Ry.

We analyze the stability of edges by comparing the edge formation energy per unit length, defined as

$$E_{\text{edge}} = \frac{1}{2L} \left[E_{\text{tot}} - N_{\text{BN}} E_{\text{BN}} - \frac{N_{\text{H}}}{2} E_{\text{H}_2} \right] \quad (1)$$

where E_{tot} , E_{BN} , and E_{H_2} are the total energy of the nanoribbon supercell, the energy of a BN pair in bulk h-BN, and the energy of the H_2 molecule; N_{BN} (N_{H}) is the number of BN pairs (H atoms) in the nanoribbon [$N_{\text{H}} = 0$ for the unpassivated BNNRs]. L is the periodic length along the ribbon axis, and the factor 2 in the denominator accounts for the two edges present per repeated unit. Since BNNRs have asymmetric edges [other than the armchair nanoribbon; for example, compare Figure 1a and b], periodic boundary condition gives rise to an artificial electric field across the width of the ribbon. Dipole correction, as implemented in the PWSCF code,²⁷ is employed to cancel this artificial field and calculate the correct total energy of the ribbon, E_{tot} . Stress along the edge or periodic direction is reported as $\sigma = V\sigma_{yy}/L$, where σ_{yy} is the diagonal element of the stress tensor in the y direction (defined along the ribbon axis) and V is the volume of the supercell. We use +ve and -ve σ to denote compressive and tensile stress, respectively. σ_{yy} is calculated using the Nielsen–Martin algorithm,²⁸ as implemented in the PWSCF code.

3. RESULTS AND DISCUSSIONS

3.1. Unpassivated BNNRs. The relaxed structures of unpassivated pristine zigzag (ZZ), pristine armchair (AC), and edge-reconstructed zigzag (ZZ_{57}) ribbons are shown in Figure 1a,b,c, respectively. The subscript 57 indicates that the edge reconstructions take place via the formation of 5–7 rings or Stone–Wales (SW) defects.²⁹ Such topological defects form in a honeycomb lattice by 90° rotation of B–N (C–C) bonds in h-BN³⁰ (graphene³¹) and are relevant for their mechanical behavior.

Table 1. Edge Formation Energy and Stress (Along the Ribbon Edge) of h-BN Nanoribbons^a

ribbon	w (Å)	E_{edge} (eV/Å)	σ (eV/Å)
ZZ	15.89	1.1813	0.30
	24.63	1.1819	0.30
	33.33	1.1825	0.32
AC	18.67	0.7587	0.25
	28.72	0.7588	0.26
	38.75	0.7590	0.27
ZZ_{57}	16.40	1.1003	-4.97
	24.80	1.1006	-5.07
	33.80	1.1008	-5.08
ZZ^{H}	18.15	0.1217	-0.21
	26.85	0.1218	-0.22
	35.54	0.1218	-0.22
AC^{H}	20.77	0.1164	-0.19
	30.84	0.1172	-0.20
	40.86	0.1177	-0.20
ZZ_{57}^{H}	18.70	0.6400	-3.70
	27.43	0.6405	-3.76
	36.16	0.6401	-3.85
ZZ_{57}^{H}	18.67	0.9309	-2.55
	27.37	0.9321	-2.57
	36.07	0.9327	-2.67

^a See eq 1 and text thereafter for the definition of E_{edge} and σ .

As mentioned previously, SW defects create energetically unfavorable homoelemental (B–B and N–N) bonds in h-BN²⁴ [see Figure 1c]. The edge formation energies, calculated using eq 1, are plotted in Figure 1d and reported in Table 1. E_{edge} (as well as σ) hardly changes as the ribbon width increases by a factor of 2. This gives us confidence that the numerical values reported here for those two parameters are indeed well converged.

We find that pristine armchair BNNR has lower edge formation energy than that of pristine zigzag. Similar to bare edge graphene nanoribbons,^{18,19} triple bonds are formed at the armrests [shown by the arrow in Figure 1b], which is evident from their bond length of 1.30 Å, 10% shorter than the sp^2 bonds of the two-dimensional h-BN. Zigzag BNNR cannot form triple bonds and have a higher formation energy due to the existence of expensive dangling bonds at the edges. As shown in Figure 1c, triple bonding (length 1.30 Å) is also observed in ZZ_{57} , which lowers the edge formation energy by 0.08 eV/Å, compared to that of pristine zigzag BNNR. Thus, we conclude that *the gain due to triple bond formation (which eliminates the dangling bonds at the edges) is more than the energy expense of unfavorable homoelemental bond creation, making ZZ_{57} relatively favorable than the pristine ZZ BNNR.* However, AC is the most stable of the BNNRs, its edge formation energy being 30% smaller than that of ZZ_{57} .

As reported in Table 1, pristine BNNRs are under little compressive stress. It is well-known that compressive stress is relieved by wrinkle formation in graphene nanoribbons.^{22,32} However, we have found that the stress is too small to show any significant out of plane deformation in pristine BNNRs. On the other hand, 5–7 reconstructions induce great tensile stress along the ribbon edge, which ensures the planarity of such BNNRs.

3.2. H-Passivated BNNRs. Nanoribbon edges can be stabilized by saturating the dangling bonds with H atoms (or any other molecule in general). Note that there is no further scope of

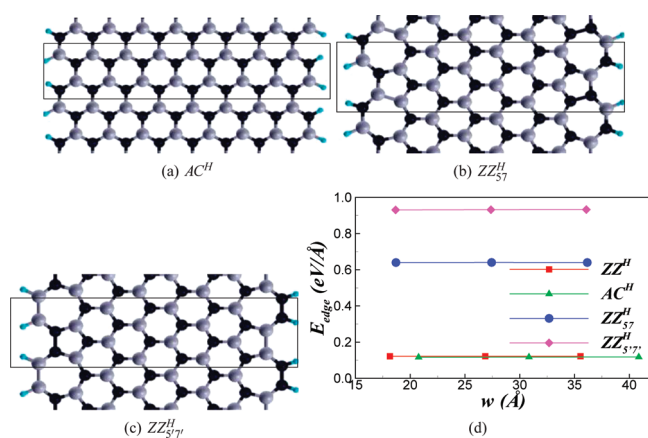


Figure 2. (a) Hydrogen passivated pristine AC^H h-BN nanoribbon. Note that H termination leads to smoother edges [compare with Figure 1b]. The shape of the pristine ZZ^H h-BN nanoribbon is very similar to that in Figure 1a. Two types of reconstructions, ZZ₅₇^H and ZZ_{57'}^H, are shown in b and c, respectively. The latter has twice the number of homoelemental B–B and N–N bonds per repeat unit than the former. (d) Edge formation energy [see eq 1 and Table 1] as a function of ribbon width w ; E_{edge} does not depend on ribbon width. H passivation stabilizes the pristine zigzag BNNR, edge formation energy being nearly equal to that of the pristine armchair BNNR, and reconstructions are unfavorable.

triple bonding for the edge atoms (of both armchair and zigzag nanoribbons), as they form sp² bonds with two neighboring B (or N) atoms and a H atom each. This eliminates the difference in the nature of chemical bonding of pristine zigzag and armchair edge atoms (explained in the previous section for unpassivated BNNRs), which is manifested in respective edge formation energies, being nearly equal in ZZ^H and AC^H [see Figure 2d and Table 1]. The superscript H is used to mark the hydrogen-terminated BNNRs.

We observe that H termination reduces the edge formation energy and changes σ to small tensile stress in pristine BNNRs [see Table 1]. Comparing Figure 2a and Figure 1b, it is clear that H termination makes the equilibrium shape of the armchair BNNR edges smoother. Bare and H-passivated pristine zigzag BNNR have a similar equilibrium edge shape, and we do not illustrate the structure of the latter in the paper. We study two types of edge reconstructions in H-passivated zigzag BNNR, ZZ₅₇^H and ZZ_{57'}^H, as shown in Figure 2b and c. Both the reconstructions induce tensile stress along the ribbon edge [see Table 1], and the resulting nanoribbons have planar geometry. We have verified that the 5'–7' reconstruction is not stable in unpassivated BNNR. Though both are made of 5–7 rings, ZZ_{57'}^H has a higher E_{edge} , albeit having a smaller σ (costing lesser strain energy) than that of ZZ₅₇^H [consult Figure 2d and Table 1]. Higher edge formation energy can be attributed to the higher linear density of unfavorable B–B and N–N bonds in the former [see Figure 2b and c; ZZ_{57'}^H has double the number of homoelemental bonds per repeat unit than ZZ₅₇^H]. Comparing the values of E_{edge} [Figure 2d and Table 1], we conclude that unlike the bare zigzag BNNR, edge reconstructions are not favorable in H-terminated zigzag BNNR; 5–7 (5'–7') reconstruction increases the edge formation energy by ~ 5.5 (8) times. Thus, *edge passivation offers a simple way to stabilize the pristine zigzag edge against 5–7 reconstructions in BNNRs.*

3.3. Comparison with Graphene Nanoribbons. Graphene and h-BN have several similarities. Not only do they have

Table 2. Magnetic Property and Band Gap (E_g) of the BNNRs Shown in Figure 1 and 2

BNNR	magnetic property	E_g (eV)
ZZ	magnetic	metallic
AC	nonmagnetic	4.8
ZZ ₅₇	nonmagnetic	2.7
ZZ ^H	nonmagnetic	5.0
AC ^H	nonmagnetic	5.5
ZZ ₅₇ ^H	nonmagnetic	2.6
ZZ _{57'} ^H	nonmagnetic	1.7

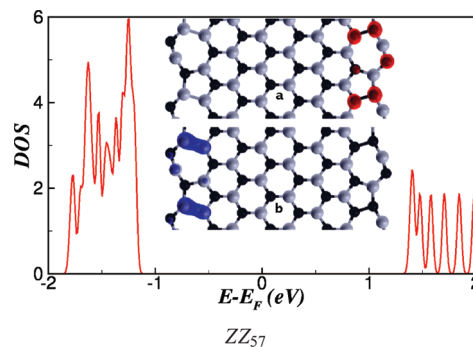


Figure 3. Unpassivated, edge reconstructed BNNR densities of state (DOS) and band decomposed electron densities of (a) the top of the valence band and (b) the bottom of the conduction band, localized mainly at the N–N and B–B bonds, created by 5–7 defects. We have used XCRYSDEN³⁵ to plot the charge densities.

identical structures (constituent atoms are arranged in a honeycomb lattice) but also their values of cohesive energy (in-plane stiffness) are within ~ 5 ($\sim 18\%$) of each other.³³ Thus, it is worth comparing the edge formation energies and reconstructions of graphene and h-BN nanoribbons.

In the case of unsaturated nanoribbons, we observe two major differences. First, the 5–7 reconstructed edge has the lowest formation energy, followed by the pristine armchair edge in graphene nanoribbons.^{18,19} As shown here, the sequence happens to be opposite in unpassivated BNNRs. Second, self-passivating 5–7 reconstruction in bare zigzag graphene is more effective than that in the bare zigzag h-BN nanoribbon. While E_{edge} of the reconstructed zigzag graphene nanoribbon is $\sim 17\%$ ^{18,19} smaller than that of the bare pristine zigzag graphene nanoribbon, we find that similar edge reconstruction leads to an energy gain of $\sim 6\%$ [see Table 1] in the unpassivated zigzag h-BN nanoribbon. Such distinctions can be understood from the formation energy of the SW defect in graphene and h-BN: ~ 5 eV³¹ and ~ 6 – 6.5 eV,³⁴ respectively. As mentioned previously, higher formation energy in the latter is due to energetically unfavorable B–B and N–N bonds, created by the SW defect in h-BN.²⁴

Due to a similar reason, between the H-saturated zigzag BNNR and graphene, edge reconstructions are more unfavorable in the former. Comparing the edge formation energies, we find that 5–7 reconstructions augment E_{edge} by -450% to 700% in H-terminated zigzag BNNR [see Table 1], larger than a 300% increase in the H-passivated zigzag graphene nanoribbon.¹⁹ Since pristine BNNRs are devoid of homoelemental bonds, they show similar behavior to graphene nanoribbons on H passivation, which reduces the edge formation energy by $\sim 90\%$ in both

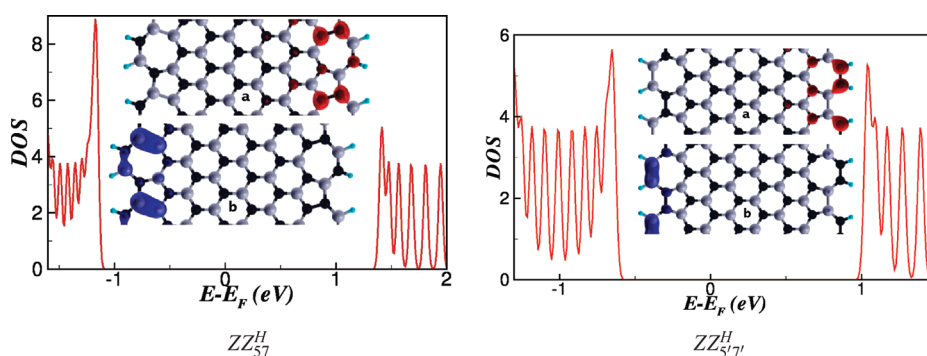


Figure 4. H-passivated, edge-reconstructed BNNR densities of state (DOS) and band decomposed electron densities. In the inset, we show charge density of (a) the top of the valence band and (b) the bottom of the conduction band, localized mainly at the N–N and B–B bonds, created by 5–7 defects. This appears to be a general phenomenon associated with 5–7 defects and also observed in unpassivated BNNRs [see Figure 3].

the systems [compare the values reported in Table 1 and by Wassmann et al.¹⁹].

3.4. Electronic and Magnetic Properties. Finally, we briefly discuss the electronic and magnetic properties of BNNRs, summarized in Table 2. Other than the unpassivated pristine zigzag, all of the BNNRs are nonmagnetic and have band gaps ranging from 1.7 to 5.5 eV. The magnetic moment originates from the dangling bonds at the edges. Our results are in good agreement with those of Barone and Peralta,¹³ where the authors have discussed the electronic and magnetic properties of pristine BNNRs in great detail. In this paper, we rather restrict ourselves to the discussion of electronic and magnetic properties of edge-reconstructed BNNRs.

We find that 5–7 reconstruction is detrimental for the edge magnetic moment. In pristine unpassivated zigzag BNNR, the magnetic moment is found to be 1 Bohr magneton per edge atom, which vanishes completely in ZZ_{57} . Similar behavior has been observed in edge-reconstructed graphene nanoribbons.^{17,22} We show the electronic densities of state (DOS) of ZZ_{57} BNNR in Figure 3. While the pristine unpassivated zigzag BNNR is metallic, 5–7 defects open up a gap of 2.7 eV. The inset of Figure 3 illustrates the charge density of the top of the valence band and bottom of the conduction band. Interestingly, they are confined at N–N and B–B bonds, created by the 5–7 defects. Thus, such sites are expected to play a major role in electronic applications and the chemical reactivity of edge-reconstructed BNNRs. As shown in Figure 4, similar behavior is observed in H-passivated, edge-reconstructed BNNRs also. While pristine H-passivated BNNR has a band gap of 5.0 eV, 5–7 edge reconstructions reduces it by at least 50% or more.

4. SUMMARY AND CONCLUSIONS

We have investigated the edge stability and emerging electronic and magnetic properties of h-BN nanoribbons. In terms of dangling bond saturations, we divide BNNRs into two groups, unpassivated and H-passivated. Among unsaturated BNNRs, armchair edges are found to have minimum formation energy, followed by ZZ_{57} . In both of the ribbons, edge atoms form triple bonds, replacing the dangling bonds and minimizing the edge formation energy in the process. Atoms located at the bare zigzag edges cannot form triple bonds, and ZZ BNNR has the highest E_{edge} due to the presence of dangling bonds. However, 5–7 rings create energetically unfavorable homoelemental B–B and N–N bonds, and edge reconstructions of bare zigzag BNNRs are not as effective as they are in bare zigzag graphene nanoribbons. Saturating

the dangling bonds at the edges by H not only minimizes the edge formation energies but also imparts stability to the pristine zigzag edge against 5–7 reconstructions.

Controlling the edge shape is important for the application of nanoribbons in devices because electronic and magnetic properties depend on the geometry of the edges. We find that 5–7 reconstructions destroy the edge magnetism observed in pristine unpassivated zigzag BNNR. In H-passivated ZZ BNNR, such reconstructions reduce the band gap. Our results show that, depending on edge shape, BNNRs can be magnetic or nonmagnetic and metallic or semiconducting, having a moderate to wide band gap. Such a rich collection of properties shows how promising h-BN nanoribbons are for application in future generation electronic devices.

Note added: After the completion of the paper, we came to know about a similar work³⁶ discussing the stability of pristine zigzag and armchair edges of h-BN sheets, and the results are in good agreement with our findings. However, the authors of the other paper did not consider edge reconstructions, which we discuss in great detail.

AUTHOR INFORMATION

Corresponding Author

*E-mail: rajdip@platinum.materials.iisc.ernet.in (R.M.); bsomnath@mrcc.iisc.ernet.in (S.B.).

ACKNOWLEDGMENT

We thank SERC, IISc for providing computational facilities. S.B. thanks Umesh V. Waghmare for useful discussions on stability and reconstructions of nanoribbon edges.

REFERENCES

- (1) Neto, A. H. C.; Guinea, F.; Peres, N. M. R.; Novoselov, K. S.; Geim, A. K. *Rev. Mod. Phys.* **2009**, *81*, 109.
- (2) Geim, A. K.; Novoselov, K. S. *Nat. Mater.* **2007**, *6*, 183–191.
- (3) Golberg, D.; Bando, Y.; Huang, Y.; Terao, T.; Mitome, M.; Tang, C.; Zhi, C. *ACS Nano* **2010**, *4*, 2979–2993.
- (4) Blase, X.; Rubio, A.; Louie, S. G.; Cohen, M. L. *Phys. Rev. B* **1995**, *51*, 6868–6875.
- (5) Park, C.-H.; Louie, S. G. *Nano Lett.* **2008**, *8*, 2200–2203. PMID: 18593205.
- (6) Ci, L.; Song, L.; Jin, C.; Jariwala, D.; Wu, D.; Li, Y.; Srivastava, A.; Wang, Z. F.; Storr, K.; Balicas, L.; Liu, F.; Ajayan, P. M. *Nat. Mater.* **2010**, *9*, 430–435.

- (7) Novoselov, K. S.; Jiang, D.; Schedin, F.; Booth, T. J.; Khotkevich, V. V.; Morozov, S. V.; Geim, A. K. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 10451–3.
- (8) Rao, C. N. R.; Sood, A. K.; Subrahmanyam, K. S.; Govindaraj, A. *Angew. Chem.* **2009**, *48*, 7752–7777.
- (9) Pacilé, D.; Meyer, J. C.; Girit, Ç. Ö.; Zettl, A. *Appl. Phys. Lett.* **2008**, *92*, 133107.
- (10) Han, W.-Q.; Wu, L.; Zhu, Y.; Watanabe, K.; Taniguchi, T. *Appl. Phys. Lett.* **2008**, *93*, 223103.
- (11) Han, M. Y.; Özyilmaz, B.; Zhang, Y.; Kim, P. *Phys. Rev. Lett.* **2007**, *98*, 206805.
- (12) Chen, Z.-G.; Zou, J.; Liu, G.; Li, F.; Wang, Y.; Wang, L.; Yuan, X.-L.; Sekiguchi, T.; Cheng, H.-M.; Lu, G. Q. *ACS Nano* **2008**, *2*, 2183–2191.
- (13) Barone, V.; Peralta, J. E. *Nano Lett.* **2008**, *8*, 2210–2214.
- (14) Chen, W.; Li, Y.; Yu, G.; Li, C.-Z.; Zhang, S. B.; Zhou, Z.; Chen, Z. *J. Am. Chem. Soc.* **2010**, *132*, 1699–1705.
- (15) Wu, X.-j.; Wu, M.-h.; Zeng, X. *Front. Phys. China* **2009**, *4*, 367–372.
- (16) Yao, W.; Yang, S. A.; Niu, Q. *Phys. Rev. Lett.* **2009**, *102*, 096801.
- (17) Huang, B.; Liu, M.; Su, N.; Wu, J.; Duan, W.; Lin Gu, B.; Liu, F. *Phys. Rev. Lett.* **2009**, *102*, 166404.
- (18) Koskinen, P.; Malola, S.; Häkkinen, H. *Phys. Rev. Lett.* **2008**, *101*, 115502.
- (19) Wassmann, T.; Seitsonen, A. P.; Saitta, A. M.; Lazzeri, M.; Mauri, F. *Phys. Rev. Lett.* **2008**, *101*, 096402.
- (20) Girit, C. O.; Meyer, J. C.; Erni, R.; Rossell, M. D.; Kisielowski, C.; Yang, L.; Park, C.-H.; Crommie, M. F.; Cohen, M. L.; Louie, S. G.; Zettl, A. *Science* **2009**, *323*, 1705–1708.
- (21) Gass, M. H.; Bangert, U.; Bleloch, A. L.; Wang, P.; Nair, R. R.; K., G. *Nature Nanotechnol.* **2008**, *3*, 676–681.
- (22) Bhowmick, S.; Waghmare, U. V. *Phys. Rev. B* **2010**, *81*, 155416.
- (23) Ding, Y.; Wang, Y.; Ni, J. *Appl. Phys. Lett.* **2009**, *94*, 233107.
- (24) Yuge, K. *Phys. Rev. B* **2009**, *79*, 144109.
- (25) Giannozzi, P.; et al. *J. Phys.: Condens. Matter* **2009**, *21*, 395502.
- (26) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- (27) Bengtsson, L. *Phys. Rev. B* **1999**, *59*, 12301–12304.
- (28) Nielsen, O. H.; Martin, R. M. *Phys. Rev. B* **1985**, *32*, 3780–3791.
- (29) Stone, A. J.; Wales, D. J. *Chem. Phys. Lett.* **1986**, *128*, 501–503.
- (30) Bettinger, H. F.; Dumitrică, T.; Scuseria, G. E.; Yakobson, B. I. *Phys. Rev. B* **2002**, *65*, 041406.
- (31) Lusk, M. T.; Carr, L. D. *Phys. Rev. Lett.* **2008**, *100*, 175503.
- (32) Shenoy, V. B.; Reddy, C. D.; Ramasubramaniam, A.; Zhang, Y. W. *Phys. Rev. Lett.* **2008**, *101*, 245501.
- (33) Topsakal, M.; Ciraci, S. *Phys. Rev. B* **2010**, *81*, 024107.
- (34) Chen, W.; Li, Y.; Yu, G.; Zhou, Z.; Chen, Z. *J. Chem. Theory Comput.* **2009**, *5*, 3088–3095.
- (35) Kokalj, A. *J. Mol. Graphics Model.* **1999**, *17*, 176–179.
- (36) Huang, B.; Lee, H.; Gu, B.; Liu, F.; Duan, W. arXiv: 1011.6010v1, 2010.

Molecular Dynamics Simulations of a Binding Intermediate between FKBP12 and a High-Affinity Ligand

Lilian Olivieri and Fabrice Gardebien*

DSIMB, INSERM, U665, Paris, F-75015, France, DSIMB, INSERM, U665, Faculté des Sciences et Technologies, Université de la Réunion, 15, avenue René Cassin, BP 7151 97715 Saint Denis Messag Cedex 09, La Réunion, France, INTS, Paris, F-75015, France, and Laboratoire de Biochimie et Génétique Moléculaire, Université de la Réunion, 15, avenue René Cassin, BP 7151 97715 Saint Denis Messag Cedex 09, La Réunion, France

S Supporting Information

ABSTRACT: We characterized a binding intermediate between the protein FKBP12 and one of its high-affinity ligands by means of molecular dynamics simulations. In such an intermediate, which is expected to form at the end-point of the bimolecular diffusional search, short-range interactions between the molecular partners may play a role in the specificity of recognition as well as in the association rate. Langevin dynamics simulations were carried out to generate the intermediate by applying an external biasing force to unbind the ligand from the protein. The intermediate was then refined by seven independent molecular dynamics simulations performed with an explicit solvent model. We found consistent results both for the structure of the protein and for the position of the ligand in the intermediate. The two carbonyl oxygens O2 and O3 of the ligand core region act as two main anchors, making permanent contacts in the intermediate. The transient contacts with the protein are made by the ligand noncore moieties whose structures and mobilities enable many alternative contacts of different types to be formed: π - π molecular overlap and weak hydrogen bonds $\text{NH}\cdots\pi$, $\text{CH}\cdots\pi$, and $\text{CH}\cdots\text{O}$. Hence, the stability of the ligand at the entrance of the protein binding pocket offers the possibility of fine-tuning a variety of short-range contacts that involve the ligand noncore moieties. Under the hypothesis that the stability of this intermediate is related to the affinity of the ligand, this binding intermediate model comes closest to explaining the role played by the noncore moieties in the affinity of this ligand. Moreover, this model also provides a plausible explanation for how structurally diverse core motifs that all share the carbonyl atoms O2 and O3 bind to FKBP12.

1. INTRODUCTION

The protein FKBP12 is a cytosolic enzyme of 12 kDa that catalyzes the peptidylprolyl *cis*-*trans* isomerization. FK506, a fungal metabolite, can bind tightly to FKBP12 with an inhibition constant of 0.6 nM (FKBP stands for FK506 binding protein).¹ Rapamycin is another ligand that has a similar inhibition constant (0.3 nM).¹ The immunosuppressive properties of rapamycin and FK506 are currently used for the treatment of transplant rejection.^{2,3} In the brain, the expression level of FKBP12 is much higher than in immune tissues, and in neurons, the FK506-bound form of the protein has been associated with neuroprotective properties.⁴ Nerve regenerative properties are also induced by FK506 or analogues and involve the isoform FKBP52, whose N-terminal FK506 binding domain is very similar in sequence and structure to FKBP12.⁵ Much effort is directed toward finding FK506 analogues with neuroprotective and neurotrophic activities but devoid of the undesirable immunosuppressive activity that is functionally associated with the ligand region responsible for calcineurin inhibition. The high-affinity ligand **8**,⁶ bound to the protein FKBP12 in Figure 1, is an example of such non-immunosuppressive ligands (the ligand is labeled as in the work of Holt et al.).

The development of inhibitors of the protein FKBP12 represents a major interest for extending the potential of many therapeutic treatments. To this end, a detailed understanding of the structure-activity relationships of the FKBP12 ligands is critically

important.^{7,8} To address this problem, we have focused on a complex formed by FKBP12 and the ligand **8** that has an inhibition constant of 10 nM, making it a good model for understanding the molecular basis of its high affinity for this protein. Instead of exploring the structure and dynamics of the bound state, however, this study was intended to find a binding intermediate (or encounter complex) that is expected to form at the end-point of the bimolecular diffusional search. In this intermediate state, short-range interactions could play a crucial role for the specificity of recognition as well as for the association rate.⁹ As the association constant for a protein-ligand complex is the ratio between the association (k_{on}) and dissociation (k_{off}) rate constants, the binding properties of a ligand are thus dependent on k_{on} and/or k_{off} . Under the hypothesis that an intermediate is formed along the binding pathway, the overall rate of association k_{on} is increased, i.e., the affinity, whenever the intermediate has a slow dissociation rate or a fast association rate for the binding.^{10,11} In a NMR study of two phosphopeptides that bind to a mutant of the N-terminal SH2 domain of PI3-K,¹² Mittag et al. have shown the existence of an intermediate state along the association pathway. Interestingly, the lifetime of the intermediate is related to the affinity of the ligand: the higher the affinity of the ligand, the longer is the lifetime of the intermediate (in this case, the slow dissociation rate of the

Received: July 16, 2010

Published: January 31, 2011

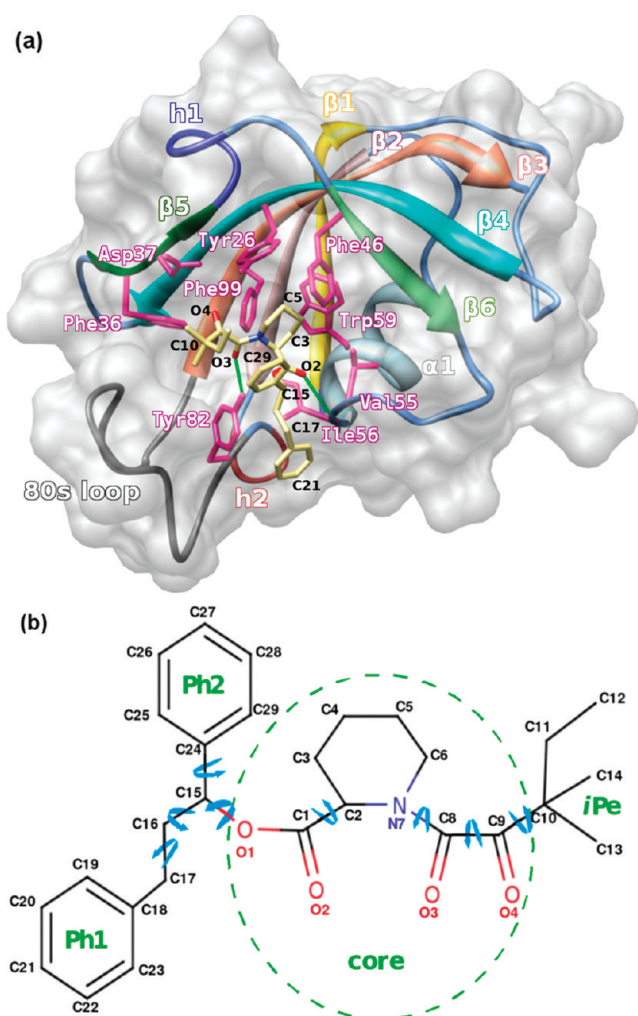


Figure 1. (a) Experimental structure of the complex between FKBP12 and the high-affinity ligand **8** (compound numbering is from ref 6). The β sheet and the α_1 helix are highlighted. The two native hydrogen bonds Ile56-NH \cdots O2 and Tyr82-OH \cdots O3 are also represented by thick green lines. (b) Sketch of the ligand **8** that indicates all of the atom numbers and the eight dihedral angles that were parametrized (highlighted by blue arrows). The core and the noncore regions *iPe* (C10 up to C14), Ph1 (C18–C23), and Ph2 (C24–C29) are also defined.

intermediate may explain the high affinity). The authors have also stressed that such intermediates may be more common in protein–ligand systems than previously anticipated. In the context of protein–protein association, an experimental kinetic study has correlated the decrease in affinity between Ras and a Ras-binding domain with the increase in the overall dissociation rate constant.¹³ Schreiber et al. have used a computer algorithm to predict the association rate between TEM1 β -lactamase and its inhibitor BLIP.¹⁴ The authors have found that charge mutations in the inhibitor increase the association rate, resulting in higher affinities.

In the case of the complex formed by FKBP12 and the high-affinity ligand **8**, the properties of a binding intermediate along the association pathway could help to delineate the essential structural features of this ligand that are important for the recognition step. Two observations made for FKBP12 and its ligands have prompted the search of a binding intermediate. The first observation is the lack of correlation between the structural features

of the various FKBP12 ligands and their measured binding affinities. For example, many core regions of high-affinity ligands, which represent the regions that strongly interact with the protein, are structurally diverse.^{6,15–17} To further complicate the structure–affinity relationships, even for the ligands that are sharing the same diketo-pipecolinic acid core (from the bond C9O4 up to the atom O1 in Figure 1b), it is difficult to explain the trends in the affinity data.⁶ The reasons are two-fold: first, the core regions of these ligands in the complexes are nearly superimposable, displaying the same contacts with the protein, and second, the noncore regions of the ligand often make few contacts with the protein.^{6,18} Hence, the structure–affinity relationships of both the core and the noncore regions of the high-affinity ligands of FKBP12 still await clarification. In this work, we address these issues by focusing on the case of the high-affinity ligand **8**.

Prior to presenting the approach chosen, let us now consider the second observation that has also motivated this work and that is related to the structure of the binding pocket of FKBP12. In the crystal structure of FKBP12–**8**, the ligand **8** is deeply buried with only 30% of its surface exposed to the solvent: the pipecolinyl ring is facing the indole ring of Trp59, and the methyl group C13 of the isopentyl moiety (*iPe*) is lying in a small adjacent cavity lined by the side chains of Tyr82, His87, Ile90, and Ile91. If one considers a rigid docking of ligand **8** to a structure of FKBP12 that is the same as the bound form of the protein, then the binding would be prevented by a large steric hindrance. For a binding mechanism that would occur in such a single step, one may therefore anticipate a large energy barrier of complexation that would make this single step process kinetically inaccessible. Indeed, a positive enthalpic term and a negative entropic contribution would result from the steric hindrance and from the constraints to fit, in only one step, the pipecolinyl ring and the *iPe* moiety of the ligand into the main pocket and the small side cavity, respectively. The results gathered from the docking of the ligand FK506 to a rigid protein FKBP12 are consistent with the above analysis.¹⁹ In this rigid docking study, the author failed to identify a geometry close to the experimentally reported structure of FKBP12–FK506 because of unfavorable steric overlaps. Moreover, these docking results indicate that the relaxation of the protein in the loop region 82–95, also termed the 80s loop, helps to accommodate the ligand in the binding pocket. Hence, during the binding process of FK506 or related ligands such as **8**, a displacement of the 80s loop would facilitate the binding to FKBP12. This hypothesis is also supported by a comparison between solution structures of unliganded FKBP12 and a crystal structure of FKBP12–FK506.²⁰ Taken together, the experimental and the docking studies^{19–21} suggest that, during the binding, the 80s loop undergoes conformational transitions to facilitate the entrance of the ligand by reducing the steric hindrance.

The successive events such as the 80s loop displacement, the first contacts with the ligand, and the formation of the fully bound structure may require only one step, all of these events being concerted; however, these may also proceed in two steps with the formation of a binding intermediate as a first binding event. This intermediate along the association pathway would allow decomposition of the process in a two-step mechanism in order to reduce the overall free energy barrier due both to steric hindrance and to an entropic penalty, as discussed above, thereby increasing the kinetic accessibility to the bound state.¹⁰ Our study has two objectives: (i) to locate a binding intermediate on the free energy surface of FKBP12 and its high-affinity ligand **8** and (ii), by

analyzing this intermediate, to provide a possible guide for understanding the role played by the ligand core and noncore regions in the specificity of recognition. A ligand loosely trapped in a proper orientation at the entrance of the protein binding pocket has translational and rotational motions that are substantially reduced, thereby allowing the ligand to fine-tune its short-range interactions with a more limited region of the protein surface. In particular, the extent and the nature of the intermolecular contacts in the intermediate may be related to its stability (large contact numbers favoring slow dissociation rate), and thus to the binding properties of the FKBP12 ligand. The analysis of this intermediate may also provide clues regarding the diversity of the core structures found in the high-affinity ligands.

To locate this intermediate, Langevin dynamics (LD)²² and an analytical model for the solvation free energy EEF1²³ were used in combination. As a preliminary step toward the search of this intermediate, we assessed the ability of this simulation protocol LD/EEF1 to reproduce accurately the experimental structure of the complex FKBP12–8. LD/EEF1 simulations were then carried out to unbind the ligand. To this end, an external force was applied to pull the ligand out of the binding pocket. This force was chosen as low as possible to ensure a minimal perturbation of the system. The structure of the intermediate state located by the LD simulations was subsequently refined by stochastic boundary molecular dynamics (SBD)²⁴ simulations that focused only on the binding region of the protein.

Under low-force conditions, forced unbinding simulations and spontaneous unbinding likely proceed through similar mechanisms.²⁵ Hence, forced unbinding simulations represent a new tool to investigate the structure–activity relationship.^{7,8,26,27} However, the present study is an original approach that focuses on the structural characteristics of a protein–ligand system in an early stage of recognition. We shall begin by describing the simulation procedure in section 2 (other computational details are provided in the Supporting Information). In section 3.1, the validation step of the LD/EEF1 methodology that was carried out on the bound state is presented. All of the results for the binding intermediate are presented and discussed in sections 3.2 and 4, respectively. These results provide a consistent picture of the binding intermediate with a well-defined position of the ligand that displays both permanent and transient contacts with the protein, thus revealing important aspects of its recognition properties in this early binding stage.

2. METHODS

2.1. Molecular Dynamics Simulations of the Bound State.

The X-ray coordinates of the complex between the protein FKBP12 and the synthetic inhibitor 8 (PDB code: 1FKG;⁶ Figure 1) were used for the simulations of the bound state. The all-atom Charmm22 force field²⁸ was used for the protein and the ligand; however, a few parameters for the ligand were calculated by following the procedure defined by Foloppe and MacKerell²⁹ (see the Supporting Information for details). All of the molecular dynamics simulations that used an implicit or explicit solvation model were performed with the CHARMM program³⁰ using the leapfrog integrator and a time step of 1 ps. SHAKE³¹ restraints were applied to the bonds containing hydrogen atoms. Langevin dynamics (LD)²² were used to simulate the bound state. To reproduce the effects of the water collision frequency at 300 K,³² the friction coefficient was chosen at 60 ps⁻¹. Effective Energy Function 1 (EEF1)²³ was chosen to compute the solvation free energy of the atoms in the LD simulations of both the bound state and

the unbinding process (see the Supporting Information for the parameters).

For the LD simulations of the bound state, the complex was energy-minimized and then heated from 0 to 300 K in steps of 25 K (for 600 ps); different random seeds for the distribution of initial velocities led to four independent simulations. The heating run was followed by an equilibration divided into two periods of 500 ps each (see the Supporting Information for details on the minimization, heating, and equilibration procedures). After the equilibration, the LD production runs were conducted with five NOE restraints between the residue pairs Glu31–Thr96 and Lys34–Ile90 and for a total simulation time of 15.5 ns: 4, 3, 5, and 3.5 ns for the four runs LD1, LD2, LD3, and LD4, respectively. It should be noted that the NOE restraints were used during the LD simulations to avoid a distortion of the 80s loop due to the absence of explicit water molecules (see the Supporting Information).

To compare the implicit solvation model with an explicit model, three independent stochastic boundary molecular dynamics (SBD)²⁴ simulations of the bound state were also carried out. For the SBD simulations of the bound state, the protein and the ligand were immersed in a sphere of water molecules of 25 Å radius (see the Supporting Information for the construction of this sphere). This sphere size is sufficient to provide at least one hydration shell around all of the regions of the complex. For the SBD runs, the complex was energy-minimized and then heated from 0 to 300 K using a step of 10 K (for 30 ps). Three independent heating runs were each followed by an equilibration divided into two periods of 300 and 150 ps, respectively (see the Supporting Information for details). After the equilibration, the SBD production runs were conducted with no restraints for 5 ns (total of 15 ns).

Prior to the unbinding simulations, we checked that the LD simulations of the bound state provided results in agreement with both the X-ray structure and the SBD simulations. The purpose in using LD simulations for the description of the bound state is two-fold: (i) to validate the force field and solvation parameters that are also used in the unbinding simulations and (ii) to obtain initial structures for the unbinding simulations. Implicit solvent simulations are required for the unbinding step, as has been recommended for unfolding or unbinding events.³³

2.2. Unbinding Simulations. The protocol used to pull the ligand out of the binding pocket and to obtain metastable states is shown in Figure 2. This protocol started with 18 structures (labeled “BS from LD”) extracted from the four LD simulations of the bound state. As the unbinding is a long-extended process,³⁴ a pulling force was added to the molecular potential energy function for each unbinding simulation. The time-dependent perturbation implemented in the Biased Molecular Dynamics (BMD)³³ method was used. It is noteworthy that, among the perturbation methods that have been recently compared, BMD is the method that causes the least perturbation in a system.³⁵ In this method, a quadratic time-dependent perturbation is introduced in the system only when the distance between the ligand atom O1 and the C α atom of Glu5 decreases (d_{RC} ; see the Supporting Information for details). Otherwise, no external perturbation is applied, and the increased distance d_{RC} is taken as the new reaction coordinate for the next time step.

In the initial BMD simulations (step 1 in Figure 2), the constant of the pulling force (parameter α) was chosen at 300 pN/Å (4.32 kcal/mol/Å²), since this value enabled the unbinding of the ligand within a time limit of 5 ns. The unbinding process was

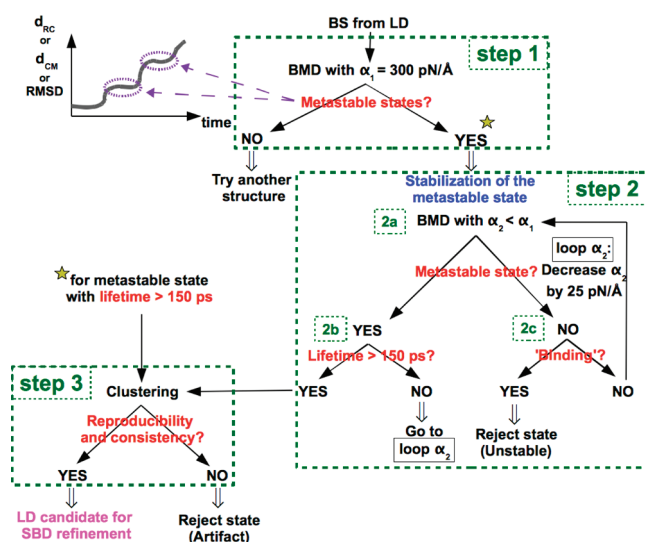


Figure 2. Protocol to find the LD candidates for the binding intermediate. Step 1: Unbinding simulations on each of the 18 structures separated by 1 ns and selected from the four LD simulations of the bound state. The star above “YES” indicates that the metastable states with a lifetime greater than 150 ps are also added to the set of simulations used in the clustering step (step 3). Step 2: In 2a, for each metastable state from step 1, BMD simulations are performed with α_2 lower than α_1 , to extend the lifetime of the metastable state found in step 1. In 2b and 2c, metastable states with lifetime > 150 ps are added to the pool used in step 3; otherwise, a lower α_2 is used for restarting the simulations. “Binding” checks whether the ligand returns to its position in the bound state (BS). Step 3: The metastable states with a lifetime greater than 150 ps are clustered into subsets that exhibit related structural features (see text). Reproducibility and consistency within a cluster are critical in finding the LD candidates for further refinement.

followed by monitoring the time evolutions of d_{RC} and of d_{CM} ; the latter represents the distance between the center of mass of the binding pocket and each of the four centers of mass of the four ligand moieties (core, *iPe*, Ph1, and Ph2). This procedure was also followed by Curcio et al. in their study of forced unbinding of fluorescein from anti-fluorescein antibody FITC-E2.²⁵ We also monitored the root-mean-square deviations (RMSDs) for both the protein and the ligand moieties (core, *iPe*, Ph1, and Ph2) with the C_α atoms aligned on those of the X-ray structure of the bound state. From the unbinding simulations (step 1 in Figure 2), many metastable states were found by relying on the time evolutions of d_{RC} , d_{CM} , and RMSDs; constant mean values for these distances and RMSDs for periods of hundreds of picoseconds constitute useful indicators.

In step 2 of the protocol, all of the metastable states found in step 1 were used to perform BMD simulations with smaller pulling forces (stage 2a in Figure 2). The states from step 1 with a lifetime greater than 150 ps were used both in stage 2a and in the clustering step of the protocol (step 3; *vide infra*). The purpose in using smaller values of α (in the range 50–250 pN/Å) in step 2 is to decrease the pulling force, thereby increasing the sampling of the potential energy surface for the ligand coordinates in the vicinity of the binding pocket. Hence, simulations with a longer lifetime were obtained that exhibited relatively stable values of d_{RC} , d_{CM} , and RMSDs for the protein and for the ligand. At stages 2b and 2c in Figure 2, the simulations with short lifetimes (<150 ps) were restarted by gradually decreasing the α_2 value (by steps of 25 pN/Å). The simulations with lifetimes greater than

Table 1. Average LD RMSDs (Å) of All of the C_α Atoms of the Protein in the Complex and Those Calculated Only for the 40s, 50s, and 80s Loops (Which Comprise the Residues 39–45, 50–56, and 82–95, Respectively) on the Basis of the Alignment of All of the C_α Atoms

residue	$\langle LD \rangle$	$\langle SBD \rangle$
all	0.86 ± 0.08	1.02 ± 0.08
39–45 (40s)	0.91 ± 0.23	1.68 ± 0.04
50–56 (50s)	0.95 ± 0.17	1.14 ± 0.59
82–95 (80s)	1.38 ± 0.32	1.57 ± 0.01

150 ps were added to the pool of states analyzed in the step 3, and those for which the ligand returned to the binding pocket were discarded.

In step 3 of Figure 2, all of the metastable states with a lifetime greater than 150 ps, obtained from both the unbinding (step 1 in Figure 2) and the simulations performed to increase the sampling (step 2 in Figure 2), were clustered into subsets. The following four criteria were used to cluster and to select the LD simulations that could model the binding intermediate. First, we checked that the trajectories were stable for periods of at least hundreds of picoseconds. Within a subset of metastable states that may represent a binding intermediate, the values derived for d_{RC} , d_{CM} , and RMSDs of the protein and of the ligand were checked to ensure that they remained constant during the course of each simulation as well as consistent among all of the simulations. For the binding intermediate that is presented in this work, the d_{CM} and RMSDs are shown in Table 2 and discussed in section 3.2. The same structural criterion was followed by Li and Daggett in the search of an intermediate state along the unfolding pathway of barnase.³⁶ Second, we checked that the structure of the protein was not significantly altered, except for the 80s loop, as it is known to be highly flexible and, more importantly, surrounds and contributes to the binding pocket. The RMSD values for the protein are presented in section 3.2 and reveal an overall low distortion, except for the 80s loop. Third, we checked that the set of LD simulations sampled a range of conformations that were structurally related and, in particular, the positions of the ligand with respect to the protein were not spread sparsely but rather formed a tight cluster. As seen in the analysis of the position of the ligand in section 3.2, a narrow ensemble of orientation is obtained for the ligand. Central to this last criterion is the reproducibility and consistency that constitute useful indicators that the results have sufficiently converged so as to provide a meaningful average picture of the intermediate. Fourth, as we hypothesize that the intermediate state is an obligatory step from the freely diffusing molecules toward the bound state, we also checked that the intermediate was structurally close to the bound state in relative separation and relative orientation between the two molecular partners. In the intermediate that was identified, the ligand core moiety is lying at 4.6 Å above its position in the native complex (section 3.2).

2.3. Refinement Procedure. Only three sets of LD simulations, the LD candidates, passed the acceptance criteria and were selected for the refinement procedure that used independent SBD production runs with no biasing forces or restraints on the system (see the Supporting Information for details). The purpose in using an explicit solvent representation to further characterize the intermediate state is two-fold: (i) to further probe the stability of the binding intermediate located using LD simulations and (ii) to refine the results since an explicit solvent model

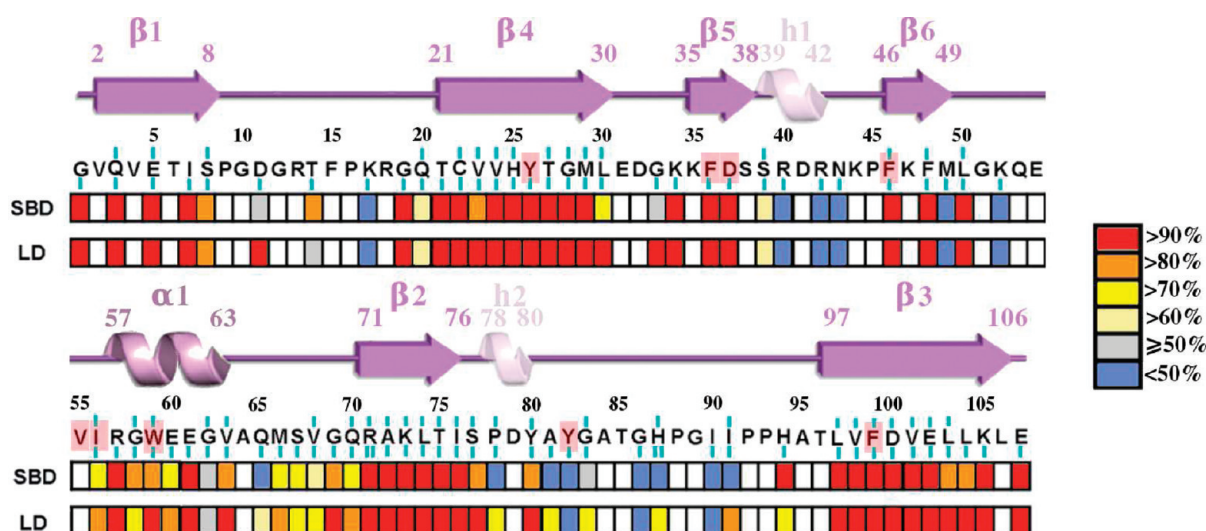


Figure 3. Contact frequencies for all of the backbone hydrogen bonds calculated as ensemble averages for the LD and SBD simulations. In the one-letter code representation of the amino acid sequence, up to three vertical bars are drawn above and below a few residues in the sequence and represent the number of native hydrogen bonds formed by their backbone atoms (through CO and/or NH). When the backbone atoms of a residue are involved in two or three hydrogen bonds, the contact frequency is calculated by averaging over these two or three contacts. For example, the color code adopted for Trp59 results from the average between the contact frequencies calculated for the two native hydrogen bonds Ile56-O \cdots Trp59-HN and Trp59-O \cdots Val63-HN. The one-letter codes of the residues that are in contact with the ligand in the crystal structure are shaded. The secondary structure elements are designated as in the work of Holt et al.⁶ This secondary structure assignment is consistent with a DSSP analysis⁵⁸ except for the regions 39–42 and 78–80, where single turn helices are predicted (helices *h1* and *h2*, respectively).

is known to provide a more realistic description of a system. For the SBD simulations of each candidate for the binding intermediate, the water molecules were added to the selected structures in a sphere of 22 Å in diameter centered on the binding site (see the Supporting Information for the construction of this sphere).

The SBD refinement step enabled us to discard two of the three LD candidates for the binding intermediate. These two states corresponded to more distant positions of the ligand from the binding pocket. And, in contrast to the binding intermediate discussed in this work, the extent of consensus within these two sets of short simulations was rather weak. Despite extensive effort to complete and refine these two sets of metastable states, the overall results did not show sufficient consensus, reproducibility, and stability. This failure to achieve sufficient consistency, either within LD runs or SBD runs or between LD and SBD results, emphasizes the importance of the third criteria used for the clustering step to provide confidence in the intermediate model derived.

It should be mentioned that the analysis of the intermediate rests on the assumption that the results from the unbinding experience are also valid for the reverse process, which is expected to be the case under the chosen conditions, i.e., low forces applied to the ligand and implicit solvent representation.^{25,33} All of the data that are reported for the bound state and the binding intermediate correspond to ensemble averages, except where otherwise indicated.

3. RESULTS

3.1. Bound State. Prior to the unbinding simulations, we tested the ability of our LD/EEF1 protocol to yield stable structures for the bound complex. The stability of the structures generated from the LD approach was evaluated by analyzing (i) the root-mean-square deviations (RMSDs) of the C_{α} atoms from

the experimental complex, (ii) the native hydrogen bonds, and (iii) the position of the ligand inside the binding pocket. Our goal in this validation step is two-fold: first, to check whether the structures generated by LD simulations are consistent with the crystal structure and, second, to compare the LD results with those of the SBD simulations that use an explicit solvent representation.

The average RMSDs of the C_{α} atoms from the experimental complex are reported in Table 1 for the LD and SBD simulations. The ensemble averages are below ~ 1 Å and are consistent with those typically seen in molecular simulations of X-ray protein structures.³⁷ Though informative, the average RMSDs of all of the C_{α} atoms provide only a global measure of the deviation from the crystal structure. Attention was also focused on the distortion of the three loops closest to the binding site, the 40s, 50s, and 80s loops (the former two loops comprise the residues 39–45 and 50–56, respectively). The RMSDs of the C_{α} atoms of these three loops are also tabulated; their values do not exceed 1.7 Å.

A total of 59 backbone hydrogen bonds were extracted from the experimental structures based on the following contact criteria: a maximum hydrogen-acceptor distance of 2.4 Å and the minimum donor-hydrogen-acceptor angle set to 90°. Figure 3 illustrates the average contact frequencies for all of the backbone hydrogen bonds in the two simulation series. Overall, the two series of LD and SBD simulations give comparable results, especially in the regions of secondary structure. As seen in Figure 3, all of the native hydrogen bonds between the five β strands are well conserved in both series of simulations (contact frequencies are above 80%). These strands form an antiparallel β sheet wrapped around an α helix of seven residues, α_1 . In the two simulation series, the structure of α_1 is preserved owing to contact frequencies higher than 75% (between Arg57–Glu61, Trp59–Val63, Ile56–Trp59, and Ile56–Glu60). In contrast, in the loop regions, four hydrogen bonds between the following pairs of residues are not conserved: Met49–Lys52 in the 50s loop,

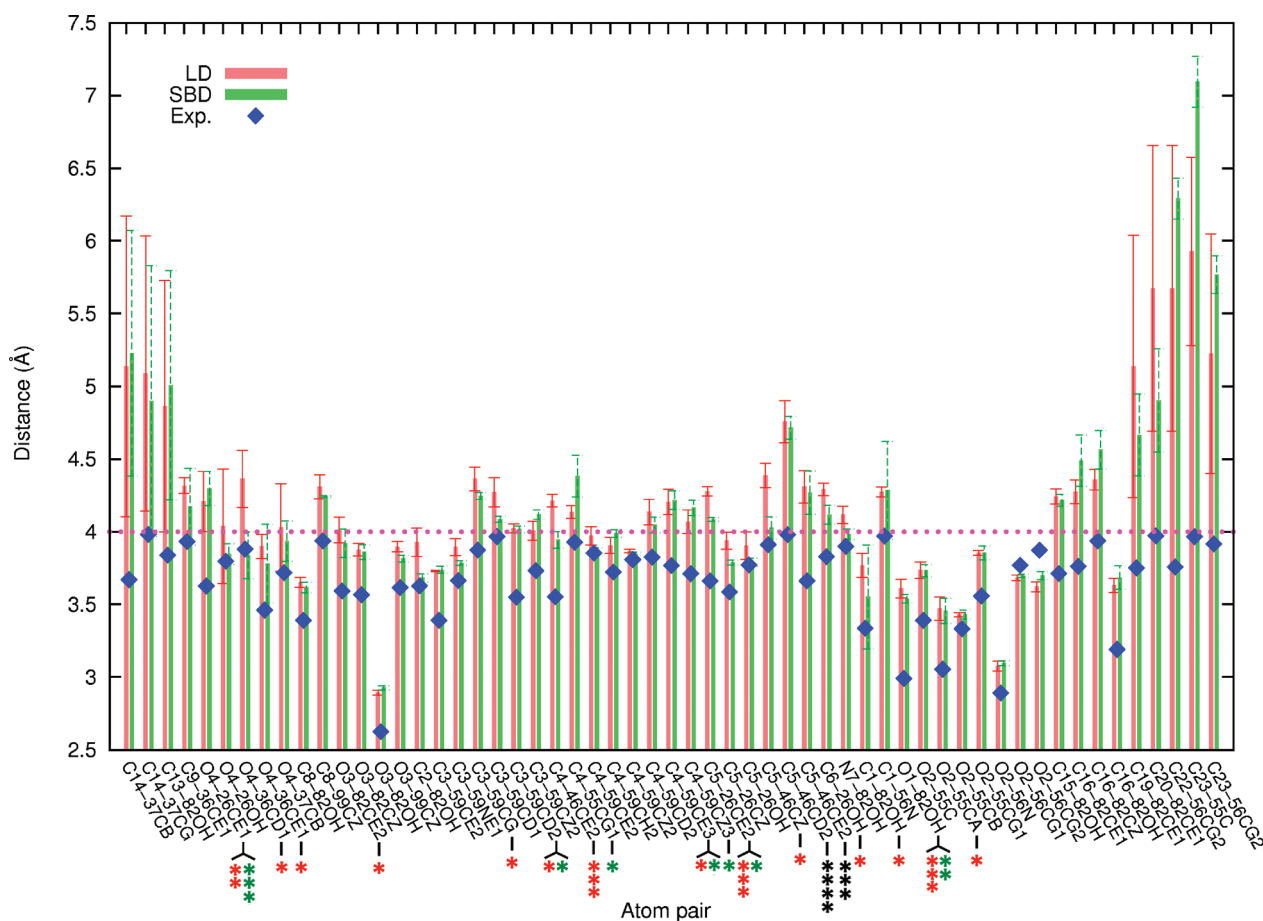


Figure 4. List of 57 separation distances between protein and ligand atoms averaged from the LD and SBD series. This list corresponds to all of the intermolecular distances found in the X-ray structure that are below 4 Å. Any average separation distance that differs from its experimental counterpart by a value found in the ranges $]0.4;0.5]$, $]0.5;0.6]$, $]0.6;0.7]$, and $]0.7;0.8]$ are indicated by one, two, three, and four stars, respectively. Red and green stars refer to the LD and SBD results, respectively, and the black stars to both.

Tyr82–Gly86 and His87–Ile90 in the 80s loop, and Lys17–Gln20. In the SBD runs only, the hydrogen bond between the residue pair Gly62–Gln65 is weakened, whereas one of the two hydrogen bonds between His87 and Ile91 in the 80s loop is lost (frequencies of 44% and 1%, respectively). Finally, no new hydrogen bonds were formed between backbone atoms in SBD simulations, whereas only one is seen in the LD simulations between Asp11–O and Arg13–NH.

Hence, only four and five of the 59 native hydrogen bonds are not reproduced by LD and SBD simulations, respectively. Two and three of these hydrogen bonds that are disrupted in LD and SBD runs, respectively, are found in the 80s loop. The failure to reproduce these few hydrogen bonds that are assumed in this region may be related to crystal packing effects. For example, in the crystal structure of the protein in complex with FK506, packing effects are known to exist for the segment 82–90.³⁸ In a quantum chemical study of the crystal structure of FKBP12–8, Nakanishi et al.³⁹ have pointed out that ligand **8** experiences packing effects by interacting with adjacent proteins in the crystal, presumably with the same residue segment 82–90 as for FK506. The authors have also derived a RMSD from the crystal structure of 0.80 Å for ligand **8** after a minimization by the quantum mechanical method FMO. This large RMSD has been ascribed to crystal packing effects. In the following, the position of the ligand inside the binding pocket is analyzed for the LD and SBD simulations.

A list of 57 atomic separation distances between the protein and the ligand were monitored to assess the stability of the ligand in the binding site. This list was obtained by considering all of the protein and ligand pairs of heavy atoms within 4 Å of each other in the X-ray structure. In the crystal structure, the diketopipercolinic acid core region makes 45 contacts with the protein; the *i*Pe group, three; and the propyl-Ph1 group, nine. The comparison of the 57 average separation distances calculated for the LD and the SBD series (Figure 4) reveals very similar results, especially for the 45 contacts made by the core region. When the separation distances for the ligand core region are compared in the two simulation series, the differences in distances are below ~ 0.25 Å (except for the distances O4···Phe36-CD1 and C5···Phe46-CZ that are longer in LD runs than in SBD runs by ~ 0.5 Å). In contrast to the core region of the ligand, large differences between the LD and the SBD results are observed for the propyl-Ph1 region (on the right of Figure 4).

In order to compare the experimental and the calculated separation distances, the magnitude of the errors that are expected in the crystal structure should be first tentatively estimated. On average, the uncertainty on the position of an atom in a X-ray structure is roughly one-fifth to one-tenth of the resolution.⁴⁰ Under this assumption, an estimated error in the range 0.4–0.8 Å can be anticipated for the separation distances. For the core region of the ligand, 16 and 9 separation distances in the LD and SBD

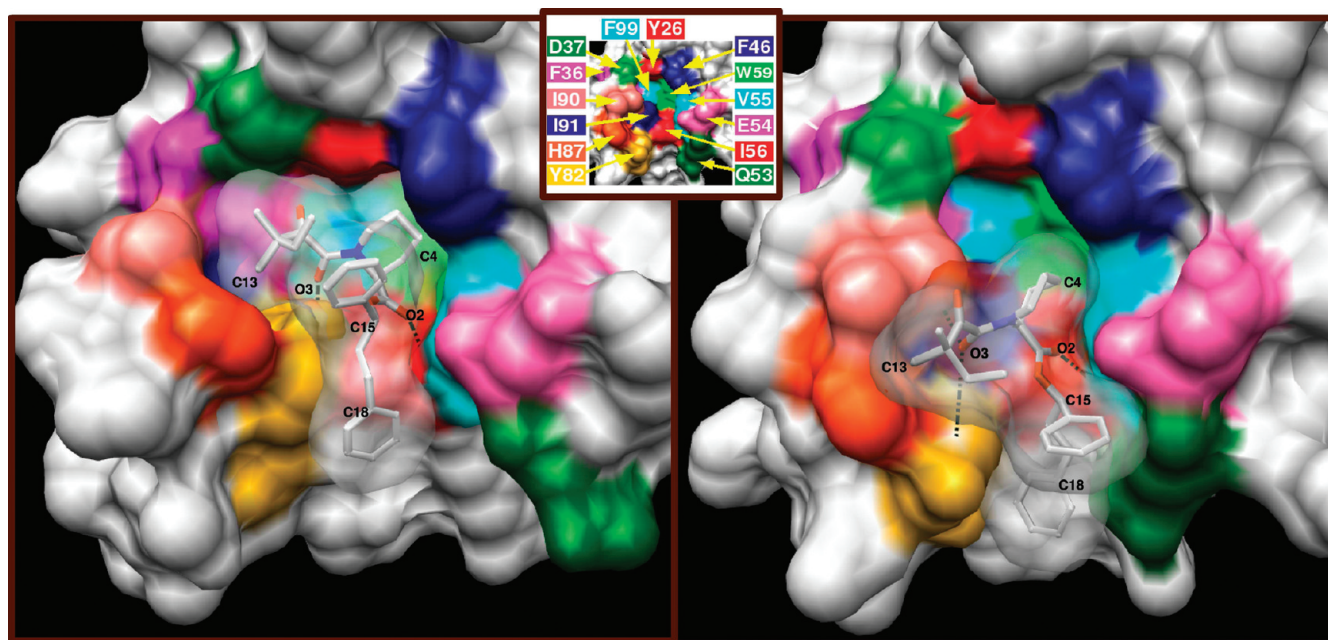


Figure 5. Comparative analysis of the position of the ligand in the binding pocket of FKBP12 in the crystal (left) and in the binding intermediate IS (right). The protein residues are color-coded according to the inset at the top of the figure. The two contacts Ile56-NH \cdots O2 and Tyr82-OH \cdots O3 are represented by dashed lines, along with the contact Ile90-CG2 \cdots O3 that is shown only for IS.

simulations, respectively, are longer than their experimental counterparts by more than 0.4 Å; the corresponding 18 distinct distances are highlighted on the x axis in Figure 4. Most of these 18 distances differ from their respective experimental values by, at most, 0.5 Å (marked with one star on the x axis). For the noncore regions *i*Pe and propyl-Ph1, large shifts from the experimental separation distances are observed; these deviations may result from the errors in the experimental positions of these exposed noncore regions that experience packing effects (*vide supra*).^{38,39} Finally, very few new contacts, with averages below 4 Å, are formed between the protein and the ligand: in the SBD simulations, each of the pairs C4 \cdots Phe46-CZ and O2 \cdots Ile56-CB have a distance of \sim 3.9 Å; in the LD simulations, the distance C25 \cdots Glu54-O is \sim 3.6 Å.

The results in terms of the RMSDs, of the native hydrogen bonds, and of the native intermolecular contacts for the LD simulations are in good agreement with the experimental data. This provides confidence in the simulation methodology that uses an implicit description of the solvent and an analytical model for the solvation free energy. Previous comparison studies between implicit and explicit solvent representations have also led to the conclusion that the former can replace the latter for many purposes.⁴¹ We have thus applied our methodology to locate a binding intermediate by pulling the ligand out of the binding pocket using a minimal force.

3.2. Binding Intermediate. Only one set of four independent LD simulations satisfied the selection criteria discussed in the section 2. These runs exhibited energetic stability and provided consistent results both for the structure of the protein and for the position of the ligand. As the goal of consistency between independent LD runs was achieved, these runs provided a reasonable guess for the subsequent refinement protocol using SBD simulations. Seven independent SBD trajectories were then obtained for a total of about 12 ns of sampling. A snapshot of the intermediate IS is

shown in Figure 5 along with the crystal structure for purposes of comparison.

In the following comparative analysis between LD and SBD results, we first check the structural integrity of the protein, the overall consistent position of the ligand, and the relative proximity of the ligand to its native binding position. Importantly, this comparative analysis will also reveal a structural consistency between these two approaches, which gives confidence in the prediction of the binding intermediate. Then, in the remainder of this section, we will discuss more specifically the SBD results obtained for the 80s loop (distortion, mobility, and hydrogen bond contact frequencies) and for the protein–ligand contacts.

Table 2 reports the RMSDs of the C_{α} atoms from the experimental complex. The ensemble average RMSDs calculated for the nonloop region of the protein are below \sim 1.1 Å for both LD and SBD runs, values that are similar to those of the protein in the bound state (Table 1). However, the RMSDs of the 80s loop indicate an overall distortion in the range \sim 3.2–3.4 Å for the two approaches. This distortion may be expected on the basis of previous works on FKBP12,^{19,20,42} as discussed hereafter. To assess its position with respect to the protein, the ligand was divided into fragments, namely, isopentyl (*i*Pe), a diketo-pipecolic acid core, and two phenyl rings (Ph1 and Ph2), as shown in Figure 1b, and the separations between the center of mass of each of these fragments and that of the protein binding pocket were checked to ensure they remained constant during the course of each simulation as well as consistent among the two sets of simulations (LD and SBD). Table 2 reports the average distances between the center of mass of the diketo-pipecolic acid core and that of the binding pocket. The ensemble average position of the core region (d_{CM}) measured from the center of the binding pocket is 7.19 and 6.37 Å in the LD and the SBD simulations, respectively. In particular, for the individual SBD simulations, these values of d_{CM} reveal a stable position of the ligand core in the individual simulation (fluctuations of \sim 0.3–0.4 Å) and a

Table 2. Average Distances (Å) between the Center of Mass of the Diketo-Pipecolic Acid Core and That of the Binding Pocket (d_{CM}) and RMSDs (Å) of the Four Ligand Moieties and of the C_{α} Atoms from the Respective X-Ray Structural Positions^a

run	pocket...core d_{CM}	RMSD of ligand				RMSD of protein			
		iPe	core	Ph1	Ph2	nonloop ^b	82–95	82–87/94–95	only 88–93
LD1	6.58 ± 0.32	8.79 ± 0.59	6.19 ± 0.42	4.45 ± 0.37	7.68 ± 0.62	1.03 ± 0.07	3.17 ± 0.13	1.66 ± 0.17	4.45 ± 0.20
LD2	7.09 ± 0.37	9.29 ± 0.58	6.69 ± 0.40	4.38 ± 0.48	8.25 ± 0.56	0.90 ± 0.06	3.27 ± 0.16	1.39 ± 0.11	4.73 ± 0.25
LD3	7.43 ± 0.31	8.51 ± 0.50	6.68 ± 0.31	3.83 ± 0.46	8.20 ± 0.54	1.26 ± 0.10	3.49 ± 0.22	3.09 ± 0.25	3.95 ± 0.24
LD4	7.67 ± 0.41	10.31 ± 0.67	7.23 ± 0.44	4.69 ± 0.47	8.82 ± 0.66	1.04 ± 0.05	3.65 ± 0.17	1.73 ± 0.17	5.21 ± 0.25
⟨LD⟩	7.19 ± 0.47	9.22 ± 0.79	6.70 ± 0.42	4.34 ± 0.36	8.24 ± 0.47	1.06 ± 0.15	3.40 ± 0.22	1.97 ± 0.76	4.58 ± 0.53
SBD1	6.41 ± 0.39	8.44 ± 0.76	6.03 ± 0.49	4.14 ± 0.61	7.10 ± 0.60	0.76 ± 0.03	3.15 ± 0.23	1.75 ± 0.30	4.34 ± 0.40
SBD2	6.48 ± 0.45	8.78 ± 0.68	6.14 ± 0.53	4.22 ± 0.61	7.07 ± 0.70	0.79 ± 0.11	3.33 ± 0.14	1.75 ± 0.30	4.65 ± 0.24
SBD3	6.36 ± 0.31	8.50 ± 0.53	5.93 ± 0.36	4.09 ± 0.43	6.95 ± 0.51	0.83 ± 0.13	3.27 ± 0.19	1.74 ± 0.31	4.56 ± 0.31
SBD4	6.42 ± 0.34	8.46 ± 0.48	6.05 ± 0.39	4.20 ± 0.48	7.39 ± 0.53	0.84 ± 0.10	3.24 ± 0.14	1.48 ± 0.16	4.64 ± 0.22
SBD5	6.44 ± 0.43	8.75 ± 0.63	6.09 ± 0.50	4.42 ± 0.53	7.04 ± 0.55	0.75 ± 0.09	3.29 ± 0.18	1.61 ± 0.22	4.66 ± 0.29
SBD6	6.36 ± 0.34	8.50 ± 0.57	5.95 ± 0.41	4.22 ± 0.47	7.14 ± 0.60	0.79 ± 0.12	3.29 ± 0.13	1.47 ± 0.16	4.73 ± 0.20
SBD7	6.13 ± 0.35	8.27 ± 0.59	5.66 ± 0.40	3.81 ± 0.64	6.70 ± 0.69	0.83 ± 0.11	3.07 ± 0.23	1.75 ± 0.38	4.20 ± 0.39
⟨SBD⟩	6.37 ± 0.11	8.53 ± 0.18	5.98 ± 0.16	4.16 ± 0.18	7.06 ± 0.21	0.80 ± 0.04	3.23 ± 0.09	1.65 ± 0.13	4.54 ± 0.20

^a For all of the RMSD calculations, the alignment was based upon the protein C_{α} atoms. ^b All C_{α} atoms but those of segment 82–95.

consistent position throughout all of the simulations (standard deviation of 0.11 Å). The RMSDs of the four ligand moieties from their respective X-ray structural positions after alignment of the protein C_{α} atoms are also tabulated. These RMSDs are large since they measure the shift in the position of the ligand moieties in IS relative to their positions in the native complex. Most important is the fact that these values are relatively close throughout each set of runs, and taken together, they indicate an overall consistent position of the ligand. For the SBD results, the low standard deviations of ~0.2 Å indicate strikingly similar positions of the ligand moieties in all of the runs. It is noteworthy that the core region of the ligand is 4.64 ± 0.11 Å away from its position in the bound complex (difference between average distances from the center of the binding pocket of 6.37 Å in IS and 1.73 Å in the bound state), which is a distance similar to the ~4.5 Å separating the two protein surfaces in the encounter complex of the barnase–barstar system.¹¹

From Figure 5, the 80s loop clearly has a configuration in IS that differs from that in the native complex. In Table 2, RMSD calculations performed for the protein and the 80s loop are also reported. Their analysis reveals that only a few residues contribute significantly to this distortion. Indeed, the backbone of segment 88–93 exhibits large deviations of ~4.5 Å for both LD and SBD results, whereas the backbone structure of the remaining segments 82–87 and 94–95 of the 80s loop is only moderately distorted (RMSD < 2 Å). Remarkably, in the short sequence 88–93 at the tip of the 80s loop, the largest structural deviation stems from Gly89 (5.5 Å). This glycine residue is located in this short sequence at the tip of the 80s loop that also contains the largest proportion of proline residues in FKBP12. In other protein–ligand systems, the location of glycine residues near the active site of proteins has been identified as crucial for ligand recognition by allowing local fluctuations or changes in loop structure that contribute to the entering of the ligand.^{43–45} In this intermediate IS, which is hypothesized to represent an early stage of binding, the distortion facilitated by a glycine residue can thus appear as an effective strategy to favor an interaction between the neighboring residue Ile90 and the diketone group of the ligand, as we will see below in the analysis of the protein–ligand contacts. Previous studies have shown the distortion of the backbone structure

Table 3. Average Separation Distances between the Centers of Mass of the Side Chains of Tyr82, His87, Ile90, and Ile91, along with Those Measured in the Crystal Structure (All Distances Are in Ångstroms)

distance	reference	average	std dev.
His87...Ile90	5.17	6.06	0.08
His87...Ile91	6.22	6.87	0.29
Ile90...Ile91	5.63	5.73	0.02
Tyr82...His87	5.50	5.44	0.15
Tyr82...Ile90	8.43	8.63	0.30
Tyr82...Ile91	4.97	5.31	0.27

around residue 89 or the importance of this region of the 80s loop for the recognition. For example, in a comparison between unliganded and various bound forms of FKBP12, a systematic difference in the backbone conformations of the region 87–89 of the protein has been pointed out, with RMSDs as large as 2 Å.⁴⁶ In the context of the formation of the ternary complex FKBP12–FK506–calcineurin, the importance of region 87–91 of the loop in the recognition of calcineurin has also been demonstrated.⁴⁷

Although the 80s loop is distorted in the intermediate IS, the hydrogen bonds found in this loop are well preserved. Six of the 10 native hydrogen bonds found in the 80s loop and in the 3₁₀ helix 78–81 have average contact frequencies greater than 55%, including the three interactions between His87, Ile90, and Ile91 at the tip of the 80s loop. For these six hydrogen bonds, the corresponding distances averaged over all of the simulations lie below 2.41 Å. Contact frequencies that are lower than average are obtained for the four remaining hydrogen bonds of the loop, which each involves a glycine residue: Gly83-O...His94-NH (41%), Ser77-NH...Gly1-O (27%), and values below 3% for both Pro78-O...Gly83-NH and Tyr82-O...Gly86-NH.

By comparing unliganded and FK506-bound conformations of FKBP12, Ivery and Weiler have inferred that the 80s loop must undergo a displacement as a unit in the binding process of the ligand FK506.²⁰ The same conclusion was drawn by Wilson et al. on the basis of a comparative analysis of X-ray structures of unbound, FK506-, and rapamycin-bound forms of FKBP12.⁴²

Table 4. Time Average and Ensemble Average Distances (Å) for the Most Persistent Contacts between the Protein and the Ligand^a

		atom pair ^b		SBD1	SBD2	SBD3	SBD4	SBD5	SBD6	SBD7	fluctuation range	⟨SBD⟩	std dev.		
protein	ligand														
Ile90	CG2	C8		3.88	3.90	3.86	3.94	3.94	3.93	3.97	0.29–0.46	3.92	0.04		
		O3		3.82	3.77	3.82	3.83	3.81	3.87	3.88	0.38–0.59	3.83	0.04		
		O4		3.86	4.06	3.79	4.02	4.10	3.95	3.99	0.37–0.53	3.97	0.11		
	CD	O4		4.76	4.05	3.94	4.03	4.17	4.41	4.41	0.49–0.72	4.25	0.29		
Ile91	CG1	C3		4.24	4.25	4.06	4.19	4.26	4.02	4.13	0.38–0.50	4.16	0.10		
His87	CD2	O3		4.33	4.82	4.91	4.16	5.53	5.40	3.83	0.59–1.03	4.71	0.63		
		C13		4.13	5.32	5.75	3.88	5.82	5.67	6.24	0.79–1.61	5.26	0.90		
		CG	C13	4.27	5.48	6.03	4.06	5.91	5.80	6.29	0.78–1.53	5.41	0.88		
	NE2	C13		4.35	5.31	5.69	4.07	5.61	5.60	6.20	0.75–1.59	5.26	0.77		
Tyr82	OH	C8		5.70	5.28	4.37	4.62	5.36	4.43	6.78	0.90–2.13	5.22	0.86		
		O3		4.94	4.58	3.58	3.94	4.62	3.66	6.05	0.99–2.21	4.48	0.87		
		O1		5.74	5.17	4.17	4.36	5.21	4.13	6.19	0.96–2.00	5.00	0.80		
		C16		5.31	4.69	4.08	4.18	4.72	4.06	5.75	0.73–1.69	4.69	0.65		
	CE2	O3	4.46	4.21	4.35	3.96	4.35	3.96	5.31	0.54–1.67	4.37	0.46			
	CZ	O3	5.11	4.80	4.15	4.37	4.90	4.22	5.99	0.63–1.78	4.79	0.64			
	O	C22	4.72	5.00	4.85	5.41	5.40	5.42	4.05	4.05	0.71–1.68	4.98	0.50		
Ile56	CG2	O2		3.45	3.52	3.45	3.48	3.52	3.40	3.48	0.25–0.34	3.47	0.04		
		CA	O2	4.07	4.16	4.05	4.08	4.13	4.01	4.06	0.26–0.37	4.08	0.05		
		N	O2	3.34	3.29	3.31	3.30	3.27	3.28	3.32	0.24–0.34	3.30	0.03		
	∠N–H	O2		88.6	104.1	90.0	96.4	105.2	90.8	86.5	13.6–17.6	94.5	7.6		
Val55	CA	O2		3.53	3.44	3.46	3.48	3.45	3.48	3.42	0.22–0.40	3.46	0.04		
		C	O2	3.43	3.52	3.40	3.47	3.52	3.40	3.35	0.25–0.37	3.44	0.07		
		O	O2	4.08	4.37	4.08	4.24	4.39	4.09	3.99	0.37–0.45	4.17	0.16		
Glu54	O	C4		4.29	4.88	4.13	4.49	4.96	4.24	4.08	0.52–0.79	4.44	0.35		
		C1		3.97	4.56	3.91	4.18	4.73	3.92	3.83	0.36–0.48	4.16	0.35		
		O1		4.04	4.61	3.99	4.21	4.84	3.99	3.97	0.38–0.61	4.24	0.35		
		O2		3.74	4.24	3.74	3.96	4.36	3.74	3.62	0.37–0.52	3.92	0.28		
		C15		3.84	4.24	3.78	3.91	4.47	3.79	3.70	0.34–0.69	3.96	0.28		
		C24		3.94	4.07	3.97	3.98	4.25	3.95	4.14	0.36–0.47	4.04	0.12		
		C25		3.75	3.49	4.46	4.66	3.46	3.50	4.21	0.33–0.87	3.93	0.50		
		C26		4.54	4.05	5.23	5.30	3.94	4.37	5.18	0.43–0.80	4.66	0.57		
		C29		4.80	5.03	4.22	3.92	5.29	5.10	4.99	0.42–0.88	4.76	0.50		
		C	C25	4.26	4.08	5.17	5.44	4.07	4.07	4.77	0.35–0.97	4.55	0.58		
		Gln53	O	C25		3.91	4.24	4.81	5.18	4.17	3.74	4.19	0.51–1.10	4.32	0.51
				C26		3.82	4.18	5.19	5.42	3.92	3.83	4.65	0.57–1.01	4.43	0.67
				C29		5.67	6.14	4.36	4.26	6.16	5.88	4.87	0.45–1.08	5.33	0.82

^a A persistent contact was considered whenever the corresponding contact frequency was higher than 50% in at least one simulation. The reported fluctuation range (Å) represents the minimum and the maximum of the fluctuations in the simulation set. One angle value (deg) is reported for the native hydrogen bond between Ile56-NH and ligand atom O2. ^b Except for Ile56 where the angle for the native hydrogen bond it formed with O2 is also reported.

Hence, the conclusions of these studies prompted us to examine whether, in the case of ligand **8**, the 80s loop underwent a displacement as a unit between the native complex and IS. Prior to this analysis, one key observation should be reported for the native complex FKBP12–**8**: the bulky side chains of the 80s loop, namely, Tyr82, His87, Ile90, and Ile91, are clustered together, forming a cylinder-like hydrophobic core whose upper edge defines a small cavity to where atom C13 of the ligand is pointing; this cavity has a horseshoe shape (on the left of Figure 5). We have thus monitored the six center of mass separations between these four side chains in IS for comparison with their values in the crystal structure. The overall averages are

reported in Table 3. All of the distances measured in IS are close to their counterparts in the native complex, with deviations less than ~0.3 Å, except for the two distances His87···Ile90 and His87···Ile91 that are slightly larger in IS by 0.9 and 0.7 Å, respectively. Taken together, the analysis of the data for the 80s loop, i.e., RMSDs and separation distances between its bulky side chains, reveals a distortion of the loop backbone conformation near the tip region 88–93 (most severe for the residue 89) and a concerted displacement of all of the bulky side chains of the loop between the bound state and IS.

In Table 4, the distances for the most persistent contacts between the protein and the ligand calculated for every simulation are

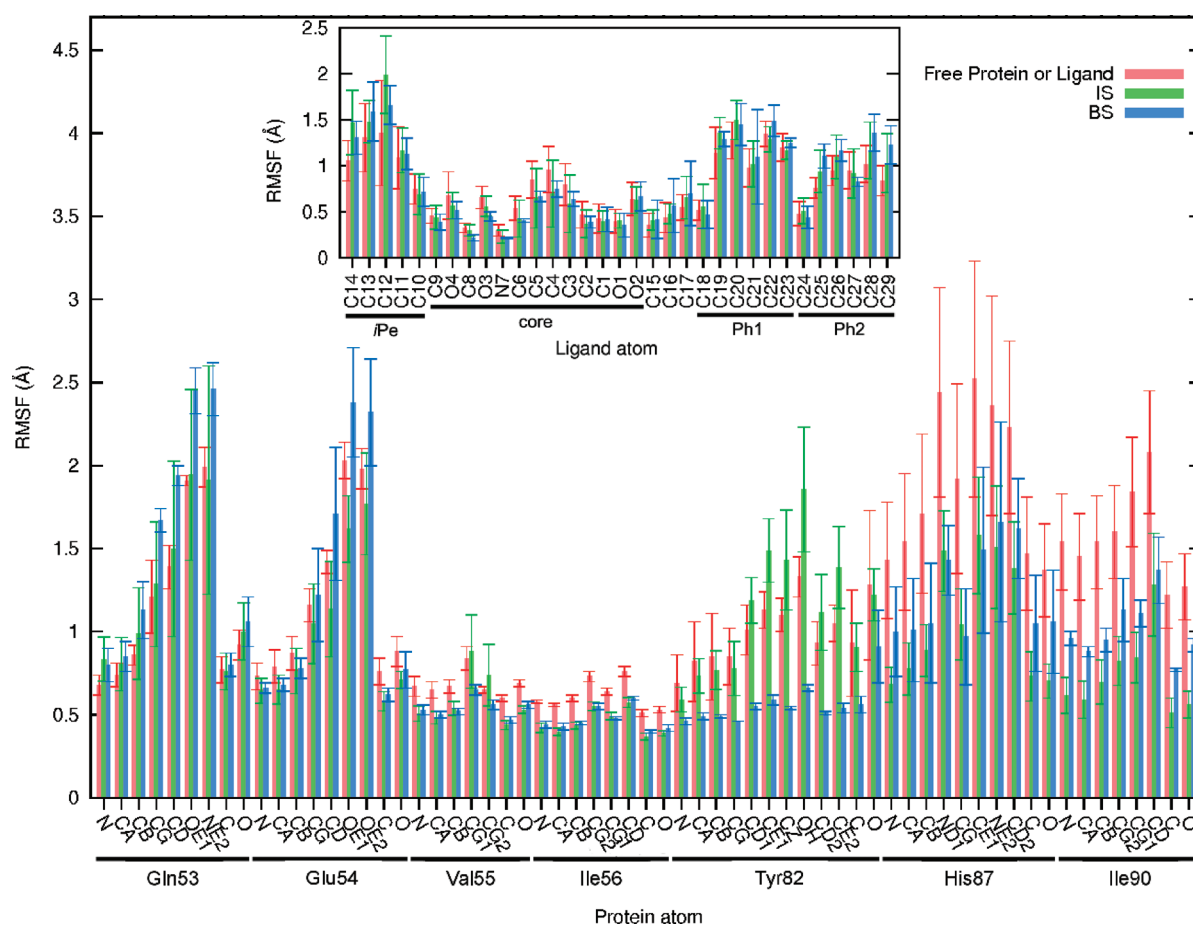


Figure 6. RMSFs calculated for the residues 53–56, 82, 87, and 90 that are in contact with the ligand in IS. The corresponding values for the bound state (BS) are also reported for comparison. The RMSFs for the ligand atoms in IS and the bound state are shown in the inset. For purposes of comparison, the RMSFs of these seven residues in the free protein and those of the free ligand (inset) were also calculated by averaging over four and five 2-ns SBD simulations, respectively.

reported, along with their ensemble average values. It is striking that only a few average contact distances lie below 4 Å: two clusters of interactions are stable, namely, between Ile90-CG2 and the ligand atoms C8, O3, and O4 (average distances are in the narrow range 3.83–3.97 Å) and between the atoms Ile56-(N,CG2) and Val55-(C,CA) of the protein and O2 in the ligand (distance range of 3.30–3.47 Å). The corresponding standard deviations are below ~ 0.1 Å, indicating overall consistent distances in the simulations (the largest standard deviation is calculated for the contact Ile90-CG2 \cdots O4, whose distance lies slightly above 4 Å in three simulations). The backbone oxygen of Glu54 also makes three contacts with distances of ~ 3.9 Å; however, they are associated with significantly higher standard deviations of 0.28 and 0.50 Å. The native hydrogen bond between O2 and Ile56-NH should more properly be viewed in IS as a van der Waals contact since the corresponding distance and angle are 3.30 ± 0.03 Å and $94 \pm 8^\circ$, respectively. For the other native hydrogen bond between O3 and Tyr82-OH, the average distance is 4.48 ± 0.87 Å. Hence, for the latter contact, a large variability in the seven average distances is observed and also high fluctuations in each simulation; this results from the high mobility of the side chain atoms of Tyr82 (as seen below in the fluctuation analysis) since atom O3 retains a very stable position as stated previously (Table 2). All other average distances in Table 4 lie mainly between 4 and 5 Å with standard deviations less than 1 Å.

In Figure 6, we have reported the RMS fluctuations of the residues that interact with the ligand (except for Ile91, which hardly interacts with the ligand since only one of its contact frequencies is higher than 50%), all of which are listed in the first column of Table 4. For purposes of comparison, the RMSFs of the unbound and bound forms of the protein are also shown, as well as those derived for the ligand itself in an explicit solvent model, in the bound complex, and in IS. The RMSFs for residues 53–56 are very similar in the three states of the protein (except for the amide and the carboxylate groups of Gln53 and Glu54, respectively). However, for the residues Tyr82, His87, and Ile90 in the 80s loop, the RMSFs vary between the three states of the protein. In going from the apoprotein to IS, the fluctuations of the side chain of Tyr82 increase by 20 to 30% (except for CB and CG), whereas those of both backbone and side-chain atoms of His87 and Ile90 are damped by 40 to 50%. Between IS and the bound complex, reversed trends are again observed between Tyr82 and the two residues His87 and Ile90: for the former residue, the RMSFs sharply decrease toward the lower values 0.5–0.7 Å, while no change or a small increase of less than 0.35 Å is observed for the latter residues. The RMSFs derived for the ligand in the three forms are surprisingly similar (with only moderately larger values for the atoms C12 and C14 in IS).

Overall, the convergence achieved in all of the simulations (i) for the ligand core position; (ii) for a few intermolecular distances

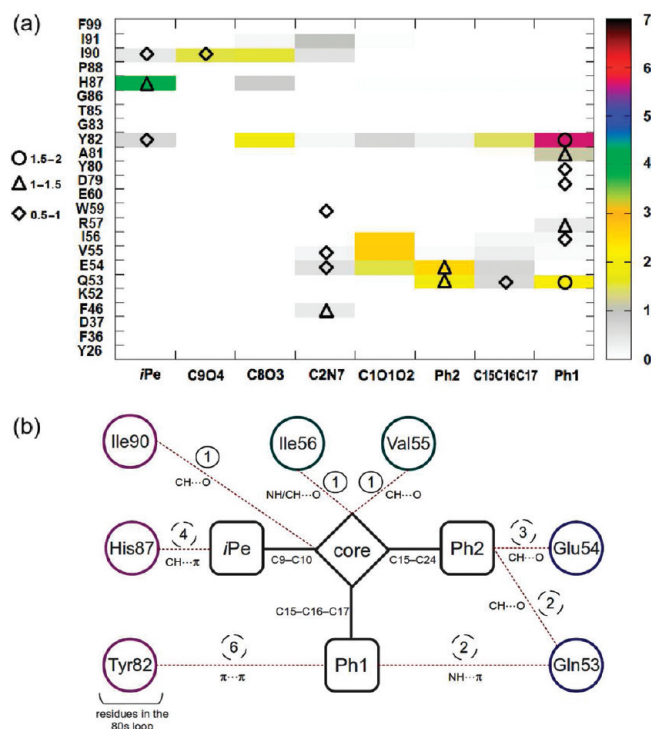


Figure 7. (a) Average number of contacts between each pair of interacting groups of atoms that consists of a ligand moiety and a protein residue (color scheme on the right) and the average number of water molecules in contact with both a ligand moiety and a protein residue (diamond, triangle, and circle on the left are for a number of bridging water molecules in the intervals $]0.5;1]$, $]1;1.5]$, and $]1.5;2]$, respectively). The core region of the ligand is further subdivided into two keto groups (C9O4 and C8O3), the pipercolinyl ring (C2N7), and the ester bond (C1O1O2); C15C16C17 refers to the propyl group. All types of contacts (cutoff of 4 Å) were first obtained for each SBD simulation; the average over the seven independent simulations was then derived. (b) Schematic representation of the main protein–ligand contacts. The numbers of permanent and transient contacts are circled in solid and dashed lines, respectively; the type of noncovalent interaction is indicated in each case. The bonds that link the three mobile moieties to the core are also indicated. Positions and distances for all moieties are not realistically represented.

that involve primarily the atom O2 and the bond C8O3 of the ligand and their binding partners Val55, Ile56, and Ile90; and (iii) for the RMSDs of the protein testifies to a convergent picture of the binding intermediate IS. However, a question arises as to how the ligand retains a stable position relative to the protein given the very few persistent contacts between the two molecules. The other intermolecular contacts are indeed showing large distance variations from simulation to simulation and are short-lived. To quantify the amounts of instantaneous short-lived contacts between the two molecules, we have calculated the total number of contacts between each pair of interacting groups of atoms that consists of a ligand moiety and a protein residue. A 4 Å distance cutoff was used for counting the interatomic contacts between each group of atoms (Figure 7a). In this figure, we can see the two previously reported clusters of contacts between Val55 or Ile56 and the ester bond C1O1O2 (where O2 is engaged in contacts), and between Ile90 and the two bonds C8O3 and C9O4 (of which O4 is responsible for the contacts with the atoms Ile90-(CG2,CD), while C9 is involved in only transient contacts with Ile90-CG2 as all its contact frequencies are well below 50%). About two

contacts are found between C8O3 and Tyr82 because of transient contacts between O3 and the three atoms Tyr82-(OH,CZ,CE2); in SBD3 and SBD6, however, a lasting van der Waals interaction Tyr82-OH...O3 is found (Table 4). It is striking that other groups of atoms also participate in intermolecular contacts, as indicated by the following averages: the *iPe* moiety makes four contacts with His87; the Ph1 moiety makes two and almost six contacts with the residues Gln53 and Tyr82, respectively; and the Ph2 ring makes about two and three contacts with Gln53 and Glu54, respectively (Figure 7b).

Figure 7a also gives the average number of water molecules that are in contact with both a ligand moiety and a protein residue. Such bridging water molecules are mainly seen in contact with the ligand moieties Ph2, *iPe*, Ph1, and the pipercolinyl ring (C2N7). These contacts are in the range 1–1.5 for *iPe* and His87, as well as for Ph2 and both Gln53 and Glu54. A slightly larger number of such contacts involve Ph1 (about two with both Gln53 and Tyr82) and the pipercolinyl ring (about one with each of the residues Phe46, Glu54, Val55, and Trp59, respectively).

A detailed analysis of all of the short-lived contacts between the protein and the ligand is useful in identifying the atomic partners and the nature of these contacts. As seen in Table 4, all of the average separation distances between the three mobile moieties of the ligand Ph1, Ph2, and *iPe* and the residues Tyr82, His87, Ile90, Gln53, and Glu54 are above 4 Å (except for C25...Glu54-O). The large averages and standard deviations seen for these separation distances result from the large RMSFs obtained not only for the side chains of residues 82, 87, 90, and 53 but also for the ligand moieties Ph1, Ph2, and *iPe* (Figure 6). The high fluctuations of *iPe*, Ph1, and Ph2 are due to low dihedral energy barriers associated with the bonds C9–C10 (1.4 kcal/mol), C17–C18 (0.5 kcal/mol), and C15–C24 (1.6 kcal/mol), respectively. The facile rotation of the Ph2 group around the bond C15–C24 enables interchangeable interactions of the segments C25–C26 and C29–C28 with the backbone oxygens of Glu54 and Gln53: these transient interactions are responsible for an average of about two CH...O contacts between Ph2 and each of these two residues (Figure 8a). For the *iPe* moiety, a pseudo-3-fold symmetry can be considered around the bond C9–C10, and similarly, the rotations about this bond exchange the atoms C11, C13, and C14 in the interaction with the imidazole ring of His87 (Figure 8b). Each of the atoms C11, C13, and C14 of the *iPe* moiety is making an average of four CH... π contacts with the atoms of the imidazole ring.

In contrast to the *iPe* or Ph2 moiety, where there are only a few easily identifiable contacts with the protein, a higher number of transient contacts is found between the Ph1 moiety and both the amide group of Gln53 and the aromatic ring of Tyr82. The distances between Ph1 and Tyr82 ring centroids as well as between the Ph1 ring centroid and each of the two atoms NE2 and OE1 of Gln53 are reported in Figure 8c. Because of the high mobility of the interacting partners, very short-lived contacts are seen during the course of the simulation (in particular, the propyl “arm” C15–C16–C17 of the ligand is very flexible). With a view to quantitatively analyzing the preferred orientations of the binding partners Ph1 and the side chains of Tyr82 and Gln53, for each simulation, the three instantaneous separation distances, such as those presented in Figure 8c, were collected into equally spaced bins of 0.25 Å width. This yields a histogram with distance bins each counting the number of occurrences of certain distances within this short distance interval of 0.25 Å. The same procedure is followed for every simulation, and the resulting averaged

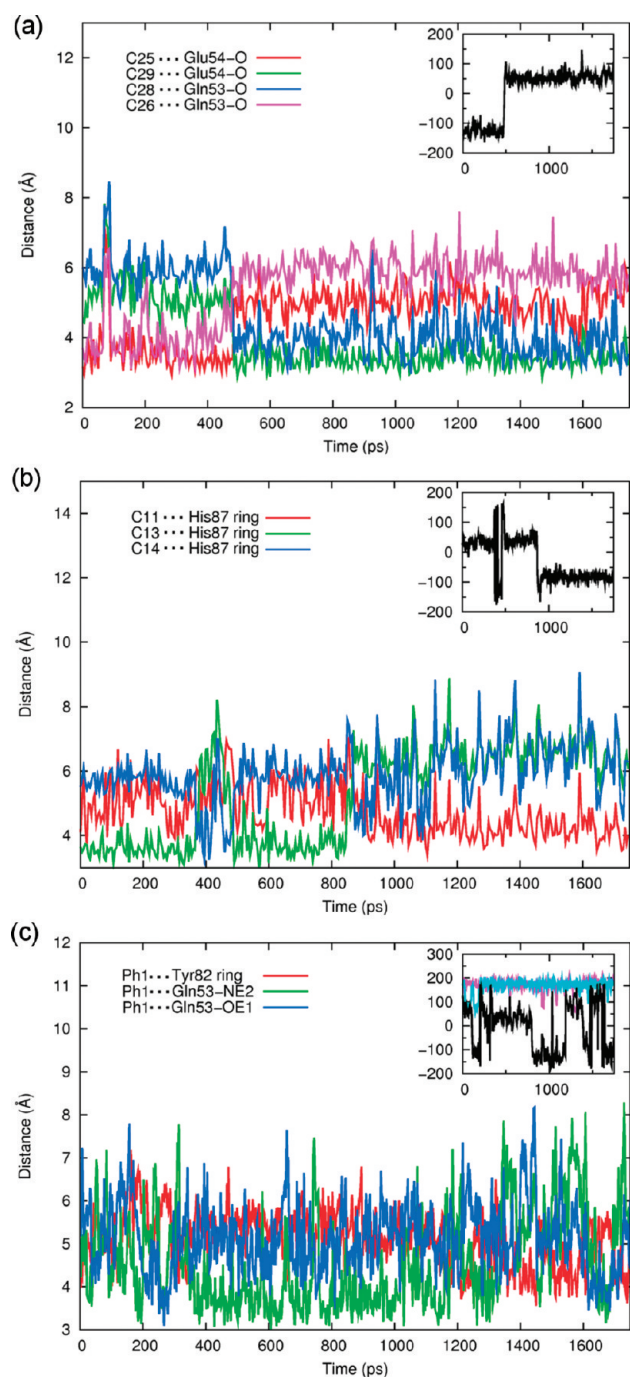


Figure 8. Distances between the noncore regions of the ligand and a few protein atoms as a function of time in the SBD simulations. (a) Distances between the edge atoms C25, C26, C28, and C29 of the ring Ph2 and the backbone oxygen atoms of Gln53 and Glu54. The inset represents the evolution of the dihedral angle around C15–C24. (b) Distances between the atoms of *iPe* and the center of the imidazole ring of His87. The inset represents the evolution of the dihedral angle around C9–C10. (c) Distances between the center of the Ph1 ring and that of the side chain ring of Tyr82, and between the Ph1 ring centroid and the two atoms NE2 and OE1 of the Gln53 side chain. The inset represents the evolution of the dihedral angles around C15–C16 (purple), C16–C17 (cyan), and C17–C18 (black).

histogram is shown in Figure 9. Between Ph1 and Tyr82 ring centroids, the most populated bin distances are found in the range

4–5.5 Å, a distance range compatible with π – π interactions and consistent with the results obtained for two interacting benzene rings (with benzene ring centroids in the range 3.75–4.02 Å and at 4.96 Å⁴⁸). Regarding the distances between Gln53-NE2 or -OE1 and the Ph1 centroid, the interactions with NE2 are predominant (NH··· π type), as reflected by the higher number of contacts than for OE1 at short distance ranges (3–4 Å).

To summarize, one striking feature of the protein in IS is the displacement as a rigid body of its 80s loop, favoring various contacts with the ligand. The segment 88–93 exhibits great flexibility, as previously observed for other FKBP12 complexes.⁴⁶ The ligand core region has a well-defined position with respect to the protein, which enables permanent anchoring contacts of the types NH···O and CH···O by the ligand atoms O2 and O3. Moreover, since the two native hydrogen bonds are not yet formed in this early binding stage, their formation thus represents a major enthalpic force driving the system toward the native complex. On the other hand, the symmetry and flexibility provided by the noncore regions enable transient contacts of various types (π – π , NH··· π , CH··· π , and CH···O), which is another important aspect of this binding intermediate (*vide infra*). The dynamic aspect of the protein–ligand contacts in IS also helps to avoid a trapped state that could impede the formation of the bound state.

4. DISCUSSION

We have characterized one intermediate state (IS) along the unbinding pathway of ligand **8** to FKBP12 by performing seven independent molecular dynamics simulations with the explicit inclusion of water molecules. The analyses of the separation between the center of mass of the binding pocket and that of the ligand core reveal a stable position of the ligand in IS. The ligand core moiety is restricted to a narrow range of positions, lying at 4.64 ± 0.11 Å above its position in the native complex. The stability of the ligand is further confirmed by the few persistent intermolecular contacts, mainly O2···Val55/Ile56 and O3···Ile90, with corresponding distances varying over limited ranges (Table 4). The ligand core position as well as the protein structure are consistently predicted by the series of simulations. These results, together with the overall agreement with the LD results, provide confidence in the model derived for this binding intermediate.

In contrast to the core moiety, the three peripheral groups (*iPe*, Ph1, and Ph2) have a high mobility that results mainly from the low dihedral transition barriers around a few bonds. Consequently, no long-lived contacts between these moieties and the protein are consistently observed in all of the simulations. Nevertheless, these groups can form many alternative contacts with the protein, owing to both their intrinsic mobility and their (pseudo)symmetry. Recently, the importance of such alternative contacts has been highlighted in a study of the configurational entropic contributions of the residue side chains upon formation of the PKA/AKAP complex.⁴⁹ The authors found that the affinity between the two proteins is increased with the number of alternative contacts available, which is associated with an increase in configurational entropy. Though their conclusions were drawn from the analysis of a native protein–protein complex, we believe that, in the intermediate IS, such alternative contacts could help the ligand to retain at least a part of the configurational entropy it has in the free state, while the nascent interactions would contribute to enthalpic gain. This assertion is further supported

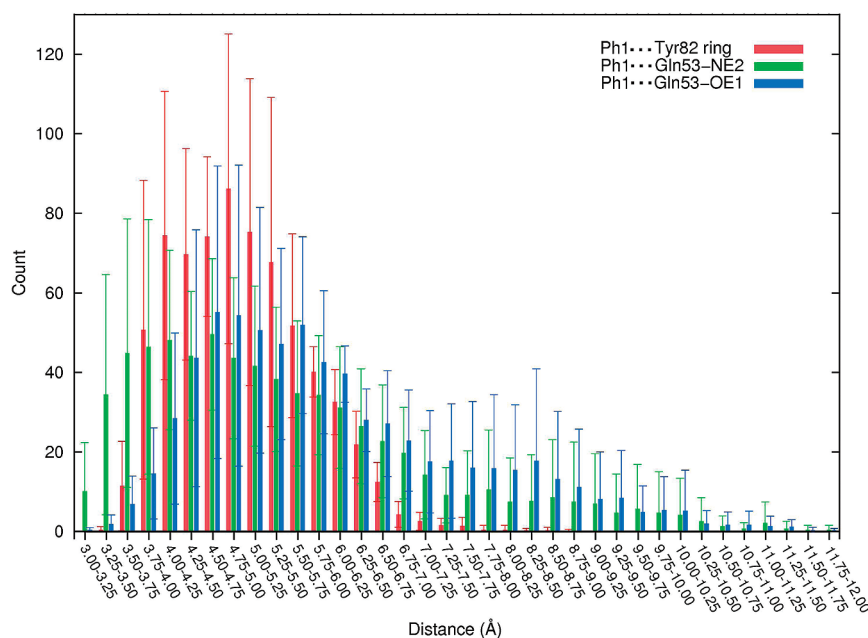


Figure 9. At regular distance intervals of 0.25 Å between 3 and 12 Å, the three distances between the Ph1 ring centroid and the Tyr82 ring centroid, Gln53-NE2, and Gln53-OE1, respectively, that fall into each interval were counted for each trajectory. The resulting numbers were averaged over the whole set of SBD trajectories to yield this histogram. On the *x* axis, the notation for the distance interval *a*–*b* indicates an interval of type [*a*; *b*].

by the fluctuations of the ligand calculated in IS that are similar to those in the free state (Figure 6). The dynamic aspect of the protein–ligand contacts also allows avoiding a trapped intermediate along the complexation pathway.

Another important aspect of the binding intermediate is the structure of the 80s loop. As discussed in the Introduction, in the bound state, ligand **8** is largely buried in the binding pocket: the pipercolinyl ring is buried in the main pocket, and the methyl group C13 is buried in the small side cavity defined by residues 82, 87, 90, and 91. Hence, the latter four residues would contribute to hindering the escape or the binding of the ligand if their bulky side chains remained in the position found in the crystal structure. A similar conclusion was drawn by Zacharias in a rigid docking study of FK506 to FKBP12, where the methyl group C35 of the ligand fills the same side cavity as C13 in **8**.¹⁹ However, in a binding mechanism that proceeds through the intermediate IS, the concerted motion observed for the bulky side chains of the 80s loop exposes the small side cavity to the incoming ligand, thereby allowing regions of the ligand to initiate the interactions with the protein. Indeed, in IS, atom O3 points to this small side cavity and forms a stable interaction with the side chain of Ile90 as well as weak interactions with the side chain of Tyr82 (Table 4 and right of Figure 5). An important consequence of such a two-step mechanism, as opposed to a one-step one, is the drastic decrease of the energy barrier of association due both to the nascent protein–ligand interactions and to the reduced steric hindrance; the latter two can be regarded as favorable enthalpic and entropic contributions, respectively. In IS, the observed displacement of the 80s loop is also consistent with the conclusion of previous comparative analyses of bound and unbound forms of FKBP12,^{20,42} as well as with NMR data that indicate at least two distinct conformations for the 80s loop of the unbound protein.²¹ The displacement of the 80s loop seen in our model IS is additionally supported by a docking study performed on a set of 20 protein–protein complexes.⁵⁰ Though this study has not dealt with the docking of small molecules to FKBP12, it is worth

mentioning that the authors have found that the regions of the protein that prevent binding because of steric hindrance are undergoing conformational transitions compatible with the recognition of the binding partner.

With a view to testing the transferability of this binding intermediate model to other ligands, we have substituted ligand **8** in IS with FK506. We thus hypothesize here that a similar intermediate exists for the binding mechanism of FK506. This larger ligand was docked to the protein pocket in IS by aligning the atoms of the common core region. FK506 was taken in its bound crystallographic geometry, and no relaxation was further attempted. Figure 10 shows the result of the superposition of the two ligands. Remarkably, ligand FK506 fits well into this model by retaining many of the contact features of ligand **8**, with very few short contacts, despite the fact that neither FK506 nor the mobile side chains of His87 and Tyr82 were allowed to relax. In particular, the distances O2...Ile56-N and O3...Ile90-CG2 are both compatible with a van der Waals interaction (3.45 and 3.57 Å, respectively, in the snapshot of Figure 10). The methyl group C35 in FK506 that points to the imidazole ring of His87 has a position equivalent to the methyl or methylene group on C10 in **8** (note that in the respective bound forms of the protein, the methyl C35 of FK506 fills the same small side cavity as the methyl C13 in **8**). In Figure 10, the cyclohexyl group (C29–C34) in FK506 has the same role as the Ph1 ring in **8**, and, as with Ph1, this group is known to be very mobile.⁵¹ The fact that some configurations taken from the simulations of IS could fit the larger ligand FK506 while still retaining the above-described structural features of ligand **8** suggests a transferability of our model to other ligand structures.

The analysis of this binding intermediate model also sheds some light on the role of the noncore regions *i*Pe, Ph1, and Ph2 in the binding process. Since these three peripheral groups form extensive contacts with a few residues in IS and thus contribute to the stability of this intermediate, we can understand how these groups enhance the binding ability of this ligand. The comparison of

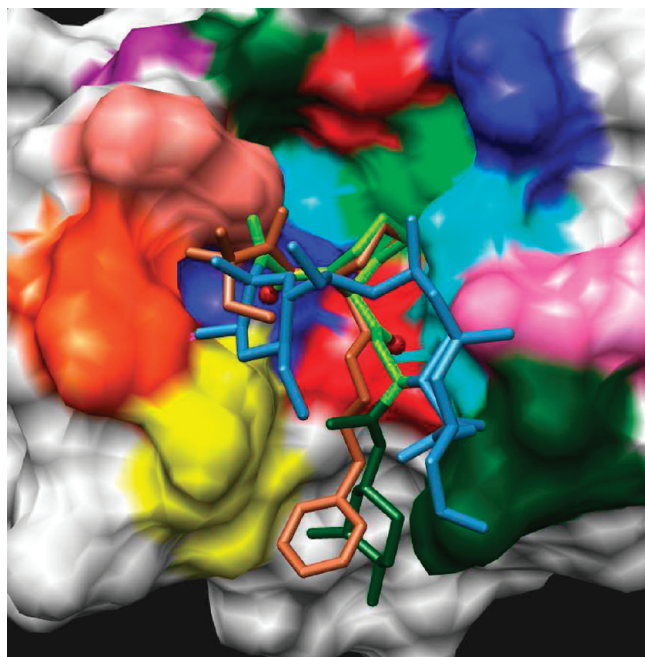


Figure 10. Snapshot that shows a superposition of the FK506 and ligand 8 core regions. The color code for the protein residues is as in Figure 5. Ligand 8 is colored brown, and oxygen atoms O2 and O3 are shown as red spheres. FK506 is colored as follows: core region (C9–C26) in light green, the cyclohexyl region (C27–C34) in dark green, and the nonbinding region or “effector” domain (C25–C10) in blue.

the affinity data for a few close analogues of ligand 8 indicates that the affinity is reduced whenever one of the three peripheral groups of 8 (*i*Pe, Ph1, or Ph2) is missing. The substitution of any of these groups also has important consequences on the affinity data, as discussed below. The *i*Pe moiety seems essential for high-affinity properties since many high-affinity ligands contain this group. Alternatively, for the ligands that have a *tert*-butyl group on C9, the affinity measured is higher than for the group *i*Pe.⁵² Other high-affinity ligands have a phenyl or methoxyphenyl derivatives instead of *i*Pe on C9.⁵³ By relying on our model and by assuming that it is transferable to other ligands, this phenyl group (or phenyl derivatives) linked to C9 would contribute to aromatic stacking with the imidazole ring of His87 in the binding intermediate analogue, thereby stabilizing this species. Regarding the Ph2 group, the close ligand analogue 5 (Figure 11a) exhibits a 10-fold decrease in affinity (110 nM vs 10 nM for 8). It should be stressed here that the role of Ph2 is not only to allow contacts in IS as seen, its presence in 8 also has a consequence on its neighboring group Ph1. Indeed, the analysis of the torsion profile around the bond O1–C15 indicates that Ph2 severely restrains the motion around O1–C15. Thus, the space that is spanned by the Ph1 ring in 8 is narrower than in the case of 5, which makes it more available for the interaction with Tyr82: the presence of the Ph2 ring on C15 in 8 therefore contributes indirectly to enhancing the π – π interactions in IS. For the analogue 9 (Figure 11a), the binding constant remains similar to that of 8 (7 vs 10 nM); this comparison suggests that the aromatic nature of the substituent on C15 is not essential. On the other hand, a substitution of Ph2 by a 1,1-dimethyl-2-propenyl group as in ligand 7 results in a 25-fold lower affinity than that of 8 (250 nM).⁶ By assuming the transferability of our intermediate model to the two previous ligand analogues, we may argue that, in contrast to the case of the

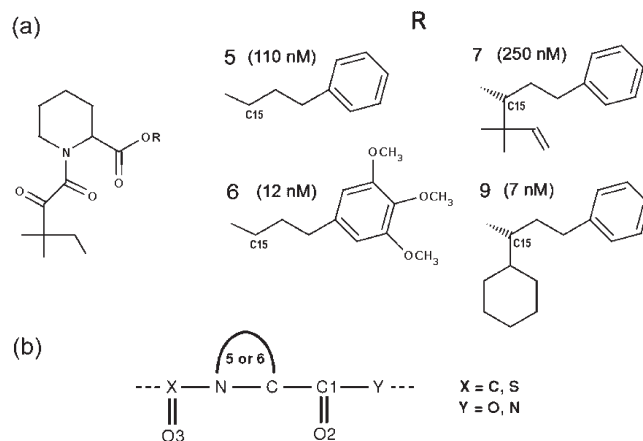


Figure 11. (a) Structures of a few ligands 5, 6, 7, and 9 with structures related to that of the ligand 8. The ligand identifiers and the K_i constants in parentheses are taken from the work of Holt et al.⁶ Atom C15 is also indicated. (b) A five- or six-membered ring is often found at the center of the core binding region of the high-affinity FKBP12 ligands (pyrrolidine or piperidine). A thiazane ring (six-membered ring with a sulfur atom) can also be found as a central motif.¹⁶ Recently, open alkyl forms for the central motif have demonstrated the ability to promote neurite outgrowth, thus suggesting that a central cyclic structure is not a prerequisite for binding to FKBP.¹⁷ An amide bond also replaces the ester bond in the context of the peptidylprolyl *cis*–*trans* isomerization.

ligand 7, in 9 the symmetry of the cyclohexyl substituent allows the formation of alternative contacts after the rotation of this group, as in the case of Ph2. Moreover, as mentioned above, Ph2 in 8 and the cyclohexyl ring in 9 are bulky groups that also indirectly enhance the interactions that Ph1 makes in IS. Hence, the number of contacts seen in the intermediate model, which determines its stability, seems to be related to the affinity of a few analogues of 8.

Regarding the Ph1 group, it is striking that almost all of the high-affinity synthetic ligands of FKBP12 contain an aromatic ring (phenyl, pyridyl, or trimethoxyphenyl ring) linked by a propyl group to the ester bond of the ligand core. The presence of this aromatic group and the length of the alkyl arm that links it to the core therefore seem crucial for the binding properties. The results for our model IS provide a guide for understanding the role played by such an aromatic group in the recognition process. First, in IS the Ph1 moiety makes on average two contacts with Gln53 and almost six others with Tyr82, Ph1 being the ligand moiety that is making the highest average number of contacts with FKBP12 (in comparison, the pipercolinyl ring is only weakly interacting with the protein in IS; see Figure 7a). The same analysis performed for Ph1 in the bound state indicates that Ph1 participates in an average of about three contacts both with His87 and with Tyr82, and one with Gly86 (data not shown). Second, in IS the RMSD values of all of the ligand moieties from their respective native positions reveal that Ph1 is the moiety closest to its native position (4.16 ± 0.18 Å in Table 2). Therefore, in IS, where the nascent interactions are formed between the molecular partners, the Ph1 ring appears critical since it not only interacts the most with the protein but is also the moiety closest to its native position. What is more, in tracing the nature of the contacts between Ph1 and Tyr82 by performing semiempirical molecular orbital calculations, a molecular orbital overlap is obtained between the aromatic rings of Tyr82 and Ph1, as illustrated in Figure 12. This molecular overlap is observed whenever the two interacting

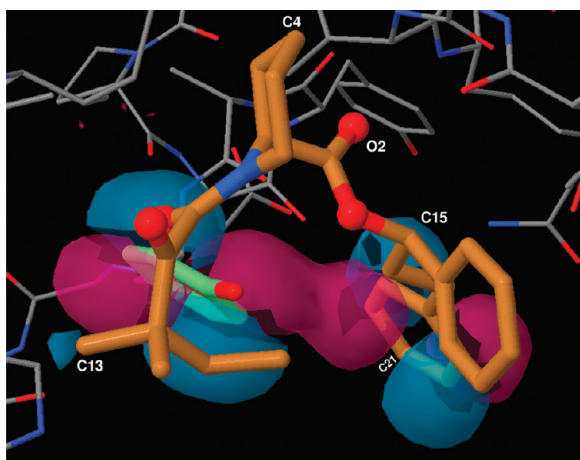


Figure 12. Molecular overlap between the Ph1 moiety of the ligand (shown in brown) and the aromatic ring of Tyr82 (in light green) calculated for IS at the semiempirical level of theory by using the MOZYME method with the PMS Hamiltonian, as implemented in the MOPAC2006 program.⁵⁹ The molecular orbital is the HOMO-40 represented with an isosurface value of 0.008.

rings exhibit a parallel displaced configuration in a series of structures taken from the SBD trajectories. This semiempirical approach has been shown to yield binding enthalpies for protein–ligand systems in nice agreement with experimental values.⁵⁴ To our knowledge, this is the first result showing that short-range electronic interactions (overlap forces) may be important in an early stage of the recognition process in a biomolecular context.

The short-range electronic interactions may also explain the 10-fold difference in K_i measured by Holt et al.⁶ for compounds **5** and **6** (Figure 11a). The two ligands **5** and **6** consist of the *i*Pe and core moieties but differ in the group linked to the ester bond: propylphenyl and (3,4,5)-trimethoxyphenylpropyl, respectively. In **6**, the three methoxy groups act as electron donors on the aromatic ring and are thus responsible for an enrichment in electronic density on both sides of the ring, as compared to the case of the nonsubstituted ring of **5**. Hence, in the hypothetical binding intermediate analogue for **6**, the out-of-plane increased density in the substituted ring could enhance the face-to-face interactions with the Tyr82 ring, as compared to the corresponding interactions in the binding intermediate analogue for **5**. Under the hypothesis that the stability of the intermediate is related to the affinity of the ligand, the stabilization brought by the stronger π – π interactions in the case of **6** may thus, at least partly, explain the difference in affinity.

From an experimental viewpoint, the importance of the aromatic character of residue 82 in the binding of the ligand can also be seen in the measures of the inhibitory activity against the mutants Y82F and Y82D of the protein. The mutation of Tyr82 to Phe lowers the affinity for FK506 and rapamycin nearly 2-fold.¹ Other experiences performed on the binding of rapamycin to yeast FKBP12 indicate a 95% loss of binding activity for the mutant Y89D;⁵⁵ Y89 corresponds to Y82 in the human form of the protein. Taken together, these results suggest that the hydrogen bond between O3 and Tyr82 is not crucial for the binding but rather the aromatic character of the side chain of residue 82. These observations are consistent with the model IS where we have found short-range electronic interactions that involve the aromatic ring of Tyr82 while the hydrogen bond with this residue is only weak at this early stage. In the hypothetical binding

intermediate analogue for FK506, the interactions of types C–H $\cdots\pi$ and O¹²–H $\cdots\pi$ between the ring C29–C34 and Tyr82 may substitute the π – π interactions found for **8** (Figure 10). For rapamycin, similar interactions may involve its ring C37–C42 and hydroxyl oxygen O¹³. Interestingly, in a more recent mutational analysis of the FKBP12 enzymatic activity, it was found that, among the 19 substitutions of the 82nd residue, the residues with an aromatic side chain (Phe, Tyr, Trp) exhibited the highest activity, along with Arg (which also has a marked aromatic character) and Pro.⁵⁶ These experimental findings therefore also emphasize the importance of the aromatic character of the 82nd residue.

Finally, the analysis of this binding intermediate also provides clues regarding the diversity of the core structures found in the high-affinity ligands. Our results reveal that atoms O2 and O3 are the only atoms that form stable contacts in the binding intermediate, acting as two anchoring points. This result suggests that the presence of these two atoms in the core region of a FKBP12 ligand may be important for the binding. Interestingly, when the structures of the high-affinity ligands of FKBP12 are compared so as to extract a common structural motif, O2 and O3 are found as recurrent atoms, as seen in Figure 11b. It is noteworthy that, in this motif, the pipercolinyl or pyrrolidinyl ring seems not to be a prerequisite for tight binding since a thiazane ring¹⁶ or even open alkyl forms for the central ring¹⁷ are also found in high-affinity ligands. In fact, the precise structure (ring or open alkyl form) of the central region of the motif in Figure 11b appears to be of less importance than that of the peripheral substituents of the core, as concluded by Zhao et al. in their study on the abilities of various FKBP ligand structures to promote neurite outgrowth.¹⁷ Interestingly, these authors have also found that the most important regions of the ligand for promoting neurite outgrowth are the group linked to the ester bond of the core (Ph1 or other substituents).¹⁷ Regarding the bond C9O4 found in the core, Orozco et al. have inferred that this group is not really important for the binding of ligands to FKBP12.⁵⁷ Their finding is also consistent with the observation that some high-affinity ligands contain only one ketone group (C8O3);⁵³ other ligands have one sulfonyl group that replaces the diketone motif.¹⁶ From the above analysis of the core region, O2 and O3 thus appear as the main recurrent atoms. If we hypothesize that the model IS is transferable to other ligand analogues, the prevalence of the atoms O2 and O3 may then be explained because of their role as anchoring points in the respective binding intermediate, as stated for ligand **8**. This result therefore helps to rationalize the observation that structurally diverse core regions that all share the atoms O2 and O3 can bind to FKBP12.

5. CONCLUSIONS

We have characterized one intermediate state (IS) along the unbinding pathway of ligand **8** to FKBP12. The ligand core region is restricted to a narrow range of positions, lying 4.6 ± 0.1 Å above its position in the native complex. Only the two carbonyl atoms of the core regions (O2 and O3) make permanent contacts in the intermediate, acting as two anchoring points. In contrast, the noncore regions have a large mobility that (i) ensures a dynamic aspect of the interactions, thereby avoiding a trapped state along the complexation pathway, and (ii) offers the possibility of fine-tuning the specificity of recognition owing to various types of interaction (π – π molecular overlap and weak hydrogen bonds N–H $\cdots\pi$, C–H $\cdots\pi$, and C–H \cdots O). Previous results on

the structure of the 80s loop are consistent with our results for the intermediate. Moreover, under the hypothesis that our binding intermediate model is transferable to other related ligands, which seems valid for FK506, this model comes closest to explaining both the diversity found for the core structures of the FKBP12 ligands and the role that the noncore regions could play in the recognition and the affinity (if we hypothesize that, for the binding to FKBP12, the slower the dissociation rate of the binding intermediate, the higher is the affinity^{10,11}). Indeed, for other high-affinity ligands of FKBP12, this model may explain the common occurrence of both the atoms O2 and O3 found in the core and the aromatic ring linked to the core by the same propyl “arm”, the latter being important for the flexibility and the contacts with Tyr82 and Gln53. A striking finding is the important role that the aromatic ring of Ph1 seems to play in the binding process by making short-range electronic interactions with Tyr82. This result is further supported by previous studies that have demonstrated the importance of the aromatic character of the 82nd residue in FKBP12.^{1,55,56} We speculate that, in other protein–ligand systems, short-range electronic interactions between an aromatic group of the ligand and an aromatic residue of the protein, as between Ph1 and Tyr82 in our model IS, might play a critical role in an early stage of the recognition process.

■ ASSOCIATED CONTENT

Supporting Information. More details on the simulation procedure are available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*Fax: +262(0)262-93-82-37. E-mail: Fabrice.Gardebien@univ-reunion.fr.

■ ACKNOWLEDGMENT

L.O. was the recipient of a doctoral fellowship from the French Ministry of Research.

■ REFERENCES

- DeCenzo, M. T.; Park, S. T.; Jarrett, B. P.; Aldape, R. A.; Futer, O.; Murcko, M. A.; Livingston, D. J. *Protein Eng.* **1996**, *9*, 173–180.
- Fung, J. J.; Starzl, T. E. *Ther. Drug Monit.* **1995**, *17*, 592–595.
- Danovitch, G. M. *Transplant. Proc.* **1999**, *31*, 2–6.
- Herdegen, T.; Fischer, G.; Gold, B. G. *Trends Pharmacol. Sci.* **2000**, *21*, 3–5.
- Davies, T. H.; Sánchez, E. R. *Int. J. Biochem. Cell Biol.* **2005**, *37*, 42–47.
- Holt, D.; Luengo, J.; Yamashita, D.; Oh, H.; Konialian, A.; Yen, H.; Rozamus, L.; Brandt, M.; Bossard, M.; Levy, M. A.; Eggleston, D. S.; Liang, J.; Schultz, L. W.; Stout, T. J.; Clardy, J. *J. Am. Chem. Soc.* **1993**, *115*, 9925–9938.
- Xu, Y.; Wang, R. *Proteins: Struct., Funct., Genet.* **2006**, *64*, 1058–1068.
- Wang, J.; Deng, Y.; Roux, B. *Biophys. J.* **2006**, *91*, 2798–2814.
- Zhang, C.; Chen, J.; DeLisi, C. *Proteins: Struct., Funct., Genet.* **1999**, *34*, 255–267.
- Ubbink, M. *FEBS Lett.* **2009**, *583*, 1060–1066.
- Schreiber, G.; Haran, G.; Zhou, H. X. *Chem. Rev.* **2009**, *109*, 839–860.
- Mittag, T.; Schaffhausen, B.; Günther, U. L. *J. Am. Chem. Soc.* **2004**, *126*, 9017–9023.
- Sydor, J. R.; Engelhard, M.; Wittinghofer, A.; Goody, R. S.; Herrmann, C. *Biochemistry* **1998**, *37*, 14292–14299.
- Selzer, T.; Albeck, S.; Schreiber, G. *Nat. Struct. Biol.* **2000**, *7*, 537–541.
- Lamb, M. L.; Jorgensen, W. L. *J. Med. Chem.* **1998**, *41*, 3928–3939.
- Sun, F.; Li, P.; Ding, Y.; Wang, L.; Bartlam, M.; Shu, C.; Shen, B.; Jiang, H.; Li, S.; Rao, Z. *Biophys. J.* **2003**, *85*, 3194–3201.
- Zhao, L.; Liu, H.; Wang, L.; Li, S. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 4385–4390.
- Babine, R. E.; Bender, S. L. *Chem. Rev.* **1997**, *97*, 1359–1472.
- Zacharias, M. *Proteins: Struct., Funct., Genet.* **2004**, *54*, 759–767.
- Ivery, M. T. G.; Weiler, L. *Bioorg. Med. Chem. Lett.* **1997**, *5*, 211–232.
- Rosen, M. K.; Schreiber, S. L. *Angew. Chem., Int. Ed.* **1992**, *31*, 384–400.
- Lemons, D. S.; Gythiel, A. *Am. J. Phys.* **1997**, *65*, 1079–1081.
- Lazaridis, T.; Karplus, M. *Proteins: Struct., Funct., Genet.* **1999**, *35*, 133–152.
- Brünger, A.; Brooks, C. L., III; Karplus, M. *Chem. Phys. Lett.* **1984**, *105*, 495–500.
- Curcio, R.; Caffisch, A.; Paci, E. *Protein Sci.* **2005**, *14*, 2499–2514.
- Jorgensen, W. *Nature* **2010**, *466*, 42–43.
- Colizzi, F.; Perozzo, R.; Scapozza, L.; Recanatini, M.; Cavalli, A. *J. Am. Chem. Soc.* **2010**, *132*, 7361–7371.
- MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarty, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
- Foloppe, N.; MacKerell, A. D., Jr. *J. Comput. Chem.* **2000**, *21*, 86–104.
- Brooks, B. R.; Brooks, C. L., III; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caffisch, 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. *J. Comput. Chem.* **2009**, *30*, 1545–1614.
- Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327–341.
- Brooks, C.; Karplus, M. *J. Chem. Phys.* **1983**, *79*, 6312–6325.
- Paci, E.; Caffisch, A.; Plückthun, A.; Karplus, M. *J. Mol. Biol.* **2001**, *314*, 589–605.
- Swegat, W.; Schlitter, J.; Krüger, P.; Wollmer, A. *Biophys. J.* **2003**, *84*, 1493–1506.
- Huang, H.; Ozkirimli, E.; Post, C. *J. Chem. Theory Comput.* **2009**, *5*, 1304–1314.
- Li, A.; Daggett, V. *J. Mol. Biol.* **1998**, *275*, 677–694.
- Caves, L. S.; Evanseck, J. D.; Karplus, M. *Protein Sci.* **1998**, *7*, 649–666.
- van Duynne, G. D.; Standaert, R. F.; Karplus, P. A.; Schreiber, S. L.; Clardy, J. *J. Mol. Biol.* **1993**, *229*, 105–124.
- Nakanishi, I.; Fedorov, D. G.; Kitauro, K. *Proteins: Struct., Funct., Genet.* **2007**, *68*, 145–158.
- Rhodes, G. A User’s Guide to Crystallographic Models. In *Crystallography Made Crystal Clear: A Guide for Users of Macromolecular Models*; Academic Press, an imprint of Elsevier Science: San Diego, CA, 2006; pp 159–186.
- Cornell, W.; Abseher, R.; Nilges, M.; Case, D. A. *J. Mol. Graph. Model.* **2001**, *19*, 136–145.
- Wilson, K. P.; Yamashita, M. M.; Sintchak, M. D.; Rotstein, S. H.; Murcko, M. A.; Boger, J.; Thomson, J. A.; Fitzgibbon, M. J.; Black, J. R.; Navia, M. A. *Acta Crystallogr.* **1995**, *D51*, 511–521.
- Peters, G. H.; Bywater, R. P. *Protein Eng.* **1999**, *12*, 747–754.

- (44) Teplyakov, A.; Sebastiao, P.; Obmolova, G.; Perrakis, A.; Brush, G. S.; Bessman, M. J.; Wilson, K. S. *EMBO J.* **1996**, *15*, 3487–3497.
- (45) Hornak, V.; Okur, A.; Rizzo, R. C.; Simmerling, C. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 915–920.
- (46) Burkhard, P.; Taylor, P.; Walkinshaw, M. D. *J. Mol. Biol.* **2000**, *295*, 953–962.
- (47) Rosen, M. K.; Yang, D.; Martin, P. K.; Schreiber, S. L. *J. Am. Chem. Soc.* **1993**, *115*, 821–822.
- (48) Castellano, R. K.; Diederich, F.; Meyer, E. A. *Angew. Chem., Int. Ed.* **2003**, *42*, 1210–1250.
- (49) Chang, C. A.; McLaughlin, W. A.; Baron, R.; Wang, W.; McCammon, J. A. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 7456–7461.
- (50) Dobbins, S. E.; Lesk, V. I.; Sternberg, M. J. E. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 10390–10395.
- (51) Lepre, C. A.; Pearlman, D. A.; Cheng, J. W.; DeCenzo, M. T.; Moore, J. M.; Livingston, D. J. *Biochemistry* **1994**, *33*, 13571–13580.
- (52) Hamilton, G. S.; Huang, W.; Connolly, M. A.; Ross, D. T.; Guo, H.; Valentine, H. L.; Suzdak, P. D.; Steiner, J. P. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 1785–1790.
- (53) Yang, W.; Rozamus, L. W.; Narula, S.; Rollins, C. T.; Yuan, R.; Andrade, L. J.; Ram, M. K.; Phillips, T. B.; van Schravendijk, M. R.; Dalgarno, D.; Clackson, T.; Holt, D. A. *J. Med. Chem.* **2000**, *43*, 1135–1142.
- (54) Nikitina, E.; Sulimov, V.; Zayets, V.; Zaitseva, N. *Int. J. Quantum Chem.* **2004**, *97*, 747–763.
- (55) Koser, P. L.; Eng, W. K.; Bossard, M. J.; McLaughlin, M. M.; Cafferkey, R.; Sathe, G. M.; Faucette, L.; Levy, M. A.; Johnson, R. K.; Bergsma, D. J.; Livi, G. P. *Gene* **1993**, *129*, 159–165.
- (56) Ikura, T.; Ito, N. *Protein Sci.* **2007**, *16*, 2618–2625.
- (57) Orozco, M.; Tirado-Rives, J.; Jorgensen, W. L. *Biochemistry* **1993**, *32*, 12864–12874.
- (58) Kabsch, W.; Sander, C. *Biopolymers* **1983**, *22*, 2577–2637.
- (59) Stewart, J. J. P. *MOPAC2006*, version 1.0; Fujitsu Limited: Tokyo, Japan, 2006.

Rubredoxin Function: Redox Behavior from Electrostatics

Ana Patricia Gamiz-Hernandez,[†] Gernot Kieseritzky,[†] Hiroshi Ishikita,[‡] and E. W. Knapp^{*,†}[†]Institute of Chemistry and Biochemistry, Department of Biology, Chemistry and Pharmacy, Freie Universität Berlin, Fabeckstrasse 36a, D-14195, Berlin, Germany[‡]Career-Path Promotion Unit for Young Life Scientists, Kyoto University, 202 Building E, Graduate School of Medicine, Yoshida-Konoe-cho, Sakyo-ku, Kyoto 606-8501, Japan

Supporting Information

ABSTRACT: Continuum electrostatic theory was applied to compute redox potentials of rubredoxin (Rd) proteins. We used multiple side chain conformers of Rd crystal structures, optimized geometries of salt bridges, mutated residues, and residues in the neighborhood of the iron–sulfur complex (FeS complex) self-consistently for given solvent pH and redox potential. The following contributions to Rd redox potentials are discussed: side chain conformations, H-bond geometries of the FeS complex, dielectric environment, charged residues, and salt bridges. We considered 15 different Rd's (of different species/strains and mutants) with available crystal structures whose redox potentials vary between -86 mV and $+31$ mV. The computed redox potentials deviated by less than 16 mV, root-mean-square deviation (RMSD), from measured values. The amide H-bond geometry is considered to be crucial for the variation of Rd redox potentials. To test this assumption, we considered 14 mutant Rd's for which we modeled the structures based on Rd from WT *Clostridium pasteurianum* (Cp) leaving the amide H-bond geometry of the FeS complex invariant. Here, we obtained an RMSD of only 14 mV with measured values demonstrating that the amide H bond geometries cannot be a major factor determining Rd redox potentials. We analyzed the factors determining the Rd redox potentials of a mesophilic and a thermophilic Rd differing by nearly 90 mV. We found that half of the difference is due to sequence and half is due to backbone variations. Albeit salt-bridge networks vary considerably between these two Rd's and are considered to be responsible for differences in thermostability, their overall influence on Rd redox potentials is small.

INTRODUCTION

Iron–sulfur proteins are ubiquitous in living systems.^{1–5} Rubredoxin (Rd) is a redox-active protein and belongs to the simplest type of iron–sulfur proteins. One of its functions is storing and transferring electrons to alkane hydroxylase from Rd reductase. Alkane hydroxylase belongs to a large class of membrane proteins of interest as biocatalysts for the production of alcohols, fatty acids, and epoxides.^{6,7} The Rd's under study have 52 to 54 amino acids, except one with 45 residues. The redox-active iron sulfur complex (FeS complex) in Rd consists of a single iron, ligated to four sulfur atoms of cysteine residues (see Figure 1). The FeS complex in Rd involves only a single iron. Thus, quantum chemical computations are facilitated considerably, since antiferromagnetic coupling appearing in multinuclear FeS complexes is absent.^{8–14} Studies on iron sulfur proteins and corresponding model systems have demonstrated that several factors are important in understanding their redox properties.^{3,5,9–13,15–25} These are the number, types, and strengths of H bonds with the FeS complex;^{3,5,13,15,17,21–25} the polarity and protonation states of residues surrounding the FeS complex;^{26,27} the degree of solvent exposure;^{16,20,21,28} and quite generally, the electrostatic environment defined by the protein–solvent boundary.^{9–12,20}

Using NMR spectroscopy, Lin et al.^{15,23} found a correlation between the redox potentials of Rd's from *Clostridium pasteurianum* (Cp) mutants and the strengths of the six amide H bonds with the FeS complex characterized by H-bond lengths. They

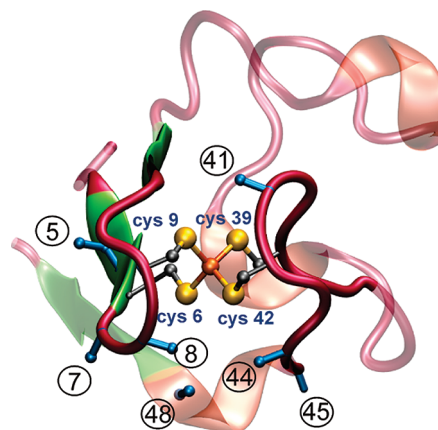


Figure 1. Rubredoxin (Rd) crystal structure from Cp (PDB code 1IRO³⁸) with 53 amino acids and the iron–sulfur complex (FeS complex) in the center. The iron is covalently bound to four cysteine sulfur atoms belonging to two different protein loops (in red). Blue C_{α} – C_{β} bonds with residue numbers depict amino acid side chains of Rd that can have an impact on the redox potential, as will be discussed in text.

inferred variations of H–S distances ranging from 2.2 Å to 2.8 Å and concluded that these are essentially responsible for the variations of redox potentials of different Rd's from mutant Cp.

Received: August 24, 2010

Published: January 21, 2011

Table 1. Sequence Comparison between Eight Rubredoxins (Rd) from Seven Different Species (Two Different Strains from *Dv*) with Available Crystal Structures: *Clostridium Pasterianum* (*Cp*),^{38,81} *Desulfovibrio Desulfuricans* (*Dd*),^{82,83} *Desulfovibrio Gigas* (*Dg*),^{84,85} *Desulfovibrio vulgaris* (*Dv*),^{86–88} *Pyrococcus Abbyssis* (*Pa*),^{25,80} *Pyrococcus Furiosus* (*Pf*),^{32,49,89} and *Pseudomonas aeruginosa* (*Ps*)⁷^a

Rd	1	11	21	31	41	51
<i>Cp</i>	MKKYTC	YIYNPEDG	DPDNGVNP	GTDFKDI	IPDDWV	CPLCGVVKDQFEVEE
<i>Pf</i> ³	-AKWVCKIC	YIYDE	DAGDPDNG	ISPGTKF	EELPDDWV	CPICGAPKSEFEKLED
<i>Dg</i>	MDIYVCTVC	YEYDPAK	DPDPSG	IKPGTKF	EELPDDWA	CPVCGASKDAFEKQ
<i>Dd</i>	MQKYV	CNVCG	YEYDPAE	HD	NVPFDQLPDDW	CCPVCGVSKDQFSPA
<i>Dv</i> ¹	MKKYVCTVC	YEYDPAE	GDNDNGV	KPGTAF	EDVPADWV	CPICGAPKSEFEPA
<i>Dv</i> ²	MKKYVCTVC	YEYDPAE	GDNDNGV	KPGTSTF	DDLPAWV	CPVCGAPKSEFEAA
<i>Pa</i> ⁴	MAKWRCKIC	YIYDE	DEGDNDNG	ISPGTKF	EELPDDWV	CPLCGAPKSEFERIE
<i>Ps</i>	MRKWQCVVC	GFIYDE	ALGLPEE	EGIPAGTRW	EDIPADWV	CPDCGVGKIDFEMIE

^a Some of the conserved residues are highlighted in different background colors: cysteines in yellow, amino acids surrounding the FeS complex (see Fig. 1) in light gray, amino acids belonging to different salt-bridge patterns in the crystal structures in dark blue/red/green (see details in Supporting Information Table S3). The residue numbers in the first line refer to Rd from *Cp*. ¹ Miyazaki strain of *Dv*. ² Hildenborough strain of *Dv*. ³ N-terminal alanine forms a salt bridge with D14. ⁴ In *Pa*, E53 is involved in two different salt bridges: a side chain of E53 with K3 and C-terminal acidic group of E53 with R51.

However, the mutations of Rd do not only vary the H-bond strengths but also conformations, charge distributions, and the electrostatic boundary between the protein and solvent, which all can shift the Rd redox potentials.^{20,21,29} One purpose of the present work is to study the influences of these different factors on Rd redox potentials.

Subtle conformational details may also influence the redox potentials of Rd-like FeS complexes.³⁰ In this context, it is worth mentioning that the FeS complex structures of Rd's vary with the redox state^{31–34} and differ from crystal structures of small Rd-like FeS-complex models.^{21,35} Such conformational changes of Rd-like FeS complexes are quite important, as steric constraints imposed by the protein alter the geometry and energetics.^{21,35} Therefore, quantum chemical computations of small FeS complexes may not faithfully describe the properties of the FeS complex in the protein, while larger models³⁶ that can represent the protein environment more faithfully may introduce new problems, if they possess several energy minima with different geometries.³⁷

One of the challenges of the Rd function is to understand that at room temperature ($T = 25\text{ }^{\circ}\text{C}$) the redox potential from the mesophilic *Cp* is at -55 mV , while it is at $+31\text{ mV}$ ³⁹ from the thermophilic *Pyrococcus furiosus* (*Pf*)³⁹ species, albeit the structures (backbone RMSD of about 0.5 \AA) and sequences (sequence similarity of 70%, sequence identity close to 60%, see Table 1) are similar. The sequences of Rd's from these two species vary in charged residues and salt bridges, which were suggested to play a role in thermostability.^{40–42} Surprisingly, the Rd redox potential from WT *Pf*, which at room temperature is $+31\text{ mV}$, assumes a value of -93 mV at $95\text{ }^{\circ}\text{C}$,³⁹ much lower and close to the room temperature value of -77 mV of Rd from WT *Cp*. Understanding the peculiarities of redox potentials of these two Rd's has been the subject of many experimental and theoretical studies during recent years.^{16,27,29,32,39,43–57}

Two main sources for temperature dependent redox potentials of proteins can be named. (i) One reason for the observed downshift of the Rd redox potential with rising temperature could be the lowering of the water dielectric constant at higher temperatures,³⁹ which at $T = 100\text{ }^{\circ}\text{C}$ is as low as $\epsilon_w = 55.5$.⁵⁸ Assuming that the thermophilic Rd structure remains invariant at high temperatures, the FeS complex does not become solvent

exposed, such that there is only a moderate decrease in dielectric screening of the FeS complex. This may destabilize the more negatively charged reduced state (total charge -2) slightly more than the oxidized state (total charge -1). Hence, we would expect a small downshift of the Rd redox potential with rising temperature. This is likely the result of subtle changes in the protonation pattern that can also influence the Rd redox potentials. (ii) A second source is certainly structural changes induced at higher temperatures. Partial unfolding of the very stable Rd from the thermophilic species *Pf* is unlikely. Opening of small cavities close to the FeS complex could allow water to penetrate and form additional H bonds with the negatively charged sulfurs of the FeS complex. Again, this would stabilize more the reduced than the oxidized state and thus upshift the redox potential with rising temperature, opposite to the trend observed in experiments. However, the six amide H bonds with the sulfurs of the FeS complex may become loose at higher temperatures. This would destabilize the reduced state more than the oxidized state, resulting in a potentially significant redox potential downshift, as observed in experiments. Albeit these temperature effects are certainly of interest, we restrict our present study to room temperature only.

To compute the Rd redox potentials, we evaluate electrostatic energies in continuum dielectric models.^{59–69} This procedure yielded agreement with experiments for protein cofactors as different as hemes,^{62,63,66,69} chlorophylls,⁶⁸ and quinones^{64,65,67} in different types of proteins. In this approach, measured redox potentials of appropriate model systems in solvents were used as references to compute the shift in redox potentials between the solvent and protein environment, evaluating the electrostatic energies of solvation. For Rd-like FeS complexes, such experimental redox potential values are not available. Therefore, we considered in a previous study²¹ FeS complex model compounds whose redox potentials were obtained by a combination of *ab initio* quantum chemical and electrostatic energy computations. This study is in line with previous work^{21,70,71} where we computed redox potentials of a larger number of organic compounds⁷⁰ and transition metal complexes,⁷¹ yielding agreement with experimental results with a root-mean-square deviation (RMSD) of 55 mV and 60 mV , respectively. For a small Rd-like FeS complex, $[\text{Fe}(\text{SCH}_2\text{CH}_3)_4]^{-2-}$ with S_4 instead of the

C_2 symmetry prevalent in Rd our computations yielded in acetonitrile a redox potential of -813 mV, as compared to the measured value of -838 mV.⁷¹ Using the same approach, we obtained for a large Rd-like FeS complex²¹ that mimics the situation in Rd faithfully with proper C_2 symmetry and H-bond pattern a redox potential of $+194$ mV in a dielectric continuum of $\epsilon_w = 80$ corresponding to water. In the present work, we considered the complete Rd protein from different species and from different mutants to study the variation of the redox potentials of the embedded FeS complex. These Rd redox potentials were computed with the newly developed extended version of the program Karlsberg+.^{72,73} In contrast to the former version of Karlsberg,⁶¹ which uses the charge model of a single protein structure (closely related to a corresponding crystal structure with optimized hydrogen atom positions), the new Karlsberg version can be used to compute pK_A values and redox potentials in proteins by optimizing not only hydrogen atom positions but also specific amino acid side chains self-consistently at a given solvent pH and redox potential.

The aim of this study is to understand how Rd proteins from different species tune their FeS complex redox potentials by varying the amino acid composition and protein conformation. We quantify how much the dielectric environment affects the redox potentials of different Rd's from mutant *Cp* by modeling the Rd structures using a polypeptide backbone and unchanged side chain conformers from the WT *Cp* crystal structure. Since the amide H bonds with the FeS complex remain invariant in these mutant Rd structures, we can study redox potential variations of factors that differ from H-bond strengths. We also investigated the role of charged amino acids and salt bridges to understand the redox potential differences between mesophilic and thermophilic Rd's. Furthermore, we consider the conformational variations observed in Rd crystal structures and evaluate the influence they have on Rd redox potentials.

METHODS

Foundations and Conditions for Electrostatic Energy Computations. All redox potentials reported in this work are based on electrostatic energies evaluated by solving numerically the linearized Poisson–Boltzmann equation:

$$\vec{\nabla} \cdot [\epsilon(\vec{r}) \vec{\nabla} \phi(\vec{r})] - \kappa^2(\vec{r}) \phi(\vec{r}) = -4\pi\rho(\vec{r}) \quad (1)$$

using the program APBS^{74,75} to obtain the electrostatic potential ϕ , where ϵ is the dielectric constant, κ is the ionic strength, and ρ is the charge distribution. The total electrostatic energy is given by refs 60 and 72:

$$\begin{aligned} \Delta G^{(n,l)} = & \sum_{\mu=1}^N x_{\mu}^{(n,l)} (g_{\mu}^{(l)} - g_{\text{ext},\mu}) \\ & + \sum_{\mu, \nu=1, \mu < \nu}^N x_{\mu}^{(n,l)} x_{\nu}^{(n,l)} \Delta \Delta w_{\mu\nu}^{(l)} + \Delta G_{\text{conf}}^{(l)} \end{aligned} \quad (2)$$

where μ and ν label the total number of N -charge variable (titratable or redox-active) groups in the protein, while n labels one of the many possible combined protonation/redox patterns and l labels the protein conformer. The zero point of energy corresponds to the reference charge state where all titratable

groups are deprotonated and all redox-active groups are reduced. The set of integers $x_{\mu}^{(n,l)}$, $\mu = 1, \dots, N$, describes the protonation/redox pattern n in protein conformer l ; $x_{\mu}^{(n,l)}$ vanishes if the group μ is deprotonated/reduced and is unity if the group μ is protonated/oxidized. The first term in eq 2 considers the electrostatic energy of individual charge variable groups μ in protein conformer l , if all other charge variable groups ($\nu: \nu \neq \mu$) are in the reference charge state. The second term in eq 2 considers the correction energy necessary, if the other charge variable groups (ν) are not in the reference charge state where $x_{\nu}^{(n,l)} = 0$. The third term in eq 2 accounts for the specific protein conformer l and vanishes if the reference conformer is adopted. The reference conformer can be arbitrarily chosen and in principle also the reference charge state, defining thus the zero point of energy of $\Delta G^{(n,l)}$. However, in the latter case, the $x_{\mu}^{(n,l)}$ in eq 2 have to be replaced by $\Delta x_{\mu}^{(n,l)} = x_{\mu}^{(n,l)} - x_{\mu, \text{ref}}^{(n,l)}$, where the $x_{\mu, \text{ref}}^{(n,l)}$ characterize the reference charge state. Since the energy eq 2 is only applied to evaluate energy differences, the choice of reference state has no influence on the results. The free energy of a proton or electron in the solvent is given by $g_{\text{ext},\mu}$ where the index μ is only needed to discriminate between the proton and electron depending on the type of charge variable group μ . If μ is a titratable group, $g_{\mu}^{(l)} = -\ln(10) RT_{298\text{K}} pK_{A, \text{intr}, \mu}^{(l)}$ and $g_{\text{ext},\mu} = -\ln(10) RT_{298\text{K}} \text{pH}$. If μ is a redox-active group, $g_{\mu}^{(l)} = F \cdot e_{\mu}^{(l)}$ and $g_{\text{ext},\mu} = F \cdot e_{\text{sol},\nu}$ where F is the Faraday constant and $e_{\mu}^{(l)}$ and $e_{\text{sol},\nu}$ are the redox potentials of redox-active group μ and the solvent, respectively. How the energy terms $g_{\mu}^{(l)}$ and $\Delta \Delta w_{\mu\nu}^{(l)}$ from eq 2 are evaluated with the help of the PB eq 1 is explained in detail in refs 60 and 72. The probability of the charge variable group μ to be in its protonated/oxidized state is computed as a Boltzmann average:

$$\begin{aligned} \langle x_{\mu} \rangle = & \frac{1}{Z} \sum_{n,l} x_{\mu}^{(n,l)} \exp\left(-\frac{\Delta G^{(n,l)}}{RT}\right) \text{ with} \\ Z = & \sum_{n,l} \exp\left(-\frac{\Delta G^{(n,l)}}{RT}\right) \end{aligned} \quad (3)$$

where the sums in eq 3 run over all conformers (l) and charge states (n). Since the number of different charge states can be enormously large, the sums are evaluated using Metropolis Monte Carlo importance sampling, as explained in more detail in ref 61.

The protein was described as a set of atomic partial charges [defining thus the charge distribution $\rho(r)$ in the PB eq 1] embedded in an inhomogeneous dielectric continuum where the dielectric constant was set to $\epsilon_p = 4$ inside the protein and $\epsilon_w = 80$ outside for bulk water. The more details of the protein are modeled explicitly, the lower is the value of the dielectric constant inside the protein volume. Considering a single or few protein conformations with a detailed atomic charge model and flexible protonation pattern, as done in the present study and in a legacy of our past studies, the dielectric constant $\epsilon_p = 4$ accounting for electronic polarizability and residual conformational variation of amino acid side chains of the protein yielded the best results. Larger values of the dielectric constant ϵ_p would become necessary if the protein model was less detailed. This is, for instance, the case for the aqueous solvent, where $\epsilon_w = 80$ is used, since no molecular details are considered. If two different electrostatic models of the same protein yield the same quality of agreement, the model with the lower value of the dielectric

constant is superior. Since one reviewer was questioning the dependence of the computed redox potentials on the protein dielectric constant, we added results for $\epsilon_p = 2$ and 8, which are discussed later, together with the results on the Rd mutants from *Cp*, and presented in detail in Figure S1 in the Supporting Information.

The boundary interface between the protein and solvent lumen was calculated by the molecular surface routine implemented in APBS using a solvent probe radius of 1.4 Å. Large enough cavities inside the protein can also be detected by this algorithm, such that inside their volumes a dielectric constant of $\epsilon_w = 80$ can be assigned. This feature becomes relevant if one considers mutants where a voluminous side chain is replaced by a smaller one. An ionic strength of potassium chloride concentration at 100 mM was implicitly included.

An alternative approach would be to use explicit water and to evaluate the energetics of the different protonation and redox pattern by molecular dynamics (MD) simulations, as done by other groups.^{76,77} This protein–water model is more detailed but requires ensemble averaging over different water conformations. The amount of CPU time needed when generating by MD simulations the necessary number of water conformations for a single combined protonation and redox pattern may be roughly the same as the time needed to solve a single linearized PB equation. However, MD simulations must be performed for each protonation pattern. In contrast, due to the additivity of electrostatic energies obtained by the linearized PB equation, it can be solved independently for the states of the individual titratable and redox-active groups. Hence, for a protein like Rd from WT *Cp*, with 20 titratable residues, 2²⁰ MD simulations are necessary, while in the present approach the PB equation needs to be solved roughly only 2×20 times. The typical accuracy for pK_A computations is about 1 pK_A unit corresponding to 60 meV.⁷² For redox potential computations, we expect an even higher accuracy, since changes of the charge pattern between the different states of a redox-active group are less localized than for protonation changes of titratable groups.

Atomic Partial Charges and Redox Potential of the Reference FeS Complex. Atomic partial charges were taken from the CHARMM22 force field^{78,79} if available. For the FeS complex, atomic partial charges were adapted from previous work where electrostatic potentials were computed with a combination of quantum chemical and electrostatic methods applied to the small Rd-like model $[\text{Fe}(\text{SCH}_2\text{CH}_3)_4]^{-2-}$ in C_2 symmetry,²¹ yielding a redox potential value of +56 mV in water ($\epsilon = 80$) and –165 mV for a continuum dielectric medium, corresponding approximately to protein environment ($\epsilon = 20$, Table 3 in ref 21). The too low value of the redox potential in the latter case is due to the absence of explicit H bonds with the sulfur atoms of the FeS complex in this small Rd-like model. The charge model of a more realistic larger Rd-like FeS complex involving C_2 symmetry and the appropriate H-bond pattern (model 4 in ref 21 with six amide H bonds), as it appears in Rd, yielded a redox potential value in water of +194 mV. The same large model embedded in a continuum dielectric of $\epsilon = 20$ that roughly approximates the protein environment, if for the protein no explicit atomic charges are used, yielded +57 mV,²¹ which is close to the interval (–87 to +39 mV) of measured Rd redox potentials. However, this large Rd-like model is too complex and in conflict with the CHARMM charge model and cannot easily be adjusted to compute electrostatic energies considering a detailed protein model with atomic charges. Therefore, we

used in the present study the smaller reference model $[\text{Fe}(\text{SCH}_2)_4]^{-2-}$ whose charges were made compatible with the CHARMM force field. For this smaller model, no measured redox potentials are available, presumably since it is not stable in solution. A detailed explanation is given in section S1 and charges are listed in Table S2 of the Supporting Information. Alternatively, using a reference redox potential value of +112 mV (56 mV larger than the redox potential computed for the small Rd-like model in a dielectric medium of $\epsilon = 80$) for the small charge adapted FeS complex in water ($[\text{Fe}(\text{SCH}_2)_4]^{-2-}$), the RMSD between computed and measured Rd redox potentials is minimized (see section S2 in the Supporting Information for a detailed discussion). The differences of the computed redox potentials with this reference redox potential are within the range of deviations to be expected from the *ab initio* computational method employed^{21,70,71} for the Rd-like FeS complexes. Since the variations in measured Rd redox potentials to be analyzed are of the same size as these differences, we prefer to use +112 mV as a reference redox potential value for the small FeS complex in our computations. Thus, we have introduced an adjustable parameter in the electrostatic energy computations.

Redox Potential Computations in Proteins with Multi-conformers. For the electrostatic computations of the Rd redox potentials, we employed a modified new version of Karlsberg+,^{72,73} which we used in the past primarily for accurate pK_A computations. Karlsberg+ combines classical electrostatic energy computations with pH-dependent conformational relaxation of salt bridges and H bonds inside proteins. The original Karlsberg+, available on the Web, tries to predict alternative atom positions for side chains of basic and acidic residues involved in salt bridges under high and low pH conditions, where these salt bridges are generally not stable any more. It does so by generating pH adapted conformations (PACs) that are obtained by self-consistent geometry optimizations of the input crystal structure combined with random changes of side chains of basic and acidic residues. In the new modified version of Karlsberg+, we consider redox rather than pH-dependent alternative side chain conformers. Accordingly, we introduced redox and pH 7 adapted conformations (RACs) for a given Rd redox state in the same spirit as we introduced PACs before to describe the pH dependence of the protonation pattern.⁷² These RACs can also be generated to model side chain conformers of mutated residues. FeS complex redox potential contributions due to specific amino acids were analyzed by setting the charges of the corresponding side chain atoms to zero while computing the Rd redox potential.

Generating Redox Potential Adapted Conformers (RACs). To compute redox and protonation equilibria in proteins, solvent pH and solvent redox adapted equilibrium protein conformers are generated from the appropriate crystal structures where hydrogen atoms were added by Karlsberg+. In summary, Karlsberg+ consists of three modules: (i) generation of PACs or RACs using a special CHARMM⁷⁸ script, (ii) computation of electrostatic energies with the Poisson–Boltzmann solver APBS,^{74,75} and (iii) Boltzmann averages of protonation and redox patterns and conformations (inherent in PACs or RACs) using the Metropolis Monte Carlo algorithm as in our legacy program Karlsberg.

This task is solved by self-consistent geometry optimization determining the most likely occupied protein conformations (with respect to protonation and redox pattern, hydrogen atoms, and side chain conformers of a selected set of residues) without

knowing its most likely protonation pattern and vice versa. To achieve this goal, Karlsberg+ performs an iterative procedure analogous to quantum mechanical self-consistent energy computations. First, Karlsberg+ calculates an initial protein protonation pattern at a high dielectric constant of $\epsilon_p = 80$ everywhere (to avoid an unphysical protonation pattern, which may easily appear initially) by adding and optimizing hydrogen atoms in the crystal structure with standard protonation (acids deprotonated and bases protonated). On the basis of this first guess of protonation pattern, Karlsberg+ performs geometry optimizations (at ambient pH and a given redox state using the CHARMM22⁷⁸ force field with $\epsilon_p = 1$ everywhere for all iteration steps) of amino acid side chains (including hydrogens) involved in salt-bridges and a selection of others that may influence the cofactor redox potential starting with a set of 30 structures with randomized side chain geometries for the considered amino acids. The resulting lowest energy global conformer of the protein is used in the next step to re-evaluate the protonation pattern by electrostatic energy computations (from now on, with $\epsilon_p = 4$ in the protein and $\epsilon_w = 80$ outside). If it changes, the geometry optimization step is repeated using the new protonation pattern. The cycle is repeated until the protonation pattern remains constant at a given pH and redox state. With this procedure, five RACs are generated for each redox state at ambient pH, starting with a different random seed in each case. Finally, Karlsberg+ computes redox titration curves of the protein using the ensemble of generated RACs with their electrostatic energies. This is done using Metropolis Monte Carlo importance sampling according to eq 3 considering the states of lowest electrostatic energy of the protonation, redox, and conformation pattern of the protein and allows calculating the redox potentials as proper Boltzmann-weighted thermodynamic averages.⁶¹

Besides the treatment of salt bridges, Karlsberg+ can now also generate RACs with geometry optimized side chains of mutated residues (and of other electrostatically relevant side chains of varying geometry), which is useful when no experimental coordinates are available for them. Since the side chains of the mutated residues possess different volumes and shapes than the side chains they are replacing, we included also the side chains of the neighbor residues (around 4 Å from the mutated residue) in the optimization procedure to generate RACs. If, for the same protein, crystal structures from two different species are available, it is instructive to build RACs that use the whole polypeptide backbone and chain conformers of all conserved residues from one protein but have inserted side chain conformers of the other protein structure for the nonconserved residues. In case side chains were found in multiple conformations in the protein crystal structure, RACs were generated for each combination of side chain conformers and used in the Boltzmann averaging.

Relaxing amino acid side chains can also be quite useful, if a protein crystal structure involves structural deficiencies. This is for instance the case for Rd crystal structures from Cp that were solved as trimers^{28,55} and tetramers,⁸⁰ where some side chains belonging to different monomers are in too close contact and needed to be relaxed.

Usage of Rubredoxin Crystal Structures and RACs. The experimental information of atomic coordinates of all considered Rd's in this study were taken from the protein database (PDB): these are the WT structures of Rd's in the oxidized state from eight different species/strains, as listed in Table 1, where the sequences are aligned. We also considered available Rd crystal structures in the reduced state from Cp (WT³¹ and

L41A mutant³⁴) and from WT Pf.³² The lengths of the six amide H bonds with the sulfur atoms of the FeS complex vary within the same Rd and between different Rd's typically between 3.34 and 3.99 Å, with an average of about 3.5 Å (see in ref 21 Supporting Information Table S1.2). More information regarding PDB codes, structure resolutions, redox states, and the literature are given in Table S3 of the Supporting Information.

We also used crystal structures of Rd's from mutant Cp (L41A,³⁴ V44L,⁵⁵ V44A,⁵⁵ V44G,⁵⁵ and V44G/G45P⁵⁵) and from mutant Pa (W4L/R5S and W4L/R5S/A44S^{25,80}). We excluded Rd structures based on NMR experiments. All Rd crystal structures from Cp (WT and mutants) were pairwise superimposed using the Kabsch algorithm⁹⁰ relative to the structure from WT Cp (1FHH) considering atoms of the FeS complex only (iron, sulfurs, and C_β atoms of the four ligating cysteines) to detect residues whose side chains are in different conformation (see Supporting Information Table S4). In four Rd crystal structures (WT Cp and V44A, V44G, and V44L mutant Cp), coordinates of the C-terminal residue (E54) are missing (53C_α-Fe distance 18.2 Å). The missing atoms were modeled yielding computed redox potentials that deviated by less than 4 mV from the corresponding values in absence of E54. Therefore, in the following, computed redox potentials of these Rd's refer to the crystal structures without a modeled C-terminal residue.

We computed the Rd redox potentials using two different sets of RACs generated by Karlsberg+: a minimal and an extended set. In the minimal set of RACs, we considered all available crystal structure data. These are crystal structures from different laboratories, for different redox states, for all monomers, if there are several monomers per unit cell and for all combinations of multiple side chain conformers (for more details, see Supporting Information Table S3). For the three Rd's, where crystal structures are available for both redox states, we generated RACs of the oxidized state also using the reduced crystal structure and vice versa. The RACs for the extended set are obtained by generating for each RAC of the minimal set all combinations of side chain geometries (1) of residues involved in salt-bridges, (2) of selected residues near the FeS complex, and (3) of mutated residues (if no corresponding crystal structures are used instead). All generated RACs were generated at constant pH = 7. Note that RACs based on a single Rd crystal structure involve only side chain and no backbone variations. However, we also used different Rd crystal structures (from different laboratories, for different redox states, different monomers in the unit cell of the same crystal) and generated for each of these structures RACs. Hence, the complete set of RACs involved also backbone variations if not otherwise stated.

Measured Rd Redox Potentials. Measured redox potentials of all considered Rd's are listed in Table S5 (Supporting Information). Redox potentials of the same protein measured by different groups vary by about 20 mV. This may be due to different measurement methods and usage of redox modifiers (like polylysine) to obtain stable square voltammetry for negatively charged proteins like Rd.^{29,91} The use of these modifiers can slightly upshift redox potentials.^{29,91} Furthermore, there are small variations of pH and temperature where Rd redox potentials were measured (see for example Rd from Pf, Table S5). Accordingly, Rd redox potentials are more positive at lower than at higher pH⁹² and decrease with higher temperatures.³⁹ In our redox potential computations, we assume standard conditions,

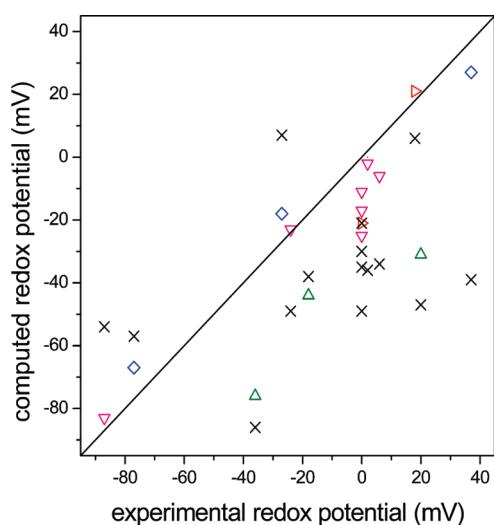


Figure 2. Comparison of measured and computed redox potentials of Rd's with 15 different sequences using the corresponding crystal structures. These are WT Rd's from eight different species/strains (*Cp*, *Pa*, *Pf*, *Ps*, *Dg*, *Dv(H)*, *Dv(M)*, and *Dd*) and seven Rd's from mutant *Cp* (V44A, V44L, V44A/G45P, V44G, and L41A) and from mutant *Pa* (W4L/R5S and W4L/R5S/A44S). The symbol "x" refers to computations based on the minimal set of RACs (see Methods, RMSD 42 mV). The open symbols (blue \diamond , red right-facing \triangle , green \triangle , and magenta ∇) refer to computations using the larger extended set of RACs (RMSD 16 mV). For the 12 Rd's marked by open triangles (red right-facing \triangle , green \triangle , and magenta ∇), crystal structures are only available for the oxidized states. For the three Rd's marked by blue \diamond , crystal structures are also available for the reduced states. The three Rd redox potentials from *Pa* denoted by green \triangle were measured in the same lab⁸⁰ using polylysine as a redox modifier that can upshift Rd redox potentials (see discussion in text). Rd redox potential values from two species (*Cp* V44A/G45P and WT *Ps*) denoted by red right-facing \triangle were not measured but estimated.⁷ Redox potentials are listed in Table S7 (Supporting Information).

i.e., pH 7 and 25 °C, and no adjustment of measured redox potentials due to the use of redox modifiers was considered.

RESULTS AND DISCUSSIONS

Overview of Computations. The evaluation of Rd redox potentials in this study consists of five parts. (1) Rd redox potentials of different species and their mutants were computed using the minimal set of RACs as described in the Methods section and Table S3 (Supporting Information). (2) We considered the same Rd's but use now the extended set of RACs as described in the Methods section. (3) To study the dependence of the Rd redox potentials on H-bond geometries, side chain conformers, and charges, we focused on mutant Rd's from a single species (*Cp*), using RACs that were based on the single Rd crystal structure from WT *Cp*. (4) We consider the Rd's from mesophilic *Cp* and thermophilic *Pf* to discuss the factors that explain the difference in their Rd redox potentials. (5) The protein backbone variation close to the FeS complex observed in Rd crystal structures from *Pa* (mutant W4L-R5S) allows for studying computationally how geometry variations of H bonds with FeS complex sulfurs influence the Rd redox potential.

1. Computed Rd Redox Potentials Based on the Minimal Set of RACs. We generated RACs according to the minimal set. In

these preliminary computations, the correlation between measured and computed Rd redox potentials is poor (RMSD of 42 mV; see "x" in Figure 2; values are listed in Supporting Information Table S7). Results of similar quality were reported before¹⁶ using solely Rd crystal structures, optimizing only hydrogen atom positions. Some improvement was obtained when the computed Rd redox potentials were based on structures from MD simulations^{16,57} However, these MD based approaches generally involve global changes in Rd structures, making it difficult to quantify the different redox potential contributions. In the present study, only moderate structural modifications of the Rd crystal structures were considered, making it easier to identify specific influences on Rd redox potentials, as shown in the next two sections, 2 and 3.

2. Computed Rd Redox Potentials Based on the Extended Larger Set of RACs. We calculated the Rd redox potentials again using the extended larger set of RACs. Here, we optimized side chains of a selected set of residues close to the FeS complex (S, 7, 8, 41, 44; see Figure 1) and of residues involved in salt bridges for the oxidized and reduced states. In studies of Rd's from *Cp* mutants,^{27–29,31,34,38,46,55} it was suggested that these selected residues influence the Rd redox potential (see also, next section). The correlation of measured and calculated Rd redox potentials of all 15 Rd's yields now an RMSD of only 16 mV (open symbols in Figure 2; values listed in Supporting Information Table S7). This is a significant improvement compared to the results obtained by optimizing hydrogen atom positions only ("x" in Figure 2). For 13 of these Rd's, the redox potentials were measured, while for two Rd's (from WT *Ps* and from V44A/G45P mutant *Cp*; red right-facing \triangle in Figure 2), only estimated values are available. Nevertheless, no difference in quality of agreement with computed values could be observed for the latter two.

The largest discrepancies are systematic deviations that occur for redox potentials of the three Rd's from *Pa* (green \triangle in Figure 2). These were measured in the same lab^{25,80} using polylysine as a redox mediator, which may have caused an upshift of the measured redox potential by about 20 mV.⁹¹ The RMSD increases slightly (17 mV), if we consider the 12 Rd's where crystal structures are available only for the oxidized states (red right-facing \triangle , green \triangle , and magenta ∇ in Figure 2). It becomes smaller (8 mV), considering only the three Rd's (blue \diamond in Figure 2) where crystal structures are available for both redox states that were used to generate the RACs and where systematic deviations of the measured redox potentials (green \triangle in Figure 2) are excluded. Furthermore, we observe small but systematic deviations for the computed redox potentials to lower values for the Rd's where crystal structures are only available for the oxidized states (open triangles in Figure 2), while this seems to be not the case for the three redox potential values obtained for Rd's where crystal structures are available for both redox states (blue \diamond in Figure 2). In fact, it can be expected that the energy of a reduced state is higher if it is computed on the basis of the structure of the oxidized state, rendering computed redox potentials more negative as compared to the measured values. An analysis of the amide H-bond lengths reveals that they are slightly shorter (by less than 0.1 Å, Supporting Information Table S6) for the reduced state than for the oxidized state. Ignoring these small structural differences in the computation of Rd redox potentials may explain this small systematic deviation.

The relatively large differences in the sequences of the considered Rd's from seven different species make it difficult to

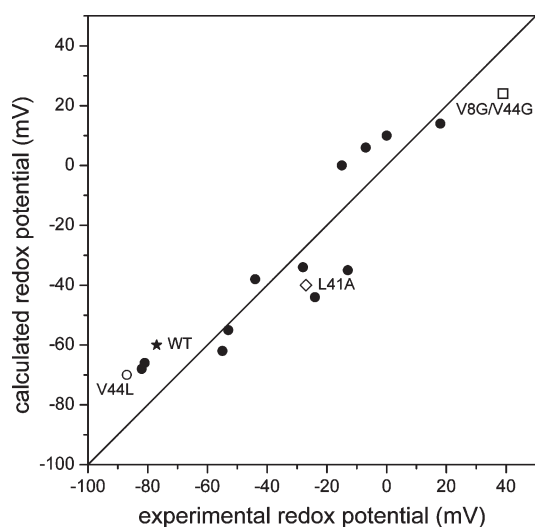


Figure 3. Comparison of 16 experimental and computed redox potentials of Rd's from WT and mutant *Cp*. Values presented in Supporting Information Table S8. On the basis of the Rd crystal structure from WT *Cp* (1IRO), the structures of 15 single and double mutants (V8A, V8D, V8G, V8I, V8L, V8R, V44A, V44G, V44I, V44L, V8G/V44G, V8I/V44G, V8I/V44I, V44A/G45P, and L41A) were modeled, and the redox potentials were computed (★, WT; ○, mutant V44L; ●, other mutants). For four mutants (V44L, V44A, V44G, and V44A/G45P), crystal structures are available. These were, however, not used to generate the RACs, except for mutant V44L, where a short side chain is exchanged by a long side chain. To reduce the number of necessary RACs, we used the rotamer of side chain L44 from the corresponding mutant crystal structure. Some mutants explicitly discussed in the text are labeled in the figure.

explain how individual amino acids may contribute to the shifts of redox potentials. Therefore, we will continue our study, focusing on the redox potentials of mutant Rd's from the same species. Nevertheless, a trend can be observed on the basis of the results presented so far: replacing small by large hydrophobic side chains in positions 8, 41, and 44 decreases the solvent accessibility of the FeS complex and consequently lowers its redox potential.²⁹

3. Redox Potentials of Mutant Rd's Based on Crystal Structures from WT *Cp*. We tried to assess the structure–function relationship between mutated residues and the FeS complex redox potential. For this purpose, we used the crystal structure of oxidized Rd from WT *Cp* (1IRO) as a scaffold to construct structures of 15 Rd's from mutant *Cp* (mutated at sequence positions 8, 41, 44, and 45) for which the FeS complex redox potentials have been measured (Table S5, Supporting Information). We generated an extended set of RACs with optimized side chains in six sequence positions (5, 8, 41, 44, 45, and 48; see Figure 1) involving mutated residues and direct neighbors of the FeS complex as well as side chains involved in salt-bridge networks (D29-K31-D32, D35-K46, and E51-K2-E53). With this extended set of RACs, computed and measured Rd redox potentials yielded a surprisingly good agreement with measurements showing an RMSD of only 14 mV (Figure 3; values listed in Supporting Information Table S8). Again, this is a considerable improvement compared to the preceding results where the minimal set of RACs was used (“×” symbols in Figure 2).

The data set of the 15 Rd's from mutant *Cp* and the corresponding WT Rd are suitable to demonstrate how the Rd

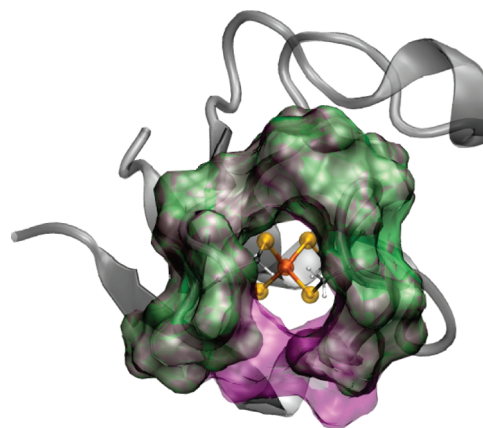


Figure 4. Comparison of solvent accessible surface for two Rd mutants from *Cp*: V8G/V44G (green) and V44L (purple). The replacement of Val in positions 8 and 44 with two Gly's in WT Rd from *Cp* exposes the FeS more to the solvent, while replacement with large amino acids like Leu isolates the FeS complex from the solvent. Rd's with a more solvent exposed FeS complex possess a more stabilized reduced state, leading to an upshift of the redox potential.

redox potentials vary, if the protein dielectric constant ϵ_p is set to 2 and 8. But, we will continue to strictly stick to $\epsilon_p = 4$, a value which we have found to be optimal and have used exclusively for more than 15 years. In both cases $\epsilon_p = 2$ and 8, the results are very poor, yielding RMSDs of 37 and 38 mV using redox potential reference values of 173 and -18 mV, respectively (for more details, see Figure S1 of the Supporting Information).

Since the protein backbone atom coordinates were kept invariant in all of these modeled Rd structures, the H-bond geometries of the FeS complex involving only backbone amide groups did not vary with mutant Rd's and redox states. Hence, the variation of Rd redox potentials can be explained by other factors than the local H-bond geometry around the FeS complex. One such important factor seems to be the FeS complex solvent accessibility and its variations with the redox states. This is corroborated by the Rd crystal structures from WT^{31,38} and L41A mutant *Cp*,³⁴ showing variations of T5, K31, I12, and L41 side chains (Supporting Information Figure S2). In the high resolution Rd crystal structure from WT *Cp* (PDB ID 1IRO), several residues (T5 P15, K31, L41, and E50) appear in multiple side chain conformations. Analog variations of residue side chains were also observed for Rd's from *Cp* mutants (Supporting Information Table S4).

The WT Rd from *Cp* possesses, at -77 mV, one of the lowest measured Rd redox potentials (★ in Figure 3). The largest Rd redox potential of $+39$ mV was obtained for the double mutant V8G/V44G from *Cp* (□ in Figure 3), while the L41A mutant from *Cp* has a redox potential lying in the center of the interval of measured Rd redox potentials (-27 mV, ◇ in Figure 3). When the bulky residue Val is replaced with Gly, as done in the double mutant V8G/V44G from *Cp*, the FeS complex becomes more solvent exposed (see Figure 4). This stabilizes the reduced state (more negatively charged than the oxidized state) more than the oxidized state, resulting in an upshift of the Rd redox potential. Dependences of the Rd redox potential on sizes of amino acid side chains of Rd's from *Cp* in positions 8 and 44 were already identified for several Rd mutants.²⁷ Accordingly, small amino acids like Gly upshift the redox potential, while large amino acids like Leu downshift the redox potential.

The most populated RACs of the oxidized and reduced states contain multiple conformers of L41 (Supporting Information Figure S3). We observed that L41 is less buried in the reduced state and thus opens a small cavity near the FeS complex. This “open” L41 conformer correlates with results of MD simulations, where a water molecule near L41 gets closer to the FeS complex in the reduced state.^{31,34} We forced the L41 side chain not to switch to the open conformer in the reduced state, which as expected downshifted the computed Rd redox potential by 34 mV. Alternative side chain rotamers for T5, T7, V8, I12, and Q48 were modeled showing only small variations of the Rd redox potential (about 10 mV), since these alternative conformations were only marginally populated in our RACs. Other multiple side chain conformers involving P15, K31, and E50 found in the Rd crystal structure from WT *Cp* (1IRO) had also practically no influence on the Rd redox potential.

The three salt bridges (K31-D29, K31-D32, and K46-D35, Supporting Information Table S3) present in the Rd crystal structure from WT *Cp* (1IRO) are also present in all RACs of Rd's from *Cp* mutants. Additional charged residues introduced in the mutants V8D and V8R do not form additional salt-bridges in the generated RACs. In both mutants, the redox potentials are upshifted compared to WT Rd (measured, 49 mV and 62 mV; computed, 26 mV and 58 mV, respectively). The upshift is surprising for the mutant V8D, since a negative charge (D8 was found to be deprotonated) is introduced, which is expected to stabilize the oxidized state of the FeS complex and, therefore, decrease its redox potential. Our computations can qualitatively explain this effect: Asp8 is solvent exposed and does not interact directly with the FeS complex that becomes more solvent exposed in this mutant, thus, upshifting the redox potential.

4. *Differences between Rd's from Cp and Pf.* Understanding redox potential variations between the Rd's from the mesophilic (*Cp*) and thermophilic (*Pf*) species has been the concern of many studies.^{16,27,28,31–33,39,44,45,48,49,51,52,56,57,93} It has been suggested that the presence of A44 in Rd from *Pf* contributes significantly to an upshift of the Rd redox potential.²⁸ However, there are variations in the Rd sequences from *Cp* and *Pf* at 24 sequence positions that may influence the redox potential. Alternatively, it was inferred from NMR spectroscopy on Rd's of WT and several mutants from *Cp* that the strength of the six amide H bonds (characterized by lengths) with the FeS complex sulfurs may have a major influence on the Rd redox potentials.^{15,23} Consequently, we compared the H-bond lengths of the crystal structures of the two Rd's from *Cp* and *Pf*. In fact, they are just slightly longer by about 0.07 (0.08) Å for the Rd from the mesophilic *Cp* as compared to the Rd from the thermophilic *Pf* in the corresponding oxidized (reduced) state (see Supporting Information Table S6). Hence, due to shorter amide H bonds, the reduced state is slightly more stabilized for the Rd from *Pf* than from *Cp*, yielding a redox potential that is higher for the former than for the latter Rd. Indeed, at room temperature, the measured Rd redox potential from *Pf* is 108 mV higher compared to the Rd from *Cp*. Hence, one might conclude that each of the six amide H bonds contributes on average a redox potential shift of 18 mV to make up the difference, which sounds reasonable. To verify this conclusion, we computed for each of the six amide H bonds the full contributions to the FeS complex redox potential by setting to zero the backbone charges of the corresponding NH and direct neighbor groups (C=O, C α H). We observed upshifts of the Rd redox potentials of on average 60 mV (Supporting Information Table S9) due to individual amide H-bond formation for both

proteins, in agreement with recent *ab initio* quantum chemical computations on FeS complex model complexes.²¹ The computed upshifts of the redox potentials were per H bond on average nearly the same for both proteins: 61.5 mV and 65 mV for the Rd's from *Cp* and *Pf*, respectively (Supporting Information Table S9). Hence, according to our computations, only 21 mV of the redox potential difference between Rd's from *Cp* and *Pf* is due to the difference in the amide H bonds, and the preceding conclusion is wrong. Obviously, there are other sequence related factors (the sequences of Rd from *Cp* and *Pf* differ in 24 positions) in the environment of the H bonds influencing the redox potentials of the two Rd's, as we will show in the following. Note that the conclusion on the limited role of the amide H bonds for the preceding case is not necessarily in conflict with our suggestion that the temperature dependence of the Rd redox potential from *Pf* may very well be explained by increasing H-bond lengths at higher temperatures, since in the latter case no sequence variations can interfere.

Without considering the peptide terminal groups and the FeS complex itself, there are three positively and 13 negatively charged residues in the WT Rd from the mesophilic *Cp*, while in the WT Rd's of the thermophilic species from *Pf* and *Pa*, there are five and six positively and 13 and 13 negatively charged residues, respectively. In the crystal structures, only part of these charged residues are involved in salt bridges. In contrast, the RACs generated for the two thermophilic Rd's from *Pf* and *Pa* notably involve more salt bridges than the corresponding crystal structures (see Table S10, Supporting Information). But, we found that these additional salt bridges have in general only a small influence on the Rd redox potentials (albeit they may be important for thermostability; Table S11, Supporting Information), such that they cannot be used to explain the redox potential differences of the Rd's from *Cp* and *Pf*. By setting to zero the atomic partial charges of all charged side chains, we observed downshifts of redox potentials by -63 mV and -38 mV for the Rd's from *Cp* and *Pf*, respectively (Table S9 Supporting Information). Hence, the different composition in charged amino acids can account for about a 25 mV difference in Rd redox potentials. The contributions of the backbone conformers are around 45 mV more positive for *Pf* than for *Cp* (values were obtained by setting the charges of backbone atoms to zero and computing the resulting shift in redox potential, see also Table S9), from which less than the half (21 mV) is due to differences in H-bond strengths. Hence, we can explain explicitly (25 mV + 45 mV) 70 mV of the total difference of 108 mV between the redox potentials of Rd's from WT *Cp* and *Pf*, although the sum of these contributions is not strictly additive. The remaining difference in the two Rd redox potentials results from a larger number of small contributions which are due to subtle variations in solvent accessibility and charge distributions that can also involve small changes in protonation pattern.

5. *Variation of the H-Bond Number in Mutant Rd's from Pa.* The crystal structure of the mutant Rd (W4L/R5S) from *Pa* (1YK5) was solved at high resolution (~ 0.7 Å) in the oxidized state. It exhibits multiple side chain conformers for six residues (five residues have two and one three conformers), which are not too distant to influence the FeS complex. In addition, the Rd crystal structure of the mutant shows for Lys7-Ile8-Cys9 two backbone conformations on which we will focus in our next modeling steps. One backbone conformation corresponds to the WT crystal structure from *Pa*; the second

backbone conformation differs from WT in the amide plane between Cys9–Ile8 that is rotated by about 20° and in the amide plane between Ile8–Lys7 that is translated by about 0.13 Å. We considered this protein backbone variation to understand how the amide H bond with the FeS complex can affect the Rd redox potential. In the Rd's from WT (W4L/R5S mutant) *Pa*, the amide H bonds formed by residues K7 and I8 exhibit N–S distances of 3.54 and 3.62 Å (3.41 and 3.79 Å), respectively. Hence, in the Rd from mutant *Pa*, one amide H bond is shorter and the other longer compared to the Rd from WT *Pa*.

For a careful case study on the influence of amide H-bond length variations, we modeled three different backbone conformations using side chain and backbone conformations of WT Rd from *Pa* except for (1) K7, (2) I8, and (3) both K7 and I8, where we used the backbone conformation of the W4L/R5S mutant Rd from *Pa* to change the corresponding amide H-bond lengths, while the torsion angles of the side chains of K7 and I8 were kept at the WT conformations. Finally, we also considered the Rd conformation (4), where we varied both amide H bonds and the side chain conformations of K7 and I8 simultaneously. Shortening of the amide H bond with K7 (case 1) yielded a redox potential upshift of 18 mV, while stretching the H bond (case 2) yielded a downshift by –21 mV. These shifts nearly cancel when varying both amide H bonds (case 3), yielding a small downshift of –7 mV. Interestingly, when also the side chain conformation was varied (case 4), a redox potential upshift of 23 mV was obtained. Thus, assuming no H-bond variation with the change in redox state, we could estimate (from cases 1 and 2) that shortening an amide H bond by 0.1 Å results in an upshift of about 13 mV. If all six amide H-bonds lengths would be stretched by about 0.1 Å and there would be no other structural variations as in the preceding case study, the small individual contributions to the Rd redox potential would add up to a relatively large downshift of about 65 mV, as we would expect for the Rd structures at higher temperatures. However, as seen from the comparison of the Rd's from *Cp* and *Pf* (see discussion in the preceding subsection 4), the influence of other sequence related factors may destroy correlations between H-bond strengths and redox potential shifts.

To investigate the influence of an S···HO H bond on Rd redox potentials, Bönisch et al.²⁵ generated the Rd triple mutant W4L/R5S/A44S from *Pa* whose Ser44 OH group forms an additional H bond with Cys6 of the FeS complex, whose length is 3.24 Å. In contrast to the corresponding double mutant W4L/R5S, the triple mutant does not involve multiple backbone conformers at Cys9. Hence, the six amide H bonds with the FeS complex remain intact. The measured redox potential upshift of the Rd triple mutant relative to the double mutant is 56 ± 20 mV.²⁵ This value is similar to our computed upshift of 39 mV using the crystal structure of the Rd triple mutant W4L/R5S/A44S from *Pa* (see Table S7, Supporting Information). In the most occupied RACs, Ser44 forms predominantly a H bond with Cys6, in agreement with the crystal structure. However, there are also RACs where Ser44 forms a H bond with Cys42, which is also a ligand of the FeS complex.

Alternatively, using the crystal structure of the Rd double mutant W4L/R5S from *Pa*, we modeled the Rd triple mutant that includes an extra O–H···S H bond. In this Rd structure, only the Ser44 side chain was optimized, which allowed estimation of the effect from this additional H bond excluding other influences. The resulting redox potential upshift of 58 mV is in

good agreement with previously computed values²¹ and experimental estimates.²⁵

CONCLUSIONS

Continuum electrostatic energy computations were applied on rubredoxin structures (RACs) generated from all information of available crystal structures to evaluate FeS complex redox potentials. The RACs are the result of limited geometry optimizations on the different crystal structures, performed self-consistently with a proper protonation pattern at a given pH and solvent redox potential. If only hydrogen atom positions were optimized self-consistently, the RMSD between measured and computed redox potentials of Rd's with 15 different sequences was 42 mV, while the measured redox potentials vary only between –86 mV and +31 mV. This poor result is enormously improved, if additional RACs were used that consider all combinations of optimized side chains for residues that are mutated or involved in salt bridges or close to the FeS complex, yielding an RMSD of only 16 mV.

Rd structures are generally very rigid. Therefore, Rd crystal structures can be at high resolution where they exhibit subtle structural variations like multioccupancies, which are used in the generated RAC structures. Using a larger number of such RACs, we can obtain higher accuracies. This is the main reason that we obtained for the computed Rd redox potentials better agreement with measured values. For molecular systems, whose structures are less well-defined, the agreement with measured redox potentials is generally of lower quality.

Although the composition of charged groups and formation of salt bridges differ considerably between the mesophilic Rd from *Cp* and the thermophilic Rd from *Pf* and are considered to be responsible for the difference in thermostability, the contributions of charged residues to the redox potential difference between the two Rd's is small (25 mV).

For 16 Rd's (WT and 15 mutants) from *Cp* where measured redox potentials are available, RACs were generated solely on the basis of the WT crystal structure not varying the backbone, but with appropriate side chain optimizations as mentioned before. In this case, the computed Rd redox potentials showed an RMSD with measured data of only 14 mV. In the RACs used for these computations, the amide H-bond lengths with the FeS complex did not vary. Hence, the excellent agreement between measured and computed redox potentials does not support the general assumption that mainly differences in amide H-bond strengths are responsible for variations in Rd redox potentials. This result is not in conflict with the experimentally observed correlation of H-bond strengths with Rd redox potentials. But, this correlation does not prove that the H-bond variation is the origin of the redox potential variations in different Rd's.

On the other hand, removing one of these amide H bonds downshifts the Rd redox potential by about 60 mV. Thus, the fact that Rd redox potentials are relatively high, lying in the interval –90 mV to +40 mV, is determined by the presence of the six amide H bonds. This was the basis to assume that the variation of the measured redox potentials in different Rd's (from different mutants and different species) should be mainly due to different H-bond strengths, where the latter are related to the H-bond lengths. These amide H-bond lengths vary in Rd crystal structures. However, when the crystal structures of WT Rd's from *Cp* and *Pf* are compared, it turns out that these H bonds are

just slightly shorter for the latter Rd and the corresponding influence on the Rd redox potential is as expected small.

We estimated that variation of an amide H-bond length by 0.1 Å can shift the Rd redox potential by 13 mV. But, this value cannot straightforwardly be used to understand the variations of redox potentials in different Rd's, since this estimate was obtained by considering that only the amide backbone atoms are varied. As we have demonstrated in the present study, such backbone changes typically go along with variations of amino acid side chains, resulting in a total shift of the redox potential, which may even go in the opposite direction.

In summary, the variations in redox potentials of different Rd's are not the result of a single factor but are due to combinations of backbone and side chain variations, which lead to subtle changes of the electrostatic boundary between the protein and solvent, rendering the FeS complex more or less solvent exposed. Furthermore, differences in composition of charged residues for Rd's from different species play some role, while variations in the salt-bridge pattern alone have no significant effect. Finally, the variation of amide H-bond strengths, which was traditionally considered to play a major role, has only a very moderate influence on the Rd redox potentials.

■ ASSOCIATED CONTENT

S **Supporting Information.** Details of the methodology, a list of Rd PDB information, measured redox potentials and references, Cp mutant conformations obtained after structural alignment, and atomic partial charges as well as figures of conformations in crystal structures and RACs. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*Fax: +493083856921. E-mail: knapp@chemie.fu-berlin.de.

■ ACKNOWLEDGMENT

We thank Dr. Galstyan for useful discussions. This work was supported by Volkswagen Foundation, JST PRESTO program (H.I.); Grant-in-Aid (H.I. 21770163) for Science Research from the Ministry of Education, Science, Sport and Culture of Japan; Special Coordination Fund for Promoting Science and Technology of MEXT (H.I.); and Takeda Science Foundation (H.I.). A.P.G.H. and G.K. would like to thank CONACyT/DAAD and German National Academic Foundation for financial support.

■ ABBREVIATIONS

Cp, *Clostridium pasteurianum*; Dd, *Desulfovibrio desulfuricans*; Dg, *Desulfovibrio gigas*; Dv, *Desulfovibrio vulgaris*; FeS complex, iron sulfur complex; MD, molecular dynamics; PB equation, Poisson–Boltzmann equation; PAC, pH adapted conformation; Pa, *Pyrococcus abyssi*; Ps, *Pseudomonas auruginosa*; Pf, *Pyrococcus furiosus*; RAC, redox adapted conformation; RMSD, root-mean-square deviation; Rd, rubredoxin; WT, wild type

■ REFERENCES

- (1) Hall, D. O.; Cammack, R.; Rao, K. K. *Origins Life Evol. Biosphere* **1974**, *5*, 363–386.
- (2) Holm, R. H.; Kennepohl, P.; Solomon, E. I. *Chem. Rev.* **1996**, *96*, 2239–2314.

- (3) Beinert, H.; Holm, R. H.; Münck, E. *Science* **1997**, *277*, 653–659.
- (4) Insande, J. *Plant Physiol. Biochem.* **1999**, *37*, 87–97.
- (5) Beinert, H. *J. Biol. Inorg. Chem.* **2000**, *5*, 2–15.
- (6) van Belien, J. B.; Neuenschwander, M.; Smits, T. H. M.; Roth, C.; Balada, S. B.; Witholt, B. *J. Bacteriol.* **2001**, *184*, 1722–1732.
- (7) Hagelueken, G.; Wiehlmann, L.; Adams, T. M.; Kolmar, H.; Heinz, D. W.; Tümmler, B.; Schubert, W. D. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 12276–12281.
- (8) Noodleman, L.; Norman, J. G., Jr.; Osborne, J. H.; Aizman, A.; Case, D. A. *J. Am. Chem. Soc.* **1985**, *107*, 3418–3426.
- (9) Mouesca, J. M.; Chen, J. L.; Noodleman, L.; Bashford, D.; Case, D. A. *J. Am. Chem. Soc.* **1994**, *116*, 11898–11914.
- (10) Noodleman, L.; Lovell, T.; Liu, T.; Himo, F.; Torres, R. A. *Curr. Opin. Chem. Biol.* **2002**, *6*, 259–273.
- (11) Torres, R. A.; Lovell, T.; Noodleman, L.; Case, D. A. *J. Am. Chem. Soc.* **2003**, *125*, 1923–1936.
- (12) Noodleman, L.; Han, W. G. *J. Biol. Inorg. Chem.* **2006**, *11*, 674–694.
- (13) Schreiner, E.; Nair, N. N.; Pollet, R.; Staemmler, V.; Marx, D. P. *Natl. Acad. Sci. U. S. A.* **2007**, *104*, 20725–20730.
- (14) Solomon, E. I.; Xie, X.; Dey, A. *Chem. Soc. Rev.* **2008**, *37*, 623–638.
- (15) Lin, I. J.; Gebel, E. B.; Machonkin, T. E.; Westler, W. M.; Markley, J. L. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 14581–14586.
- (16) Stephens, P. J.; Jollie, D. R.; Warshel, A. *Chem. Rev.* **1996**, *96*, 2491–2531.
- (17) Hunsicker-Wang, L. M.; Heine, A.; Chen, Y.; Luna, E. P.; Todaro, T.; Fee, J. A. *Biochemistry* **2003**, *42*, 7303–7317.
- (18) Venkateswara, R.; Holm, R. H. *Chem. Rev.* **2004**, *104*, 527–559.
- (19) Gunner, M. R.; Mao, J.; Song, Y.; Kim, J. *Biochim. Biophys. Acta* **2006**, *1757*, 942–968.
- (20) Dey, A.; Jenney, F. E., Jr.; Adams, M. W. W.; Babini, E.; Takahashi, Y.; Fukuyama, K.; Hodgson, K. O.; Hedman, B.; Solomon, E. I. *Science* **2007**, *318*, 1464–1468.
- (21) Gamiz-Hernandez, A. P.; Galstyan, A. S.; Knapp, E. W. *J. Chem. Theory Comput.* **2009**, *5*, 2898–2908.
- (22) Adman, E. T.; Watenpaugh, K. D.; Jensen, L. H. *Proc. Natl. Acad. Sci. U. S. A.* **1975**, *72*, 4854–4858.
- (23) Lin, I. J.; Gebel, E. B.; Machonkin, T. E.; Westler, W. M.; Markley, J. L. *J. Am. Chem. Soc.* **2003**, *125*, 1464–1465.
- (24) Solomon, E. I.; Gorelsky, S. I.; Dey, A. *J. Comput. Chem.* **2006**, *27*, 1415–1428.
- (25) Bönsch, H.; Ladenstein, R. *J. Biol. Inorg. Chem.* **2007**, *12*, 1163–1171.
- (26) Zheng, Q.; Smith, E. T.; Kurtz, D. M., Jr.; Scott, R. A. *Inorg. Chim. Acta* **1996**, *242*, 245–251.
- (27) Zheng, H.; Kellog, S. J.; Erickson, A. E.; Dubauskie, N. A.; Smith, E. T. *J. Biol. Inorg. Chem.* **2003**, *8*, 12–18.
- (28) Eidsness, M. K.; Burden, A. E.; Richie, K. A.; Kurtz, D. M.; Scott, R. A.; Smith, E. T.; Ichiye, T.; Beard, B.; Min, T.; Kang, C. *Biochemistry* **1999**, *38*, 14803–14809.
- (29) Xiao, Z.; Maher, M. J.; Cross, M.; Bond, C. S.; Guss, J. M.; Wedd, A. G. *J. Biol. Inorg. Chem.* **2000**, *5*, 75–84.
- (30) Kennepohl, P.; Solomon, E. I. *Inorg. Chem.* **2003**, *42*, 689–695.
- (31) Min, T.; Ergenekan, C. E.; Eidsness, M. K.; Ichiye, T.; Kang, C. *Protein Sci.* **2001**, *10*, 613–621.
- (32) Day, M. W.; Hsu, B. T.; Joshuaor, L.; Park, J. B.; Zhou, Z. H.; Adams, M. W.; Rees, D. C. *Protein Sci.* **1992**, *1*, 1494–1507.
- (33) Yelle, R. B.; Park, N.-S.; Ichiye, T. *Proteins* **1995**, *22*, 154–167.
- (34) Park, I. Y.; Youn, B.; Harley, J. L.; Eidsness, M. K.; Smith, E.; Ichiye, T.; Kang, C. *J. Biol. Inorg. Chem.* **2004**, *9*, 423–428.
- (35) Millar, M.; Lee, J. F.; O'Sullivan, T.; Koch, S. A.; Fikar, R. *Inorg. Chim. Acta* **1996**, *243*, 333–343.
- (36) Ullmann, G. M.; Noodleman, L.; Case, D. A. *J. Biol. Inorg. Chem.* **2002**, *7*, 623–639.
- (37) Siegbahn, P. E. M.; Himo, F. *J. Biol. Inorg. Chem.* **2009**, *14*, 643–651.

- (38) Dauter, Z.; Wilson, K. S.; Sieker, L. C.; Moulis, J.-M.; Meyer, J. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 8836–8840.
- (39) Gillès de Pélichy, L. D.; Smith, E. T. *Biochemistry* **1999**, *38*, 7874–7880.
- (40) Xiao, L.; Honig, B. J. *Mol. Biol.* **1999**, *289*, 1435–1444.
- (41) Pace, C. N. *Nat. Struct. Biol.* **2000**, *7*, 345–346.
- (42) Perl, D.; Mueller, U.; Heinemann, U.; Schmid, F. X. *Nat. Struct. Biol.* **2000**, *7*, 380–383.
- (43) Blake, P. R.; Park, J. B.; Bryant, F. O.; Aono, S.; Magnuson, J. K.; Eccleston, E.; Howard, J. B.; Summers, M. F.; Adams, M. W. W. *Biochemistry* **1991**, *30*, 10885–10895.
- (44) Bradley, E. A.; Stewart, D. E.; Adams, M. W.; Wampler, J. E. *Protein Sci.* **1993**, *2* (4), 650–665.
- (45) Swartz, P. D.; Ichiye, T. *Biochemistry* **1996**, *35*, 13772–13779.
- (46) Zeng, Q.; Smith, E. T.; Kurtz, D. M., Jr.; Scott, R. A. *Inorg. Chim. Acta* **1996**, *242*, 245–251.
- (47) Lazardis, T.; Lee, I.; Karplus, M. *Protein Sci.* **1997**, *6*, 2589–2605.
- (48) Eidsness, M. K.; Richie, K. A.; Burden, A. E.; Kurtz, D. M. J.; Scott, R. A. *Biochemistry* **1997**, *36* (34), 10406–10413.
- (49) Bau, R.; Rees, D. C.; Kurtz, D. M.; Scott, R. A.; Huang, H.; Adams, M. W.; Eidsness, M. K. *J. Biol. Inorg. Chem.* **1998**, *3*, 484–493.
- (50) Bertini, I.; Kurtz, D. M., Jr.; Eidsness, M. K.; Liu, G.; Luchinat, C.; Rosato, A.; Scott, R. A. *J. Biol. Inorg. Chem.* **1998**, *3*, 401–410.
- (51) Grottesi, A.; Ceruso, M.-A.; Colosimo, A.; Di Nola, A. *Proteins: Struct., Funct., Genet.* **2002**, *46*, 287–294.
- (52) Bonomi, F.; Burden, A. E.; Eidsness, M. K.; Fessas, D.; Iametti, S.; Kurtz, D. M. J.; Mazzini, S.; Scott, R. A.; Zeng, Q. *J. Biol. Inorg. Chem.* **2002**, *7*, 427–436.
- (53) Bougault, C. M.; Eidsness, M. K.; Prestegard, J. H. *Biochemistry* **2003**, *42*, 4357–4372.
- (54) Dolan, E. A.; Yelle, R. B.; Beck, B. W.; Fischer, J. T.; Ichiye, T. *Biophys. J.* **2004**, *86*, 2030–2036.
- (55) Park, I. Y.; Eidsness, M. K.; Lin, I. J.; Gebel, E. B.; Youn, B.; Harley, J. L.; Machonkin, E. E.; Frederick, R. O.; Markley, J. L.; Smith, E. T.; Ichiye, T.; Kang, C. *Proteins* **2004**, *57*, 618–625.
- (56) Bonomi, F.; Eidsness, M. K.; Iametti, S.; Kurtz, D. M. J.; Mazzini, S.; Morleo, A. *J. Biol. Inorg. Chem.* **2004**, *9* (3), 297–306.
- (57) Sulpizi, M.; Raugei, S.; Vondele, J. V.; Carloni, P.; Sprik, M. *J. Phys. Chem. B* **2007**, *111*, 3969–3976.
- (58) *CRC Handbook of Chemistry and Physics*, 84th ed; Press, C., Ed.; CRC Press: Boca Raton, FL, 1992.
- (59) Rabenstein, B.; Ullmann, G. M.; Knapp, E. W. *Eur. Biophys. J.* **1998**, *27*, 626–637.
- (60) Ullmann, G. M.; Knapp, E. W. *Eur. Biophys. J.* **1999**, *28*, 533–551.
- (61) Rabenstein, B.; Knapp, E. W. *Biophys. J.* **2001**, *80*, 1141–1150.
- (62) Popovic, D. M.; Zaric, S. D.; Rabenstein, B.; Knapp, E. W. *J. Am. Chem. Soc.* **2001**, *123*, 6040–6053.
- (63) Voigt, P.; Knapp, E. W. *J. Biol. Chem.* **2003**, *278*, 51993–52001.
- (64) Ishikita, H.; Morra, G.; Knapp, E. W. *Biochemistry* **2003**, *42*, 3882–3892.
- (65) Ishikita, H.; Knapp, E. W. *J. Am. Chem. Soc.* **2004**, *126*, 8059–8064.
- (66) Ishikita, H.; Knapp, E. W. *FEBS Lett.* **2005**, *579*, 3190–3194.
- (67) Ishikita, H.; Knapp, E. W. *J. Am. Chem. Soc.* **2005**, *127*, 14714–14720.
- (68) Ishikita, H.; Saenger, W.; Biesiadka, J.; Loll, B.; Knapp, E. W. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 9855–9860.
- (69) Gamiz-Hernandez, A. P.; Kieseritzky, G.; Demir-Kavuk, O.; Galstyan, A. S.; Knapp, E. W. *ChemPhysChem* **2010**, *11*, 1196–1206.
- (70) Schmidt am Busch, M.; Knapp, E. W. *J. Am. Chem. Soc.* **2005**, *127* (45), 15730–15737.
- (71) Galstyan, A. S.; Knapp, E. W. *J. Comput. Chem.* **2009**, *30*, 203–211.
- (72) Kieseritzky, G.; Knapp, E. W. *Proteins: Struct., Funct., Bioinf.* **2008**, *71*, 1335–1348.
- (73) Kieseritzky, G.; Knapp, E. W. *J. Comput. Chem.* **2008**, *29* (15), 2575–2581.
- (74) Holst, M.; Saied, F. J. *Comput. Chem.* **1993**, *14*, 105–113.
- (75) Baker, N.; Sept, D.; Joseph, S.; Holst, M.; McCammon, J. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 10037–10041.
- (76) Baptista, A. M.; Teixeira, V. H.; Soares, C. M. *J. Chem. Phys.* **2002**, *117*, 4184–4200.
- (77) Lee, M. S.; F., R. S., Jr.; C., L. B., III. *Proteins: Struct., Funct., Bioinf.* **2004**, *56*, 738–752.
- (78) Brooks, B. R.; Brucoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.
- (79) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem B* **1998**, *102*, 3586–3616.
- (80) Bönisch, H.; Schmidt, C. L.; Bianco, P.; Ladenstein, R. *Acta Crystallogr., Sect. D* **2005**, *61*, 990–1004.
- (81) Watenpaugh, K. D.; Sieker, L. C.; Jensen, L. H. *J. Mol. Biol.* **1980**, *138*, 615–633.
- (82) Sieker, L. C.; Stenkamp, R. E.; Jensen, L. H.; Prickril, B.; LeGall, J. *FEBS Lett.* **1986**, *208*, 73–76.
- (83) Stenkamp, R. E.; Sieker, L. C.; Jensen, L. H. *Proteins* **1990**, *8*, 352–264.
- (84) Frey, M.; Sieker, L.; Payan, F.; Haser, R.; Bruschi, M.; Pepe, G.; Legall, J. *J. Mol. Biol.* **1987**, *197*, 525–541.
- (85) Chen, C. J.; Lin, Y. H.; Huang, Y. C.; Liu, M. Y. *Biochem. Biophys. Res. Commun.* **2006**, *349*, 79–90.
- (86) Adman, E. T.; Sieker, L. C.; Jensen, L. H. *J. Mol. Biol.* **1991**, *217*, 337–352.
- (87) Dauter, Z.; Sieker, L. C.; Wilson, K. S. *Acta Crystallogr., Sect. B* **1992**, *48*, 42–59.
- (88) Misaki, S.; Morimoto, Y.; Ogata, M.; Yagi, T.; Higuchi, Y.; Yasuoka, N. *Acta Crystallogr., Sect. D* **1999**, *55*, 408–413.
- (89) Kurihara, K.; Tanaka, I.; Chatake, T.; Adams, M. W. W.; Moiseeva, N.; Bau, R.; Niimura, N. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 11215–11220.
- (90) Kabsch, W. *Acta Crystallogr., Sect. A* **1976**, *32*, 922–923.
- (91) Xiao, Z.; Lavery, M. J.; Bond, A. M.; Wedd, A. G. *Electrochem. Commun.* **1999**, *1*, 309–314.
- (92) Adams, M. W. *Adv. Inorg. Chem.* **1992**, *38*, 341–396.
- (93) Swartz, P. D.; Beck, B. W.; Ichiye, T. *Biophys. J.* **1996**, *71*, 2958–2969.

A Coarse-Grained Model for Molecular Dynamics Simulations of Native Cellulose

Jakob Wohlert* and Lars A. Berglund

Wallenberg Wood Science Center, Royal Institute of Technology, SE-10044 Stockholm, Sweden

Supporting Information

ABSTRACT: We have constructed a coarse-grained model of crystalline cellulose to be used in molecular dynamics simulations. Using cellobiose from the recently published MARTINI coarse-grained force field for carbohydrates [Lopez, C. A. et al. *J. Chem. Theory Comput.* **2009**, *5*, 3195–3210] as a starting point, we have reparameterized the nonbonded interactions to reproduce the partitioning free energies between water and cyclohexane for a series of cellooligomers, cellobiose through cellopentaose. By extrapolating the model to longer cellooligomers, and by assigning special cellulose–cellulose nonbonded interactions, we obtain a model which gives a stable, ordered structure in water that closely resembles the crystal structure of cellulose I β . Furthermore, the resulting model is compatible with an existing coarse-grained force field for proteins. This is demonstrated by a simulation of the motion of the carbohydrate-binding domain of the fungal cellulase Cel7A from *Trichoderma reesei* on a crystalline cellulose surface. The diffusion coefficient at room temperature is calculated at $D_l = 3.1 \times 10^{-11} \text{ cm}^2 \text{ s}^{-1}$, which is in good agreement with experimental numbers.

INTRODUCTION

Cellulose, the 1–4 linked polymer of β -D-glucopyranose (see Figure 1), is synthesized in nature in slender fibrils, whose lateral dimensions typically range from 3 nm up to 5 nm, depending on the synthesizing species. These fibrils are usually assembled into larger structures, called fibril aggregates. Within the fibrils, the glucan chains are packed in a specific crystal form, named cellulose I, or native cellulose, which is known from X-ray crystallography and neutron diffraction.^{1,2} Furthermore, there are two distinct allomorphs of native cellulose,³ namely, cellulose I α and cellulose I β , with the former being dominant in bacterial and algal cellulose and the latter being dominant in cellulose produced in higher plants.

Cellulose, being the most abundant, naturally occurring polymer on earth, has in recent years been the target of great, renewed interest. On one hand, cellulose-based biocomposites offer the potential to replace petroleum-based plastics and composites.^{4,5} The use of cellulose nanofibrils extracted from plant cell walls is of particular interest since the potential for a significantly extended property range is vast for nanostructured cellulose materials.⁶ On the other hand, cellulose offers potential as an important feedstock for large-scale production of bioethanol by enzymatic and chemical degradation. Great efforts are directed toward a better understanding of the enzymatic pathways, and the engineering of both enzymes and cellulosic material, with the goal of achieving better turnover rates.^{7,8}

Molecular dynamics computer simulation has been shown to be an excellent tool to contribute to a molecular-level understanding of crystalline cellulose, its structure and its dynamics, as well as its interactions with solvents and other biomolecules.^{9–14} At present, there exist several force fields that are being used in simulations of cellulose, such as GROMOS 45a4,¹⁵ CHARMM36,¹⁶ PCFF,¹⁷ and GLYCAM06,¹⁸ to mention a few. These are all atomistic force fields, meaning that all atoms, except for nonpolar

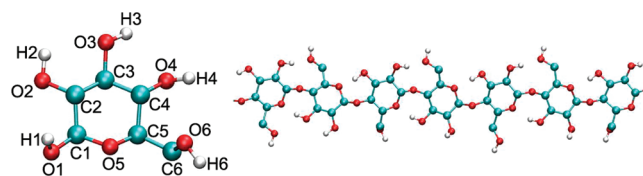


Figure 1. β -D-Glucose with the atom names used in this paper (left) and a cellulose chain (right).

hydrogen atoms in the case of GROMOS, are represented explicitly by one interaction site, making the simulations limited to fairly small systems and short simulation times.

Coarse-grained (CG) modeling is an alternative to all-atom (AA) modeling, wherein some of the atomic detail is averaged out, so that much longer time and length scales come within reach. During recent years, coarse-grained approaches have been highly successful for a wide range of biomolecular systems such as proteins, lipid bilayers, and carbohydrates.¹⁹

There are a few different approaches to construct CG models. Typically, a number of atoms are grouped together to form a CG superatom, or bead. The beads interact through effective pairwise potentials, which are constructed to reproduce relevant properties from the atomistic representation. This can be achieved in different ways, e.g., using Monte Carlo schemes to optimize potential parameters,²⁰ or through force-matching procedures.²¹ These procedures generally lead to models that are highly specialized, meaning that they are usually very good at describing the system for which they are originally developed, but not easily transferable to other systems.

The MARTINI force field²² was parametrized using a different approach. Here, a number of predefined CG beads interact via

Received: August 27, 2010

Published: January 26, 2011

fixed set of pairwise potentials. They are used to construct small molecular building blocks, which are calibrated against thermodynamic data, in particular partitioning coefficients between polar and apolar phases. This leads to a model that can be extended to include new molecular species, in a fairly straightforward way, while at the same time retaining its internal consistency and compatibility.

There exist a couple of CG models for carbohydrates in the literature. The M3B model by Molinero and Goddard²⁰ was developed to simulate malto-oligosaccharides in solution, using Boltzmann inversion to construct potentials for the bonded interaction and a Monte Carlo scheme to optimize the non-bonded interactions which are described with Morse potentials. The model reproduces thermodynamic data for glucose very well and gives a helical structure for amylose which is stable for tens of nanoseconds. Liu et al.²¹ have constructed a CG model for glucose in solution using a force-fitting procedure. Their model shows good structural properties, such as radial distribution functions, as well as some thermodynamic properties, like isothermal compressibility. Recently, the MARTINI force field was extended to include carbohydrates as well.²³ The extended model provides parameters for the monosaccharides glucose and fructose, a number of disaccharides, including cellobiose, and the oligosaccharides maltoheptose (a short strand of amylose) and laminaraheptabiose (Curdlan).

Bu et al.²⁴ used the M3B model as a starting point to construct a model which, to the best of our knowledge, is the only CG model for crystalline cellulose in the literature. They refitted the bonded interaction terms and scaled the nonbonded terms in order to get a crystal structure that closely matches that of cellulose I β . The model was used to study the interactions between a crystalline cellulose surface and the carbohydrate-binding domain (CBD) of a fungal cellulase, which was represented in atomic detail, together with an implicit water model. They show that it is indeed possible to get a stable cellulose crystal using only three interaction sites per glucose residue. However, since different force fields were mixed in their study, the results they obtain for cellulose–protein interactions are somewhat questionable. For instance, they report a diffusion coefficient for the CBD which is around 3 orders of magnitude too large compared to experimental values.²⁵

To really benefit from a CG approach, it is desirable to have a model that is internally consistent. Furthermore, treating the system as whole at a CG level will also benefit from that simulation time, and length scales can be significantly extended. To that end, we have developed a model for crystalline cellulose based on the MARTINI parameter set.²² By using the same basic methodology in the parametrization process, i.e., ensuring that the partitioning of the cellulose chains between polar and apolar phases are correct, we obtain a model that is inherently compatible with the existing models for proteins, lipids, carbohydrates, and solvents. This is demonstrated by simulating the diffusion of the CBD from cellobiohydrolase I (Cel7A) from *Trichoderma reesei* on a crystalline cellulose surface.

METHODS

Model. We took as our starting point the recently published MARTINI force field for carbohydrates.²³ The MARTINI parameter set²² includes 18 predefined standard particle types, with different levels of polarity, which interact through pairwise Lennard-Jones potentials. Each particle, or bead, is used to represent

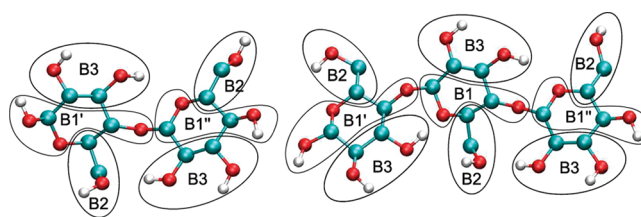


Figure 2. Mapping of atoms to CG beads for cellobiose and cellotriose, following López et al.²³

three or four heavy atoms. Cellobiose is represented by six beads (see Figure 2), three beads per glucose unit, where each bead is positioned at the center of mass of its constituting AA atoms. To decide which bead types to use, we chose to parametrize our model using the partitioning between pure water and cyclohexane for a series of oligomers, cellobiose through cellopentaose, with the published model for cellobiose²³ as the starting point. Since there are, to the best of our knowledge, no experimental data for the partitioning of longer cellooligomers, we chose to use simulation data from an atomistic model as our reference data set. Bonded interactions were adapted directly from ref 23, with the exception of the dihedral angles, as described in the Results section.

Simulation Details. The systems considered in this present work are the following:

1. Cellopentaose, cellotetraose, cellotriose, and cellobiose in pure water and in cyclohexane, using both atomistic and coarse-grained modeling. These simulations were used for the parametrization of the nonbonded interactions, i.e., choice of bead types for the CG model. The size of the simulation box was roughly $3.6 \times 3.6 \times 3.6$ nm in all cases.
2. A cellulose crystal in water, in a CG representation, which was used to optimize cellulose–cellulose interactions. The crystal consisted of 36 chains, each chain 40 glucose units in length, and was solvated in a box with dimensions $6.5 \times 6.5 \times 25.0$ nm.
3. A CG model of a crystalline cellulose surface with the carbohydrate-binding domain of Cel7A attached to it, in solution.

All simulations were performed using GROMACS 4.0,²⁶ in a NPT ensemble. The temperature was maintained at 300 K, unless stated otherwise, using stochastic velocity rescaling,²⁷ and the pressure was kept at 1 atm using a Parrinello–Rahman barostat.²⁸ For the case of the crystal simulations, pressure scaling was applied in the lateral (x/y) directions only. Molecular graphics were produced using VMD.²⁹

Coarse-Grained Simulations. For the coarse-grained simulations, we used the standard settings for the MARTINI parameter set. Nonbonded interactions were made to go smoothly to zero between 0.8 and 1.2 nm using a shift function. Integration was performed using a leapfrog algorithm with a 20 fs time step. Coordinates for the CBD were taken from Kraulis et al.³⁰ (PDB id: 1CBH) and were converted to a CG description using the MARTINI conversion tools. To preserve the structure of the CBD in solution, an elastic network approach was used. All pairs of backbone beads that were separated by a distance between 0.5 and 0.9 nm were restrained around that value using a harmonic potential with force constant $500 \text{ kJ mol}^{-1} \text{ nm}^{-2}$. It has been shown that the elastic network potentials can be optimized to obtain a model that is better at preserving the internal motions of the protein,³¹ but for our purposes this simple approach is

sufficient. Finally, parameters for the solvents, water and cyclohexane, were taken from the MARTINI force field.²²

Atomistic Simulations. The atomistic simulations used the GROMOS united atoms force fields G45a3³² for cyclohexane and G45a4¹⁵ for carbohydrates, together with SPC water.³³ Despite GROMOS being a united-atoms force field, which means that aliphatic hydrogens are not modeled explicitly, the atomistic simulations will be referred to as AA simulations throughout this manuscript. The nonbonded interactions were handled with a twin-range cutoff approach. Inside 0.8 nm, the nonbonded interactions were updated every step, and between 0.8 nm and an outer cutoff of 1.4 nm they were updated once every 10 steps, which is the same frequency as the neighbor list updates. To account for long-range electrostatic interactions, a reaction field correction was applied with a relative permittivity of 66 in water. The cyclohexane model has no explicit charges, and for that reason no long-range correction is needed. The basic time step used was 2 fs. All bonds were kept at their equilibrium values using P-LINCS.³⁴

Partitioning Free Energies. The partitioning free energy of a solute between water and cyclohexane, $\Delta\Delta G_{wc}$, can be calculated from the difference between the solvation free energies of the solute in the respective solvents, ΔG_w and ΔG_c . The partitioning free energy is related to the partitioning coefficient P_{wc} through $\Delta\Delta G_{wc} = -k_B T \log P_{wc}$, where k_B is Boltzmann's constant and T is the temperature. To this end, ΔG_w and ΔG_c were calculated using thermodynamic integration (TI), in which the solute-solvent interactions are controlled by a coupling parameter, λ . The coupling parameter λ was varied between 0 (all interactions intact) and 1 (fully decoupled state) in 25 discrete steps. Note however that, for AA simulations in water, the decoupling was actually performed in twice that number of steps, first the electrostatic and then the Lennard-Jones interactions. To avoid singularities when the Van der Waal's radii approach zero, the Lennard-Jones interactions were gradually switched to a soft-core potential with $\alpha = 0.5$, $p = 1$, and $\sigma = 0.3$. Intermolecular interactions of the solute were not coupled, meaning that the fully decoupled state corresponds to the gas phase of the solute, and the resulting energies are proper solvation free energies. For each value of λ , a 5 ns simulation was run. The output from these simulations is the derivative of G with respect to λ , $dG/d\lambda$, as a function of λ , which enables ΔG to be calculated using numerical integration over all λ 's. For a more detailed description of TI, and the meaning of the soft-core parameters, we refer the reader to the GROMACS manual.

RESULTS AND DISCUSSION

The corresponding bead types are listed in Table 1. Note the difference between terminal and nonterminal residues, and also between the two different terminal residues, where B1' is used at the reducing end and B1'' is used at the nonreducing end. Longer cellooligomers are constructed by repeated insertion of non-terminal residues.

Partitioning. As already mentioned, our starting point for the parametrization was the model for cellobiose from López et al.²³ However, this model turned out to be overall too polar, giving a partition free energy between water and cyclohexane, $\Delta\Delta G_{wc}$, of 70 kJ mol⁻¹, which is off by 50% compared to the reference value of 44 kJ mol⁻¹, calculated at an AA level. To overcome this, we settled on the model shown in Figure 2 and Table 1, which has the same basic features of the cellobiose from López et al.,²³ in

Table 1. Beads and Optimized Bead Types for the Model in Figure 2^a

bead	atoms	mass (amu)	optimized bead type	original bead type ^b
B1	C1, C4, O4, O5	58.0368	Na (neutral, h-bond acceptor)	
B1'	C1, O1, H1, C4, O4, O5	75.0442	P1 (polar)	P2
B1''	C1, C4, O4, H4, O5	59.0448	P1 (polar)	P2
B2	C5, C6, O6, H6	44.0534	P1 (polar)	P1
B3	C2, O2, H2, C3, O3, H3	60.0528	P2 (more polar)	P4

^a See the original MARTINI ref 22 for their definitions and interactions. ^b From ref 23.

Table 2. Partitioning Free Energies of Cellooligomers between Water and Cyclohexane in kJ mol^{-1a}

	ΔG_w^{CG}	ΔG_c^{CG}	$\Delta\Delta G_{wc}^{CG}$	ΔG_w^{AA}	ΔG_c^{AA}	$\Delta\Delta G_{wc}^{AA}$
cellobiose	-104	-61	43	-103	-59	44
cellotriose	-146	-88	58	-147	-86	61
cellotetraose	-189	-115	74	-183	-113	70
cellopentaose	-229	-139	90	-223	-140	83

^a Statistical errors are approximately 3 kJ mol⁻¹ for the CG simulations, and 5 kJ mol⁻¹ for AA.

that each glucose unit consists of three polar beads, one of which is more polar than the others, representing the two hydroxyl groups on C2 and C3. This model gives a $\Delta\Delta G_{wc}$ of 43 kJ mol⁻¹, in good agreement with the AA result (see Table 2). When one more glucose unit is attached to a cellobiose to make a cellotriose, the middle residue has one hydroxyl group less than the terminal residues. Consequently, the bead representing the glycosidic linkage in our model for cellotriose has a less polar bead type for residue number two than corresponding beads in residues one and three. The remaining beads in that residue are the same as in the terminal residues. The model that gives the best correspondence to the reference data is shown in Table 1. Longer cellooligomers are modeled according to the same scheme, using slightly different representations of terminal and nonterminal residues. Table 2 shows that our model matches the reference data for the whole series of cellooligomers investigated in this work, cellobiose through cellopentaose, very closely.

It is somewhat surprising that our model for cellobiose is different from the one in López et al.²³ After all, they should represent the same molecule, using the same force field. While the original model was parametrized against the partitioning between water and water-saturated octanol, we used water and cyclohexane instead, which may explain a large part of this difference. One reason for our choice is that we could not reproduce the results from López et al.²³ for octanol in the AA representation; in fact, we did not even get the correct sign for $\Delta\Delta G$. The reason for this might be that the water/octanol mixture, with a solute, is a fairly complicated system, which is hard to bring to convergence. Cyclohexane, being a much simpler system, does not suffer from this to the same extent. We also note a small difference in ΔG for cellobiose in water, in the AA representation, between ref 23 and the present result. This is perhaps surprising, since the same force fields are used in both

studies. It is however possible that this difference can be explained by small differences in temperature and simulation protocols.

Cellulose Crystal Structure. The next step was to investigate our model's ability to represent crystalline cellulose. To this end, a model crystal was constructed by replicating the experimental unit cell for cellulose $I\beta$, which is known from X-ray crystallography and neutron fiber diffraction,¹ in its three principal directions, creating a structure consisting of 36 chains (6×6), 40 glucose units long. The crystal was constructed with the two staggered faces, corresponding to the (110) and (1-10) crystallographic planes, facing out. This structure was then converted to a CG representation using the mapping in Figure 2 and placed in an orthorhombic simulation box filled with CG water. After energy minimization, the system was run for approximately 100 ns. The system quickly drifted away from the initial crystal structure and ended up in a structure in which the cellulose chains were stacked directly on top of each other, rather than the staggered arrangement of the native crystal. This structure was then stable for the rest of the simulation. This behavior is not very surprising considering that all interactions between the chains in this model are fairly attractive. In a stacked arrangement, beads of the same type can interact with each other in an optimal way. In an all-atom representation of cellulose, there exist forces that make a slightly offset arrangement more optimal, such as the ability to form hydrogen bonds between sheets, but these depend on details that the CG model evidently is too coarse to capture. To overcome the problem with artificial stacking, we introduced a repulsive component to the cellulose–cellulose interaction. In practice, we made all interactions between beads of type B1 strongly repulsive, while at the same time interactions between the side-chain beads B2 and B3 were made more attractive. Interactions between cellulose bead types and all other bead types were left untouched. The new cellulose–cellulose interaction matrix is shown in Table 4. That the optimized potential for the oligomers does not give a reasonable crystal structure is perhaps a bit disappointing, but not at all unexpected. As a matter of fact, this is analogous to that when modeling proteins using MARTINI; restraints are often needed to maintain the structural integrity of the protein, which, in a way, means making the structure a model parameter. Since a realistic representation of the structure, of both cellulose and proteins, is crucial for modeling their interactions with their respective surroundings, this is a necessary compromise. However, the fact that the internal energetics of the cellulose crystal are not correct is important to keep in mind when choosing suitable problems for our model.

Even though the interactions between the cellobiomers and the solvents (water and cyclohexane) are the same after the reparameterization as before, it is of course possible that the new intramolecular interactions have an impact on the solvation free energies and the partitioning. For that reason, a new series of simulations was performed for the CG model, identical to the ones described in the previous section with the exception that they were using the new interaction matrix. The resulting partitioning free energies are the same as those in Table 2, within the given error range.

Another feature of the resulting structure was that it developed a quite pronounced twist. This twist stemmed from the equilibrium dihedral angle between successive glucose units not being 180° in the original parameter set. Twists have been observed previously, using both CG²⁴ and AA⁹ modeling, and also in

Table 3. Energy Minima of the Dihedral Angles between the Side-Chain Beads of Two Neighboring Cellulose Monomers, along the B1–B1 Bond, for the Present Model and the Original Reference²³

	present model	original
B2–B2	180°	-150°
B3–B3	180°	-150°
B2–B3	0°	30°

Table 4. Interaction Matrix for Cellulose–Cellulose Interactions^a

	B1	B2	B3
B1	super repulsive	almost attractive	almost attractive
B2		attractive	attractive
B3			attractive

^a See ref 22 for definitions of the interactions.

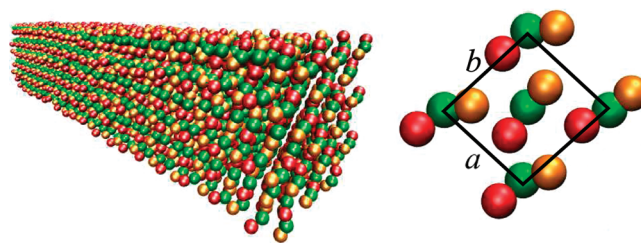


Figure 3. Snapshot of a stable coarse-grained cellulose crystal structure (left) and a representation of the cellulose $I\beta$ crystal unit cell (right). The crystallographic axes a and b are also indicated in the figure. The third crystal axis (c) is directed parallel to the fibril long axis, which is pointing out from the figure to the right. CG beads of type B1 are green, beads of type B2 are orange, and beads of type B3 are red. Water beads are omitted for clarity.

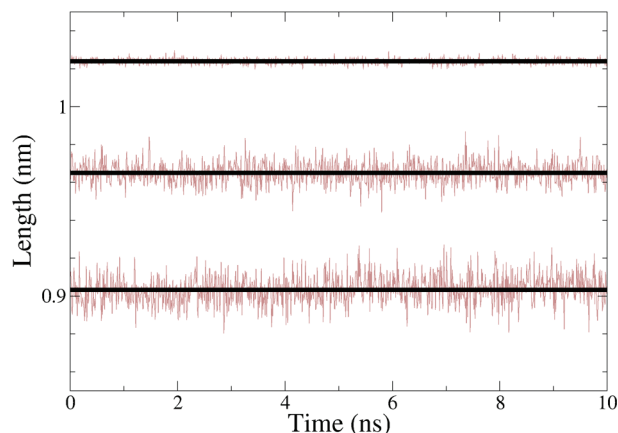
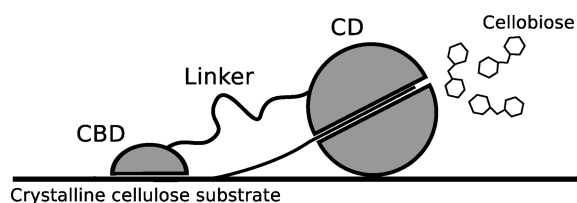
experiments using both X-ray¹ and microscopic³⁵ methods. There seem to be no consensus about the nature of these twists, or how large they are. Since the predictive power of our model concerning cellulose structure is very limited, and moreover, since the twist is very easy to relate to a single model parameter, we chose to redefine the dihedral angles for the side chains around the glycosidic bonds (see Table 3), to force the cellulose chains into a flat conformation, as in ref 1.

As can be seen in Figure 3, the resulting structure is an ordered structure which retains the nearly quadratic cross-section of the initial conformation. It is possible to compare this structure to that of cellulose $I\beta$. A close examination of the structure shows that the unit cell parameters a and b (see Figure 3), which are predominantly a result of the nonbonded parameters, are on the high side (see Table 5). The third crystal axis, c , which is directed parallel to the chain direction and thus is determined by the B1–B1 bond distance, is in very good agreement with the experimental value. Even the unit cell angles are fairly close to the X-ray results. The time evolutions of the unit cell parameters a , b , and c are shown in Figure 4. They are stable over several nanoseconds, and their fluctuations are fairly small.

An attempt was made to fine-tune the interactions to see if the unit cell parameters could be improved. However, any other combination tried than the one in Table 4 only disrupted the desired staggered conformation. Consequently, we decided to

Table 5. Unit Cell Parameters of the CG Model, Compared to X-Ray and Neutron Diffraction Data for Cellulose I β ¹

	cellulose I β	CG model
<i>a</i> (nm)	0.78	0.90
<i>b</i> (nm)	0.82	0.97
<i>c</i> (nm)	1.04	1.02
α (deg)	90	90
β (deg)	90	90
γ (deg)	96.5	92

**Figure 4.** Time evolution of the unit cell parameters, from bottom to top, *a*, *b*, and *c* and their respective mean values.**Figure 5.** An illustration of the fungal cellulase Cel7A acting on a crystalline cellulose surface. The cellulase is composed of three distinct domains: a carbohydrate-binding domain (CBD) and a catalytic domain (CD) interconnected by a flexible linker peptide. During procession of the cellulase along the substrate, a cellulose chain is fed into the catalytic tunnel where the glycosidic bonds are hydrolyzed, with cellobiose as the end product.

keep the interactions from Table 4 and accept deviations from the experimental unit cell parameters.

Diffusion of a CBD on a Cellulose Surface. The carbohydrate-binding domain (CBD) of the fungal cellulase Cel7A from *Trichoderma reesei* is responsible for anchoring the cellulase to the cellulose substrate. It is attached to the catalytic domain (CD) by a flexible linker peptide (see Figure 5), and it has been shown to be essential for effective binding of the cellulase to cellulose.³⁶ There has also been speculation about whether the CBD also can assist the CD more directly by facilitating the lifting of the cellulose chains from the surface,³⁷ but recent experimental data confirm that its main purpose is to increase the local concentration of CDs on the cellulose surface.³⁸ However, computer simulations suggest that it might be involved directly in the recognition of loose chain ends.³⁹

A common motif for all fungal CBDs is a cellulose-binding surface featuring three solvent-exposed aromatic residues, tyrosines in the case of Cel7A, which has been shown to be critical for the binding of the CBD to cellulose.³⁶ The spacing of these residues coincides with the spacing of the glucose units in cellulose, and it has therefore been proposed that the CBD binds to the cellulose surface essentially by favorable van der Waals interactions between the aromatic residues and the pyranose rings exposed on the flat surfaces of crystalline cellulose.^{36,40} This has been partly confirmed by computer simulations that indicate the existence of localized binding grooves on the cellulose surface,⁴¹ which makes it probable that the motion of the CBD on the surface takes place in discrete steps of about 0.5 nm in length.

Jervis et al.²⁵ give a value for the diffusion constant of a bacterial family 2CBD of 2×10^{-11} to 1.2×10^{-10} cm² s⁻¹ at room temperature, which translates to an approximate rate constant of 4.8×10^4 s⁻¹ to 2.9×10^5 s⁻¹ assuming diffusion on a hexagonal lattice with a lattice constant of 0.5 nm. This rate is very fast compared to the intrinsic rate of the CD, which has been measured to be approximately 3.5 nm s⁻¹ for the case of Cel7A,³⁸ suggesting that surface diffusion of the CBD does not limit cellulase activity.^{25,42} Still, this rate is extremely slow compared to the time scales normally accessible in computer simulations. Even though the diffusion constant is for a different family, it is the only one that is currently available in the literature. We will assume that it is a reasonable estimate also for the diffusion constant of the CBD of Cel7A.

As mentioned above, it has been suggested that the flat face of the cellulose crystal (corresponding to the (110) crystallographic plane in the case of cellulose I α and (100) in the case of cellulose I β) is the most likely to which the CBD will attach.³⁶ Starting from the X-ray structure of crystalline cellulose I β ¹, we built a surface consisting of two layers of cellulose chains. Each layer was composed of 10 chains, and each chain consisted of 16 glucose units. The surface was made periodic in all directions, with each chain covalently bonded to its own periodic image in the chain direction, effectively mimicking a surface infinite in size. Using only two layers of cellulose saves computational time, but at the same time it makes the structure unstable. To prevent the crystal structure from being disrupted, one of the two layers, the “bottom” one, was made subject to harmonic restraints in all directions with a force constant of 10^3 kJ mol⁻¹ nm². This was enough to stabilize the layer on top of it as well, even in the high temperature simulations described below. Next, the CBD was placed on top of the surface, with the three tyrosines mentioned above facing down, toward the cellulose surface. Finally, the system was solvated using a water layer approximately 5-nm-thick. Since the cutoff for interactions is only 1.2 nm, this is sufficient to prevent any interactions between surfaces and their periodic images and interactions between the CBD and surfaces to which it is not directly attached. This system was then simulated for 1 μ s, during which the CBD after some small initial displacement sat completely still in one spot. Figure 6 shows the system after equilibration. As can be seen, the alignment of the three tyrosines is not perfectly parallel to the direction of the chains, which has also been noted in atomistic models.⁴³

A step rate of 10^5 s⁻¹ means that we would expect the CBD to take one step every 10 μ s. Indeed, the equilibrated structure is stable for several microseconds. A simple way of speeding up the dynamics of the system is to raise the temperature. For this particular setup, 380 K seemed to be a lower limit to be able to see sufficient motion of the CBD during the time scales accessible. For

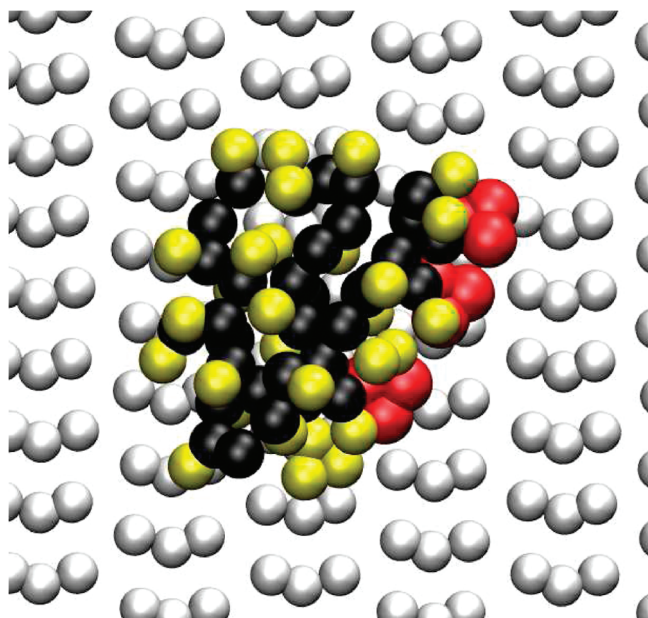


Figure 6. Top view of the CG model of a CBD on a cellulose surface. Backbone beads are black, side chains yellow, and the important tyrosine residues are colored red. The cellulose surface is light gray. Water is omitted for clarity. The chain direction is running vertically in the figure.

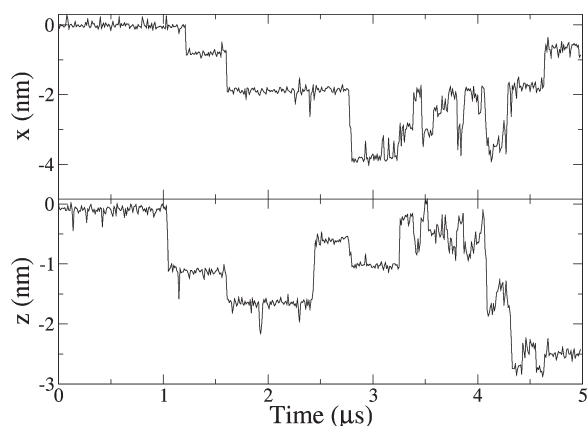


Figure 7. Position of the center of mass of the CBD as a function of simulation time at 410 K. The motion does, to a large degree, occur in discrete steps.

this reason, we ran several 5 μs simulations at elevated temperatures, between 380 and 450 K in 10 K increments. Of course, for a real system, such high temperatures would be devastating for both the crystalline cellulose and the CBD, but this is fortunately not necessarily the case in simulations. As explained above, the position restraints imposed on the cellulose were sufficient to preserve its structure even at the highest temperatures used. At the same time, the elastic network type restraints do the same thing for the CBD. For this reason, we assume that the main effect of raising the temperature is to speed up the dynamics.

Figure 7 shows the positions of the center of mass of the CBD during the simulation at 410 K. From the figure, it is clear that the motion to a high degree takes place in steps which are more or less, but certainly not always, multiples of 0.5 nm. Why the step lengths differ between steps becomes evident from looking at the trajectory in a molecular graphics representation. The motion of

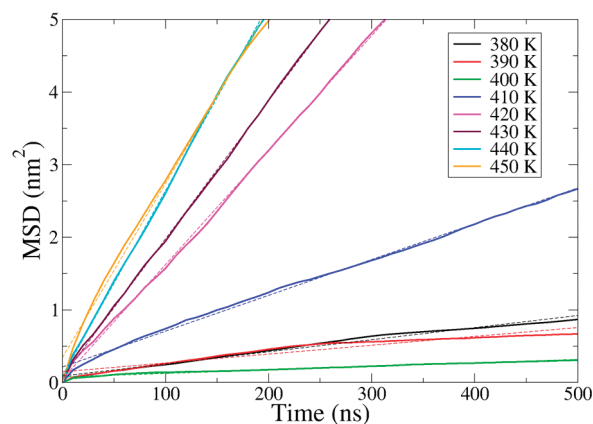


Figure 8. Calculated lateral mean square displacements of the CBD (solid) along with linear fits which were used to calculate D_l .

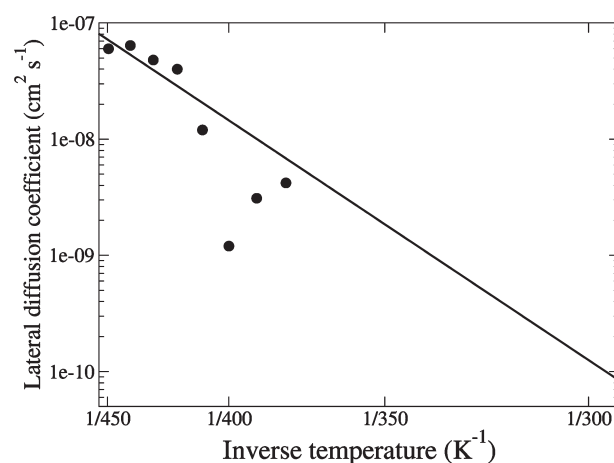


Figure 9. Linear regression, using least-squares fitting, of $\log(D_l)$ plotted against $1/T$, where D_l is calculated from the lateral mean square displacement at different temperatures using eq 1.

the CBD does not only involve pure translations but also rotations of the CBD around the surface normal, and the combination of translations and rotations between stable states leads to a variable step length.

The lateral diffusion constant D_l was calculated from the mean square displacements (MSD) of the CBD, using the Einstein relation in two dimensions:

$$D_l = \lim_{x \rightarrow \infty} \frac{1}{4} \frac{d}{dt} \langle [\mathbf{r}(t + t_0) - \mathbf{r}(t_0)]^2 \rangle_{t_0} \quad (1)$$

Figure 8 shows the calculated lateral MSD along with a least-squares fit of eq 1 to the data. Only the first 500 ns of the MSD were used for fitting due to the statistics quickly becoming too poor at longer times, especially at the lower temperatures.

Assuming stepwise diffusion where the diffusion rate follows the simple Arrhenius' law, we can write

$$D_l \propto \exp\{-\Delta G/k_B T\} \quad (2)$$

where ΔG is the free energy barrier between stable states. This means that the diffusion constant at room temperature can be obtained from a linear regression of $\log[D_l(T)]$ plotted against $1/T$. From the data presented in Figure 9, we calculate the diffusion coefficient at room temperature to be $D_l = 1.2 \times 10^{-10} \text{ cm}^2 \text{ s}^{-1}$.

This is in very good agreement with the experimental value²⁵ stated above. Furthermore, the slope of the fit in Figure 9 gives the average barrier height between states, $\Delta G = 48 \text{ kJ mol}^{-1}$, which of course poses a significant obstacle to the motion.

Until now, we have not mentioned the fact that time scales in CG simulations generally differ from time scales in atomistic simulations. This is due to the fact that coarse-graining tends to smoothen the energy landscape, which leads to faster dynamics. By calculating the self-diffusion of water, time in the MARTINI model has been found to run about 4 times faster than in atomistic models.²² This means that the diffusion coefficients calculated above should be scaled by a factor 1/4, which leads to $D_1 = 3.1 \times 10^{-11} \text{ cm}^2 \text{ s}^{-1}$ at room temperature, which is still within the experimental range.

CONCLUSIONS

We have constructed a coarse-grained molecular model of cellulose, consisting of three interaction sites per glucose residue, by optimizing the nonbonded parameters against the partitioning of a set of cellobioses between water and cyclohexane, using a recently published model for cellobiose²³ as our starting point. The reference values for the partitioning free energies, which were obtained from atomistic simulations, were reproduced for a series of cellobioses, cellobiose through cellopentaose. This indicates that the resulting model has the right balance between polar and apolar parts. The internal (cellulose–cellulose) nonbonded potentials were modified in order to obtain a crystal structure similar to the one for cellulose I β obtained by X-ray crystallography.¹ This means that the model probably does not correctly represent the internal energetics of the cellulose crystal. However, this was deemed a necessary compromise since a good representation of the crystal structure is important when modeling interactions between crystalline cellulose and its surroundings, e.g., solvents and other biomolecules such as proteins.

By building our model within the framework provided by the MARTINI force field,²² we get a model that is inherently compatible with the existing model for proteins.⁴⁴ This was demonstrated by simulating the diffusive motion of the carbohydrate binding domain from *Trichoderma reesei* on a crystalline cellulose surface in explicit water. The calculated lateral diffusion coefficient is within the experimental range,²⁵ which is a clear indicator that the interactions between the protein model and the present model for cellulose are realistic.

Many applications directly involving cellulose are dependent on the interplay between crystalline cellulose and a large array of other molecules, e.g., solvents and proteins, as well as both synthetic and naturally occurring polymers. The present model can, in that context, provide useful information on a molecular level, which may be used to, for instance, optimize the controlled degradation of cellulose or help in tailoring new materials based on renewable resources.

ASSOCIATED CONTENT

S Supporting Information. A complete GROMACS topology and coordinates for the CG cellulose crystal described in this work are provided. This information is available free of charge via the Internet at <http://pubs.acs.org/>.

AUTHOR INFORMATION

Corresponding Author

*E-mail: jacke@kth.se.

ACKNOWLEDGMENT

Computational resources were kindly provided by PDC Center for High Performance Computing under project title *Computer Modeling of Crystalline Cellulose*. The authors thank Malin Bergenstråhle for helpful discussions.

REFERENCES

- (1) Nishiyama, Y.; Langan, P.; Chanzy, H. *J. Am. Chem. Soc.* **2002**, *124*, 9074–9082.
- (2) Nishiyama, Y.; Sugiyama, J.; Chanzy, H.; Langan, P. *J. Am. Chem. Soc.* **2003**, *125*, 14300–14306.
- (3) Atalla, R. H.; VanderHart, D. L. *Science* **1984**, *223*, 283–285.
- (4) Samir, M. A. S. A.; Alloin, F.; Dufresne, A. *Biomacromolecules* **2005**, *6*, 612–626.
- (5) Berglund, L. A.; Peijs, T. *MRS Bull.* **2010**, *35*, 201–207.
- (6) Eichhorn, S. J.; Dufresne, A.; Aranguren, M. *J. Mater. Sci.* **2010**, *45*, 1–33.
- (7) Ragauskas, A. J.; Williams, C. K.; Davison, B. H.; Britovsek, G.; Cairney, J.; Eckert, C. A.; Frederick, W. J., Jr.; Hallett, J. P.; Leak, D. J.; Liotta, C. L.; Mielenz, J. R.; Murphy, R.; Templer, R.; Tschaplinski, T. *Science* **2006**, *311*, 484–489.
- (8) Himmel, M. E.; Ding, S.-Y.; Johnson, D. K.; Adney, W. S.; Nimlos, M. R.; Brady, J. W.; Foust, T. D. *Science* **2007**, *315*, 804–807.
- (9) Matthews, J. F.; Scopeç, C. E.; Mason, P. E.; Zuccato, P.; Torget, R. W.; Sugiyama, J.; Himmel, M. E.; Brady, J. W. *Carbohydr. Res.* **2006**, *341*, 138–152.
- (10) Bergenstråhle, M.; Berglund, L. A.; Mazeau, K. *J. Phys. Chem. B* **2007**, *111*, 9138–9145.
- (11) Bergenstråhle, M.; Wohler, J.; Larsson, P. T.; Mazeau, K.; Berglund, L. A. *J. Phys. Chem. B* **2008**, *112*, 2590–2595.
- (12) Mazeau, K.; Rivet, A. *Biomacromolecules* **2008**, *9*, 1352–1354.
- (13) Zhong, L.; Matthews, J. F.; Crowley, M. F.; Rignall, T.; Talón, C.; Cleary, J. M.; Walker, R. C.; Chukkappalli, G.; McCabe, C.; Nimlos, M. R.; C., L. B., III; Himmel, M. E.; Brady, J. W. *Cellulose* **2008**, *15*, 261–273.
- (14) Nishiyama, Y.; Johnson, G. P.; French, A. D.; Forsyth, V. T.; Langan, P. *Biomacromolecules* **2008**, *9*, 3133–3140.
- (15) Lins, R. D.; Hünenberger, P. H. *J. Comput. Chem.* **2005**, *26*, 1400–1412.
- (16) Guvench, O.; Greene, S. N.; Kamath, G.; Brady, J. W.; Venable, R. M.; Pastor, R. W.; MacKerell, A. D., Jr. *J. Comput. Chem.* **2008**, *29*, 2543–2564.
- (17) Sun, H.; Mumby, S. J.; Maple, J. R.; Hagler, A. T. *J. Am. Chem. Soc.* **1994**, *116*, 2978–2987.
- (18) Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; González-Outeiriño, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. *J. Comput. Chem.* **2007**, *29*, 622–655.
- (19) Voth, G. A. *Coarse-graining of condensed phase and biomolecular systems*; CRC Press: Boca Raton, FL, 2009.
- (20) Molinero, V.; Goddard, W. A., III. *J. Phys. Chem. B* **2004**, *108*, 1414–1427.
- (21) Liu, P.; Izvekov, S.; Voth, G. A. *J. Phys. Chem. B* **2007**, *111*, 11566–11575.
- (22) Marrink, S. J.; Risselada, H. J.; Yemov, S.; Tieleman, D. P.; de Vries, A. H. *J. Phys. Chem. B* **2007**, *111*, 7812–7824.
- (23) López, C. A.; Rzepiel, A. J.; de Vries, A. H.; Dijkhuizen, L.; Hünenberger, P. H.; Marrink, S. J. *J. Chem. Theory Comput.* **2009**, *9*, 3195–3210.
- (24) Bu, L.; Beckham, G. T.; Crowley, M. F.; Chang, C. H.; Matthews, J. F.; Bomble, Y. J.; Adney, W. S.; Himmel, M. E.; Nimlos, M. R. *J. Phys. Chem. B* **2009**, *113*, 10994–11002.
- (25) Jervis, E. J.; Haynes, C. A.; Kilburn, D. G. *J. Biol. Chem.* **1997**, *272*, 24016–24023.
- (26) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.
- (27) G. Bussi, D. D.; Parrinello, M. *J. Chem. Phys.* **2007**, *126*, 014101.

- (28) Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981**, *52*, 7182–7190.
- (29) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33–38.
- (30) Kraulis, P. J.; Clore, G. M.; Nilges, M.; Jones, T. A.; Pettersson, G.; Knowles, J.; Gronenborn, A. M. *Biochemistry* **1989**, *28*, 7241–7257.
- (31) Periole, X.; Cavalli, M.; Marrink, S.; Ceruso, M. A. *J. Chem. Theory Comput.* **2009**, *5*, 2531–2543.
- (32) Schuler, L. D.; Daura, X.; van Gunsteren, W. F. *J. Comput. Chem.* **2001**, *22*, 1205–1218.
- (33) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces*; Pullman, B. E., Ed.; Riedel: Dordrecht, The Netherlands, 1981; p 331.
- (34) Hess, B. *J. Chem. Theory Comput.* **2008**, *4*, 116–122.
- (35) Hanley, S. J.; Revol, J.-F.; Godbout, L.; Gray, D. G. *Cellulose* **1997**, *4*, 209–220.
- (36) Lehtiö, J.; Sugiyama, J.; Gustafsson, M.; Fransson, L.; Linder, M.; Teeri, T. T. *Proc. Natl. Acad. Sci.* **2003**, *100*, 484–489.
- (37) Ruovinen, J.; Bergfors, T.; Teeri, T.; Knowles, J. K. C.; Jones, T. A. *Science* **1990**, *249*, 380–386.
- (38) Igarashi, K.; Koivula, A.; Wada, M.; Kimura, S.; Penttilä, M.; Samejima, M. *J. Biol. Chem.* **2009**, *284*, 36186–36190.
- (39) Nimlos, M. R.; Matthews, J. F.; Crowley, M. F.; Walker, R. C.; Chukkapalli, G.; Brady, J. W.; Adney, W. S.; Cleary, J. M.; Zhong, L.; Himmel, M. E. *Protein Eng. Des. Sel.* **2007**, *20*, 179–187.
- (40) Reinikainen, T.; Ruohonen, L.; Nevanen, T.; Laaksonen, L.; Kraulis, P.; Jones, T. A.; Knowles, J. K. C.; Teeri, T. T. *Proteins* **1992**, *14*, 475–482.
- (41) Beckham, G. T.; Matthews, J. F.; Bomble, Y. J.; Bu, L.; Adney, W. S.; Himmel, M. E.; Nimlos, M. R.; Crowley, M. F. *J. Phys. Chem. B* **2010**, *114*, 1447–1453.
- (42) Ting, C. L.; Makarov, D. E.; Wang, Z. *J. Phys. Chem. B* **2009**, *113*, 4970–4977.
- (43) Zhong, L. H.; Matthews, J. F.; Hansen, P. I.; Crowley, M. F.; Cleary, J. M.; Walker, R. C.; Nimlos, M. R.; Brooks, C. L.; Adney, W. S.; Himmel, M. E.; Brady, J. W. *Carbohydr. Res.* **2010**, *344*, 1984–1992.
- (44) Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S. J. *J. Chem. Theory Comput.* **2008**, *4*, 819–834.

On the Convergence of QM/MM Energies

LiHong Hu,^{†,‡} Pär Söderhjelm,[§] and Ulf Ryde^{*,†}[†]Department of Theoretical Chemistry, Lund University, Chemical Centre, P.O. Box 124, SE-221 00 Lund, Sweden[‡]Faculty of Chemistry, North-east Normal University, Changchun, 130024, P. R. China[§]Department of Chemistry and Applied Biosciences - Computational Science, ETH Zürich, Via Giuseppe Buffi 13, CH-6900 Lugano, Switzerland Supporting Information

ABSTRACT: We have studied the convergence of QM/MM calculations with respect to the size of the QM system. We study a proton transfer between a first-sphere cysteine ligand and a second-sphere histidine group in [Ni,Fe] hydrogenase and use a 446-atom model of the protein, treated purely with QM methods as a reference. We have tested 12 different ways to redistribute charges close to the junctions (to avoid overpolarization of the QM system), but once the junctions are moved away from the active site, there is little need to redistribute the charges. We have tested 13 different variants of QM/MM approaches, including two schemes to correct errors caused by the truncation of the QM system. However, we see little gain from such correction schemes; on the contrary, they are sensitive to the charge-redistribution scheme and may cause large errors if charges are close to the junctions. In fact, the best results were obtained with a mechanical embedding approach that does not employ any correction scheme and ignores polarization. It gives a mean unsigned error for 40 QM systems of different sizes of 7 kJ/mol with a maximum error of 28 kJ/mol. The errors can be significantly decreased if bonds between the QM and MM system (junctions) are moved one residue away from all active-site residues. Then, most QM/MM variants give mean unsigned errors of 5–9 kJ/mol, maximum errors of 16–35 kJ/mol, and only five to seven residues give an error of over 5 kJ/mol. In general, QM/MM calculations converge faster with system size than pure QM calculations.

INTRODUCTION

During the past two decades, quantum mechanical (QM) calculations have been established as an attractive and competitive complement to experiments to study biochemical reactions.^{1–4} However, there is still no consensus on how such calculations are best performed. In principle, two schools have arisen. In the first, which we will call QM-only in the following, a small part (20–200 atoms) is cut out of the macromolecule of interest, typically the active site and a few nearby residues.^{5,6} This system is studied with QM methods, whereas the rest of the macromolecule is either ignored or, more commonly, modeled as a featureless continuum, characterized by a dielectric constant of ~ 4 . It is typically necessary to fix a number of atoms at the periphery of the QM system to model steric restrictions of the macromolecule. Entropic effects can be modeled by a harmonic model, based on calculated vibrational frequencies.

The alternative approach is to include the whole macromolecule in the calculation by the use of combined QM and molecular mechanics (MM) methods, the QM/MM approach.^{7,8} In this approach, a central system of a similar size to that of the QM-only approach is treated by QM methods, whereas the rest of the macromolecule as well as some explicit solvent molecules are modeled by MM methods. The advantage with this approach is of course that the whole macromolecule is explicitly modeled and that free energies can be calculated by free-energy perturbations or related approaches.^{9–11} On the other hand, the size of the system makes the method more expensive, and it becomes hard to control the conformation of the MM system.

Unfortunately, there are few direct comparisons of the two approaches. Ochsenfeld and Sumowski have studied proton transfer within a 32-residue polypeptide and a 1637-atom model of triose isomerase and shown that the QM/MM approach converges appreciably faster than the QM-only approach with respect to the size of the QM system, although the convergence is quite slow for both approaches.¹² For example, with 299 QM atoms, the errors for the QM-only and QM/MM approaches were 43 and 12 kJ/mol, respectively, and even with 1092 QM atoms, the error in the QM-only approach was still 6 kJ/mol. We have observed a similar slow convergence of the QM-only approach for QM systems up to 696 atoms.¹³ Even worse, different ways to select what residues to include in the QM system gave widely different results. In fact, after the addition of 40 residues, there was still a difference of 60 kJ/mol if the residues were added according to their distance to the QM system or if they were added according to their energy components in a QM/MM free-energy perturbation approach.¹⁴ This gives a quite pessimistic view of the use of QM-only methods for the study of biochemical systems.

On the other hand, QM/MM calculations also show a quite slow convergence with respect to the size of the QM system. For example, it has been shown that the errors in QM/MM forces are sizable for QM regions with a radius of up to 9 Å and that QM/MM free energies change by 12 kJ/mol when increasing the size of the QM system from 3 to 5 Å.¹⁵

Received: September 17, 2010

Published: February 07, 2011

In this paper, we supplement our QM-only study by an investigation of the corresponding convergence of the QM/MM approach. Thus, we use a 446-atom model of [Ni,Fe] hydrogenase from our previous investigation¹³ as a reference and investigate how the QM/MM energy converges toward the QM energy of the full system as the QM system in the QM/MM treatment is enlarged. We concentrate on the energy of a reaction that takes place inside the QM system, in accordance with the common use of QM/MM methods. Therefore, errors arise primarily from two sources. The first is the treatment of the surroundings by an MM potential, rather than by a more accurate QM method. The second problem with the QM/MM approach is that it is typically necessary to truncate the QM system by cutting some (normally C–C) bonds in the macromolecule. This is a well-known problem in QM/MM that has been much discussed.^{7,8,16–19} There are several ways to solve it, but the most simple and common one is to truncate the QM system with hydrogen atoms. If the surrounding macromolecule is included in the QM calculations as a point-charge model, there is a risk of overpolarization, owing to the fact that some point charges are located close to the hydrogen junction atoms. This has been much discussed, and several solutions have been tested and compared.^{17,19–23} In general, the results are varying, and large errors (60–100 kJ/mol) are occasionally found, especially for charged systems.

We also address another related question that has received much less attention, viz, whether the errors caused by the junctions can be corrected. We test two such correction schemes, one available in the ONIOM approach²⁴ and another used in the QTCP approach,¹¹ and also some variants of them. Moreover, we test several variants of mechanical embedding (i.e., when electrostatic interactions between the QM and MM systems are treated at the MM level).

We study the energy of a simple proton-transfer reaction between a bridging Cys ligand and a second-sphere His residue in [Ni,Fe] hydrogenase. Although the proton moves only 0.97 Å and the structure of the surroundings hardly changes, it has been shown that this reaction is very sensitive to the surroundings, changing the reaction energy from ~0 kJ/mol in a vacuum to ~80 kJ/mol in the protein.¹⁴ Therefore, it provides a sensitive test case for the modeling of effects of the surroundings in biochemical reactions.

METHODS

QM/MM Calculations with Electrostatic Embedding. In the QM/MM approach,^{7,8} a small but important part of the total system (called system 1 or the QM system) is treated by quantum mechanics (QM), whereas the rest (called system 2 or the MM system) is treated by molecular mechanics (MM). The QM/MM calculations have been performed with the program COMQUM,^{25–27} which is a modular combination of the QM software Turbomole 5.10²⁸ and the MM software Amber 9.²⁹

Special attention is needed when there are covalent bonds between the QM and MM systems. Many approaches have been suggested to treat such junctions, e.g., by truncating the QM system by certain link atoms or by using localized orbitals at the junctions.^{7,8} We have employed the simplest and most widely used approach, the hydrogen link-atom approach, in which the QM system simply is truncated with hydrogen atoms. To simplify the discussion, we will use the following nomenclature, illustrated in Figure 1:^{8,22} The hydrogen link-atom is called HL,

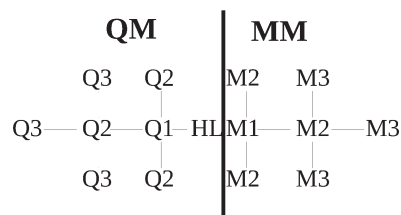


Figure 1. Illustration of the partitioning between the QM and MM systems and the naming of the atoms.

whereas the MM atom it replaces is called M1. The QM atom directly connected to HL is Q1. QM atoms directly bound to Q1 are called Q2, and those directly bound to Q2 are called Q3, and so on. Likewise, MM atoms directly connected to M1 are called M2; those directly bound to the M2 atoms are called M3 and so on. Sometimes, we divide all atoms into three systems, viz, those in the MM system, excluding M1 (M), the HL atoms (J), and the rest of the QM atoms (Q).

In principle, the HL atoms will introduce additional degrees of freedom. To avoid this, the HL atoms are placed along the Q1–M1 bond, with a Q1–HL bond length ($r_{\text{Q1-HL}}^{\text{QM0}}$) that is proportional to the Q1–M1 bond length ($r_{\text{Q1-M1}}^{\text{MM0}}$) according to

$$r_{\text{Q1-HL}} = r_{\text{Q1-M1}} \frac{r_{\text{Q1-HL}}^{\text{QM0}}}{r_{\text{Q1-M1}}^{\text{MM0}}} \quad (1)$$

where $r_{\text{Q1-M1}}^{\text{MM0}}$ is the equilibrium Q1–M1 bond length in the MM force field used and $r_{\text{Q1-HL}}^{\text{QM0}}$ is the optimum length of the Q1–HL bond optimized with the QM method and basis sets used.²⁶ Thereby, the HL and M1 atoms can be considered to be the same atom, albeit with different positions (and sometimes also different charges) in the calculations with system 1 alone or with both systems 1 and 2. The QM/MM forces are calculated with the help of the chain rule.^{27,30}

Most calculations in this paper have been calculated with electrostatic embedding (EE),^{7,8} meaning that a point-charge model of the MM system is included in the QM calculations, so that the QM system is polarized by the MM system. Thus, the total QM/MM energy is calculated from

$$E_{\text{QM/MM}}^{\text{EE}} = E_{1+\text{ptch2}}^{\text{QM, HL}} + E_{12,\text{no1el}}^{\text{MM, M1}} - E_{1,\text{no1el}}^{\text{MM, HL}} \quad (2)$$

where the three terms on the right-hand side are the QM energy of the QM system with HL atoms, including the point-charge model of the MM system; the MM energy of all atoms (with M1, rather than HL atoms), but with the charges of the QM system zeroed; and the MM energy of the QM system (again with HL atoms), with zeroed charges. The latter term is needed to cancel the MM term of the QM system from the second term, to avoid double counting. Likewise, the charges of the QM atoms are zeroed in order to avoid double counting of the electrostatic interactions between the QM and MM systems. The self-energy of the point-charge model is excluded from the QM term (this energy is instead included in the $E_{12,\text{no1el}}^{\text{MM,M1}}$ term).

Charge-Redistribution Schemes. When using electrostatic embedding, it is not fully clear what atoms should be included in the point-charge model of the MM system. In particular, it is unclear whether the charges of the M1 atoms should be included or not. If HL and M1 are considered to be the same atom, it is evident that M1 should not be included. This becomes even clearer if you do not use a hydrogen link atom but rather a

reparametrized atom that behaves like a carbon atom^{19,31–33} — then the HL and M1 atoms will overlap.

However, most discussions and developments have started from the assumption that the M1 charge should also be included.^{22,23} The reason for this is probably that many QM/MM methods are based on force fields with charge groups (e.g., OPLS and CHARMM^{34,35}), i.e., where small chemical groups, like a CH₂ unit, have a neutral charge. Then, it is natural to also include the M1 charge, to keep the charge neutrality.²²

On the other hand, the distance between HL and M1 is quite short, ~ 0.5 Å, which can lead to a significant overpolarization of the QM system. In fact, even the distance between HL and M2 is quite short, 1.3–1.7 Å. Therefore, it is common to exclude or redistribute some point charges in the QM calculations.^{7,8,22,23} In this paper, we have tested six different approaches, which were chosen among those that have given the best results in previous tests:^{19,22,23}

- Z0: All charges are included, including those on the M1 atoms.
- Z1: The charges of the M1 atoms are excluded.
- Z2: The charges of the M1 and M2 atoms are excluded.
- Z3: The charges of the M1, M2, and M3 atoms are excluded.
- RCD (the redistributed charge and dipole method²²): The charges of the M1 atoms are redistributed over all of the M2 atoms, keeping the bond dipole constant by adding a compensating charge at the bond midpoint between each M1 and M2 atom. If we let q_0 be the original charge on M1 divided by the number of M2 atoms, then the charge on the bond midpoint will be $2q_0$, whereas q_0 will be subtracted from the charge on each M2 atom.²²
- CS (the charge shift scheme^{36,37}): Similar to RCD, in that the M1 charge is redistributed, keeping the bond dipole by compensating charges. However, q_0 is added to the charges on M2 atoms, and two point charges are placed on each side of M2 along the M1–M2 bond. We place the two point charges at 6% of the M1–M2 bond length from the M2 atom with charges $\pm 50/6 q_0$. This is slightly different from the implementation of this approach in ChemShell, where the distance is only approximately 6% to allow the charges to be truncated after four decimals (to ensure numerical stability in geometry optimizations; P. Sherwood, personal communication; in this paper, we do not change the geometries).

As mentioned above, the Z0, RCD, and CS schemes were designed for MM force fields that use charge groups, so that the part of each junction residue that is not in the QM system (including the M1 atom) has a net integer charge. This is not the case with the Amber force field³⁸ we are using, for which only the full residue has an integer charge. In order to test the importance of charge groups, we used two different approaches: In the first, we simply used the original charges, although they do not sum up to an integer. This approach is denoted by the six abbreviations noted above (Z0, Z1, Z2, Z3, RCD, and CS).

Alternatively, we changed the charge on the M1 atom so that the sum of the charges of the atoms in the junction residue that are not in the QM system becomes an integer (zero in all cases tested here, except for the carboxy-terminal His-591).¹⁹ Then, each of the charge-distribution approaches was performed as described above. These balanced approaches will be called BZ0, BRCD, and BCS in the following.

The Z1, Z2, and Z3 approaches are not affected by this redistribution of the M1 charge. For these, we instead tested to redistribute the sum of the deleted charges evenly on the other

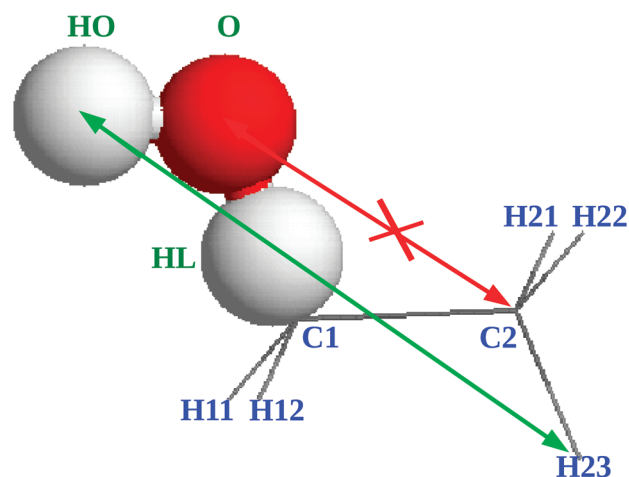


Figure 2. Illustration of the excluded-atom problem, using ethanol as an example, with the HOH moiety as the QM system and the rest as the MM system. The HO and H23 atoms are separated by four bonds, and therefore their interaction is included fully in both the MM and QM calculations. However, the O and C2 atoms are separated by only two bonds, and their interaction is therefore excluded in the MM calculations, whereas it is included in the QM calculations.

MM atoms in that residue. These approaches will be called DZ1, DZ2, and DZ3. The DZ2 approach is the default in COMQUM.²⁶ Finally, for Z2, we also tested to set the sum of the remaining MM charges in the residue to zero by adding the same increment to all charges. This approach will be called NZ2 below. All of these approaches are implemented in our local software *changeqmm*, which generates the point-charge file from the MM topology file.

Energy Correction Schemes. Covalent junctions between the QM and MM systems inevitably introduce an unphysical perturbation of the system. The question then naturally arises whether this error can be corrected. Strangely enough, this important question has been much less discussed than possible charge-redistribution schemes. The errors caused by introducing a link atom are of three types:

1. The HL atom is placed in the wrong position compared to the real M1 atom.
2. The HL atom is a hydrogen atom, rather than the correct M1 (typically carbon) atom, and thus it will have incorrect MM parameters, in particular an incorrect charge.
3. Electrostatics are treated inconsistently around the junction: In most macromolecular MM force fields, non-bonded interactions between atoms that are directly bonded or that are separated by two covalent bonds (1–2 and 1–3 interactions) are excluded, whereas interactions between atoms connected by three covalent bonds (1–4 interactions) are scaled down. However, the QM software does not know about such exclusion rules and includes all electrostatic interactions between the QM atoms and the point charges. This is illustrated in Figure 2 for a simple model system, consisting of ethanol, in which we use a HOH QM model of the alcohol group, whereas the rest is treated at the MM level. In the QM calculations, the three QM atoms (HO, O, and HL) interact fully with the MM charges of all seven MM atoms (C1, C2, and H11–H23), giving 21 electrostatic terms. Among these, only the three HO–H21/H22/H23 interactions are also

present in the full MM treatment of ethanol. The six HO–C2/H11/H12 and O–H21/H22/H23 interactions should be scaled down (e.g., by a factor of 1.2 for electrostatics and 2.0 for van der Waals interactions in the Amber force field), whereas all of the others should be excluded.

In principle, all of these errors can be corrected by the general approach of eq 2. For example, if the MM parameters of the HL and M1 atoms are chosen wisely, there will be a cancellation of all bonded and van der Waals terms involving HL (between $E_{1+\text{ptch2}}^{\text{QM,HL}}$ and $E_{1,\text{no1el}}^{\text{MM,HL}}$) so that the total QM/MM energy corresponds to calculations with M1 atoms only. However, for the electrostatic interactions, no such cancellation is obtained with eq 2. This is the case for standard COMQUM and also in the calculations by Lin and Truhlar.²² This seems to be the case also in most other QM/MM software, although the details of the implementations are seldom discussed. This approach will simply be called electrostatic embedding (EE) in the following.

However, two approaches have been suggested to also correct the electrostatic interactions: the QM to QM/MM correction in the QTCP (QM/MM thermodynamic cycle perturbation) approach¹¹ and the implementation of electrostatic embedding in ONIOM.²⁴ Both approaches assume that the charge distribution of the QM system can be accurately described by a point-charge model. In the QTCP approach, the following correction factor is added to eq 2:

$$E_{\text{corr}}^{\text{QTCP}} = \sum_{i \in \text{QM with M1}, j \in \text{MM}} \frac{f_{ij} Q'_i q_j}{4\pi\epsilon_0 r_{ij}} - \sum_{i \in \text{QM with HL}, j \in \text{ptch}} \frac{Q_i q'_j}{4\pi\epsilon_0 r_{ij}} \quad (3)$$

where Q_i are charges fitted to the QM electrostatic potential (ESP charges^{11,39,40}) for the QM system, including the HL atoms (these charges change when the QM system or the charge-redistribution scheme changes); Q'_i are the same ESP charges, except that the charge on the M1 atom has been modified to be a charge typical for a carbon atom and to give an integer net charge of all atoms (see below); q_j are the standard MM charges for the MM atoms (always the same charges); q'_j are the point charges, i.e., the q_j charges, but possibly modified by a charge-redistribution scheme; f_{ij} is a scaling factor for MM exclusion rules (for the Amber force field, used in the present calculations, $f_{ij} = 0$ for atoms separated by one or two bonds, $f_{ij} = 0.5$ for atoms separated by two bonds, and $f_{ij} = 1$ otherwise; note that this factor is present only in the first term, not in the second); and r_{ij} is the distance between atoms i and j . Note that the coordinates of the junction atoms are those of M1 in the first term but those of HL in the second term. The philosophy behind this correction is that the second sum should remove the effect of the wrong positions and charges of the HL atoms, as well as remove the 1–2, 1–3, and 1–4 interactions in the $E_{1+\text{ptch2}}^{\text{QM,HL}}$ term in eq 2, using an ESP-charge description of the QM system. Then, the first sum should introduce these terms again, but with the correct positions and charges of the M1 atoms, and with correct exclusion rules. All other interactions between QM and MM, which do not involve HL and are more than three bonds apart, are identical in the two sums and therefore cancel in eq 3. If we use the division of the total system into three parts (Q, J, and M), this provides corrected energies for the Q–M and J–M interactions at a MM approximation. In the following, we will call this the QTCP correction.

In the implementation of electrostatic embedding with ONIOM, Morokuma and co-workers took this idea one step further by also correcting the Q–J and J–J interactions at the MM level.²⁴ This is done by adding the following correction term to the QTCP-corrected results:

$$E_{\text{corr}}^{\text{ONIOM}} = \sum_{i \in \text{J with M1}, j \in \text{QM}, i \neq j} \frac{f_{ij} Q'_i Q'_j}{4\pi\epsilon_0 r_{ij}} - \sum_{i \in \text{J with HL}, j \in \text{QM}, i \neq j} \frac{f_{ij} Q_i Q_j}{4\pi\epsilon_0 r_{ij}} \quad (4)$$

The two sums run over the same atom pairs, but the first sum uses coordinates and charges of M1 atoms, whereas the second term uses instead coordinates and charges of the HL atoms. In fact, the total QM/MM energy with both the QTCP and ONIOM correction can simply be written as

$$E_{\text{QM/MM}}^{\text{EE, ONIOM}} = E_{1+\text{ptch2}}^{\text{QM,HL}} + E_{12}^{\text{MM,M1}} - E_{1+\text{ptch2}}^{\text{MM,HL}} \quad (5)$$

Here, the first term on the right-hand side appears already in eq 2. The second term is the standard MM energy of a full system with M1 coordinates and charges, whereas the last term is the MM energy of the QM system, with HL coordinates and charges, and including the point charge model of the MM system as a separate molecule (i.e., without applying exclusion rules for the QM/MM cross terms). By this simple approach, all errors introduced by the junctions are corrected, provided that the ESP charges give a proper description of the charge distribution in the QM system and the MM approximation is accurate enough to describe the difference between the HL and M1 atoms. We will call this the ONIOM correction in the following. We are not aware of any previous comparison of these approaches.

We will see below that the QTCP corrections sometimes become too large because the ESP charges on the HL atoms become strange, owing to overpolarization by the point-charge model. Therefore, we also tested to obtain the ESP charges from a wave function that is calculated without the point-charge model (i.e., in a vacuum). These charges will be called Q_0 and Q'_0 in the following, and the corresponding corrected results will be called QTCP₀ and ONIOM₀. Note that in these corrected energies, we still use the $E_{1+\text{ptch2}}^{\text{QM,HL}}$ term, so that the MM system still polarizes the QM system in the energy; it is only when calculating the ESP charges for the QM system that the point-charge model is excluded. This requires an extra set of QM wave-function calculations for each system.

Mechanical Embedding. Another way to partly correct the use of HL atoms is to use mechanical embedding (ME),^{7,8,24} although this approach is normally not introduced with this explicit aim. ME implies that the QM calculations are performed in a vacuum (i.e., without any point charges), giving $E_1^{\text{QM,HL}}$. This energy is the QM-only energy of the isolated QM system, and it will be called QM below. If it is combined with the two MM energy terms in eq 2, we obtain a QM/MM energy that does not contain any electrostatic interactions between the QM and MM systems. We will call this QM+vdW:

$$E_{\text{QM/MM}}^{\text{QM+vdW}} = E_1^{\text{QM,HL}} + E_{12,\text{no1el}}^{\text{MM,M1}} - E_{1,\text{no1el}}^{\text{MM,HL}} \quad (6)$$

The electrostatic interaction between the QM and MM systems can be introduced by calculating it at the MM level. This can be done by using the same two MM terms as in eq 5 (i.e., without

zeroing the charges of the QM system):

$$E_{\text{QM/MM}}^{\text{ME}} = E_1^{\text{QM, HL}} + E_{12}^{\text{MM, M1}} - E_1^{\text{MM, HL}} \quad (7)$$

This is the standard form of ME, e.g., used in the ONIOM approach.¹⁹ The natural choices of charges for the QM system are the Q_0' and Q_0 charges for the second and third terms in eq 7, respectively, because the wave function used in the first (QM) term is obtained without any point-charge model. This approach will be called ME below. We have tested to use the Q_0' charges also for the $E_1^{\text{MM, HL}}$ term, which we call ME₁.

It should be noted that these two approaches by construction include the ONIOM correction in eq 4 (i.e., the Q–J and J–J corrections for the HL atoms). To estimate the size of this correction, we have also considered a ME variant, in which this correction term is excluded:

$$E_{\text{QM/MM}}^{\text{ME0}} = E_{\text{QM/MM}}^{\text{QM+vdW}} + \sum_{i \in \text{QM} \cup \text{MM with M1}, j \in \text{MM}} \frac{f_{ij} Q_i' q_j}{4\pi\epsilon_0 r_{ij}} \quad (8)$$

i.e., where all electrostatic MM interactions between one atom in any of the two systems and one atom in the MM system with M1 positions and charges (i.e., Q–M, J–M, and M–M interactions; eq 3) have been added to $E_{\text{QM/MM}}^{\text{QM+vdW}}$. This approach will be called ME₀ below.

A problem with the ME approach is that it completely ignores the polarization of the QM system by the MM system. A simple way to partly fix this problem is to calculate the ESP charges with a wave function polarized by a point-charge model (i.e., to use the Q' and Q charges instead of Q_0' and Q_0 charges in eq 7). We call such an approach ME'. Like QTCP₀ and ONIOM₀, it requires an extra set of wave function calculations. Moreover, the results will (slightly) depend on the charge-redistribution scheme used.

Unfortunately, such an approach is not fully consistent, because the cost of polarization is not included in the energy. However, this cost can be included in a linear-response approximation by simply taking the average of the ME' and ME₀ energies:

$$E_{\text{QM/MM}}^{\text{MEav}} = \frac{E_{\text{QM/MM}}^{\text{ME0}} + E_{\text{QM/MM}}^{\text{ME}'}}{2} \quad (9)$$

which we will call ME_{av} in the following.

We have also calculated a sixth variant of ME, in which we instead take the average of the ME' and QM+vdW energies. This energy, which we call ME_{scal}, simply scales down the electrostatic interaction energy by a factor of 2, which could be considered as a primitive model of the polarization of the MM system. The rationale for this is that it has frequently been observed that electrostatic interaction energies are overestimated by QM/MM.^{14,41,42}

A problem with all of the previous ME methods is that they are sensitive to the stability of the ESP charges used for the QM system. In particular, for a large QM system, any variation of the charges on the boundary of the QM system will make large contributions to the energy difference between various states, because they are close to the MM system. Such variation can have three causes. First, the charge redistribution that occurs in the center of the QM system, i.e., in the actual chemical reaction, may induce changes further out (polarization). Second, the charge-derivation scheme may be unstable so that small changes in the charge redistribution due to polarization lead to large changes in

the charges. Third, the wave function optimization itself may be unstable so that the two calculations on the reactant and product states end up in different local minima with respect to a remote (typically not covalently linked) part of the system, which consequently acquires different charges. Whereas the first effect is clearly desirable (it makes the treatment of polarization more self-consistent), the other two are artifacts.

A simple way to see if the desired effect is dominating is to eliminate all three effects and see if the results get worse. To this end, we tested a seventh ME method, in which the QM charges for the two studied states were forced to be identical (by averaging over the two reactants, except for the central core). Because equal charges do not contribute to the energy difference, this method in practice only includes QM/MM interactions between the central core and the MM system, whereas QM interactions are considered within the full QM system. Thus, all indirect effects of polarization are ignored. We call this direct method ME_{dir}. All possible sizes of the central core were tested, but the best results were obtained with the smallest 46-atom QM system. Consequently, all results in the tables are obtained with that selection.

All methods are summarized in Table 1. We will see that the various ME methods provide a convenient way to test the various correction terms used in this article.

Computational Details. As a test case, we use a 446-atom model of the active site of [Ni,Fe] hydrogenase. It contains the central core, consisting of [(CH₃COOH)(CH₃S)₂Ni(CH₃S)₂-Fe(CO)(CN)₂(C₃N₂H₅)⁻ as a 46-atom model of the [Ni,Fe] active site with four Cys ligands (Cys-72, 75, 543, and 546), as well as the second-sphere groups of His-79 and Glu-25 (Figure 3). To this system, we then have added 40 models of amino acids, according to their energy contribution in a QM/MM free-energy study.¹⁴ Some of the added groups are covalently connected to the original model; these are shown in detail in Figure S1 (Supporting Information). Other groups are separated and more distant (up to 16 Å from the 46-atom QM system; cf. Figure 3).

In a second set of calculations, the same 40 groups were added, although in a different order: First, the original QM residues were capped with CH₃CONH– and –CONHCH₃ groups of the surrounding backbone. Next, this backbone was extended by one further CH₃CONH– and –CONHCH₃ group on each side. However, note that the full 446-atom system does not contain all such groups (because only the 40 groups with the largest QM/MM free energy components were included; cf. Figure S1). If the backbone groups are named after the residue containing the N atom, the following backbone residues are included in the first system: Cys-72, Cys-75, His-79, Ala-80, Cys-543, Ile-544, Cys-546, and Gly-547. The second system included the backbone of Ala-71, Val-78, and Pro-542. After that, the remaining 29 groups were added one by one, according to their error in the EE calculations of set 1 with the BCS charge-redistribution scheme. These groups are not covalently connected to the other residues, except for Arg-70 and Gln-69 (the latter is added in two parts, one consisting of the backbone CH₃CONH– group and the other the inner part of the side chain. In addition, the carboxy-terminal His-549 is divided into two groups: the negatively charged backbone belongs to one group (called His-549), whereas the neutral side chain is a ligand of the Mg site, which is added as a single group, consisting of Mg²⁺, this imidazole group, three water molecules, the side chain of Glu-53, and the backbone CO group of Leu-495.

Table 1. Description of the 14 Methods Discussed^a

method	eq	QM	vdW	elstat	QTCP corr	Oniom corr	Q_{QM}	Q_j
QM	before 6	vac						
QM+vdW	6	vac	yes					
ME ₀	8	vac	yes	MM			vac	M1
ME	7	vac	yes	MM		yes	vac	M1
ME ₁	after 7	vac	yes	MM		yes	vac	HL
ME'	after 8	vac	yes	MM			pol	M1
ME _{av}	9	vac	yes	MM			(vac+pol)/2	M1
ME _{scal}	after 9	vac	yes	0.5MM			pol/2	0.5M1
ME _{dir}	7	vac	yes	MM			vac, only core	M1
EE	2	ptch	yes	QM			pol(QM)	HL
QTCP ₀	after 5	ptch	yes	QM	yes		vac(corr)	M1
ONIOM ₀	after 5	ptch	yes	QM	yes	yes	vac(corr)	M1
QTCP	3	ptch	yes	QM	yes		pol	M1
ONIOM	4, 5	ptch	yes	QM	yes	yes	pol	M1

^aThe QM calculations can either be performed in a vacuum (vac; mechanical embedding) or with a point-charge model of the surroundings (ptch; electrostatic embedding). Van der Waals interactions (vdW) can be included or not. The electrostatic interactions between the QM and MM systems (elstat) can be ignored, calculated by molecular mechanics (MM), or calculated by a point-charge model in the QM calculations (QM). Errors introduced by the junction atom can be corrected by either the QTCP correction in eq 3 or the Oniom correction in eq 4. The charges of the QM atoms (Q_{QM}) can be obtained in a vacuum (vac) or with a wavefunction polarized by a point-charge model (pol). They can also be averaged, scaled down, or corrected. Finally, the charges of the junction atoms (Q_j) can be those of the HL or the M1 atoms.

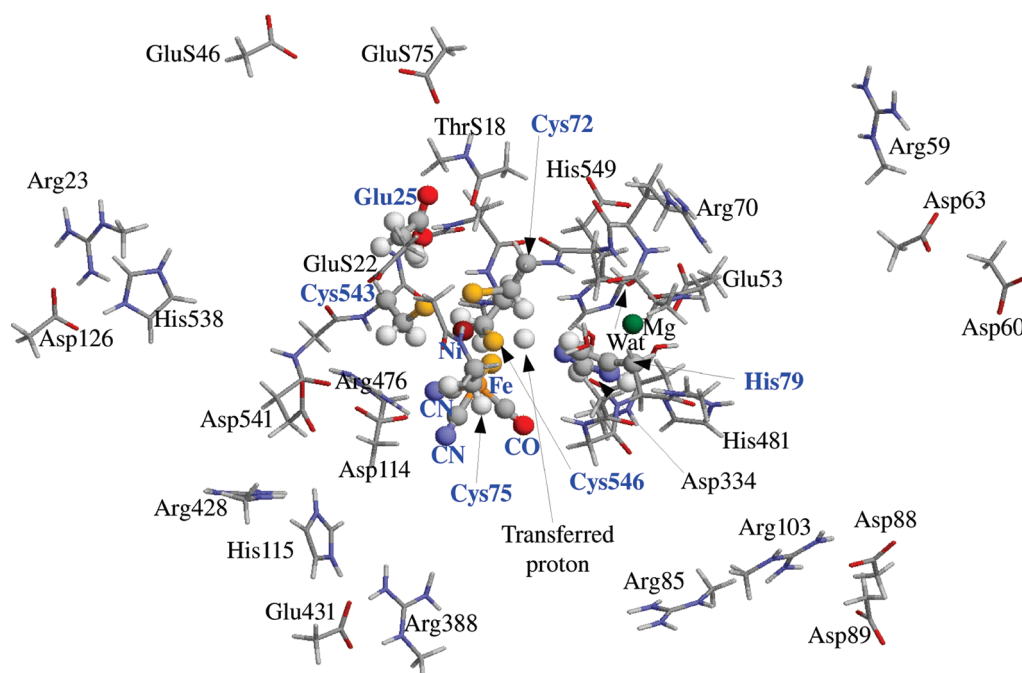


Figure 3. Atoms included in the 446-atom test system. The Mg ligand Gln540 is mainly hidden behind His79. Note that there are four water molecules in the calculation: one at the arrow, one just to the left of the label, and two just to the right of the label. All are Mg ligands, except the upper to the right of the label. The smallest (46-atom) quantum system is shown in balls and sticks and with blue bold-face labels.

We study the energy of the simple proton-transfer reaction shown in Figure 3. It involves the transfer of a proton from the S^γ atom of one of the bridging cysteine ligands (Cys-546 in *Desulfovibrio fructosovorans*) to the N^{ε2} atom of a second-sphere histidine ligand (His-79). We have calculated the energy difference between the form in which the proton resides on Cys-546 (called the HID state) and the form in which the proton resides on His-79 (called the HIP state). As a reference value, we use the QM energy for the full 446 model calculated in a vacuum,

48.3 kJ/mol.¹³ All methods give this energy for the largest QM system, because then no residues remain in the MM system. The aim of this investigation is to see how the QM/MM results converge toward this value as more and more residues are moved from the MM system to the QM system.

All calculations were performed with density-functional theory, using the Becke-1988–Perdew-1986 functional^{43,44} and the def2-SV(P) basis sets.⁴⁵ The calculations were sped up by expanding the Coulomb interactions in auxiliary basis sets, the

resolution-of-identity approximation.^{46,47} The Fe and Ni ions were assumed to be in the low-spin +II oxidation state, giving a closed-shell singlet state of the full system.^{13,14} All calculations were performed on exactly the same two structures (one for the HID state and one for the HIP state; thus no QM/MM geometry optimization was performed in this investigation), taken from QM/MM structures, obtained with the 46-atom QM system (i.e., only these 46 QM atoms had different positions in the HID and HIP states, whereas all of the other 400 atoms had the same positions in the two states).¹⁴

The full 446-atom model consists of hydrogen-atom capped amino acid fragments (Figures 3 and S1 Supporting Information). Therefore, standard MM parameters could not be directly used for the MM system. Instead, we started from a full MM model of the whole protein, described by the Amber 1999 force field.^{38,48} This system was truncated to the 446-atom model, filling all broken bonds with a hydrogen atom at a C–H distance of 1.101 Å. All internal parameters (bonds, angles, and dihedrals) that were not affected by the truncation were kept at the Amber 1999 force field. All angles and dihedrals involving the HL atoms were set to the corresponding parameters for the M1 atom, whereas the bonds involving the HL atom had an equilibrium distance of 1.101 Å and a force constant given by²⁷

$$k_{\text{Q1-HL}} = k_{\text{Q1-M1}} \frac{1.101^2}{r_{\text{Q1-M1}}^{\text{MM0}}} \quad (10)$$

The Lennard-Jones parameters for the HL atoms were the same as for the Amber HC atom type (hydrogen bound to carbon). Finally, the MM charges (q_j ; charges above) were determined individually for each covalently connected fragment of the 446-atom model by ESP charges, calculated with the Merz–Kollman scheme,³⁹ as implemented in Turbomole.²⁸ For the large fragment involving the Ni and Fe ions, charges outside the smallest QM system were averaged between those of the HID and the HIP states. These charges were always used for the MM system, either as charges for the MM system in the mechanical-embedding calculations or as point charges in the QM calculations. They were always the same in all calculations (besides possible adaptations according to the charge-redistribution schemes).

For atoms in the QM system, new charges were calculated from the wave function for all atoms in each QM system and charge-redistribution scheme (Q_i ; charges above). They were Merz–Kollman ESP charges³⁹ obtained from a QM calculation of the entire QM system (not only on fragments as for the MM charges). In the case when the wave function was polarized by the point charges, the point charges were omitted in the ESP calculations, without reoptimizing the wave function. These charges were used to describe the QM system in the QTCP and ONIOM energy-correction schemes and also in the mechanical-embedding calculations. To obtain charges for the M1 atoms (Q_i ; charges above), the charges of the HL atoms were adapted so that the full system had the correct integer charge (so that they get a size typical for carbon atoms, rather than for hydrogen atoms).²⁷ For full amino acids, this adaptation is unambiguous, because each amino acid has an integer charge. However, in the present calculations, with QM charges calculated for fragments that may be connected, the fragments do not always have an integer charge (owing to charge transfer between the fragments; this charge transfer can be extensive, up to 0.5e between ionic pairs). Therefore, the adaptation sometimes becomes ambiguous. We solved this problem by simply partitioning the ambiguous

charge equally between all fragments with junctions. This partitioning can be based on either the QM (Q_i) or MM (q_j) charges on the M1 atoms, but this gave little difference in the final energies. The presented results are based on the QM M1 charges.

Finally, a third set of calculations was performed, based on a 12 178-atom model of [Ni,Fe] hydrogenase, solvated in a spherical system (35 Å radius). The setup of this system has been described before.^{14,42} In these calculations, all 12 178 atoms were included in the MM calculations, whereas 46 to 446 atoms were included in the QM system. The groups were moved between the QM and MM system in the same order as in the first set of calculations, described above, and the QM systems were identical. Charges of the MM atoms were taken from the Amber-1999 force field,³⁸ except for the metal sites, for which the charges were taken from QM calculations.^{14,42} The MM charges were the same for the HID and HIP states. Only the Z1 charge-redistribution scheme was tested.

RESULT AND DISCUSSION

Comparison of QM/MM Methods. In this paper, we study how well QM/MM calculations reproduce a QM vacuum calculation for a 446-atom model of the active site of [Ni,Fe] hydrogenase. We study the energy of a simple proton transfer from the S γ atom of the Ni ligand Cys-546 to the N ϵ^2 atom of the second-sphere group His-79. We have systematically studied this reaction with QM/MM methods, in which we increase the QM system from 46 to 446 atoms by moving 40 groups from the MM to the QM system. The smallest QM system contains the Ni²⁺ and Fe²⁺ ions and their ligands, as well as the proton acceptor and another second-sphere ligand that also shares a proton with a Cys ligand. All amino acids have been truncated in a standard way; i.e., Cys is modeled by CH₃S⁻, His by imidazole, Asp and Glu by acetate, Lys by methylamine, and Arg by methylguanidine.

QM calculations with the full 446-atom model give a reaction energy of 48 kJ/mol in favor of the HIP state.¹³ If the QM/MM calculations were perfect, they would always give this energy, irrespectively of the number of groups in the QM system. We have tested QM-only calculations, as well as 13 different variants of QM/MM with mechanical embedding (ME) or electrostatic embedding (EE), using 12 different charge-redistribution schemes (to avoid overpolarization close to the covalent junctions between the QM and MM systems) and two different ways to correct errors introduced by the junctions. Finally, we have tested to add the 40 residues to the QM system in two different ways. In the first set of calculations, the 40 groups were added to the QM system in the order of the size of their contributions to the QM/MM free energy difference of this reaction.¹⁴ These calculations showed that the largest errors come from junctions directly connected to the active site. Therefore, we constructed a second set of calculations, in which first the backbone of all residues involved in the active site were added to the QM system (including CH₃CO– and –NHCH₃ groups from the neighboring residues), then the backbone of all neighboring residues, and finally, the remaining 29 residues were added in the order of their QM/MM error with the BCS charge-redistribution scheme with the EE method (Figure S1 in the Supporting Information shows details of the residues).

All of these calculations give a large amount of data that are presented in Figures S2 and S3 in the Supporting Information. Here, we will summarize the results and extract the most interesting conclusions.

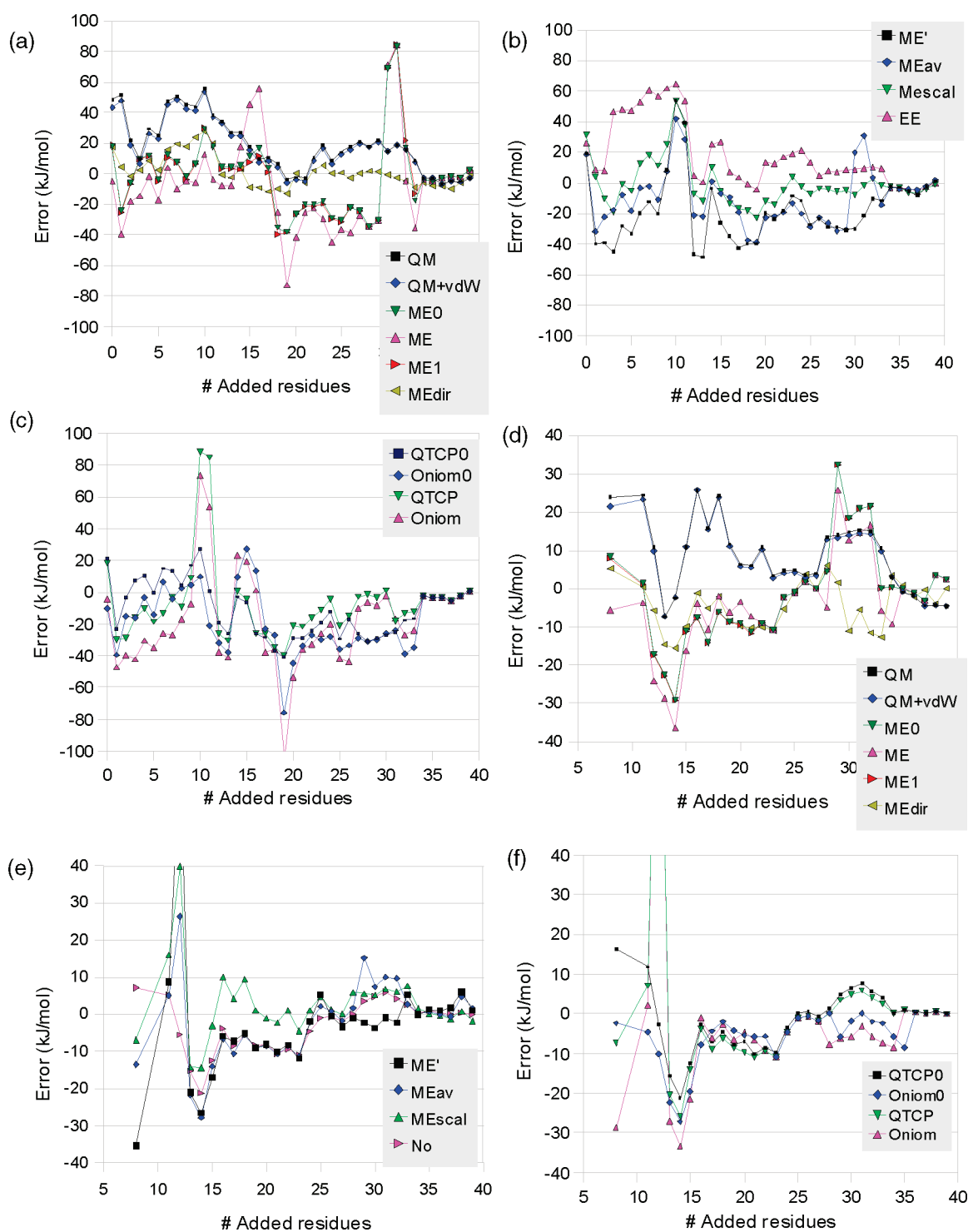


Figure 4. Errors for the 14 methods tested for the two sets (a, b, c, set 1; d, e, f, set 2) of calculations (using the Z1 charge-redistribution scheme when applicable; results for the other charge-redistribution schemes are shown in Figures S2 and S3 in the Supporting Information). Note that the QM+vdW method almost entirely coincides with QM and ME1 with ME.

First, we show in Figure 4 the error (compared to the QM calculation with 446 atoms, i.e. 48 kJ/mol) for the 40 (set 1) or 30 (set 2) individual calculations for all 13 methods, using the Z1 charge-redistribution scheme, when applicable. It can be seen that the error is smaller on average in set 2 but that the variation is large among the various calculations. Therefore, we will mainly discuss the results in statistical terms, using the mean signed error

(MSE), the mean unsigned error (MUE), and maximum unsigned error (max) among the 40 or 30 calculations for each method and charge-redistribution scheme.

We will start with discussing the 12 different charge-redistribution schemes tested (note that the QM, QM+vdW, ME₀, ME, ME₁, and ME_{dir} methods are not affected by the charge-redistribution schemes, because they do not use any point-charge

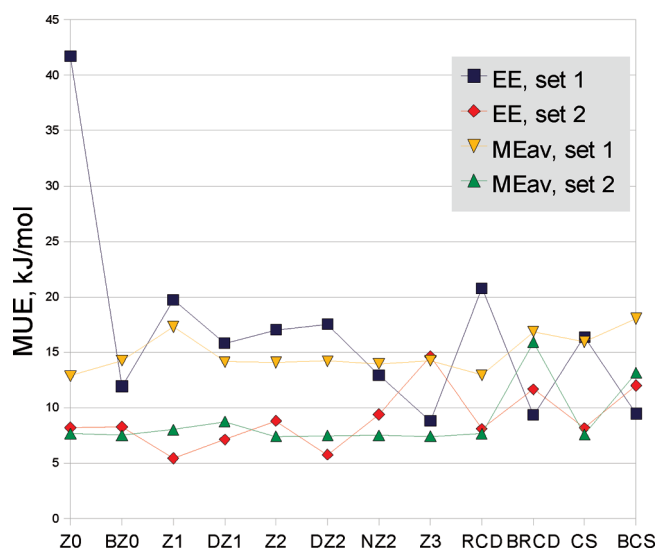


Figure 5. The mean unsigned error (MUE) of two methods, EE and ME_{av} , as a function of the charge-redistribution scheme for the two sets of calculations.

FS model). In Figure 5, the mean unsigned errors (MUEs) for the two sets of calculations and two representative QM/MM methods, EE and ME_{av} are shown. Many methods (in particular EE, QTCP₀, and ONIOM₀) give large errors with Z0, indicating that it is inappropriate to include M1 point charges in the QM calculations.

Moreover, for a small number of the BZ0, Z1, DZ1 (only set 2), BRCD, and BCS calculations, ME' and in particular QTCP and ONIOM give very large errors (100–430 kJ/mol; cf. Figure 4). The reason for this is that owing to the near-by point charges, the HID and HIP calculations end up in a different electronic states, giving different ESP charges for the QM atoms, leading to large QM–MM interaction energies, and especially QTCP and ONIOM energy corrections. Therefore, these combinations of methods and charge-redistribution schemes should be avoided. The ME_{av} and ME_{scal} methods are calculated as the average between ME' and either the ME_0 or QM+vdW energies. Therefore, these methods are also slightly affected by this problem for ME' , but the effect is quite small, as can be seen in Figure 5. The EE results are not affected by this problem, because it never uses any point-charge description of the QM system. Likewise, the ME_0 , QTCP₀, and ONIOM₀ methods are not affected by this problem because the charge model of the QM system is obtained without the point-charge model of the MM system (i.e., in a vacuum).

From Figure 5, it can be seen that the ME_{av} method gives quite stable results, once the charge-redistribution schemes affected by the above problem are disregarded, with a MUE of ~14 kJ/mol for the first set and ~8 kJ/mol for set 2. It can also be seen that the variation is somewhat larger for set 1, indicating that the choice of charge-redistribution scheme matters mainly for junctions closest to the reaction. Most of the other QM/MM methods give similar variations among the various charge-redistribution schemes. The EE method is the prime exception to this rule, showing quite large variations among the charge-redistribution schemes. For set 1, it gives a clear improvement for all charge-neutralization schemes (i.e., BZ0, DZ1, BRCD, and BCS give better results than Z1, RCD, and CS, respectively). The best results are obtained with the Z3 charge-redistribution scheme.

Unfortunately, set 2 shows the opposite results, with the best results with Z1 and DZ2 and the worst results with Z3, BRCD, and BCS. This indicates either that the variation observed is only random, or that two different effects are present, viz, the effect of junctions, for which Z3, BRCD, and BCS give the best result, and other approximations in QM/MM, which the Z1 and DZ2 methods seem to treat better. Considering that most of the charge-redistribution schemes do not attempt any physical correction of the problems involved but simply delete charges (which sometimes can be involved in important hydrogen bonds with the QM system, especially the backbone NH and CO groups) and that some of the schemes involve the questionable use of the charge on the M1 atoms, in the following we simply discuss the average over all 12 charge-redistribution schemes (for the methods affected by the occasional problem with QM charges, we also discuss the results when the affected methods are omitted from the averages).

The average performance of all methods for the two sets of calculations are summarized in Table 2. It can be seen that for set 1, the ME_{dir} method gives the lowest MUE (7 kJ/mol), which is a clear improvement from the QM-only approach (21 kJ/mol). It also gives the lowest maximum error (28 kJ/mol) and the fastest convergence to 10 and 20 kJ/mol errors. On the other hand, several other methods give a faster convergence to an error of less than 5 kJ/mol. The results do not change qualitatively if the four charge-redistribution schemes that give occasional spurious QM charges are omitted, but the values change somewhat and the variation among the various charge-redistribution schemes decreases, as can be seen from the standard deviations (σ_{MUE} , σ_{Max} , and σ_{MS} in Table 2).

For set 2, the results are somewhat more varying, because many methods give results of a similar quality. If all charge-redistribution schemes are considered, the QTCP₀ approach gives the lowest MUE (6 kJ/mol, but the ME_{dir} , ONIOM₀, and ME_{scal} methods give similar results within 1 kJ/mol), whereas ME_{dir} gives the lowest maximum error (16 kJ/mol and therefore is always converged to 20 kJ/mol). On the other hand, EE gives the fastest convergence to 5 kJ/mol, and ME_{scal} gives the fastest convergence to 10 kJ/mol. If the four problematic charge-redistribution schemes are omitted from the average, ME_{scal} gives the lowest MUE, 5 kJ/mol.

Thus, we can conclude that the ME_{dir} approach gives outstanding results for set 1 and also among the best results for set 2. This is surprising, because it suggests that all efforts of including polarization and correctly treating electrostatics around the junctions are meaningless unless a stable QM method is used; otherwise, nonphysical variations in the electrostatic description of the QM system (charges in the ME methods; charge density in the EE methods) will dominate the errors. To confirm this, we repeated the ME_{dir} calculations with all possible choices of the subsystem in which the charges were allowed to vary. Indeed, we found that the smallest QM system gave the best results and more generally that all subsystems that gave a low MUE were small. However, taking this size reduction to the extreme by using the same charges for all atoms (e.g., also the reacting atoms) gave larger errors (MUE = 22 and 19 kJ/mol for the two sets).

Energy Contributions. Further understanding of the various methods and the components involved can be gained by a pairwise comparison of the methods. The average magnitudes of these differences are shown in Table 3 for both sets. In Figure 6, this information has been combined with the mean signed and unsigned errors from Table 2 to provide a pictorial

Table 2. The Performance of the Various Methods^a

	<i>n</i>	set 1									set 2										
		MUE	max	MSE	stdev	n5	n10	n20	σ_{MUE}	σ_{max}	σ_{MSE}	MUE	max	MSE	stdev	n5	n10	n20	σ_{MUE}	σ_{max}	σ_{MSE}
QM	1	20.5	56.1	18.8	18.1	38	32	29				10.1	25.7	8.4	9.3	33	33	18			
QM+vdW	1	19.2	53.8	17.2	17.4	38	32	29				9.6	25.8	7.9	9.0	33	32	18			
ME ₀	1	18.2	83.2	-1.5	25.2	33	33	31				9.3	32.5	-1.6	13.0	32	32	32			
ME	1	23.0	84.4	-7.9	30.7	33	33	32				9.1	36.5	-3.9	12.5	34	32	29			
ME ₁	1	18.2	84.2	-2.2	25.5	33	33	32				9.4	32.5	-1.6	13.1	32	32	32			
ME _{dir}	1	7.4	28.3	1.4	10.2	38	19	10				5.9	15.5	-4.5	6.1	33	33	0			
ME'	12	18.0	64.6	-9.2	22.4	37	33.0	30.7	3.8	39.4	4.4	9.3	49.2	-1.4	14.3	38.0	24.5	17.7	5.4	44.6	5.3
ME _{av}	12	14.9	43.5	-5.4	18.7	33	33.0	31.0	1.7	18.0	2.2	8.8	35.0	-1.5	12.1	32.3	30.3	17.7	2.7	19.2	2.6
ME _{scal}	12	12.9	45.8	4.0	16.9	37	23.7	16.8	2.6	10.9	2.2	6.3	30.6	3.3	9.4	33.2	20.1	9.5	2.7	21.4	2.6
EE	12	16.8	57.4	7.2	21.4	37	27.7	19.2	8.8	31.4	7.6	9.0	32.5	-7.3	10.8	28.4	24.3	16.4	2.7	8.5	2.9
QTCP ₀	12	20.6	58.3	-6.4	23.3	37	32.8	29.7	7.4	29.7	9.6	5.6	22.2	-1.3	8.0	32.7	22.9	9.1	0.9	4.7	0.7
ONIOM ₀	12	23.2	81.2	-13.0	24.7	37	33.0	33.0	5.9	16.8	9.6	6.1	26.5	-5.0	7.7	34.9	26.0	17.8	0.5	3.8	0.7
QTCP	12	23.0	109.8	-6.8	35.0	37	33.0	30.4	16.0	132.3	8.7	15.1	109.2	2.4	28.4	30.8	24.8	17.8	19.5	149.6	16.4
ONIOM	12	32.1	154.2	-18.7	41.3	37	33.0	32.9	17.0	134.6	9.2	15.0	107.7	-0.7	27.1	34.0	25.3	18.4	17.5	134.5	15.4
EE	11	14.0	47.5	4.9	17.9	37	27.0	18.0	4.1	12.1	5.9										
ME'	8	16.0	49.5	-9.3	18.2	37	33.0	30.5	2.0	19.0	2.8	6.6	22.9	-3.5	8.5	38.0	23.0	13.9	0.3	2.5	0.1
ME _{av}	8	14.1	35.8	-5.4	17.0	33	33.0	31.0	1.0	4.7	1.4	7.5	25.5	-2.6	9.7	32.0	29.5	14.0	0.1	1.3	0.1
ME _{scal}	8	11.9	40.1	3.9	15.5	37	21.4	14.0	0.4	2.2	1.4	5.0	19.1	2.2	6.9	33.0	17.5	2.0	0.1	1.1	0.1
QTCP	8	15.4	46.2	-9.6	17.0	37	33.0	30.9	3.3	12.1	2.3	6.4	27.0	-4.3	8.4	30.0	23.0	14.1	0.8	4.6	1.0
ONIOM	8	23.9	94.6	-21.5	22.0	37	33.0	32.9	3.3	12.0	3.9	7.2	32.3	-7.1	8.6	34.0	23.6	14.9	0.9	4.1	0.9

^a The table lists the mean signed and unsigned error (MSE and MUE), the maximum error (max), the standard deviation of the MSE (stdev), and the number of residues that need to be in the QM system before the error is always smaller than 5, 10, and 20 kJ/mol, respectively (n5, n10, and n20). If applicable, the presented results are the average over all 12 charge-redistribution schemes. *n* is the number of values in each average, and σ_{MUE} , σ_{max} , and σ_{MSE} are the standard deviations of the MUE, max, and MSE values in these averages. For some methods, one or four charge-redistribution schemes are omitted from these averages in the lower part of the table. These are Z0 for EE in set 1 and Z1, BRCD, BCS, and Z0 (set 1) or DZ1 (set 2) for the ME', ME_{av}, ME_{scal}, QTCP, and ONIOM. The best value in each column is marked in bold face.

representation of how the various methods are related and how the effects behave (results are shown only for set 1).

The first column in Table 3 (vdW) shows the difference between the QM+vdW and QM methods, which estimates the effect of the van der Waals interactions (always involving an ONIOM-type correction for the HL atoms). It can be seen that the effect of the van der Waals interactions is small, 1–2 kJ/mol for the two sets, but it leads to a slight improvement in most quality estimates (Table 2).

Next, we estimate the effect of the electrostatic interactions by taking the difference between the ME₀ and QM+vdW methods (the electrostatic interactions are calculated at the MM level, using charges of the QM system obtained in vacuum). From Table 3 (column ele), it can be seen that this gives a large effect, 26 kJ/mol in set 1 and 13 kJ/mol in set 2. It gives a small improvement in the MUEs, but a very strong improvement in MSE (to -2 kJ/mol). On the other hand, the maximum error increases, showing that the variation of the results becomes larger, and in many cases, the results actually become worse. The direct electrostatic effect of the MM system on the reaction (i.e., the difference between the ME_{dir} and QM+vdw methods; column ele_{dir}) is significantly smaller (16 kJ/mol) for set 1 but equally large (13 kJ/mol) for set 2. This means that there are large indirect effects of charge redistribution whenever there are junctions close to the reaction. As discussed above, the direct electrostatic effect reduces the MUEs significantly.

The ME method differs from ME₀ only in that an ONIOM correction for the erroneous position and charge of HL of the Q–J and J–J interactions is attempted. From Table 3 (column

juncorr), it can be seen that this amounts to a quite large correction in set 1 (12 kJ/mol) but a rather small correction in set 2 (where there are only a few junctions; 4 kJ/mol). From Table 2, it can be seen that in general, this correction leads to worse rather than improved results. Thus, the point-charge model of the QM system does not seem to be accurate enough to make such corrections.

ME₁ is identical to the ME method, except that only the position of the HL atom is corrected, not the charges. From Table 3 (column q_{HL}), it can be seen that this makes a large difference. In fact, the results become very similar to those of the ME₀ method, indicating that the main effect of the ONIOM correction comes from the charge rather than the position of the HL atoms.

In the ME' approach, the charges of the QM system are obtained from a wave function polarized by the MM system (the point charges). Therefore, they are polarized by the MM system. This approach gives results of a similar quality to those of the ME₀ method (with unpolarized QM charges; Table 2), but the individual results (Table 3, column me') differ by 23 kJ/mol for set 1 and by 7 kJ/mol for set 2, showing that it makes a major difference whether the charges are calculated by a polarized or a vacuum wave function.

The ME' approach is inconsistent, because the cost of the polarization of the QM system is not considered. This cost is included in the ME_{av} method, which is the average of ME' and ME₀. From Table 3, it can be seen that the effect of the consistent polarization (column pol) is rather small: 8 kJ/mol for set 1 and 3 kJ/mol for set 2. Thus, the effect of polarization is 3–4 times

Table 3. Effects of Various Components of the Tested Methods, Presented As the Difference in Mean Unsigned Error between Two Methods^a

	vdW	ele	ele _{dir}	juncorr	q _{HL}	me'	pol	scal	elpol	qmeff	qtcp	oniom	polc1	polc2
	set 1													
Z0						23.1	7.8	19.3	36.7	34.9	20.9		30.1	37.8
BZ0						20.2	7.4	17.6	14.2	16.3	26.5		8.9	15.4
Z1						24.7	10.7	17.9	11.3	29.9	29.8		12.6	14.9
DZ1						21.3	6.5	18.1	21.3	13.8	15.0		3.9	7.2
Z2						22.1	7.5	18.3	13.6	14.1	17.5		7.0	8.8
DZ2						22.3	7.4	18.8	11.8	21.0	23.4		7.2	8.7
NZ2						21.6	7.5	18.6	9.7	16.8	19.8		7.8	10.2
Z3						23.0	8.5	18.6	19.2	10.1	12.7		6.5	8.9
RCD						24.9	8.0	20.2	25.2	15.9	17.0		12.1	19.2
BRCD						27.1	11.8	21.8	16.2	18.1	19.4		56.6	61.8
CS						27.9	10.4	20.6	8.2	25.5	15.4		27.3	33.2
BCS						30.9	13.4	22.0	16.7	18.2	17.1		62.2	66.5
Av	1.6	26.1	16.2	11.6	11.8	24.1	8.9	19.3	17.0	19.6	19.5	11.8	20.2	24.4
Av8	1.6	26.1	16.2	11.6	11.8	22.9	7.9	18.8	15.4	16.7	18.4	11.8	10.1	13.9
	set 2													
Z0						7.2	3.2	9.0	16.6	5.0	4.9		4.7	4.8
Z1						10.1	5.0	10.4	12.2	3.9	1.4		6.7	9.6
DZ1						8.3	3.8	9.0	15.3	4.9	5.1		6.7	6.3
Z2						7.5	3.3	9.0	16.3	5.1	6.1		3.4	4.0
DZ2						7.2	3.2	9.0	13.2	2.9	2.9		2.4	2.8
NZ2						6.7	3.1	8.8	14.4	4.2	6.2		5.1	5.3
Z3						7.2	2.9	8.9	23.5	12.0	12.5		4.7	4.5
DZ3						7.1	3.1	8.9	16.3	4.7	5.8		2.2	2.5
RCD						7.0	3.2	8.9	16.5	4.9	5.8		3.6	3.8
BRCD						18.8	8.7	12.3	17.7	14.7	10.9		68.4	64.6
CS						6.9	3.1	8.9	16.6	5.1	6.1		3.2	3.5
BCS						13.1	5.7	10.1	18.6	12.2	11.2		39.6	35.6
Av	0.5	13.2	13.3	3.9	3.9	8.9	4.0	9.4	16.4	6.6	6.6	4.9	12.6	12.3
Av8	0.5	13.2	13.3	3.9	3.9	7.1	3.1	8.9	16.7	5.5	6.3	4.9	3.7	3.9

^aThe components are $\text{vdW} = \text{QM} + \text{vdW} - \text{QM}$; $\text{ele} = \text{ME}_0 - \text{QM} + \text{vdW}$; $\text{ele}_{\text{dir}} = \text{ME}_{\text{dir}} - \text{QM} + \text{vdW}$; $\text{juncorr} = \text{ME} - \text{ME}_0$; $q_{\text{HL}} = \text{ME}_1 - \text{ME}$; $\text{me}' = \text{ME}' - \text{ME}$; $\text{pol} = \text{ME}_{\text{av}} - \text{ME}_0$; $\text{scal} = \text{ME}_{\text{scal}} - \text{ME}_0$; $\text{elpol} = \text{EE} - \text{QM} + \text{vdW}$; $\text{qmeff} = \text{EE} - \text{ME}_{\text{av}}$; $\text{qtcp} = \text{QTCP}_0 - \text{EE}$; $\text{oniom} = \text{ONIOM}_0 - \text{QTCP}_0$; $\text{polc1} = \text{QTCP} - \text{QTCP}_0$; $\text{polc2} = \text{ONIOM} - \text{ONIOM}_0$. Av is the average over all charge-redistribution schemes, whereas in Av8, Z1, BRCD, BCS, and Z0 (set 1) or DZ1 (set 2) are omitted from the average. Note that the first four contributions, as well as the ONIOM contribution, are independent of the charge-redistribution scheme and therefore are only listed in the average rows.

smaller than the electrostatic effect for this system. From Table 2, it can be seen that ME_{av} gives a consistent improvement in all quality measures, except MSE, compared to ME' .

Even better results are obtained if ME' is averaged with $\text{QM} + \text{vdW}$ instead. This essentially means that the electrostatic interactions are scaled down by a factor of 2. This is not a physically consistent method, but it is instead based on the common observation that electrostatic interactions seem to be overestimated in QM/MM calculations.^{41,49} A possible explanation for this overestimation is that electrostatic embedding actually is inconsistent (in contrast to mechanical and polarized embedding): In EE, the QM system is polarizable, whereas the MM system is not polarizable (like mixing polarizable and nonpolarizable force fields in MM calculations). This probably leads to an exaggeration of the (polarized) charges of the QM system, compared to those of the MM system. Several attempts have been made to cure this problem by using special van der Waals parameters for the QM system.^{49–52} The effect of the scaling (column scal in Table 3) is appreciably larger

than that of the polarization, 19 kJ/mol for set 1 and 9 kJ/mol for set 2.

In the next column of Table 3 (elpol), we compare the results of the EE and $\text{QM} + \text{vdW}$ methods. In EE, both the electrostatics and the polarization are calculated at the QM level, without any attempt to correct errors arising from the junction atoms. This is the standard way to do QM/MM (with electrostatic embedding). As expected from the previous results, the difference between EE and $\text{QM} + \text{vdW}$ is quite large, 15–17 kJ/mol. For the second set, this is quite close to the sum of the ele and pol effects, but for the first set, elpol is appreciably smaller. The EE method has the largest variation among the tested methods. Typically, it gives a slight improvement over $\text{QM} + \text{vdW}$.

EE and ME_{av} both include the same terms, although they are calculated at the QM level for EE and partly at the MM level for ME_{av} . Quite unexpectedly, the effect of this change (column qmeff) is quite large in both sets of calculations, 17 kJ/mol for set 1 and 6 kJ/mol for set 2. Apparently, there are instabilities in the calculations caused mainly by the junctions. The results in Table 2

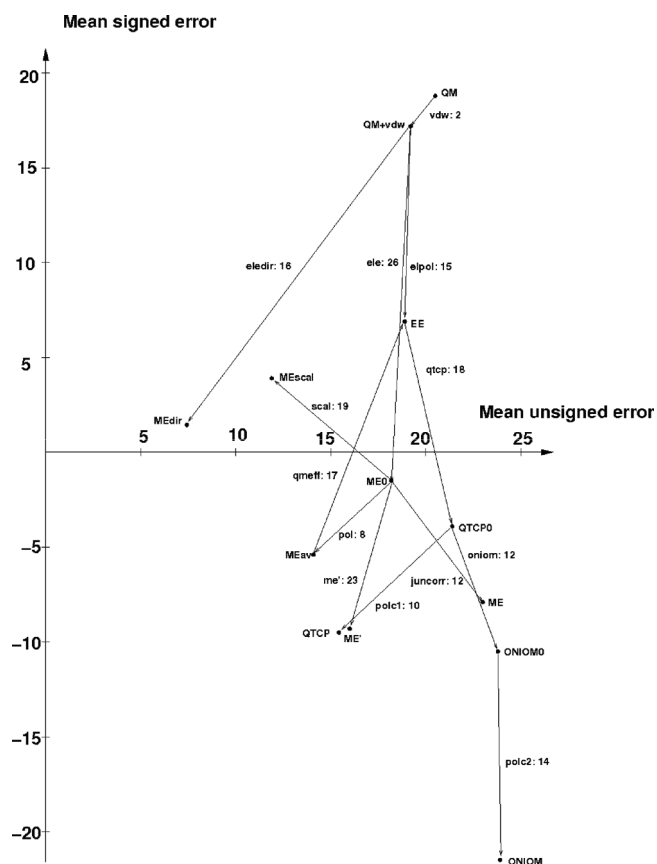


Figure 6. Position of each method in the space described by the mean unsigned error (MUE) and the mean signed error (MSE) for set 1, with the value at each arrow giving the size of the effect as described in Table 2. All values are average energies in kJ/mol over the various charge-redistribution schemes. For all methods where charge redistributions occur, the results shown are Av8, i.e., omitting Z1, BRCD, BCS, and Z0.

quite clearly show that the gain of doing electrostatics and polarization by QM is more than canceled by the errors caused by the junctions, so that the ME_{av} method actually gives the better results.

Finally, we tried to improve the results of EE by using the QTCP and ONIOM corrections. From Table 2, it can be seen that both corrections are strongly affected by the occasional problems with the QM charges for some charge-redistribution schemes. Therefore, we first discuss the results obtained with QM charges based on a vacuum wave function (QTCP₀ and ONIOM₀).

The former correction (column qtcp in Table 3) is in general quite large, 18 kJ/mol for set 1 and 6 kJ/mol for set 2. It is close to the qmeff term, showing that this effect dominates the difference between EE and ME_{av} (note that, by construction, the QTCP correction is avoided in all types of ME methods). For set 1, QTCP₀ gives results that are slightly worse than those of EE, whereas for set 2, QTCP₀ provides a clear improvement, giving the lowest MUE if all charge-redistribution schemes are considered. This indicates that the QTCP correction is advantageous, except when the junctions are too close to the center of the studied reaction.

The ONIOM correction on the other hand is problematic. It gives a correction that is very similar to what is observed with

mechanical embedding (compare columns juncorr and oniom in Table 3), 12 kJ/mol with set 1 and 5 kJ/mol with set 2. ONIOM₀ always gives worse results than QTCP₀, and it gives no or only a marginal improvement over the EE method. Again, we have to conclude that the QM charge model is not accurate enough to support a Q–J and J–J interaction correction for the HL atoms.

Finally, we compare QTCP with QTCP₀ and ONIOM with ONIOM₀ in the last two columns in Table 3, i.e., the effect of calculating the QM charges with a polarized or a vacuum wave function. It can be seen that in general, the effect is rather small, 10–14 kJ/mol with set 1 and 4 kJ/mol with set 2. However, for a few cases, very different results are obtained, as discussed above.

Residue Contributions. Some further understanding can be gained by studying the contributions from each of the added groups. A proper QM/MM method should give the same results, no matter if a group is treated by QM or by MM. Therefore, the difference in QM/MM energy between two calculations that differ only in that one group is moved from the QM system to the MM system can be considered as the QM/MM error caused by that group. Such group contributions, using the Z2 charge-redistribution scheme (if applicable), are listed in Table 4 for set 1 and Table 5 for set 2 (these tables also show in which order the residues are added in the two sets). It can be seen that, for the electrostatic embedding methods and set 1, residues that give large errors are typically those directly connected with junctions to the active site (type A) or the first neighbors (type N), although the contributions vary quite extensively among the various methods. With two exceptions, these residues give errors of 6–32 kJ/mol with the QTCP₀ correction. For set 2, most of the junction problems are collected in the first two contributions (A and N) so that the remaining contributions are much lower. However, there are also problematic residues that are not covalently connected to the active site, e.g. Glu-S22 and Arg-428. For many residues, the energy corrections are large (e.g., up to 38 kJ/mol for QTCP), and they make the errors smaller or larger in a rather random manner. Again, the ONIOM correction works poorly, giving maximum errors of 72 and 20 kJ/mol for the two sets (ignoring the A contribution for set 2), compared to 32 and 13 kJ/mol for QTCP.

For the mechanical embedding methods, the results are more unpredictable, with very large errors for some groups, e.g., Asp-63, Ile-544, and His-549. At first sight, this seems to indicate that interactions with these groups are poorly described by MM and thus that the mechanical embedding approach fails. However, the results with the ME_{dir} method are significantly better, with the maximum error reduced from ~100 to 20 kJ/mol for set 1 (~30 to 16 kJ/mol for set 2). This indicates that the problem is again related to instabilities in the description of the outer part of the QM system, leading to random errors when the interactions with the MM system are strong. In contrast, when only the direct interactions are considered, the MM description proves to be highly useful and gradually gets more accurate as the distance from the reaction center increases.

Calculations with the Whole Protein. It is conceivable that the calculations with the 446-atom model of [Ni,Fe] hydrogenase will overestimate the electrostatic interactions between the distant charged groups and the active site, although our previous results in a vacuum and a continuum solvent with a dielectric constant of 4 differed by only 14 kJ/mol for the 446-atom model.¹³ Therefore, we have also performed a set of calculations in which the protein and the surrounding solvent (in total 12 178 atoms) were included in all calculations, i.e., a more typical

Table 4. Residue Contributions (in kJ/mol) to the Error in the QM/MM Calculations for Set 1, Using the Charge-Redistribution Scheme Z2^a

#	residue	QM	QM+vdW	ME ₀	ME	ME ₁	ME _{dir}	ME'	ME _{av}	ME _{sca}	EE	QTCP ₀	Oniom ₀	QTCP	Oniom	dist	type	#J
1	Ile 544	3.6	4.8	-41.9	-35.3	-43.3	-13.5	-29.6	-35.7	-12.4	5.6	-38.2	-22.2	-32.0	-21.3	1.8	A	5
2	Arg 476	-30.1	-29.3	18.3	22.1	18.5	-6.3	11.3	14.8	-9.0	-2.1	20.1	24.0	9.0	17.1	2.2	Ch	6
3	Cys 546	-12.0	-12.1	15.2	3.3	16.1	3.8	-7.0	4.1	-9.5	-27.8	16.6	4.7	-8.2	-23.3	1.7	A	6
4	Asp 114	19.9	19.9	2.6	13.1	1.6	6.5	4.3	3.4	12.1	1.7	2.9	13.3	3.8	10.1	3.3	Ch	5
5	Cys 75	-4.5	-3.7	-15.8	-15.7	-16.4	-6.3	-9.3	-12.6	-6.5	4.3	-10.1	-10.0	-12.1	-6.7	1.6	A	5
6	Asp 541	22.5	22.4	16.2	21.1	16.0	13.2	-5.2	5.5	8.6	17.5	12.0	16.8	-2.7	0.1	4.8	Ch	4
7	Cys 72	3.3	3.4	-4.7	-13.8	-3.4	3.9	3.5	-0.6	3.4	-10.7	-6.7	-15.8	-0.2	-9.5	1.6	A	5
8	Gln 69	-5.6	-5.8	-10.0	5.1	-10.8	-1.9	1.3	-4.3	-2.3	12.1	-6.3	8.8	-0.6	14.0	3.2	O	6
9	Cys 543	-1.4	-1.2	8.9	-1.1	9.8	6.4	11.8	10.4	5.3	-17.8	5.1	-4.9	11.8	1.4	1.6	A	8
10	Glu S22	12.1	12.2	22.7	18.1	23.4	4.4	10.7	16.7	11.4	15.6	20.8	16.2	17.6	11.4	5.6	Ch	8
11	His 481	-18.3	-17.0	-11.6	-16.1	-10.9	-8.8	-11.3	-11.5	-14.1	-7.7	-16.2	-20.6	-6.9	-10.7	2.0	O	8
12	His 79	-3.7	-3.9	-13.6	-4.6	-15.6	-20.2	-16.0	-14.8	-10.0	-11.1	-20.7	-11.7	-17.5	-2.2	1.6	A	8
13	Wat	-7.5	-8.0	0.5	0.5	0.2	-1.7	-2.3	-0.9	-5.2	-4.2	-7.5	-7.5	-4.7	-3.1	2.7	O	8
14	Pro 542	0.0	-0.1	0.7	25.8	-0.5	5.9	3.0	1.8	1.5	-6.8	-1.6	23.6	6.1	44.2	3.5	N	8
15	Mg 558	-9.2	-9.3	6.5	27.4	4.9	-12.7	-19.1	-6.3	-14.2	2.1	3.6	24.6	-1.6	-4.9	4.2	Ch	6
16	Arg 428	-7.7	-7.7	4.6	10.2	4.3	-0.3	-25.4	-10.4	-16.5	-15.8	-15.3	-9.7	-23.9	-20.0	7.8	Ch	7
17	Arg 70	0.8	0.8	-12.5	-46.2	-11.1	-2.2	15.3	1.4	8.1	-7.1	-3.9	-37.5	-0.9	-38.0	7.3	Ch	7
18	His 115	-4.5	-4.6	-39.0	-34.7	-40.2	1.8	5.3	-16.8	0.4	-2.3	-2.5	1.8	-2.3	7.2	6.4	Ch	8
19	His 538	-9.9	-9.9	-3.1	-47.7	1.2	-3.4	-6.3	-4.7	-8.1	-1.2	-4.1	-48.7	-6.9	-71.9	8.8	Ch	8
20	Ala 80	1.6	1.7	12.1	31.0	11.6	13.8	3.7	7.9	2.7	-12.9	11.8	30.7	6.0	42.8	2.9	A	8
21	Arg 103	-0.3	-0.3	6.0	16.4	4.9	-6.1	-6.7	-0.4	-3.5	0.2	-4.0	6.4	2.3	20.2	12.0	Ch	8
22	Glu S46	12.7	12.7	0.2	2.7	0.0	7.4	13.3	6.8	13.0	4.1	11.1	13.5	3.6	4.5	8.7	Ch	8
23	Glu S75	8.3	8.3	1.3	-6.2	1.9	3.4	2.3	1.8	5.3	0.8	-0.8	-8.3	1.2	-2.1	7.0	Ch	8
24	Asp 88	-10.4	-10.4	-10.3	-15.4	-9.3	-4.6	-6.0	-8.2	-8.2	2.1	5.1	0.0	4.8	3.8	13.9	Ch	8
25	Arg 85	6.1	6.1	-1.5	7.6	-2.5	0.1	-2.0	-1.7	2.1	-6.5	-13.9	-4.8	-12.4	-7.9	9.6	Ch	8
26	Val 78	3.6	3.8	7.9	-1.9	9.0	-3.6	8.3	8.1	6.0	27.2	8.8	-1.0	10.4	-4.6	4.4	N	8
27	Ala 71	2.8	4.1	-1.6	11.8	-1.8	2.8	-2.4	-2.0	0.8	0.0	-4.4	9.0	-2.4	16.9	4.2	N	7
28	Arg 23	-2.9	-2.9	-10.4	-7.7	-10.5	1.8	-3.4	-6.9	-3.1	-0.3	1.2	4.0	-2.7	-0.2	11.9	Ch	3
29	Asp 126	3.8	3.8	4.3	3.7	4.4	-0.6	1.9	3.1	2.9	0.2	-8.9	-9.5	-0.8	-0.9	14.0	Ch	3
30	Asp 63	-6.7	-6.7	99.3	101.8	99.8	-1.6	-6.3	46.5	-6.5	6.5	18.0	20.5	1.0	-0.3	13.0	Ch	3
31	Gln 69'	4.0	4.0	13.9	13.6	14.8	-2.7	-0.6	6.7	1.7	-0.1	-0.9	-1.2	-5.3	1.9	3.1	O	3
32	His 549	-2.8	-2.8	-66.2	-88.9	-62.4	-1.5	8.0	-29.1	2.6	0.3	-1.3	-24.0	5.9	-11.2	7.4	Ch	2
33	Glu 334	-7.8	-7.8	-34.5	-30.9	-34.7	-4.5	3.1	-15.7	-2.3	-1.2	2.9	6.5	1.7	4.8	9.8	Ch	1
34	Gly 547	-11.3	-10.8	14.0	31.9	9.4	3.5	16.6	15.3	2.9	-7.2	12.1	30.0	18.3	30.3	1.7	A	1
35	Thr S18	0.2	0.2	-1.4	-1.3	-1.4	-2.1	-0.3	-0.9	-0.1	-0.6	-0.6	-0.5	-0.6	-0.6	3.4	O	0
36	Arg 388	-3.9	-3.9	2.1	2.1	2.1	0.2	-2.6	-0.2	-3.2	-0.2	-0.2	-0.3	-0.2	-0.2	6.4	Ch	0
37	Arg 59	1.6	1.6	1.5	1.4	1.5	-2.6	-2.5	-0.5	-0.5	-1.8	-1.8	-1.8	-1.8	-1.8	13.8	Ch	0
38	Asp 89	-0.3	-0.3	-0.8	-0.8	-0.8	4.2	5.5	2.4	2.6	2.7	2.6	2.7	2.7	2.6	13.8	Ch	0
39	Asp 60	2.3	2.3	4.6	4.6	4.6	6.6	3.6	4.1	2.9	3.1	3.1	3.1	3.1	3.0	16.3	Ch	0
40	Glu 431	3.1	3.1	-2.4	-2.4	-2.4	-0.7	-1.2	-1.8	0.9	-0.4	-0.3	-0.4	-0.4	-0.3	8.8	Ch	0

^aDist is the shortest distance (in Å) of that residue to the smallest (46-atom) QM system. Type is the type of residue, either A, part of the active site; N, a neighbour to the active site; Ch, a charged residue; or O, another residue. The residue number (#) indicates the calculation in which this residue appears in the QM system for the first time. Likewise, the number of junctions (#J) is the number of junction atoms before this residue is moved to the QM system.

QM/MM setup. As before, we studied how the energy of the proton-transfer reaction changed as the QM system was systematically increased from 46 to 446 atoms in 40 steps. Unfortunately, it is no longer evident what energy to use as a reference, because all methods give different results also with the 446-atom QM system (because the MM system is no longer empty; note also that the calculation with the largest QM system now contains 49 junctions to the protein). In order not to bias the results toward any certain method, we decided to use the result obtained for each

method with the 446-atom QM system as the reference (i.e., different references for all different methods). This reference value is 48 kJ/mol for QM and QM+vdW (i.e., without any electrostatics), 24–28 kJ/mol for the ONIOM and QTCP approaches, which we showed above give poor results, but 52–75 kJ/mol for the other methods (average 65 kJ/mol).

The calculations were performed for all 13 different QM/MM variants, but only for set 1 and only for the Z1 charge-redistribution scheme. The results are shown in Table 6. It can be seen that

Table 5. Residue Contributions (in kJ/mol) to the Error in the QM/MM Calculations for the Second Set of Calculations Using the Z2 Charge-Redistribution Scheme, if Applicable^a

#	residue	QM	QM+vdW	ME ₀	ME	ME ₁	ME _{dir}	ME'	ME _{av}	ME _{scal}	EE	QTCP ₀	Oniom ₀	QTCP	Oniom	Dist	# J
8	A	-24.1	-21.6	-9.6	-1.1	-10.5	-12.8	-17.6	-13.6	-19.6	-52.2	4.7	17.8	-29.8	-14.8	1.6	5
11	N	0.5	1.8	-6.9	1.9	-6.7	-4.8	-3.1	-5.0	-0.7	13.0	10.4	13.0	-3.5	9.5	3.5	4
12	Arg 428	-13.5	-13.5	-18.5	-20.4	-18.6	-6.1	-18.4	-18.4	-15.9	-9.9	-29.8	-21.0	-12.8	-14.0	7.8	3
13	Hid 481	-18.4	-17.1	-5.4	-4.6	-5.4	-8.9	-10.4	-7.9	-13.8	-8.8	-13.5	-12.9	-9.3	-9.0	2.0	3
14	Arg 85	5.0	5.0	-6.5	-7.8	-6.5	-1.0	-4.6	-5.5	0.2	-11.9	-11.7	-11.0	-4.4	-4.8	9.6	3
15	Glu S22	13.3	13.3	17.9	20.1	17.9	5.5	4.4	11.2	8.9	7.4	13.8	12.6	9.0	7.8	5.6	3
16	Asp 541	14.8	14.8	3.6	12.5	3.7	8.8	11.1	7.4	13.0	0.0	-2.1	0.1	9.0	20.0	4.8	3
17	Wat	-9.8	-10.3	-6.6	-6.7	-6.6	-4.0	-5.4	-6.0	-7.8	-3.9	9.9	17.8	-3.9	-3.8	2.7	2
18	Glu S46	8.4	8.4	8.1	8.7	8.0	3.1	7.6	7.9	8.0	3.0	2.5	2.5	-0.1	-0.8	8.7	2
19	Hip 115	-12.7	-12.7	-2.8	-4.3	-2.6	-6.4	-5.1	-3.9	-8.9	-1.2	1.7	2.3	0.9	1.5	6.4	2
20	Gln 69	-5.2	-5.4	-0.2	2.7	-0.6	-1.4	0.6	0.2	-2.4	14.6	7.9	6.1	-2.6	-0.3	3.2	2
21	Gln 69'	-0.3	-0.3	-2.4	-3.8	-2.3	-0.4	-2.2	-2.3	-1.3	0.2	-13.3	-10.5	-1.8	-3.5	3.1	4
22	Arg 70	4.7	4.7	2.4	-1.7	2.7	0.3	2.0	2.2	3.4	0.7	-1.7	-3.1	1.5	-2.6	7.3	1
23	Hip 538	-7.4	-7.4	-1.8	-1.8	-1.8	-0.9	-3.7	-2.8	-5.6	-1.5	-0.8	-4.8	-1.6	-1.5	8.8	0
24	Asp 60	1.2	1.2	8.5	8.5	8.5	5.5	7.4	8.0	4.3	6.2	6.2	6.2	6.2	6.2	16.3	0
25	Asp 89	0.3	0.3	1.2	1.2	1.2	4.7	6.9	4.1	3.6	3.6	3.7	3.6	3.7	3.6	13.8	0
26	Asp 88	-1.4	-1.4	3.0	3.0	3.0	4.3	-2.8	0.1	-2.1	0.3	0.2	0.2	0.2	0.3	13.9	0
27	Arg 59	0.5	0.5	-2.0	-2.0	-2.0	-3.7	-2.9	-2.4	-1.2	-0.9	-0.9	-0.9	-0.9	-1.0	13.8	0
28	Mg 558	9.5	9.4	4.5	-4.8	4.6	6.0	1.8	3.2	5.6	1.7	1.9	1.7	1.7	-6.4	4.2	0
29	Glu S75	0.5	0.5	28.0	30.6	27.9	-4.4	-1.3	13.3	-0.4	3.8	3.6	-5.8	3.5	1.5	7.0	1
30	Asp 114	0.7	0.7	-14.1	-13.2	-14.0	-12.7	-0.6	-7.4	0.0	1.3	2.0	4.5	1.2	0.4	3.3	1
31	Asp 63	0.5	0.5	2.5	2.0	2.5	5.6	2.6	2.5	1.5	0.9	1.0	1.9	1.1	3.2	13.0	1
32	Arg 103	-0.3	-0.3	0.6	1.9	0.6	-6.1	-1.1	-0.2	-0.7	-1.6	-1.7	-2.2	-1.8	-3.3	12.0	1
33	Glu 334	-4.5	-4.5	-21.4	-22.5	-21.4	-1.1	7.6	-6.9	1.6	-1.7	-1.4	-0.1	-1.6	-1.6	9.8	1
34	Arg 476	-7.6	-6.8	0.3	-3.4	0.3	16.2	-5.9	-2.8	-6.3	-2.4	-4.3	-3.6	-2.2	-1.2	2.2	1
35	Hid 549	-3.9	-3.9	-0.4	9.3	-0.4	-2.6	1.4	0.5	-1.2	0.7	0.3	-3.5	0.8	10.1	7.4	1
36	Thr S18	-0.8	-0.9	-1.1	-1.1	-1.1	-3.1	-0.3	-0.7	-0.6	-0.7	-0.5	9.1	-0.7	-0.8	3.4	0
37	Arg 23	-2.7	-2.7	-2.0	-2.1	-2.0	2.0	0.8	-0.6	-1.0	-0.2	-0.2	-0.2	-0.2	-0.2	11.9	0
38	Glu 431	0.1	0.1	6.7	6.7	6.7	-3.7	4.2	5.5	2.1	0.3	0.3	0.3	0.3	0.3	8.8	0
39	Arg 388	-0.1	-0.1	-1.2	-1.3	-1.2	4.0	-5.0	-3.1	-2.6	-0.5	-0.5	-0.5	-0.5	-0.5	6.4	0
40	Asp 126	4.5	4.5	-2.4	-2.4	-2.4	6.0	-1.0	-1.7	1.8	0.0	0.0	0.0	0.0	0.0	14.0	0

^a Dist is the shortest distance (in Å) of that residue to the smallest (46-atom) QM system. The first two systems involve the residues of the original QM system (Cys-72, Cys-75, His-79, Ala-80, Cys-543, Ile-544, Cys-546, and Gly-547; A) and the neighbouring residues (Ala-71, Val-78, and Pro-542; N). The residue number (#) indicates the calculation in which this residue appears in the QM system for the first time. Likewise, the number of junctions (#J) is the number of junction atoms before this residue is moved to the QM system.

they are quite similar to those obtained with the 446-atom model. In particular, the results are not significantly improved by the inclusion of the surrounding. On the contrary, the MUE is worse for all methods, in particular for ME'. On the other hand, the MSE is closer to zero for five of the methods (ME, ME', ME_{av}, QTCP₀, and ONIOM), and the maximum errors are reduced for seven methods.

ME_{dir} still gives the lowest MUE among the 13 QM/MM methods (16 kJ/mol), but it is only 3 kJ/mol lower than for ME₁. The reason for the poorer results of ME_{dir} in the full-protein calculations is the use of standard Amber charges for the MM system. If instead the QM-based MM charges from the 446-atom system are used, exactly the same result as for set 1 is obtained, but the sum of the MM charges for the whole protein is no longer integer (because the QM-based charges apply to a system truncated by hydrogen atoms). QTCP₀ gives the lowest maximum error (43 kJ/mol), but ME_{dir} gives essentially the same result. ME_{av} gives the lowest MSE, whereas QM+vdW gives the lowest standard deviation and EE the fastest convergence. Thus,

the results with the 12 178-atom model show that the conclusions based on the smaller model are not significantly changed.

CONCLUSIONS

In this paper, we study the accuracy of QM/MM calculations with respect to the size of the QM system. As a test case, we use a simple proton transfer between a first-sphere Cys ligand and a second-sphere His group in a 446-atom model of [Ni,Fe] hydrogenase. As the reference, we use the QM results obtained for the whole model.¹³ We study how the QM/MM results vary when the QM system is enlarged systematically. Many QM/MM variants are available.^{7,8,22–24} We have tested QM-only and 13 different variants of QM/MM, including both standard mechanical and electrostatic embedding, as well as several new variants.

We show that many groups provide sizable (up to over 100 kJ/mol) contributions to the error in the QM/MM energies and that there is a large variation between the results obtained with different sizes of the QM system (cf. Figure 4). Thus, it is not enough to study only a few sizes of the QM system as in previous

Table 6. The Performance of the Various Methods for the QM/MM Calculations Including a 12 178-Atom Model of the Protein and the Surrounding Water^a

	MUE	max	MSE	stdev	n5	n10	n20	reference
QM	20.5	56.1	18.8	18.1	38	32	29	48.3
QM+vdW	19.2	53.7	17.2	17.4	38	32	29	47.9
ME ₀	19.2	52.8	8.9	21.6	39	33	33	67.7
ME	23.2	83.5	-6.1	28.6	39	39	35	52.2
ME ₁	19.6	54.6	9.3	22.0	39	33	31	66.0
ME _{dir}	16.1	43.1	3.1	19.1	38	34	33	67.5
ME'	43.8	83.1	-7.9	50.1	34	34	32	72.4
ME _{av}	27.7	58.4	0.5	31.8	34	33	32	70.1
ME _{scal}	26.2	63.8	4.7	30.6	36	32	32	60.1
EE	26.4	68.3	25.7	19.6	39	26	25	57.4
QTCP ₀	21.0	42.7	2.2	25.0	39	33	32	74.5
ONIOM ₀	27.3	73.0	-12.8	31.2	39	39	33	59.0
QTCP	29.5	82.8	-9.3	35.6	39	39	32	28.3
ONIOM	39.2	103.3	-8.5	48.9	36	34	34	24.4

^a The quality measures are the same as in Table 2, but calculations were performed only with the Z1 charge-redistribution scheme (if applicable) and only for set 1. Reference is the reference value for the proton-transfer reaction energy (in kJ/mol) for the calculation with the 446-atom QM system.

investigations if statistically significant results are to be obtained. There are at least three sources of these errors: the point-charge model of the MM group, the improper polarization of the QM system (the MM system is not polarizable), and the approximations in treating the electrostatics around the junctions. The total QM/MM error is the sum of all of these contributions for all groups in the MM system, most of which partly cancel; therefore, it is hard to reach any general conclusions, unless a large number of calculations are performed, to give proper statistics. This paper is based on 7140 separate QM/MM calculations, using 14 different methods, 12 different charge-redistribution schemes, and two different ways to systematically move 30 or 40 different groups from the MM to the QM system. This is appreciably more than in previous studies.^{12,15,17,19–23} By systematically adding different components and corrections, we can estimate the size of the various terms and the accuracy of the various corrections. We have arrived at several interesting conclusions.

First, we show that various charge-redistribution schemes give similar results for most QM/MM methods. With junctions close to the active site, the Z0 scheme gives poor results, and for the ME', QTCP, and ONIOM methods, problems with the QM charges are observed for the BRCD and BCS schemes, as well as occasionally also for Z1, DZ1, and Z0. The standard EE method seems to be more sensitive to the charge-redistribution scheme than the other methods, but opposing results are obtained if junctions are close to the active site or not (set 1 and set 2). In general, there does not seem to be any consistent gain of using any charge-redistribution scheme, and we tend to recommend the Z1 scheme (i.e., include all point charges, besides that of M1, which is already included in the QM system as the HL atom), because it is the most simple method and there is no risk that any important interactions are omitted, as for the Z2 and Z3 schemes.

The best results in this investigation are obtained with the ME_{dir} method, which uses mechanical embedding but forces the charges used for the QM system to be identical for the reactant

and product states except for a fairly small inner core, even when the QM system itself grows. It gives MUEs of 7 and 6 kJ/mol, MSEs of 1 and 4 kJ/mol, and maximum errors of 28 and 16 kJ/mol for the two sets of calculations, respectively. This is appreciably better than the QM-only calculations, by factors of 2–3 for both MUEs and maximum errors.

The second best method is the ME_{scal} method, which uses the average of the QM+vdW method and the ME' method, in which the charges of the QM system are obtained from a wave function polarized by MM point charges. Unfortunately, it is not physically consistent but only motivated by the frequent observation that QM/MM methods with electrostatic embedding seem to overestimate electrostatic interactions.^{41,49} The ME' method itself, as well as the physically more consistent ME_{av} method, also give better results than QM-only, but with slightly worse MUEs and MSEs than ME_{scal}.

Among the methods with electrostatic embedding, standard EE and QTCP₀ give the best results, the latter especially if junctions are moved away from the active site. However, it is notable that this investigation indicates that mechanical embedding actually gives better results than electrostatic embedding, contrary to the common consensus that electrostatic embedding is a better approximation, because it includes the polarization of the QM system by the MM system.^{7,8} An advantage of mechanical embedding is that it avoids the junction problem for the electrostatic interactions between the QM and MM systems, because this term is calculated with M1 atoms and charges. Thus, it provides an alternative to the QTCP correction in eq 3 and it avoids the risk of overpolarization (and therefore no charge-redistribution is needed). Apparently, the errors introduced by letting the truncated QM system be polarized by a point-charge model are as severe as the omission of the polarization of the QM system, and the errors can only be accurately corrected if the junctions are far from the active site (i.e., QTCP₀ in set 2). The reason why previous investigations have given better results with electrostatic embedding^{12,19,22,23} is that the results depend on the tested system, i.e., the relative importance of polarization and the junction errors, but also that we have in this investigation developed new variants of mechanical embedding (ME_{dir} and ME_{scal}) that give the best results.

Moreover, the results show that the ONIOM correction of the Q–J and J–J interactions within the QM system, using an ESP-charge model of the QM system, does not work properly. On the contrary, it consistently gives worse results than without this correction, both with mechanical embedding (i.e., ME compared to ME₀) and electrostatic embedding (ONIOM or ONIOM₀ compared to EE, QTCP, or QTCP₀). Apparently, the ESP-charge model of the QM system is too poor to allow for such a correction.

Finally, we show that the largest errors in QM/MM come from junctions between QM and MM close to the active site. In the second set of calculations, such junctions are avoided, and from Table 2, it can be seen that the results of all methods are strongly improved. Of course, in protein applications, junctions cannot be fully avoided.

Thus, our results lead to the following practical recommendations for QM/MM studies: First, junctions should always be moved into the backbone of the previous and next residue of each active-site group. In fact, the use of junctions, even far from the active site, should be minimized. Second, one should pay attention to the stability of the QM calculations with respect to the various reaction states calculated. Small perturbations may lead to nonphysical differences in the charge distribution in the

outer regions of the QM system, which although having a small effect on the QM energies can have a large effect on the QM/MM interaction energy because of the close proximity of the MM charges. This problem seems to be important for both mechanical and electrostatic embedding. When using mechanical embedding, the problem is amplified by instabilities in the charge-derivation scheme. On the other hand, we have devised a simple way of reducing this problem by using fixed charges outside a core system (the ME_{dir} method), whereas there seems to be no such quick fix to the electrostatic embedding approach.

It should be noted that the junction problem is not avoided by the ME_{dir} method. On the contrary, this method can be considered as the most basic QM/MM scheme, totally ignoring polarization and treating junctions in a simple *ad hoc* manner. However, any approach to treat polarization and junctions in a consistent way must be compared thoroughly to this basic approximation.

In our previous investigation of the same [Ni,Fe] hydrogenase system, we showed that it was very hard to obtain an accurate estimate of the proton-transfer energy in the protein with a QM-only approach.¹³ For example, even if 40 groups from the surrounding protein were included (over 400 atoms), there was still a 60 kJ/mol difference if these groups were selected as those closest to the active site or if they were selected from those giving the largest energy contributions in a QM/MM calculation. In this paper, we show that the QM/MM approach also has convergence problems. However, once the junctions are moved away from the active-site residues, the results become quite stable, with MUEs and maximum errors down to 5 and 16 kJ/mol, respectively. These results provide an estimate of the expected accuracy of the QM/MM approach. Thus, with wisely selected junctions, the QM/MM approach provides an appreciably faster convergence with respect to the size of the QM system than the QM-only approach.

Finally, we have also checked that the results do not change qualitatively if a 12 178-atom model of the protein and surrounding solvent is included in the calculations. It should be noted that the reference results obtained with the largest QM system should not be considered as an accurate estimate of the true energy for this proton transfer in [Ni,Fe] hydrogenase. For such an estimate, better methods (e.g., including dispersion) and basis sets should be used, and dynamical effects need to be included. In a future publication, we will discuss how accurate estimates of protein reactions may be obtained by combining QM/MM calculations with large QM calculations.

■ ASSOCIATED CONTENT

Supporting Information. The chemical structure of groups directly connected to the active site, as well as the results of all methods and charge-redistribution schemes for sets 1 and 2. This information is available free of charge via the Internet at <http://pubs.acs.org/>.

■ AUTHOR INFORMATION

Corresponding Author

*Tel.: +46-46 2224502. Fax: +46-46 2228648. E-mail: Ulf.Ryde@teokem.lu.se.

■ ACKNOWLEDGMENT

This investigation has been supported by grants from the Swedish research council and from the Research School in

Pharmaceutical Science. It has also been supported by computer resources of Lunarc at Lund University and HPC2N at Umeå University.

■ REFERENCES

- (1) Siegbahn, P. E. M.; Blomberg, M. R. A. *Annu. Rev. Phys. Chem.* **1999**, *50*, 221–249.
- (2) Siegbahn, P. E. M.; Blomberg, M. R. A. *Chem. Rev.* **2000**, *100*, 421–438.
- (3) Ramos, M. J.; Fernandes, P. A. *Acc. Chem. Res.* **2008**, *41*, 689–698.
- (4) Lonsdale, R.; Ranaghan, K. E.; Mulholland, A. J. *Chem. Commun.* **2010**, *46*, 2354–2372.
- (5) Siegbahn, P. E. M.; Borowski, T. *Acc. Chem. Res.* **2006**, *39*, 729–738.
- (6) Himo, F. *Theor. Chim. Acta* **2006**, *116*, 232–240.
- (7) Lin, H.; Truhlar, D. G. *Theor. Chem. Acc.* **2007**, *117*, 185–199.
- (8) Senn, H. M.; Thiel, W. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198–1229.
- (9) Hu, H.; Yang, W. *Annu. Rev. Phys. Chem.* **2008**, *59*, 573–601.
- (10) Gao, J. L.; Truhlar, D. G. *Annu. Rev. Phys. Chem.* **2002**, *53*, 467–505.
- (11) Rod, T. H.; Ryde, U. *J. Chem. Theory Comput.* **2005**, *1*, 1240–1251.
- (12) Sumowski, C. V.; Ochsenfeld, C. *J. Phys. Chem. A* **2009**, *113*, 11734–11741.
- (13) Hu, L.; Eliasson, J.; Heimdal, J.; Ryde, U. *J. Phys. Chem. A* **2009**, *113*, 11793–11800.
- (14) Kaukonen, M.; Söderhjelm, P.; Heimdal, J.; Ryde, U. *J. Chem. Theory Comput.* **2008**, *4*, 985–1001.
- (15) Solt, I.; Kulhanek, P.; Simon, I.; Winfield, S.; Payne, M. C.; Csanyi, G.; Fuxreiter, M. *J. Phys. Chem. B* **2009**, *113*, 5728–5735.
- (16) Reuter, N.; Dejaegere, A.; Maigret, B.; Karplus, M. *J. Phys. Chem. A* **2000**, *104*, 1720–1735.
- (17) Amara, P.; Field, M. J. *Theor. Chem. Acc.* **2003**, *109*, 43–52.
- (18) Rodriguez, A.; Oliva, C.; Gonzalez, M.; van der Kamp, M.; Mulholland, A. J. *J. Phys. Chem. B* **2007**, *111*, 12909–12915.
- (19) Wang, B.; Truhlar, D. G. *J. Chem. Theory Comput.* **2010**, *6*, 359–369.
- (20) Lennartz, C.; Schäfer, A.; Terstegen, F.; Thiel, W. *J. Phys. Chem. B* **2002**, *106*, 1758–1767.
- (21) Das, D.; Eurenus, K. P.; Billings, E. M.; Sherwood, P.; Chatfield, D. C.; Hodosecek, M.; Brooks, B. R. *J. Chem. Phys.* **2002**, *117*, 10534–10547.
- (22) Lin, H.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 3991–4004.
- (23) König, P. H.; Hoffman, M.; Trauenheim, T.; Cui, Q. *J. Phys. Chem. B* **2005**, *109*, 9082–9095.
- (24) Vreven, T.; Byun, K. S.; Komáromi, I.; Dapprich, S.; Montgomery, J. A.; Morokuma, K.; Frisch, M. J. *J. Chem. Theory Comput.* **2006**, *2*, 815–826.
- (25) Ryde, U. *Protein Sci.* **1995**, *4*, 1124–1132.
- (26) Ryde, U. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 153–164.
- (27) Ryde, U.; Olsson, M. H. M. *Int. J. Quant. Chem.* **2001**, *81*, 335–347.
- (28) TURBOMOLE, V6.1 2009; University of Karlsruhe and Forschungszentrum Karlsruhe GmbH: Karlsruhe, Germany, 1989–2007. Available from <http://www.turbomole.com> (accessed Jan 2011).
- (29) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, T. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Pearlman, D. A.; Crowley, M.; Walker, R. C.; Zhang, W.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Wong, K. F.; Paesani, F.; Wu, X.; Brozell, S.; Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Mathews, D. H.; Schafmeister, C.; Ross, W. S.; Kollman, P. A. *AMBER 9*; University of California: San Francisco, CA, 2006.
- (30) Maseras, F.; Morokuma, K. *J. Comput. Chem.* **1995**, *16*, 1170–1179.
- (31) Antes, I.; Thiel, W. *J. Phys. Chem. A* **1999**, *103*, 9290–9295.

- (32) Zhang, Y.; Lee, T.-S.; Yang, W. J. *Chem. Phys.* **1999**, *110*, 46–54.
- (33) DiLabio, G. A.; Hurley, M. M.; Christiansen, P. A. J. *Chem. Phys.* **2002**, *116*, 9578–9584.
- (34) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.
- (35) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
- (36) Sherwood, P.; de Vries, A. H.; Collins, S. J.; Greatbanks, S. P.; Burton, N. A.; Vincent, M. A.; Hillier, I. H. *Faraday Discuss.* **1997**, *106*, 79–92.
- (37) Sherwood, P.; de Vries, A. H.; Guest, M. F.; Schreckenbach, G.; Catlow, C. R. A.; French, S. A.; Sokol, A. A.; Bromley, S. T.; Thiel, W.; Turner, A. J.; Billeter, S.; Terstegen, F.; Thiel, S.; Kendrick, J.; Rogers, S. C.; Casci, J.; Watson, M.; King, F.; Karlsen, E.; Sjøvoll, M.; Fahmi, A.; Schäfer, A.; Lennartz, C. *THEOCHEM* **2003**, *632*, 1–28.
- (38) 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.
- (39) Besler, B. H.; Merz, K. M.; Kollman, P. A. J. *Comput. Chem.* **1990**, *11*, 431–439.
- (40) Sigfridsson, E.; Ryde, U. *J. Comput. Chem.* **1998**, *19*, 377–395.
- (41) Källrot, N.; Nilsson, K.; Rasmussen, T.; Ryde, U. *Int. J. Quant. Chem.* **2005**, *102*, 520–541.
- (42) Söderhjelm, P.; Ryde, U. *J. Mol. Struct. THEOCHEM* **2006**, *770*, 199–219.
- (43) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.
- (44) Perdew, J. P. *Phys. Rev. B* **1986**, *33*, 8822–8824.
- (45) Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
- (46) Eichkorn, K.; Treutler, O.; Öhm, H.; Häser, M.; Ahlrichs, R. *Chem. Phys. Lett.* **1995**, *240*, 283–290.
- (47) Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. *Theor. Chem. Acc.* **1997**, *97*, 119–124.
- (48) Wang, J.; Cieplak, P.; Kollman, P. A. J. *Comput. Chem.* **2000**, *21*, 1049–1074.
- (49) Tu, Y.; Laaksonen, A. *J. Chem. Phys.* **1999**, *111*, 7519–7525.
- (50) Freindorf, M.; Gao, J. *J. Comput. Chem.* **1996**, *17*, 386–395.
- (51) Murphy, R. B.; Philipp, D. M.; Friesner, R. A. *J. Comput. Chem.* **2000**, *21*, 1442–1457.
- (52) Freindorf, M.; Shao, Y.; Furlani, T. R.; Kong, J. *J. Comput. Chem.* **2005**, *26*, 1270–1278.

On-the-Fly Identification of Conformational Substates from Molecular Dynamics Simulations

Arvind Ramanathan,^{†,§} Ji Oh Yoo,[‡] and Christopher J. Langmead^{*,†,‡}

[†]Lane Center for Computational Biology, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, United States

[‡]Computer Science Department, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, United States

 Supporting Information

ABSTRACT: We recently introduced a new method for discovering, characterizing, and monitoring spatiotemporal patterns in the conformational fluctuations in molecular dynamics simulation data (*J. Comput. Biol.* **2010**, *17* (3), 309–324). Significantly, our method, called Dynamic Tensor Analysis (DTA), can be performed as the simulation is progressing. It is therefore well-suited to analyzing long timescale simulations, which are critical for studying biologically relevant motions but may be too large for traditional analysis methods. In this paper, we demonstrate that the patterns discovered by DTA often correspond to functionally important conformational substates. In particular, we apply DTA to a 150 ns simulation of ubiquitin and discover patterns that provide unique insights into ubiquitin's ability to bind multiple substrates. Moreover, we take advantage of DTA's ability to identify patterns on different timescales and investigate how fast positional fluctuations may modulate slower, large-scale motions in functionally important regions. Our findings here suggest that DTA is well-suited to organizing, visualizing, and analyzing very large trajectories and discovering conformational substates.

INTRODUCTION

Proteins are intrinsically dynamic and exist in an ensemble of interconverting conformations. This ensemble can be partitioned into subsets of conformations, called *conformational substates*,^{1,2} with similar structures and internal energies. The study of these conformational substates and their relevance to biological function remains an active area of research.^{3–5} Multiple experimental techniques have been used to demonstrate the presence of conformational substates in proteins, including nuclear magnetic resonance (NMR),⁶ neutron spectroscopy,^{7,8} and X-ray crystallography.⁵ These conformational substates, in turn, have provided valuable insights into biological function, such as enzyme catalysis^{5,9,10} and protein folding.¹¹ However, experimental techniques are presently incapable of resolving many of the structural and dynamical details of individual conformational substates, or explaining how proteins transition between them.

Theoretical and computational modeling can effectively enhance our understanding about conformational substates by providing detailed atomistic information about their structural and dynamical features.³ Molecular dynamics (MD) and Monte Carlo (MC) simulations are perhaps the most commonly used computational techniques for characterizing conformational dynamics in proteins.¹² Recent advances in hardware and software (e.g., refs 13–19) have dramatically decreased the cost of MD and MC simulations, and it is now possible to investigate conformational dynamics on microsecond and millisecond timescales. Such long simulations are well-suited to identifying and studying the conformational substates relevant to biological function. At the same time, the concomitant increase in the size of the resulting data sets make them more challenging to analyze and interpret. Thus, there is a need for new, automated methods that help biologists identify and characterize conformational substates, and the transitions between them.

Several techniques have been developed in order to identify conformational substates in MD data. These techniques are

usually performed in an *offline* fashion (i.e., after all of the data are collected) and produce low-dimensional models of the data. Principal Component Analysis (PCA)^{20–22}-based analyses, for example, have been used to describe conformational substates and their relevance in both folding pathways²³ and ligand/substrate binding.²⁴ More recently, Lange and Grubmüller introduced full correlation analysis,²⁵ which can capture both linear and nonlinear correlated motions from MD simulations. Another approach to characterizing conformational substates is the use of Markov State Models (MSMs), which partitions the conformational space sampled by the simulation into kinetically accessible substates.²⁶ Note that while each of these approaches detect and characterize conformational substates, they are performed only after simulations have completed. This is important because any analysis algorithm that runs in time polynomial in the number of simulation frames (e.g., clustering^{27–29}) will face a serious computational challenge when presented with long timescale simulation data.

To overcome the aforementioned limitations, we have recently introduced an *online* algorithm to monitor and characterize collective distance fluctuations in protein simulations as they are progressing.^{30,31} This algorithm, which performs *dynamic tensor analysis* (DTA),³² represents the MD simulation trajectory as a time-evolving stream of multidimensional tensors. We have previously shown^{30,31} that DTA can (a) identify constrained and flexible regions in a protein, (b) characterize the conformational couplings that exist between different parts of a protein, and (c) detect time points where collective behavior may have significantly changed. Our method is also flexible enough to allow the end-user to track specific structural features such as hydrogen bonds or hydrophobic interactions as they vary over time and to detect collective behavior as simulations are progressing.

Received: September 17, 2010

Published: February 10, 2011

In this paper, we will demonstrate the ability of DTA to identify and characterize conformational substates in an online fashion by applying it to a 150 ns simulation of protein ubiquitin. We compare and contrast the conformational substates discovered from our technique to those identified using simple metrics (e.g., root-mean squared deviations [RMSD]) and offline PCA-based methods. Dynamic tensor analysis reveals the presence of several well-defined conformational substates that are not directly evident from RMSD-based metrics. The conformational substates we identify are distinct in their collective fluctuations and directly relevant to substrate binding. Finally, we will demonstrate that it is possible to analyze the conformational landscape on different timescales using DTA. Our multiscale analysis of the ubiquitin trajectory provides unique insights into how the motions of the binding regions may be modulated to achieve optimal binding conformation. These experiments demonstrate the utility of DTA as an approach for identifying and characterizing conformational substates from molecular dynamics simulation data.

METHODS

Molecular Dynamics Simulations. Detailed molecular dynamics simulations were performed on human ubiquitin (PDB id: 1UBQ). The initial crystal structure was processed using the Maestro software (Schrodinger Inc.), and the OPLS/AA force-field^{33,34} was used for simulations. The protonation state of each residue was determined at pH 7.0, and missing protons were added. The structure was then placed in a pre-equilibrated rectangular box of water, parametrized using SPC,³⁵ such that the distance between the protein and the boundaries of the box was at least 10 Å. The final box size was 51.3 Å × 51.3 Å × 51.3 Å. Prior to equilibration, the solvent and proteins were energy minimized using both steepest descent (50 steps) and a conjugate gradient until the overall root-mean-square (RMS) of the gradients was less than 0.25 kcal/mol/Å. The system was then equilibrated using a standard protocol involving multiple steps of energy minimization and small MD runs to allow the solvent molecules and then the solute atoms to relax. Temperature ramps were used to gradually bring the system to 300 K. Next, an NPT (constant number of particles N; constant pressure P; constant temperature T ensemble) simulation at 300 K was performed to make sure that the system was stable. A small simulation run for 1.2 ns using an NVE (constant number of particles N; constant volume V; constant energy E) ensemble was then performed to allow the system to fully equilibrate.

The production run was performed using Desmond,³⁶ under an NVE ensemble with periodic boundary conditions. The RESPA integrator was used for solving Newton's laws of motion. Hydrogen atoms were constrained via the SHAKE algorithm.³⁷ Long-range electrostatics were computed using the particle mesh ewald (PME)³⁸ technique. The production run lasted a total of 150 ns (excluding the initial 1.2 ns run), with frames being stored every 10 ps. A total of 15 000 snapshots were stored and used for further analysis.

Dynamic Tensor Analysis for Protein Simulations. Recently, our group introduced a novel online analysis tool for mining spatiotemporal patterns from MD simulations.^{30,31} Our technique encodes the MD trajectory as a time-ordered sequence of tensors. Tensors are a generalization of matrices beyond two dimensions and can be used to encode both spatial and temporal dynamics. In contrast, PCA takes as input a covariance matrix,

which is a time-averaged representation of the data. Our method then performs Dynamic Tensor Analysis (DTA) to identify and track spatiotemporal patterns in the data. DTA was first introduced in the context of analyzing streams of router data from computer networks.^{32,39}

A detailed description of the algorithm is provided in Ramanathan et al.³¹ Here, we summarize briefly how DTA works (see flowchart in Figure 1) and then describe how conformational substates are identified. Given a collection of tensors $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_T$, each of dimension $n_1 \times n_2 \times \dots \times n_M$, DTA will find orthogonal matrices \mathbf{U}_i , one for each dimension, that minimizes the *dynamical deviation*, η , which is defined as follows:

$$\eta = \sum_{t=0}^T \left\| \mathbf{X}_t - \mathbf{X}_t \prod_{i=1}^M \times_i (\mathbf{U}_i \mathbf{U}_i^T) \right\|_F^2 \quad (1)$$

Here, $\|\mathbf{X}\|_F^2$ is the square of the *Frobenius norm* of tensor \mathbf{X} , which is defined as

$$\|\mathbf{X}\|_F^2 = \sum_{i_1=1}^{n_1} \dots \sum_{i_M=1}^{n_M} \mathbf{X}(i_1, \dots, i_M)^2 \quad (2)$$

and is equivalent to the sum of the inner product operation in matrices.

Informally, eq 1 is the difference between the actual data, \mathbf{X}_t , and the approximation of \mathbf{X}_t in the space spanned by orthogonal matrices \mathbf{U}_i , denoted by $\mathbf{X}_t \prod_{i=1}^M \times_i (\mathbf{U}_i \mathbf{U}_i^T)$. In computer science and machine learning literature, η is often referred to as the error of reconstruction.⁴⁰ Here, $\mathbf{Y}_t = \mathbf{X}_t \prod_{i=1}^M \times_i \mathbf{U}_i$ is called the *core tensor*. The tensor–matrix multiplication operator, $\mathbf{X}_t \prod_{i=1}^M \times_i \mathbf{U}_i$, is defined as

$$\mathbf{X} \prod_{i=1}^M \times_i \mathbf{U}_i = \mathbf{X} \times_1 \mathbf{U}_1 \dots \times_M \mathbf{U}_M \quad (3)$$

As illustrated in the flow-chart (Figure 1), DTA takes as input (i) the new incoming tensor \mathbf{X}_t such that $1 \leq t \leq T$, (ii) the eigenvalues $\mathbf{S}_i^{(t-1)}|_{i=1}^M$, and (iii) the eigenvectors $\mathbf{U}_i^{(t-1)}|_{i=1}^M$ computed from the preceding call to DTA on tensor \mathbf{X}_{t-1} . If there are no previous eigenvalues/eigenvectors (i.e., at $t = 0$), then the only input is the first tensor, and the eigenvalues/eigenvectors will be computed for use in subsequent calls to DTA.

The algorithm proceeds by minimizing the variance in every dimension i , for $1 \leq i \leq M$. First, the input tensor is *unfolded* (or *matricized*) along the selected dimension d . Given $\mathbf{X}_t \in \mathbf{R}^{n_1 \times \dots \times n_M}$, unfolding in dimension d involves constructing the $(\prod_{i \neq d} n_i) \times n_d$ matrix $\mathbf{X}_{(d)}$ such that each row is a vector in \mathbf{R}^{n_d} obtained by holding d fixed and varying the remaining indices. Next, the variance matrix associated with dimension d from the previous call to DTA, $\mathbf{C}_d^{(t-1)}$, is reconstructed using the eigenvectors and eigenvalues from the preceding tensor, \mathbf{X}_{t-1} . The variance of the unfolded incoming tensor is, by definition, $\mathbf{X}_{(d)}^T \mathbf{X}_{(d)}$. At every time step, the variance estimates are updated according to the rule

$$\mathbf{C}_d^{(t)} \leftarrow \lambda \mathbf{C}_d^{(t-1)} + \mathbf{X}_{(d)}^T \mathbf{X}_{(d)} \quad (4)$$

where, λ is called the *memory parameter*. This parameter controls the degree to which previous tensors influence the current estimate of the variance. When $\lambda = 0$, only the present tensor at time t is considered to be relevant, and all of the past tensors are ignored for updating the variance matrix. By restricting our attention to $\lambda = 0$, we will be capturing an instantaneous description of the landscape.

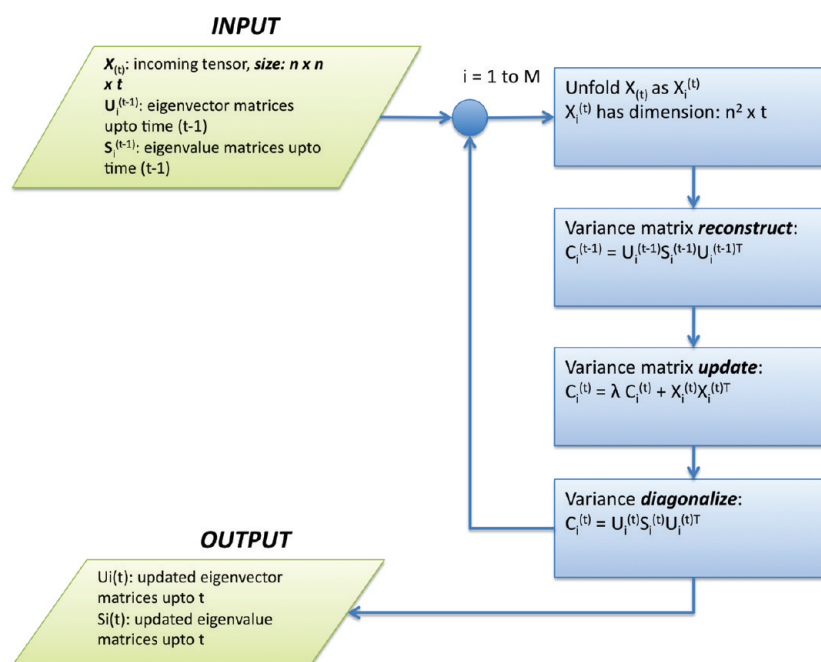


Figure 1. Schematic of dynamic tensor analysis used to capture spatiotemporal correlations from proteins simulations.

When $\lambda = 1$, all of the previous tensors are considered relevant and are used to estimate the variance matrix. Finally, DTA diagonalizes the variance matrix, resulting in an updated set of eigenvalues and eigenvectors that capture the dynamical behavior observed in the simulations observed thus far.

Identification of Conformational Substates Using Dynamical Deviations (η). Tensors provide a convenient means for encoding the collective dynamics observed in *windows* over the data. For example, in our experiments, we constructed tensors encoding the dynamics of the pairwise distances between the C^α atoms over windows of w sequential snapshots. Here, w is a parameter set by the user which can be adjusted to analyze different timescales. The resulting tensors thus had dimensions $n \times n \times w$ where n is the number of residues.

DTA tracks the evolution of the collective dynamics by updating the various covariance matrices C_d according to eq 4. These covariance matrices are then used to update the estimates of the orthogonal matrices U_i , which reveal underlying patterns within the data. We have demonstrated previously^{30,31} that it is possible to gain insights into dynamically coupled regions by clustering the U_i matrices.

The magnitude of any changes in the collective dynamics can be quantified by calculating the dynamical deviations η according to eq 1. Intuitively, any significant increase in η indicates that the collective motions have changed substantially. Such changes may be due to a transition between two different conformational substates. To detect such transitions, we monitor the empirical mean and standard deviation of η as the simulation is running. The instantaneous dynamical deviation threshold, η_{it} is defined as follows:

$$\eta_t \geq \text{mean}(e_i|_{i=1}^t) + \alpha \times \text{std}(e_i|_{i=1}^t) \quad (5)$$

where $\eta_i|_{i=1}^t$ refers to the dynamical deviation up to time t , and α is an arbitrary positive constant. In our experiments, we set α to 2 (i.e., the second standard deviation). Thus, eq 5 can be used to automatically segment the MD trajectory into *dynamical segments*. As will be shown in the Results section, these segments

correspond to different conformational substates characterized by different collective fluctuations. That is, spikes in the dynamical deviation are associated with the transition between conformational substates.

Principal Component Analysis (PCA). *Comparing Offline PCA with Online DTA.* We performed PCA on the C^α distance covariance matrix,⁴¹ \mathbf{D} , defined as

$$D_{ij} = \langle (d_i - \langle d_i \rangle) (d_j - \langle d_j \rangle) \rangle = \mathbf{V} \Lambda_D \mathbf{V}^T \quad (6)$$

where d_i and d_j represent the pairwise distances of C^α atoms. The quantities within $\langle \dots \rangle$ are average distances. The distance covariance matrix \mathbf{D} was then diagonalized to obtain a set of eigenvectors \mathbf{V} and eigenvalues λ_D . Modes were sorted according to their amplitudes in λ_D .

Note that \mathbf{D} is a $m \times m$ matrix, where $m = n(n-1)/2$ and n is the number of C^α atoms in the protein. DTA, in contrast, models distance fluctuations using $n \times n$ matrices. In order to compare the results between PCA and DTA, therefore, it is necessary to construct a reduced representation of each eigenvector \mathbf{v}_i . To do this, we used the procedure described in Abseher and Nilges.⁴¹ Here, the eigenvectors \mathbf{v}_i are first mapped to a symmetric rank 2 matrix and then reduced to an n dimensional vector by summing the squares of the entries along a row:

$$v_i^{\text{red}} = \sum_{k=1}^N (v_{ki})^2 \quad (7)$$

This procedure allows one to accumulate the eigenvector components corresponding to distances from a common C^α atom. The resulting vectors are normalized and then used to compare the distance fluctuations measured by DTA and PCA.

Comparing Collective Fluctuations between Substates. PCA was also used to compare the dynamic segments obtained via DTA and RMSD. To do this, we performed PCA on the C^α atoms of the ensemble of structures in each segment. After removing the translational and rotational motions, the covariance matrix

C^{PCA} was built using

$$C_{ij}^{\text{PCA}} = (\mathbf{x}_i - \langle \mathbf{x}_i \rangle)(\mathbf{x}_j - \langle \mathbf{x}_j \rangle) = \mathbf{U}_{\text{PCA}} \mathbf{\Lambda} \mathbf{U}_{\text{PCA}}^T \quad (8)$$

where \mathbf{x}_i and \mathbf{x}_j represent the positions of C^α atoms of residues i and j , respectively, and the $\langle \dots \rangle$ is the average positional deviation. \mathbf{U}_{PCA} represents the eigenvectors, and $\mathbf{\Lambda}$ represents the eigenvalues (amplitudes of fluctuations) determined via PCA.

We measured the difference in the subspaces spanned by the segments by computing the normalized overlap⁴² between the top 10 eigenvectors. The normalized overlap s between two subspaces A and B is defined as

$$s(A, B) = \frac{\sqrt{\text{tr}[(A^{1/2} - B^{1/2})^2]}}{\sqrt{\text{tr}(A) + \text{tr}(B)}} \quad (9)$$

The value of s can vary from a minimum of zero (no overlap) to a maximum of 1 (identical subspace). We note that the overlap was only used to compare subspaces, and not to test for convergence because short timescale windows are not expected to have converged.

We also compared the subspaces by computing the inner products of the respective eigenvectors. Here, we examined the top 10 eigenvectors (as in previous work⁴³), which account for more than 70% of the overall variance. The inner product between eigenvectors measures the similarity between the direction of the large-scale fluctuations.

RESULTS

In previous work, we demonstrated that MD simulations of ubiquitin can accurately capture its behavior on microsecond timescales and that the motions revealed using quasi-harmonic analysis are functionally relevant.⁴³ In this section, we will summarize how DTA can be applied to characterize conformational substates in a protein simulation as it is progressing. All experiments were performed on tensors tracking the pairwise distances between C^α atoms over varying window sizes, as described below.

Comparison of Ubiquitin Dynamics. We first demonstrate that the Desmond simulation used in this paper sufficiently captures the inherent dynamics of ubiquitin by comparing it to previous experimental and computational results. We note that we carried out the simulation on the entire protein (residues 1–76). However, for the purposes of analysis, we have used residues 2–70. It is known that residues 1 and 71–76 are quite flexible, undergoing large scale fluctuations, and this may affect the interpretation of our results.

Figure 2 shows the root-mean squared fluctuations (RMSF) determined for C^α atoms of ubiquitin from residues 2 to 70. For comparison, these RMSFs are compared to the same quantities obtained from (a) a previously reported 0.5 μs simulation of ubiquitin obtained from multiple initial structures using AMBER as described in ref 43, (b) an NMR ensemble determined on the microsecond timescale (PDB code: 2K39),⁴⁴ and (c) 44 crystal structures obtained from the PDB. Casual inspection reveals that the four curves share significant overlap in terms of the flexible/constrained regions in ubiquitin. The overall correlation between the Desmond simulation and NMR ensemble, for example, is 0.8, whereas the correlation between the AMBER simulation and the NMR ensemble is 0.7. This observation leads us to believe that the Desmond simulations sample the overall conformation space

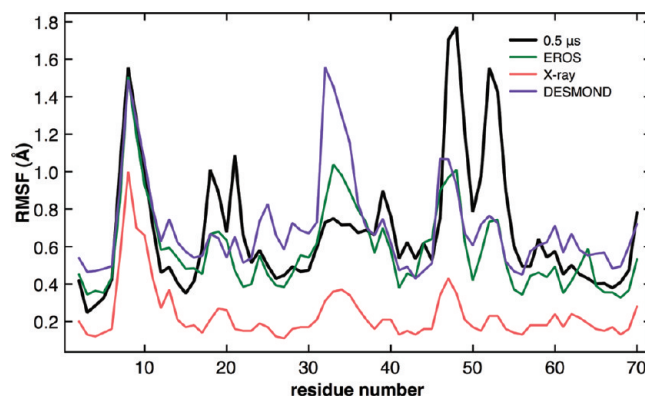


Figure 2. Comparison of Desmond to other ensembles, indicating largely similar fluctuation profiles across different forcefields and experimental ensembles. Our Desmond simulations of 1UBQ (150 ns) are most similar to the EROS (NMR) ensemble. Notice that the peaks in the RMSF curve coincide in Desmond, EROS, and a 0.5 μs ensemble determined by molecular dynamics simulations using the AMBER suite of tools.⁴³ The X-ray ensemble, as determined from 44 crystal structures, also indicates a similar fluctuation profile, albeit with a smaller magnitude of fluctuations. See correlation plots determined from AMBER, EROS, and Desmond in the Supporting Information.

of ubiquitin quite well and can be used for analyzing possible conformational substates.

Using RMSD to Identify Conformational Substates. A simple and straightforward approach to monitoring MD simulations is to compute root-mean squared deviations (RMSD) from an initial structure. RMSD measures the average distance between the backbones of two superimposed structures. We will show that DTA reveals different conformational substates than RMSD. The time-evolution of RMSD over the entire simulation is illustrated in Figure 3A. The average RMSD was computed with a window size of 10 snapshots ($w = 10$), spanning an interval of 0.1 ns. The average RMSD over the course of the 150 ns simulation was 2.2 Å, with a standard deviation of 0.56 Å. A visual inspection of the plot reveals that there are two points along the trajectory where significant structural changes occur. The first change occurs at approximately 42.0 ns and the second at 87.5 ns.

As shown in Table 1, the segments identified using RMSD are diverse in terms of their geometric properties. Closer inspection reveals that ubiquitin remains fairly stable during the first 42 ns of the simulation. However, from 42 ns to about 87.5 ns, the simulation exhibits some large-scale fluctuations involving the α_1 helix and β_1 – β_2 loops (Figure 3B). These motions are important in the context of ubiquitin binding.^{43,44} In the last segment (87.5–150 ns), ubiquitin returns to a more native-like conformation with conformational changes in the β_3 – β_4 loop and α_2 helix (Figure 3B).

RMSD is a measure of average structural deviations and does not necessarily provide insights into collective motions (i.e., whether the motions of different regions are coupled or independent). To compare the collective fluctuations across the three different substates identified using RMSD, we used PCA (described in the Methods section) to compare the subspaces spanned by each of these segments. As shown in Table 2, the pairwise overlap between segments is quite high. Moreover, the average overlap between each segment and the entire trajectory is 0.78, indicating that there is little difference in the collective fluctuations within each segment. This is further illustrated by the high overlap in the inner products of the top 10 eigenvectors (Supporting Information Figure S1).

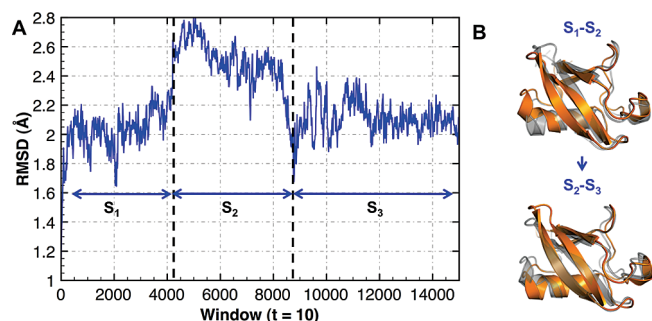


Figure 3. Tracking root-mean squared deviations (RMSD) indicates the presence of (at least) three segments. The RMSD profile from the MD simulation can be segmented into three parts: S_1 (0–42.0 ns), S_2 (42.0 ns + 45.5 ns), and S_3 (87.5 ns + 62.5 ns). Representative structures from each of the segments are compared on the right. In S_1 and S_2 , we observe large-scale changes involving β_1 – β_2 and β_3 – β_4 loops as well as the C-terminal end of the α_1 helix. Minor conformational changes are observed in β_2 – α_1 and β_4 – α_2 loops. In S_2 and S_3 , large-scale conformational changes are again observed in β_1 – β_2 and β_3 – β_4 loops as well as the C-terminal end of α_1 . Note that with this transition, the protein comes back to a conformation that is more or less similar (in the RMSD sense) to the conformations in S_1 .

Table 1. Summary of Segments Determined by Tracking RMSD^a

no.	RMSD (Å)	scaled energy	time duration (ns)
S_1	2.0 ± 0.2	-0.128 ± 0.95	42.0
S_2	2.6 ± 0.4	0.123 ± 1.14	45.5
S_3	2.3 ± 0.3	-0.015 ± 0.91	62.5
total			150.0

^a Each column shows the macroscopic geometric and energetic properties of segments. Scaled energy is defined as the sum of pairwise electrostatic and van der Waals interactions that have been normalized to have unit variance.^{45,46} The time durations represent the length of the respective segments.

Table 2. Summary of Overlaps in Subspaces Spanned by RMSD^a

no.	S_2	S_3	all
S_1	0.79	0.80	0.75
S_2		0.63	0.78
S_3			0.91

^a Each column compares the subspace overlap between the the segments identified (S_1 , S_2 , and S_3). Normalized overlaps (eq 9) are computed as outlined in the Methods section. The final column represents the entire 150 ns trajectory. A summary of the inner products determined from the RMSD-based segmentation is provided in Figure S1 in the Supporting Information.

Next, we examined the scaled internal energy of the protein^{45,46} in each segment. Internal energy, in this context, is defined as the sum of all nonbonded electrostatic and van der Waals interactions.⁴⁵ The internal energy values are normalized to have unit variance for ease of interpretation. The correlation coefficient between the RMSD profile (Figure 3) and the total energy is low ($R = 0.26$), as are the correlations between the RMSD profile and electrostatics ($R = 0.24$) and van der Waals ($R = 0.06$). Additionally, as illustrated in Table 1, the average

scaled energies between the segments are fairly similar and have high standard deviations. We conclude that using RMSD as a metric for segmenting the trajectory into conformational sub-states results in a suboptimal energetic and dynamical description of the conformational landscape. As we will show in the subsequent subsections, DTA allows one to effectively overcome these limitations and identify conformational sub-states that also correspond to jumps in internal energy of the protein as simulations are progressing.

Comparing DTA with PCA. Figure 4 and Table 3 present the results of comparing normalized fluctuation encoded in the top five eigenvectors identified by DTA and PCA. The colored boxes in Figure 4 enclose regions with large distance fluctuations, including flexible loop regions β_1 – β_2 (orange), β_3 – β_4 (green), α_1 (magenta), and β_2 – α_1 (light blue). Notice that PCA and DTA show similar fluctuations. A quantitative comparison of the Spearman's rank correlation coefficients (Table 3) indicates that the similarity between some modes is statistically significant ($p < 0.05$). However, it also reveals some subtle differences between DTA and PCA. For example, a comparison of the top two modes from DTA indicates the greatest variance is associated with the entirety of α_1 , but PCA detects motions only along the C-terminal end of α_1 (see modes A and B in Figure 4). The windowing aspect of DTA allows one to detect such hidden correlations which are not directly evident from PCA techniques. Thus, while PCA can pursue only the extent of spatial fluctuations (because the temporal dimension is averaged out), DTA can reveal correlations that include the temporal dimension. As we will show in the next section, the inclusion of the temporal dimension can also affect how the simulation can be segmented into different conformational sub-states.

The distance space PCA eigenvectors can be used to partition the landscape into conformational sub-states.^{21–23,41} For this purpose, we use the projections from the distance space PCA to identify regions of the trajectory that show large deviations from the average behavior observed. The projections from PCA are computed *offline* using

$$q_i(t) = (\mathbf{d}(t) - \langle \mathbf{d} \rangle) \mathbf{v}_i \quad (10)$$

where $q_i(t)$ represents the projection of conformation at time t , the first quantity on the right-hand side of the equation relates to the deviations in the pairwise distances, and the second term represents the eigenvector determined from eq 6. As shown in Figure 5, the top two eigenvectors partition the ubiquitin simulation into three sub-states (identified by the ellipses shown in the figure). The top two eigenvectors contribute to over 55% of the overall variance in the simulation, and hence the large-scale distance fluctuations observed can be considered a consequence of these two eigenvectors. The top two eigenvectors describe the fluctuations of the α_1 helix and the flexible loops β_1 – β_2 and β_3 – β_4 .

The temporal evolution of the projections from the top two eigenvectors (see Supporting Information Figure S2) shows a large change around 42.0 ns of the simulation, followed by a gradual relaxation that begins at around 87.5 ns. The temporal evolution of the projections therefore closely follows the overall conformational changes depicted by the RMSD plots shown in Figure 3A. Further, as shown in Figure 5, the segments identified via PCA largely follow the RMSD partitioning of the conformational landscape. This is to be expected, since PCA pursues the variance (or extent of fluctuations) and therefore blindly chases the largest deviations observed in the simulation. However, as we have

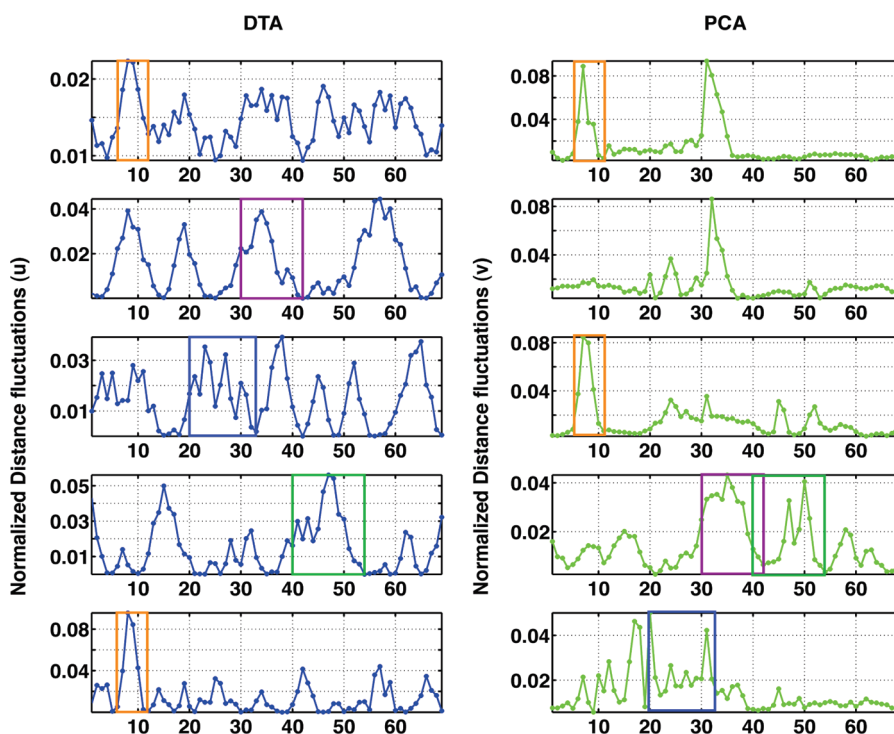


Figure 4. Comparing the online DTA with offline distance space PCA indicates similar fluctuations. Shown here are the normalized distance fluctuations from the top five eigenvectors from the online DTA performed over the entire trajectory and the offline distance space PCA. Note that while individual amplitudes across the top five modes might differ, the same regions in ubiquitin (highlighted in similar colored boxes across the plots) undergo similar fluctuations in both methods. A quantitative comparison between the modes also indicates the same (see Table 3).

Table 3. Similarity in Normalized Distance Fluctuations Determined from DTA (rows) and Distance Space PCA (columns) for the Top Five Eigenvectors from Each Method^a

DTA/PCA	A	B	C	D	E
A	0.247 (4.06e-02)		0.303 (1.17e-02)	0.525 (5.03e-06)	
B	0.420 (3.65e-04)	0.242 (4.49e-02)	0.343 (4.04e-03)	0.255 (3.45e-02)	
C			0.249 (3.96e-02)		
D				0.323 (6.95e-03)	
E				0.247 (4.09e-02)	0.214 (7.71e-02)

^a Only correlation values with p values (within brackets) that are significant (<0.05) are shown here.

recently pointed out,³¹ the changes in these conformations need not necessarily correlate with energetic changes. We conclude that, for our simulations, PCA- and RMSD-based segmentations result in similar descriptions of the landscape.

DTA Segments the Conformational Landscape into Conformational Substates. In this section, we describe how dynamical deviations (η ; described in the Methods section and eq 1) can be used to identify conformational substates as the simulations are running. We will first demonstrate how dynamical deviations η can be used to segment the trajectory by applying a threshold. We will then examine how our interpretation of the landscape changes as we examine different timescales (by increasing w).

The dynamical deviation η quantifies how much the previous window differs from the current window, in terms of its dynamical behavior. Spikes in η can therefore be used to segment the trajectory. For example, there are four obvious segments in Figure 6A, which uses a window size of 0.1 ns ($w = 10$). Each segment can be further partitioned according to η , as illustrated in Figure 6B. This is consistent with the view of a hierarchical conformational landscape.²

The process of segmenting the trajectory can be automated by applying a threshold η_t , as defined in eq 5. The time-evolution of η_t is shown as a red continuous line in Figure 6A. Note that η_t tends to rise for short periods of time and then stabilize as the simulation progresses. The rise and stabilization in η_t can be attributed to two aspects in the collective dynamics of the protein: (a) a period of fast changes in the fluctuations, indicated by the gray shaded regions in Figure 6A, followed by (b) a stable dynamical regime in which fluctuations are much less pronounced.

To compare the collective dynamics in the identified segments, we computed both the normalized overlaps (eq 9) and inner products between the top 10 eigenvectors determined from each of these segments. As shown in Figure 6C, the eigenvalue spectrum shows considerable difference between the segments identified. Table 4, which shows the overlaps between the subspace spanned by the top 10 eigenvectors in each segment, also illustrates that the maximal agreement between the subspaces is only about 0.64, which confirms that the motions between these dynamical segments are different. A second and more direct line

of evidence comes by examining the inner product matrices of each segment (see Supporting Information Figure S3), which

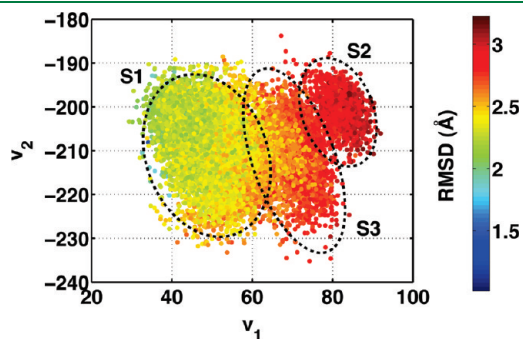


Figure 5. Projections from the top two eigenvectors determined from distance space PCA reveal three conformational substates closely following the RMSD based segmentation. Shown here are the projections of the top two eigenvectors determined from distance space PCA. Each conformation from the simulation (a total of 15 000 conformations) is colored with the individual RMSD values determined with respect to a single reference structure. Note that the segments marked (S_1 , S_2 , and S_3) correspond to the segments identified from Figure 3.

further confirms that the collective motions are quite different in each of these segments.

Now that we have quantified the extent of the dissimilarity in the collective motions between each dynamical segment, we examine whether these segments share conformations that show similarity in their internal energy distributions. Within each dynamical segment identified by DTA (Figure 6), we measured the mean internal energy and standard deviations and summarize the same in Table 5. DTA-defined dynamical segments are better separated in terms of average energy with relatively smaller standard deviations in overall energy. To better illustrate how DTA performs with respect to identifying isoenergetic substates, we plot the internal energy as shown in Figure 7. Note that the transitions (i.e., the peaks in Figure 6) correspond well with changes in the internal energy. Further, the transition between CS_3 and CS_4 shows a significant change in the internal energy values, which is not true of the transitions between the segments identified using RMSD or PCA (the segments identified by RMSD and PCA are labeled as S_1 – S_3 in Figure 7). Thus, in our experiments, DTA segmented the trajectory into conformational substates that exhibit more energetic homogeneity than those identified via RMSD or PCA. Moreover, the DTA substates exhibit significant differences in terms of their collective motions.

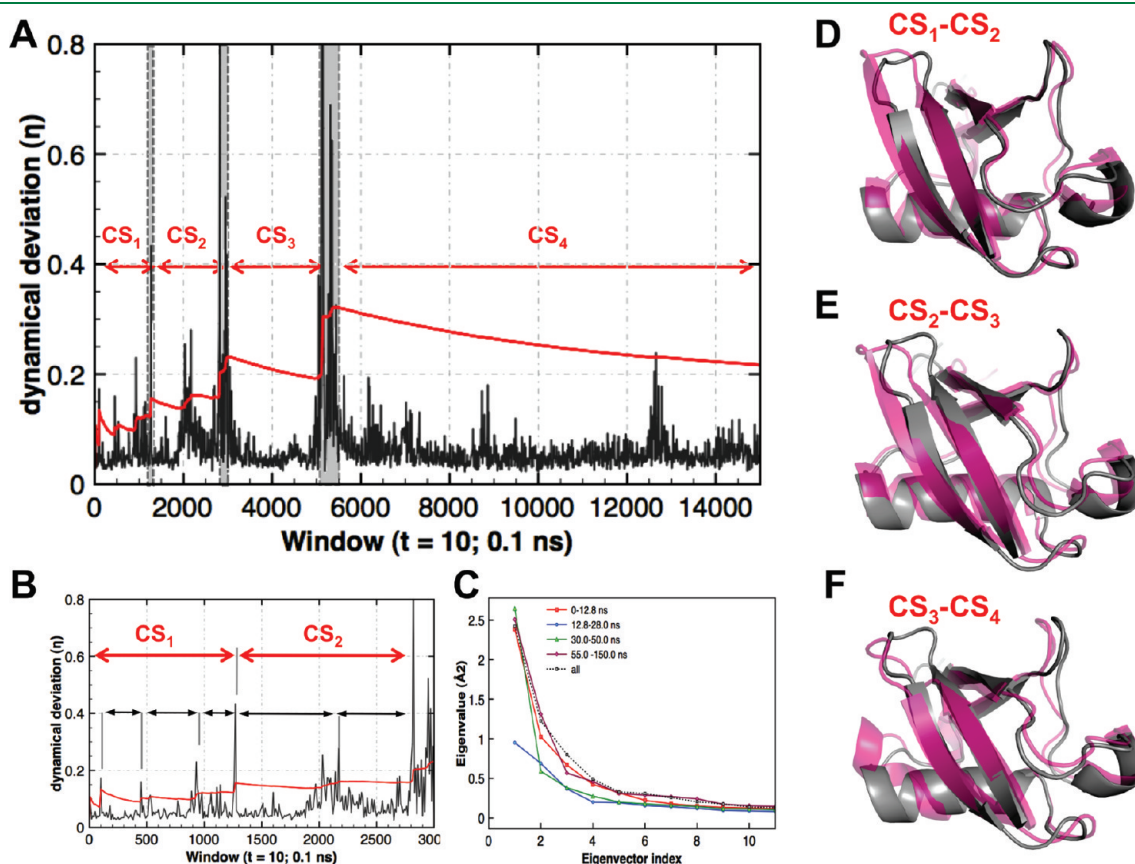


Figure 6. Tracking dynamical deviations (η) indicates the presence of four conformational substates. The η profile indicates the presence of sharp peaks along the simulation. The red line shows the second standard deviation intervals for η . (A) Peaks with significant changes in η indicate substantial changes in collective conformational dynamics, indicating that the protein has jumped into a new substate. In between transitions, there are phases where the collective behavior shows significant changes. This may be indicative of transition states involved in altering the collective behavior of the protein. (B) Each substate is formed by additional substates leading to the hierarchical organization of the landscape. (C) Comparison of the eigenspectrum from each of the dynamical segments. (D–F) Structural changes along the trajectory are summarized, highlighting the significant structural changes involved in the protein.

Table 4. Summary of Overlaps in Subspaces Spanned by DTA at $w = 10$ ($t = 0.1$ ns)^a

no.	$CS_2^{0.1}$	$CS_3^{0.1}$	$CS_4^{0.1}$	all
$CS_1^{0.1}$	0.64	0.50	0.65	0.63
$CS_2^{0.1}$		0.54	0.58	0.60
$CS_3^{0.1}$			0.55	0.63
$CS_4^{0.1}$				0.87

^a Each column compares the subspace overlap between the the segments identified (CS_1 , CS_2 , CS_3 , and CS_4). Normalized overlaps (eq 9) are computed as outlined in the Methods section. The final column represents the entire 150 ns trajectory. A comparison of the inner products determined via PCA for each of the segments is shown in Figure S2 in the Supporting Information.

Table 5. Summary of CSs Determined by DTA^a

no.	scaled energy	time duration (ns)
$CS_1^{0.1}$	-0.082 ± 1.06	$12.8 + 0.1$
$CS_2^{0.1}$	-0.254 ± 0.98	$15.1 + 0.9$
$CS_3^{0.1}$	-0.436 ± 0.99	$22.0 + 4.5$
$CS_4^{0.1}$	0.072 ± 0.94	94.6
total		150.0

^a Each column shows the macroscopic geometric and energetic properties of segments. Scaled energy is computed as outlined in the Methods section.^{45,46} The time durations represent the length of the respective segments, followed by the interval of time indicated by gray shaded regions highlighted in Figure 6.

The distinct substates identified by DTA are related to ubiquitin's function. We considered the structural changes along each of the segments (CS_1 – CS_4). Between CS_1 and CS_2 , the dominant conformational change involved is the rearrangement of the binding loop β_1 – β_2 and a slight conformational change involving both α_1 and α_2 regions of the protein. The rest of the protein in this substate does not show any significant conformational change. In the transition between CS_2 and CS_3 , one can observe the significant structural changes involved in β_1 – β_2 as well as changes in the orientation of the β_5 strand and β_2 – α_1 loop regions of ubiquitin. In the transition between CS_3 and CS_4 , the largest structural change involves the bending of the C-terminal end of α_1 as well as significant rearrangements in β_1 – β_2 and β_3 – β_4 . These conformational changes have a direct implication in binding. As described in previous experimental⁴⁴ and computational studies,⁴³ the conformational changes in each of the conformational substates is unique and related to the movements of β_1 – β_2 and β_3 – β_4 loops, both of which form important interactions with ubiquitin's natural substrates. These changes occur throughout the simulation. However, there is only one segment where significant changes occur in α_1 , implying that this motion may be much slower than the fluctuations associated with β_1 – β_2 and β_3 – β_4 .

The fact that DTA tracks the temporal evolution of the covariance matrices (eq 4) produces qualitatively different segments than those identified via RMSD and PCA. In particular, DTA segments the trajectory based on changes in collective fluctuations, whereas PCA and RMSD segment according to large structural changes. Tables 2 and 4 support this distinction by showing that the average overlap between the various segments is low. We conclude that the inclusion of the temporal dimension provides additional information that is not accessible to PCA- and

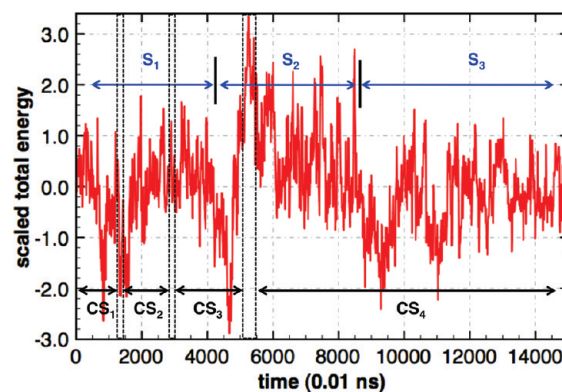


Figure 7. Tracking changes in total internal energy of the protein reveals conformational substates. The total internal energy of the protein computed from Desmond is shown by tracking windows of size $t = 10$. The segments (S_1 – S_3) identified from Figure 3 are shown in blue. Observe that although there is some correspondence with the overall trends in RMSD values (Figure 3), the overall correlation between RMSD and energy values is quite low (0.26). The substates partitioned via DTA (Figure 6) are shown in black below. Note that the DTA-based segmentation captures the transition between CS_3 and CS_4 , whereas the RMSD-based segmentation does not.

RMSD based methods. Further, as we will show in the next subsection, DTA can be used to examine the landscape on multiple timescales by varying the timescale parameter (w).

Effect of Increasing Time Scales on Identifying Conformational Substates. Note that in our analysis so far, we set the timescale parameter to $w = 10$ snapshots (0.1 ns). In this section, we demonstrate how DTA can be used to detect conformational transitions on longer timescales. We note that the timescale parameter is set by the user and that it is possible, in principle, to perform DTA on multiple timescales simultaneously (e.g., using multiple processors). For the remainder of this section, superscripts will be used to identify the timescale associated with each substate according to the timescale. For example, CS_2^1 denotes the second conformational substate on a 1.0 ns timescale. This will facilitate the comparison between substates on different timescales.

We also considered time windows of $w = 100$ (1.0 ns) and $w = 500$ (5.0 ns). The use of three different time windows ($w = 10$, $w = 100$, and $w = 500$) allows us to resolve the landscape on different timescales. The analysis on the 1 ns timescale has three dynamical segments (CS_1^1 – CS_3^1). These segments are diverse, as evidenced by the eigenspectrum (Figure 8) and the low overlaps between the subspaces (Table 6). The analysis on the 5 ns timescale has two major dynamical segments (CS_1^5 – CS_2^5 ; see Figure 9B). Pairwise comparisons of the η values across timescales are shown in Figure 9. Notice that the locations of peaks in the η values occasionally coincide (see shaded rectangles), but there are peaks that occur on only one timescale (see black arrows).

Spatially, the changes in the collective fluctuations in CS_1^1 – CS_3^1 tend to be localized near β_1 – β_2 . On the 1 ns timescale, however, the collective fluctuations in CS_1^1 – CS_3^1 are concentrated in the β_1 – β_2 turn, α_1 , and the β_2 – α_1 loop. On the 5 ns timescale, the collective fluctuations in CS_1^5 – CS_2^5 involve α_1 .

We also examined the changes in collective motions by comparing the covariance matrices (see Figure 10). Here, the covariances on the three timescales were normalized to have unit variance. Note that between 0.1 and 1.0 ns timescales, the differences in the covariance matrices are more global (left panel

in Figure 10A), spanning multiple regions of the protein. A similar observation can also be made for the comparison between 0.1 and 5.0 ns timescales (middle panel of Figure 10A). These changes are concentrated along the functionally relevant regions

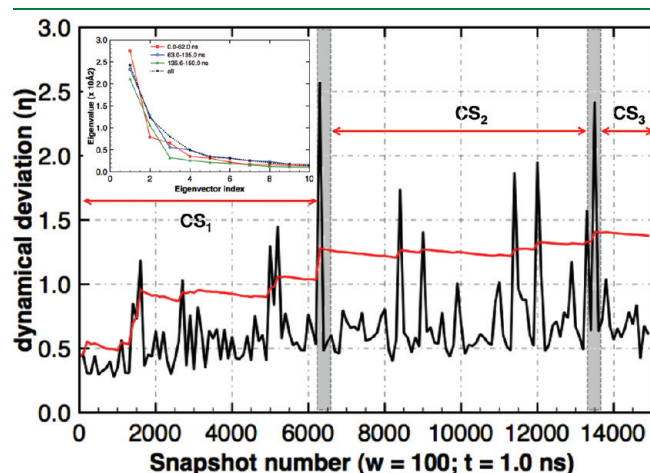


Figure 8. CSs on longer timescales reveal unique dynamical fluctuations in ubiquitin. The plot shows the η as a function of the time window ($w = 100$; 1.0 ns). On longer timescales, we observe larger η and only three segments (CS_1^1 – CS_3^1). Within each segment, we observe smaller spikes, indicating the presence of smaller substates. Gray shaded regions indicate segments where large changes are observed in η , indicating dynamical transition points. The inset shows the comparison of the eigenspectrum determined from the segments here. A comparison of the subspaces is shown in Table 6.

Table 6. Summary of Overlaps in Subspaces Spanned by DTA at $w = 100$; ($t = 1.0$ ns)^a

no.	$CS_2^{1,0}$	$CS_3^{1,0}$	all
$CS_1^{1,0}$	0.64	0.52	0.76
$CS_2^{1,0}$		0.65	0.85
$CS_3^{1,0}$			0.62

^a Each column compares the subspace overlap between the the segments identified ($CS_1^{1,0ns}$, $CS_2^{1,0ns}$ and $CS_3^{1,0ns}$). Normalized overlaps (eq 9) are computed as outlined in the Methods section. The final column represents the entire 150 ns trajectory.

of the protein. In particular, changes in collective fluctuations on 1.0 ns timescales are reduced along β_1 – β_2 with respect to α_1 and β_4 – α_2 regions (Figure 10B). However, collective changes are enhanced across several regions in the protein including β_2 – α_1 , β_3 – β_4 , and α_2 – β_5 , indicating that, on longer timescales, these correlations become more pronounced (Figure 10C). Note that between 0.1 and 5.0 ns timescales, only the correlations become more pronounced—the regions identified to be flexible within ubiquitin are still the same. This emphasizes the inherent flexibility in ubiquitin that is present even on smaller timescales, which is observed on longer timescales, albeit with higher amplitudes. It is also clear from the two plots that one can identify individual residues that undergo changes in their distance fluctuations with respect to the rest of the protein. In this case, we observe Gln40 and Asp22 to undergo changes in their motions.

A comparison of the covariance matrices at 1.0 and 5.0 ns does not yield significant changes. As seen from the right-hand panel of Figure 10A, the color bar shows relatively smaller localized changes in the covariances. This is to be expected since the timescales are roughly on the same order. However, we do observe several localized changes in the protein's motions, notably along the flexible β_1 – β_2 region of the protein. This region tends to undergo fast fluctuations ($O(\text{ps})$) and, hence, is clearly visible in the difference plots. A second localized fluctuation occurs in the α_1 – β_3 region; however, it is of much smaller amplitude. Thus, depending on the timescale at which distance fluctuations are monitored, DTA can provide a succinct and unique resolution of the conformational landscape.

DISCUSSION

DTA Overcomes Limitations of Readily Available Observables from MD Simulations. Popular measures for monitoring MD simulations include RMSD, radius of gyration, kinetic/potential/total energy, velocity, pressure, and temperature. These observables have been traditionally used to monitor the “health” of the simulation and identify events that affect its quality. While these observables are certainly valuable, they do not (and cannot) track how concerted changes to a group of atoms or residues within a protein affect its dynamical behavior as the simulations are progressing. Tracking concerted, collective changes in the dynamical behavior of a protein needs a suitable measure that is sensitive enough not only to capture large

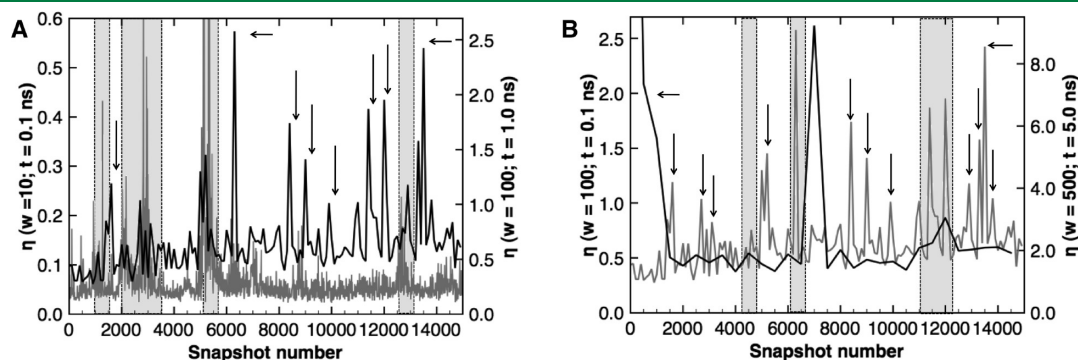


Figure 9. Comparing η on different timescales reveals differences in the resolution of the conformational landscape of ubiquitin. Panel A compares the landscape on timescales of 0.1 and 1.0 ns. Panel B compares the landscape at 1.0 and 5.0 ns resolutions. Note that there are two axes used here. The plots show whether there are overlaps in η on different timescales. Shaded rectangles are used to highlight regions that show close correspondence to changes in collective behavior on both timescales being compared. Black arrows show the time points that are present on the faster timescale but not on the longer timescale.

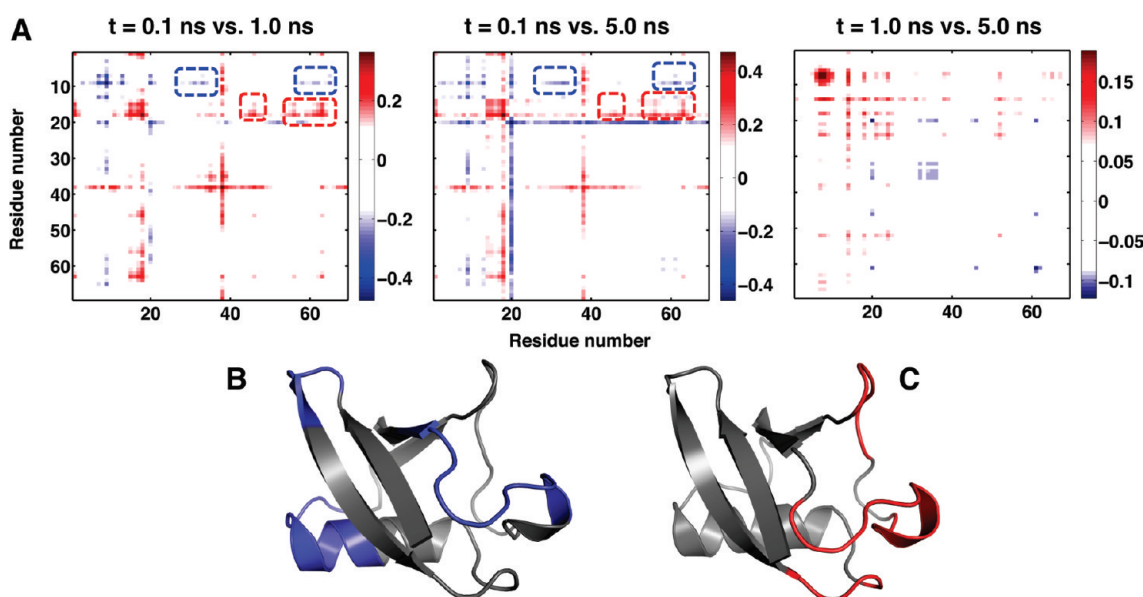


Figure 10. Comparison of covariance of distance fluctuations on longer timescales, indicating change along functionally relevant regions in ubiquitin. (A) The difference in the covariance determined at three different temporal resolutions, namely, $w = 10$ (0.1 ns), $w = 100$ (1.0 ns), and $w = 500$ (5.0 ns). Although the covariance on each timescale implicates similar flexible regions in ubiquitin (see the Supporting Information), the differences reveal subtle yet important dynamical changes. The regions showing significant differences are highlighted using a rounded rectangle in each plot. Notice that there are regions of both increased and decreased covariance. (B) Regions that show a significant decrease in covariances on longer timescales. (C) Regions that show a significant increase in covariances. The regions highlighted are important for ubiquitin's binding motions (see the Results section for more details). Also note that, between the 1.0 and 5.0 ns timescales, the differences in the covariance matrices are highly localized—indicating that the overall resolution of the landscape has not changed very much.

conformational changes but also to characterize subtle, yet functionally relevant changes. In this paper, we have used one such measure, namely, dynamical deviations, η (eq 1), to organize the conformational landscape in terms of substates that share similar collective behaviors and energies.

Recently, a technique called TimeScapes based on coarse-graining MD trajectories using side-chain contacts was implemented to track time-dependent conformational changes in protein simulations.⁴⁷ TimeScapes is able to model conformational shifts along the simulation representing secondary and tertiary structures corresponding to functionally important transitions. Furthermore, it was also able to identify basins and transitions based on an activity measure. The approach is similar in terms of analyzing simulations online. However, unlike DTA, TimeScapes does not allow for tracking collective behavioral changes over simulations. Our approach is complementary and allows one to track any feature (see below) from a MD simulation for changes in collective behavior. In previous work,⁴⁸ we have also demonstrated the ability of DTA to quantify changes in contact maps during MD simulations.

It must be noted that while DTA overcomes the limitations posed by simple observable measures from MD simulations, it is dependent on the representation used in describing collective behavior. For our representation, we have used C^α -distance matrices as a means to monitor changes in collective motions. It is entirely possible to use other distance/geometric measures such as hydrogen bond/hydrophobic interactions,⁴⁹ dihedral/torsion angles,⁵⁰ and energy measures such as electrostatics/van der Waals/potential/kinetic energy^{45,46} to track collective changes in these measures. It is of interest to see how well changes in geometric measures are correlated with changes in energy measures and will be pursued in a future study using a fairly straightforward extension of DTA.⁵¹

The energetic description obtained via DTA, especially from the dynamical segments, is much better than the RMSD based segmentation of the landscape. Although one should not expect to find energetically homogeneous substates directly from DTA, it is at least encouraging to note that there is some energetic similarities within a dynamical segment. Thus, DTA represents a step in the right direction toward understanding the complex spatiotemporal dependencies that might exist within the conformational and energetic landscape.

Spatiotemporal Insights into Conformational Landscape.

A valuable utility of DTA in monitoring collective behavior is that it can capture both spatial and temporal changes in collective behavior over the course of a MD simulation. This has some bearing on our understanding of the complex conformational landscape. The conformational landscape that is sampled by the MD simulation can be divided into phases: a *stable* phase in which the η does not change very much and a *dynamic* phase in which the η shows significant changes in its behavior. The stable phases are indicated by relatively small changes in η , whereas the dynamic phases involve large changes in η . The dynamic phase where η shows an increase involves significant rearrangements in the overall conformation before stabilizing with minor conformational changes, dominated by localized motions in side chains and corresponding changes in C^α positions.

In our simulations of ubiquitin, we observed that at a temporal resolution of 0.1 ns, there are numerous changes that occur rather suddenly over the 150 ns. Though these changes involve the functionally relevant binding regions (β_1 - β_2 , β_3 - β_4 , and C-terminal tip of α_1), in these small time windows, we do not observe any significant correlations between the functionally important regions in ubiquitin. However, on longer timescales (1.0 and 5.0 ns respectively), we find there is a small correlation between β_1 - β_2 and β_3 - β_4 (Figure 10A; observe the correlation

between residues 7 and 9 with 46 and 48). The emergence of such correlations can have some consequence in interpreting functional relevance of these motions in binding. Both $\beta_1-\beta_2$ and $\beta_3-\beta_4$ may need to be precisely positioned in order to form the interactions needed to bind its substrate. While it remains to be seen if such coordinated motions are a prerequisite for interface formation,⁵² we note that the correlations between the binding regions occur only along specific time points in our simulations. Such coordinated motions may also play a role in protein folding pathways where secondary structures might need to interact before sampling native state configurations.⁵³ The ability of DTA to pick up such correlated changes along folding pathways will be studied in the future.

Further, a comparison of DTA with offline PCA reveals the additional insights obtained by using online techniques. The fluctuations (large-scale motions) described by DTA and PCA are quite similar. However, they do differ in terms of the correlated motions that are depicted. On the basis of the time resolution used in DTA, these correlations between different parts of the protein can change. PCA-based techniques operate on time-averaged covariance matrices and, hence, cannot detect subtle changes that may occur over the course of a simulation. DTA, however, includes correlations that arise on different timescales. This provides a unique viewpoint for interpreting the conformational landscape: changes in the correlations between spatially separated parts of the protein on different timescales may provide insights into how these distal couplings arise.

CONCLUSIONS

In this paper, we have demonstrated the utility of dynamic tensor analysis (DTA) to identify and characterize dynamical segments along a MD trajectory that share similar geometric and energetic properties. That these dynamical segments may form a starting point to identify conformational substates has several interesting pointers for future studies. For one, the application of DTA to the study of protein folding pathways to distinguish successful (folded) versus unsuccessful (unfolded) states is already underway. It is also useful to investigate if one can correlate multiple tensor streams including geometric and energetic properties across simulations, which we plan to pursue using an existing approach.⁵¹ We hope that the availability of such tools can be valuable in processing extant data sets, such as those in Dymeomics,⁵⁴ and those that will be produced by Anton.¹³

The ability to identify and relate protein motions at different temporal scales opens up opportunities for characterizing protein landscape in the spirit of previous work by Frauenfelder and co-workers.⁵⁵ Further, it will also provide an integrating platform for combining studies that simultaneously includes both spatial and temporal aspects of the complex conformational landscape, which is indeed thought to be a requirement to fully characterize a protein's conformational landscape. Furthermore, in the context of complex biological functions such as enzyme catalysis,⁵⁶ it is believed that such a holistic description of the landscape can be valuable.

Implementation and Availability. DTA is implemented in both Python⁵⁷ and Matlab.⁵⁸ The Matlab code is available on request from the authors. The DTA implementation in Python is part of a package called pyTensor⁴⁹ and is hosted at <http://code.google.com/p/pytensor/>. The source code and package are available for download for free. The package has been implemented such that it can be easily adapted to read a variety of inputs generated from MD simulations. Processing outputs from MD simulations using

custom-written python scripts are also available upon request. Currently, our custom-written scripts can read distance and energy values from MD simulations (written for a variety of packages including AMBER and Desmond). The DTA package additionally implements a subset of the tensor toolbox,^{59,60} which can be used for further development. The Python version of the code can be integrated into any number of available packages including HiMach⁶¹ and Biskit.⁶²

ASSOCIATED CONTENT

S Supporting Information. Six additional figures and their descriptions are available. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*Phone: (412) 268 7571. Fax: (412) 268 5576. E-mail: cjl@cs.cmu.edu.

Present Addresses

⁵Computational Biology Institute, Computer Science Research, Computer Science and Mathematics

ACKNOWLEDGMENT

The authors thank Dr. Pratul K. Agarwal for providing access to the ubiquitin simulations and critically reading the manuscript and commenting on it.

REFERENCES

- (1) Frauenfelder, H.; Petsko, G. A.; Tsernoglou, D. *Nature* **1979**, *280*, 558–563.
- (2) Frauenfelder, H.; Parak, F.; Young, R. D. *Annu. Rev. Biophys. Chem.* **1988**, *17*, 451–479.
- (3) Henzler-Wildman, K.; Kern, D. *Nature* **2007**, *450*, 964–972.
- (4) Boehr, D. D.; Nussinov, R.; Wright, P. E. *Nat. Chem. Biol.* **2009**, *5*, 789–796.
- (5) Fraser, J.; Clarkson, M.; Degnan, S.; Erion, R.; Kern, D.; Alber, T. *Nature* **2009**, *462*, 669–673.
- (6) Boehr, D. D.; Dyson, H. J.; Wright, P. E. *Science* **2006**, *313*, 1638–1642.
- (7) Zaccai, G. *Science* **2000**, *288*, 1604–1607.
- (8) Fitter, J. *Biophys. J.* **2003**, *84*, 3924–3930.
- (9) Eisenmesser, E. Z.; Bosco, D. A.; Akke, M.; Kern, D. *Science* **2002**, *295*, 1520–1523.
- (10) Eisenmesser, E. Z.; Millet, O.; Labeikovsky, W.; Korzhnev, D.; Bosco, D.; Skalicky, J.; Kay, L.; Kern, D. *Nature* **2005**, *438*, 117–121.
- (11) Balbach, J.; Forge, V.; van Nuland, N. A. J.; Winder, S. L.; Hore, P. J.; Dobson, C. M. *Nat. Struct. Mol. Biol.* **1995**, *2*, 865–870.
- (12) Karplus, M.; McCammon, J. A. *Nat. Struct. Biol.* **2002**, *9*, 646–652.
- (13) Shaw, D. E.; et al. *SIGARCH Comput. Archit. News* **2007**, *35*, 1–12.
- (14) Stone, J. E.; Phillips, J.; Freddolino, P. L.; Hardy, D. J.; Trabuco, L.; Schulten, K. *J. Comput. Chem.* **2007**, *28*, 2618–2640.
- (15) Anderson, J. A.; Lorenz, C. D.; Travesset, A. *J. Comput. Phys.* **2008**, *227*, 5342–5359.
- (16) Friedrichs, M. S.; Eastman, P.; Vaidyanathan, V.; Houston, M.; LeGrand, S.; Beberg, A. L.; Ensign, D. L.; Bruns, C. M.; Pande, V. S. *J. Comput. Chem.* **2009**, *30*, 864–872.
- (17) Hampton, S.; Agarwal, P. K.; Alam, S. R.; Crozier, P. S. Towards In *Proceedings of the International Conference on High Performance Computing & Simulation*; Smari, W. A., McIntire, J. P., Eds.; HPCS'10; IEEE: Piscataway, NJ, 2010; pp 98–107.

- (18) Bowers, K. J.; Dror, R. O.; Shaw, D. E. *J. Comput. Phys.* **2007**, *221*, 303–329.
- (19) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.
- (20) Jolliffe, I. T. *Principal Component Analysis*, 2nd ed.; Springer-Verlag New York, Inc.: New York, 2002; Springer Series in Statistics.
- (21) Karplus, M.; Kushick, J. N. *Macromolecules* **1981**, *14*, 325–332.
- (22) Amadei, A.; Lissen, A. B. M.; Berendsen, H. J. C. *Proteins* **1993**, *17*, 412–425.
- (23) Materese, C. K.; Goldmon, C. C.; Papoian, G. A. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 10659–10664.
- (24) Okazaki, K.; Takada, S. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 11182–11187.
- (25) Lange, O.; Grubmuller, H. *Proteins* **2007**, *70*, 1294–1312.
- (26) Bowman, G. R.; Beauchamp, K. A.; Boxer, G.; Pande, V. S. *J. Chem. Phys.* **2009**, *131*, 124101.
- (27) Shao, J.; Tanner, S.; Thompson, N.; Cheatham, T. J. *J. Chem. Theory Comput.* **2007**, *3*, 2312–2334.
- (28) Frickenhaus, S.; Kannan, S.; Zacharias, M. J. *Comput. Chem.* **2009**, *30*, 479–492.
- (29) Daura, X.; van Gunsteren, W. F.; Mark, A. E. *Proteins* **1999**, *34*, 269–280.
- (30) Ramanathan, A.; Agarwal, P.; Kurnikova, M.; Langmead, C. In *Research in Computational Molecular Biology*; Batzoglou, S., Ed.; Springer: Berlin, 2009; Vol. 5541; Lecture Notes in Computer Science, pp 138–154.
- (31) Ramanathan, A.; Agarwal, P. K.; Kurnikova, M.; Langmead, C. *J. J. Comput. Biol.* **2010**, *17*, 309–324.
- (32) Sun, J.; Tao, D.; Faloutsos, C. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*; Eliassi-Rad, T., Ungar, L., Craven, M., Gunopulos, D., Eds.; KDD '06; ACM: New York, 2006; pp 374–383.
- (33) Jorgensen, W.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657–1666.
- (34) Jorgensen, W.; Maxwell, D.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.
- (35) Berweger, C. D.; van Gunsteren, W. F.; Müller-Plathe, F. *Chem. Phys. Lett.* **1995**, *232*, 429–436.
- (36) Bowers, K. J.; Chow, E.; Xu, H.; Dror, R. O.; Eastwood, M. P.; Gregersen, B. A.; Klepeis, J. L.; Kolossvary, I.; Moraes, M. A.; Sacerdoti, F. D.; Salmon, J. K.; Shan, Y.; Shaw, D. E. *SC 2006 Conference, Proceedings of the ACM/IEEE*; IEEE Computer Society: Los Alamitos, CA, 2006; p 43.
- (37) Krautler, V.; van Gunsteren, W. F.; Hünenberger, P. J. *Comput. Chem.* **2001**, *22*, 501–508.
- (38) Shan, Y.; Klepeis, J. L.; Eastwood, M. P.; Dror, R. O.; Shaw, D. E. *J. Chem. Phys.* **2005**, *122*, 054101.
- (39) Papadimitriou, S.; Sun, J.; Faloutsos, C. In *Proceedings of the 31st International Conference on Very Large Data Bases, Trondheim, Norway*; Böhm, K., Jensen, C. S., Haas, L. M., Kersten, M. L., Larson, P.-Å., Ooi, B. C., Eds.; ACM: New York, 2005; Vol. 31, pp 697–708.
- (40) Smilde, A.; Bro, R.; Geladi, P. *Multi-way Analysis: Applications in the Chemical Sciences*; J. Wiley and Sons, Ltd.: West Sussex, England, 2004.
- (41) Abseher, R.; Nilges, M. *J. Mol. Biol.* **1998**, *279*, 911–920.
- (42) Hess, B. *Phys. Rev. E* **2000**, *62*, 8438–8448.
- (43) Ramanathan, A.; Agarwal, P. K. *J. Phys. Chem. B* **2009**, *113*, 11169–11180.
- (44) Lange, O. F.; Lakomek, N.-A.; Fares, C.; Schroder, G. F.; Walter, K. F. A.; Becker, S.; Meiler, J.; Grubmuller, H.; Griesinger, C.; de Groot, B. L. *Science* **2008**, *320*, 1471–1475.
- (45) Kong, Y.; Karplus, M. *Structure* **2007**, *15*, 611–623.
- (46) Kong, Y.; Karplus, M. *Proteins* **2009**, *74*, 145–154.
- (47) Wriggers, W.; Stafford, K. A.; Shan, Y.; Piana, S.; Maragakis, P.; Lindorff-Larsen, K.; Miller, P. J.; Gullingsrud, J.; Rendleman, C. A.; Eastwood, M. P.; Dror, R. O.; Shaw, D. E. *J. Chem. Theory Comput.* **2009**, *5*, 2595–2605.
- (48) Ramanathan, A.; Agarwal, P. K.; Langmead, C. J. *Using tensor analysis to characterize contact-map dynamics in proteins*; Technical Report CMU-CS-08-10, Carnegie Mellon University: Pittsburgh, PA, 2008.
- (49) Yoo, J. O.; Ramanathan, A.; Langmead, C. J. *PyTensor: A Python based Tensor Library*; Technical Report CMU-CS-10-102; Carnegie Mellon University: Pittsburgh, PA, 2010.
- (50) Maisuradze, G. G.; Leitner, D. *Proteins* **2007**, *67*, 569–578.
- (51) Sun, J.; Papadimitriou, S.; Yu, P. S. In *Learning from Data Streams: Processing Techniques in Sensor Networks*; Gama, J., Gaber, M. M., Eds.; Springer: New York, 2007; Chapter 11, pp 165–184.
- (52) Yogurtcu, O. N.; Erdemli, S. B.; Nussinov, R.; Turkay, M.; Keskin, O. *Biophys. J.* **2008**, *94*, 3475–3485.
- (53) Narzi, D.; Daidone, I.; Amadei, A.; Di Nola, A. *J. Chem. Theory Comput.* **2008**, *4*, 1940–1948.
- (54) van der Kamp, M. W.; Schaeffer, R. D.; Jonsson, A. L.; Scouras, A. D.; Simms, A. M.; Toofanny, R. D.; Benson, N. C.; Anderson, P. C.; Merkley, E. D.; Rysavy, S.; Bromley, D.; Beck, D. A.; Daggett, V. *Structure* **2010**, *18*, 423–435.
- (55) Frauenfelder, H.; Sligar, S.; Wolynes, P. *Science* **1991**, *254*, 1598–1603.
- (56) Agarwal, P. K. *Microb. Cell Fact.* **2006**, *5*, 2.
- (57) van Rossum, G. *Python Reference Manual*; Technical Report CS-R9526; Centrum voor Wiskunde en Informatica (CWI): Amsterdam, 1995.
- (58) *MATLAB*, R2009a; Mathworks: Natick, MA, 2009.
- (59) Bader, B.; Kolda, T. *ACM T. Math. Software* **2006**, *32*, 635–653.
- (60) Bader, B.; Kolda, T. *SIAM J. Sci. Comput.* **2007**, *30*, 205–231.
- (61) Tu, T.; Rendleman, C. A.; Borhani, D. W.; Dror, R. O.; Gullingsrud, J.; Jensen, M. O.; Klepeis, J. L.; Maragakis, P.; Miller, P.; Stafford, K. A.; Shaw, D. E. A scalable parallel framework for analyzing terascale molecular dynamics simulation trajectories; In *Proceedings of ACM/IEEE Conference on Supercomputing*; SC'08; IEEE: Piscataway, NJ, 2008; pp 56:1–12.
- (62) Grünberg, R.; Nilges, M.; Leckner, J. *Bioinformatics* **2007**, *23*, 769–770.

Formal Estimation of Errors in Computed Absolute Interaction Energies of Protein–Ligand Complexes

John C. Faver,[†] Mark L. Benson,[†] Xiao He,[†] Benjamin P. Roberts,[†] Bing Wang,[†] Michael S. Marshall,[‡] Matthew R. Kennedy,[‡] C. David Sherrill,[‡] and Kenneth M. Merz, Jr.^{*,†}

[†]Quantum Theory Project, The University of Florida, 2328 New Physics Building, P.O. Box 118435, Gainesville, Florida 32611-8435, United States

[‡]Center for Computational Molecular Science and Technology, School of Chemistry and Biochemistry, and School of Computational Science and Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0400, United States

S Supporting Information

ABSTRACT: A largely unsolved problem in computational biochemistry is the accurate prediction of binding affinities of small ligands to protein receptors. We present a detailed analysis of the systematic and random errors present in computational methods through the use of error probability density functions, specifically for computed interaction energies between chemical fragments comprising a protein–ligand complex. An HIV–II protease crystal structure with a bound ligand (indinavir) was chosen as a model protein–ligand complex. The complex was decomposed into 21 interacting fragment pairs, which were studied using a number of computational methods. The chemically accurate complete basis set coupled cluster theory (CCSD(T)/CBS) interaction energies were used as reference values to generate our error estimates. In our analysis, we observed significant systematic and random errors in most methods, which was surprising, especially for parametrized classical and semiempirical quantum mechanical calculations. After propagating these fragment-based error estimates over the entire protein–ligand complex, our total error estimates for many methods are large compared to the experimentally determined free energy of binding. Thus, we conclude that statistical error analysis is a necessary addition to any scoring function attempting to produce reliable binding affinity predictions.

INTRODUCTION

One of the most challenging problems in computational biochemistry is the accurate prediction of the binding affinity of a small ligand to a protein receptor.^{1,2} This problem is challenging because of the role active site water plays in the binding process,³ the large number of conformational states of the protein and ligand,⁴ tautomeric states of the ligand,⁵ and the computation of accurate interaction energies for each of the hydrogen bonding and nonpolar contacts at the protein–ligand interface.⁶ In order to compute an accurate binding free energy, each of these factors needs to be reliably modeled to chemical accuracy (± 1 kcal/mol relative to an experimental measurement) or beyond.⁶ Reaching acceptable levels of accuracy in the computational modeling of large biomolecular systems is indeed a challenging effort. Dill has even proposed that a protein with 100 amino acid residues should have each one modeled to within 0.1 kcal/mol accuracy to yield an acceptable level of error for the entire protein.⁷

Over the years, many different types of scoring functions have been developed and used to model protein–ligand interactions, ranging from empirical scores to physics-based functions, with some even incorporating quantum mechanical calculations.^{8–11} The most popular docking and scoring algorithms have been extensively reviewed, and their success rates are variable across different score functions and different protein–ligand systems.^{2,12–16} Because of this inconsistency, the notion of consensus scoring, in which several different docking/scoring programs are used to collectively rank ligands by binding affinity, emerged.^{17,18} Consensus scoring generally yields better ordering of ligands by

binding affinity, largely due to the cancellation of errors from each individual scoring function.¹⁹ Free-energy perturbation methods (FEP), which are often used to predict differences in binding affinity between two similar ligands, also have been successful largely due to the cancellation of errors.^{20–23}

The evaluation of error in docking and scoring studies, and indeed, in computational biology in general, has been rather limited. Nonetheless, error analysis can be used to clarify the ability of a method to reach a defined outcome (e.g., the computation of relative free energies of binding). For example, simple statistical error analysis can provide insight into how computed binding affinity errors are diminished in consensus scoring and how relative free energies obtained from free-energy perturbation calculations have reduced uncertainties. More importantly, it can form a hypothesis by which improvements in the desired outcome can be realized.

Errors involved in any kind of calculation or measurement can be categorized as being systematic or random. Systematic errors are predictable in both sign and magnitude, while random errors are unpredictable in both sign and magnitude. It is important to note that the propagation of these types of errors differs in an important way—systematic errors accumulate as a simple sum (eq 1), while random errors propagate as the square root of the sum of squares²⁴ (eq 2). The systematic error sum over individual sources of error is simple to correct for, since it

Received: September 29, 2010

Published: February 11, 2011

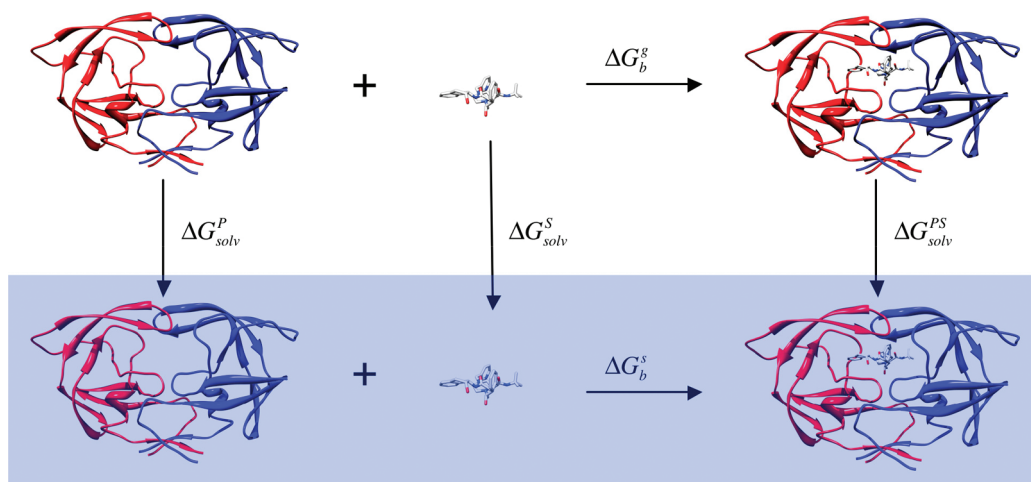


Figure 1. Thermodynamic cycle used to estimate the free energy of binding of a drug molecule to a protein receptor. PS = protein–substrate complex, P = protein, and S = substrate.

represents a predictable shift in the measured (or calculated) value from the true value. Furthermore, systematic errors largely vanish (due to error cancellation) when differences in measurements are calculated. On the other hand, the random error of eq 2 represents a maximum accumulation of individual random error sources. Large random error is a consequence of using a very imprecise method of measurement.

$$\text{Error}_{\text{Systematic}} = \text{Err}_1 + \text{Err}_2 + \text{Err}_3 + \dots \quad (1)$$

$$\text{Error}_{\text{Random}} = [(\text{Err}_1)^2 + (\text{Err}_2)^2 + (\text{Err}_3)^2 + \dots]^{1/2} \quad (2)$$

Another important tool in statistical error analysis is the construction of probability density functions (pdfs). A series of repeated measurements can be used to construct a pdf for a given type of measurement, with which one is able to infer characteristics such as the mean, standard deviation, variance, skewness, and other properties. For example, the widely applicable Gaussian distribution (eq 3) can be used to make inferences about both systematic and random errors. The mean error (μ) in a set of measurements represents a systematic shift in the pdf, while the variance about the mean error (σ^2) determines the “width” of the distribution due to the accumulation of random errors.

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} \quad (3)$$

With these ideas in mind, the success of consensus scoring can readily be rationalized in the terminology of error analysis. As more scoring functions are included in the consensus score, the predicted value of the binding affinity for a given ligand approaches the mean value across all scoring function predictions. As we approach the mean binding affinity, random error from individual scoring functions is minimized, systematic errors largely cancel, and ranking ligands by binding affinity is much more successful. When free energy perturbation methods are used to calculate relative free energies of binding, usually only one functional group interaction is modified between the ligands being compared. This essentially reduces the dimensionality of the error sources to one, which decreases random error. Since these methods only calculate differences in energies, the random error in the final answer is usually very small, and systematic errors mostly cancel.

An application of simple error analysis to the calculation of interaction energies of protein–ligand systems has been outlined in a Gedanken experiment by Merz.⁶ It involves decomposing a protein–ligand complex into independent, chemically meaningful fragment pairs. For example, a ligand may have a hydroxyl moiety hydrogen bound to the backbone carbonyl of a residue in the binding pocket of the receptor, while also containing a hydrophobic carbon chain several angstroms away exhibiting favorable van der Waals interactions with a phenylalanine side chain.²⁵ These are clearly distinguishable interacting fragments both contributing to the total binding affinity of the whole ligand. However, it is well-known that functional groups are not completely independent entities with respect to total free energy of binding. The concept of cooperativity (also termed additivity/nonadditivity) has been described and examined in several protein–ligand systems.^{7,26} The free energy changes of individual interacting fragments cannot simply be summed to produce a global free energy of binding for a protein–ligand complex. However, evidence from statistical mechanics²⁶ and isothermal titration calorimetry experiments²⁷ suggests that enthalpy changes *are* additive across several interactions in a protein–ligand complex. Even so, in some cases, strong secondary electrostatic interactions can arise in systems with multiple proximally close polar groups (DNA base pairs,²⁸ for example) and can introduce nonadditivity in the computation of enthalpy of interaction.

As discussed in Merz’s Gedanken experiment and elsewhere, through the use of the thermodynamic cycle outlined in Figure 1, the free energy change of a ligand upon binding to a protein receptor can be obtained. The final expression (eq 4) contains an electronic interaction energy term, an enthalpy correction term, an entropy term, and a change in the solvation free energy term.

$$\Delta G_b^s = \Delta E + \Delta H_{\text{corr}} - T\Delta S + \Delta\Delta G_{\text{Solv}} \quad (4)$$

$$\Delta E = E_{\text{PS}} - E_{\text{P}} - E_{\text{S}} \quad (5)$$

$$\Delta H = H_{\text{PS}} - H_{\text{P}} - H_{\text{S}} \quad (6)$$

$$T\Delta S = T\Delta S_{\text{PS}} - T\Delta S_{\text{P}} - T\Delta S_{\text{S}} \quad (7)$$

$$\Delta\Delta G_{\text{Solv}} = \Delta G_{\text{Solv}}^{\text{PS}} - \Delta G_{\text{Solv}}^{\text{P}} - \Delta G_{\text{Solv}}^{\text{S}} \quad (8)$$

The present work concentrates only on errors from the electronic interaction energy, ΔE . In doing so, we assume that the errors from ΔH_{corr} , ΔS , and $\Delta \Delta G_{\text{Solv}}$ are zero (which is certainly not the case, but serves as a computational expediency). We refer to such an error estimate as the best-case scenario error, $\text{BCS}_{\text{error}}$, because it ignores the error from these three remaining terms of eq 4. For the sake of simplicity, we assume here as a first approximation that errors in the calculated interaction energies are purely random (even for variational quantum-based methods since the variational principle does not apply to differences between ground state energies). The random errors in the calculation of interaction energies of chemical fragments 1, 2, and 3, etc. now propagate as

$$\text{BCS}_{\text{error}} = [(\Delta E_{\text{calc}}^1 - \Delta E_{\text{ref}}^1)^2 + (\Delta E_{\text{calc}}^2 - \Delta E_{\text{ref}}^2)^2 + (\Delta E_{\text{calc}}^3 - \Delta E_{\text{ref}}^3)^2 + \dots]^{1/2} \quad (9)$$

By neglecting the random error from the enthalpy, entropy, and solvation energies, the overall propagated error (eq 9) is a lower bound to the free energy error estimate and therefore can be thought of as the best-case scenario estimate. In addition to calculating the $\text{BCS}_{\text{error}}$, we also decompose errors from computational methods into systematic and random portions by using Gaussian error probability density functions. By using error probability density functions, we show that it is possible to remove the estimated systematic error for each interaction and thereby decrease the estimation of the remaining random error.

It is well-known that enthalpy and entropy tend to act against one another in free energy calculations in a phenomenon called enthalpy–entropy compensation. These two terms often have opposite signs and similar magnitudes and thus largely cancel each other in eq 4. In terms of error estimation, however, the magnitudes of random errors in each term are not known. We can only estimate their probable range on the basis of studies such as the current one. When calculating overall random error bars in the free energy, estimated errors from the individual terms must be propagated according to eq 2, which only increases with the addition of terms. In addition, while it is true that the errors in any of the three remaining terms of eq 4 could cancel favorably with a portion of the predicted error in ΔE , by exploiting this, we would be calculating the (sometimes) correct ΔG_{bind} for clearly the wrong reasons. The goal of molecular docking is to routinely predict binding free energies both accurately and precisely, so the error should be diminished in each of the terms of eq 4. As a side note, any calculation of the entropy term depends on the potential energy surface of the system and thus depends on enthalpy measurements. Because of this, errors in enthalpy estimations can lead to distorted potential energy surfaces and have unpredictable effects on the entropy term.

Clearly, experimental interaction energies/enthalpies would be desirable to use as a reference for these error functions, but in most cases this will not be possible. As a substitute, we use converged quantum mechanical methods such as CCSD(T) with complete basis set extrapolations (CBS) to provide the reference energy we need to make meaningful error estimates.^{29,30} Such methods have been known to routinely achieve what is termed “chemical accuracy” or the ability to compute energies to within ± 1 kcal/mol of experimental observations.³¹

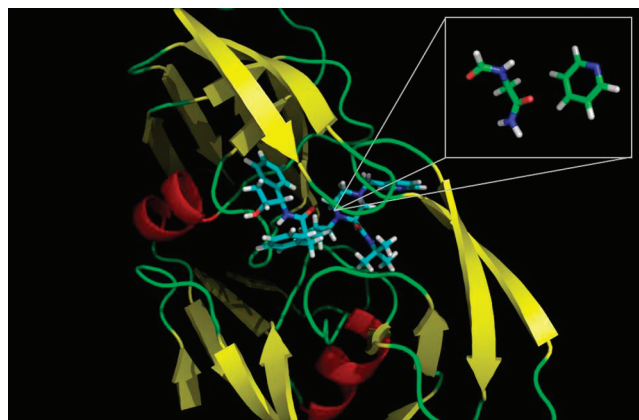


Figure 2. Crystal structure of HIV–II protease bound to indinavir (PDB ID: 1HSG). The inner window displays an example of the fragment systems studied, which includes a glycine residue interacting with a pyridine ring in the ligand.

Prior to presenting the results, it is worth hypothesizing what the expected error types would be for computed interaction energies of hydrogen bonded and nonpolar complexes at different levels of theory. Highly parametric methods like semiempirical or force field methods would be expected, on the face of it, to have very significant random errors rather than systematic errors. On the other hand, quantum chemical methods like Hartree–Fock (HF) or density functional theory (DFT), which do not correct for dispersion, would be expected to have very large systematic errors and relatively smaller random errors. One might suppose that more sophisticated correlated methods such as second-order perturbation theory (MP2) would have smaller systematic errors than HF or DFT; however, MP2 is known to exhibit significant overbinding of dispersion-bound complexes (especially with larger basis sets), and hence MP2 may also exhibit significant systematic errors. Because MP2 performs better for hydrogen-bonding interactions than for dispersion-bound interactions, these errors may also appear to have a significant random component. Interestingly, our analysis shows that systematic errors are a significant component in all cases, which arises because all methods examined tend to give too weak interaction energies with respect to our reference level of theory. If the computed interaction energy errors were better centered at the zero error point, then error cancellation would possibly be beneficial in the computation of the total interaction energy, but this was found to not be the case for the model systems examined. The good news is that the systematic error can be estimated and corrected for while reducing the remaining random error, and this then becomes the main challenge in improving the prediction of interaction energies.

In order to demonstrate our error hypothesis on a real system, we have chosen to examine the HIV–II/indinavir crystal structure³² in detail (PDBID: 1HSG, see Figure 2). This system was chosen because of its large number and diversity of chemical contacts. In addition, the experimentally observed binding free energy has been reported^{33,34} to be -12.8 kcal/mol, indicating an overall strong level of interaction between the protein and indinavir. We have decomposed the protein–ligand complex into 21 distinct, chemically important fragment-based interactions (Figure 3). Each pair of interacting fragments was evaluated in terms of gas-phase electronic interaction energy with a number of different computational methods including force field methods,

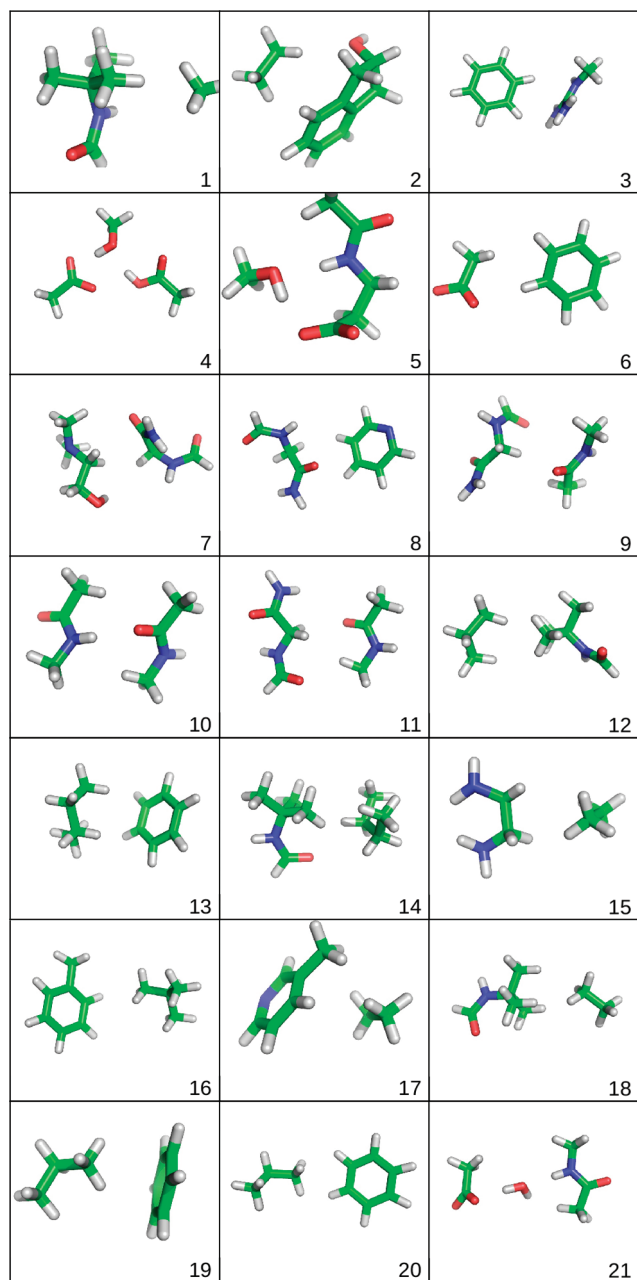


Figure 3. The 21 chemical fragment pairs comprising the 1HSG protein/ligand complex used in this study.

semiempirical quantum mechanical methods, density functional methods, and *ab initio* wave function methods. By comparing the calculated interaction energy of each fragment pair for a given method to a reference, one can calculate the BCS_{error} and generate pdfs to describe the error for the full protein–ligand interaction.

METHODS

The HIV II–protease/indinavir complex structure was obtained from the protein databank (PDBID: 1HSG).³² Hydrogen atoms were added to the structure with the program Reduce,³⁵ followed by optimization of their positions with the AMBER FF99SB force field.³⁶ Close contacts (less than 3 Å) in the resulting structure were highlighted in the visualization program

Chimera,³⁷ which were then manually examined. In many cases, one member of an interacting fragment pair would be adjacent to a polarizing group such as the peptide bond along the protein backbone. These polarizing groups, although not directly interacting with the other fragment partner, were included in the fragment structures to model the polarization effect on the fragments. Any time an aromatic fragment was involved, the entire aromatic ring was included.

After manually determining all of the chemically important fragment pairs, the fragments were generated by cutting the covalent bonds joining them to the remainder of the molecule and replacing the bonds with linker hydrogen atoms. The hydrogen bond distances were set to 1.1 Å for carbon–hydrogen bonds and 1.0 Å for nitrogen–hydrogen bonds. The resulting 21 interacting fragment pair structures are displayed in Figure 3, and their interaction energies can be found in the Supporting Information, available free of charge via the Internet at <http://pubs.acs.org>, or for an interactive version visit <http://crunch.qtp.ufl.edu/~faver/error/1hsg>.

The gas-phase electronic interaction energies of the 21 independent chemically meaningful fragment complexes (Figure 3) were calculated using the following computational methods: the generalized AMBER force field (GAFF);³⁸ AMOEBA;³⁹ MMFF;⁴⁰ AM1;⁴¹ AM1-FS1;⁴² PM3;⁴³ PM6;⁴⁴ PM6-DH2;⁴⁵ PDDG;⁴⁶ Hartree–Fock (HF); second-order Møller–Plesset perturbation theory (MP2); the M06-L density functional;⁴⁷ B97-D;⁴⁸ and coupled-cluster²⁹ with single, double, and perturbative triple excitations (CCSD(T)). Basis sets for the quantum mechanical methods ranged from the 6-31G* Pople-type basis set⁴⁹ to the aug-cc-pVXZ (X = D, T, and Q—hereafter referred to as aXZ) Dunning type basis sets.⁵⁰ In addition, extrapolation to the complete basis set (CBS) limit was also used.³⁰ The CCSD(T) calculations were evaluated using the “heavy augmented” haDZ and haTZ basis sets, which place diffuse functions only on heavy (non-hydrogen) atoms.^{51,52} In our error analysis, the CCSD(T)/CBS energies were used as our reference, which were evaluated as

$$\Delta E_{\text{CCSD(T)/CBS}} = \Delta \text{CCSD(T)} + \Delta E_{\text{MP2/CBS[aTZ-aQZ]}} \quad (10)$$

$$\begin{aligned} \Delta \text{CCSD(T)} = & \Delta E_{\text{CCSD(T)/CBS[haDZ-haTZ]}} \\ & - \Delta E_{\text{MP2/CBS[haDZ-haTZ]}} \end{aligned} \quad (11)$$

where the notation CBS[aTZ-aQZ] represents the complete basis set extrapolation from energies calculated with the aTZ and aQZ basis sets. In each quantum-based method, basis set superposition error (BSSE) was accounted for using the counterpoise correction. The Hartree–Fock and MP2 interaction energies were calculated with the Gaussian 03 program.⁵³ The M06-L and PM6 interaction energies were calculated with Gaussian 09.⁵⁴ AM1, PM3, and PDDG interaction energies were evaluated with the DivCon program.⁵⁵ The GAFF energies were calculated using the AMBER 10 suite.⁵⁶ Most of the CCSD(T) computations were performed using MOLPRO.⁵⁷ The larger CCSD(T) computations were performed using NWChem⁵⁸ on the Jaguar supercomputer at Oak Ridge National Laboratory. MMFF calculations were performed with the Schrödinger suite.⁵⁹

RESULTS AND DISCUSSION

Our analysis of various error scoring functions for the set of computational methods across all fragment pairs is shown in

Table 1. Error Analysis of the Absolute Interaction Energies from Various Methods^a

method	BCS _{error}	ME (μ)	Var (σ^2)	$N\mu$	$(N\sigma^2)^{1/2}$	RMSE	max err
GAFF	7.41	0.94	1.66	19.74	5.90	1.62	4.35
AMOEBA ^b	2.58	0.36	0.24	7.56	2.24	0.75	2.23
MMFF	20.81	-0.61	20.25	-12.81	20.62	4.54	11.43
AM1	16.39	2.57	5.89	53.97	11.12	3.58	10.69
AM1-FS1	11.44	0.03	6.23	0.63	11.44	2.50	7.97
PM3	14.70	1.82	6.67	38.22	11.84	3.21	8.24
PDDG	20.46	1.69	16.29	35.49	18.50	4.47	14.8
PM6	10.94	1.56	3.13	32.76	8.11	2.39	6.99
PM6-DH2	4.89	-0.25	1.11	-5.25	4.83	1.07	2.94
B97-D/ TZVP	2.53	0.05	0.30	1.05	2.51	0.55	2.12
M06-L/ 6-31G*	10.94	0.44	0.48	9.24	3.17	0.83	1.92
aTZ	3.25	0.62	0.11	13.02	1.52	0.71	1.16
HF/ 6-31G*	15.22	3.18	0.87	66.78	4.27	3.32	4.91
aDZ	16.25	3.33	1.42	69.93	5.46	3.55	6.68
aTZ	16.22	3.32	1.42	69.72	5.46	3.54	6.77
aQZ	16.16	3.31	1.39	69.51	5.40	3.53	6.70
MP2/ 6-31G*	7.45	1.58	0.13	33.18	1.65	1.63	2.35
aDZ	2.89	0.41	0.22	8.61	2.15	0.63	1.83
aTZ	1.35	0.09	0.08	1.89	1.30	0.30	0.77
aQZ	1.02	-0.02	0.05	-0.42	1.02	0.22	0.53
CBS[aDZ-aTZ]	1.07	-0.04	0.05	-0.84	1.02	0.23	0.57
CBS[aTZ-aQZ]	1.02	-0.05	0.04	-1.05	0.92	0.22	0.56
CCSD(T)/ haDZ	3.80	0.69	0.20	14.49	2.05	0.83	2.08
haTZ	1.28	0.23	0.03	4.83	0.74	0.28	0.74
CBS[haDZ-haTZ] ^c	0.28	0.03	0.00	0.63	0.00	0.06	0.14

^a Various scores (based on kcal/mol units) measuring the error of gas-phase electronic interaction energies of 21 independent chemical fragments involved in the binding of indinavir to IHSG using CCSD(T)/CBS calculations as a reference. The error scores are (from left to right): BCS_{error} (eq 9), mean error, variance about the mean error, overall expected systematic error, overall expected random error, root mean square error, and maximum error. ^b AMOEBA results were calculated for 12 of the 21 fragments, due to lack of parameters for atoms in the remaining nine systems. ^c haDZ-haTZ extrapolation to the CCSD(T) complete basis set limit; this approximates our benchmark CCSD(T)/CBS values but lacks the MP2 estimate of basis set effects beyond haTZ.

Table 1. The second column of Table 1, the BCS random error bar, is the evaluation of eq 9 and gives a measure of error assuming that all errors in the electronic interaction energies are purely random. If this were true, it would suggest that, for example, PM3 total interaction energies for ligand binding can be off by as much as 14.7 kcal/mol, and we would be left with no reliable statistical way to narrow this error window. This would of course be unacceptable because this error bar is actually larger in magnitude than the experimental binding affinity itself (-12.8 kcal/mol). The random error bar for GAFF is about half as large in magnitude but still represents a large portion of the experimental binding affinity. When quantum effects are accounted for, e.g., MP2/CBS, we are able to narrow the range of accumulated random errors down to ~1 kcal/mol if we use the complete basis set extrapolations.

Upon examination of the mean errors over the 21 fragment systems (see column 3 of Table 1), it is apparent that there is a significant systematic error component in the majority of the methods studied. Further evidence is shown in Figure 4, which

plots the interaction energy deviations from the CCSD(T)/CBS reference for each fragment pair for a subset of the methods examined. Interestingly, most of the calculated interaction energies are less strongly binding than those calculated with the CCSD(T)/CBS reference, indicating far less random error character in the computed interaction energy errors. Systematic errors can be distinguished from the random sources of error by choosing an appropriate probability distribution function to describe the measurement data. If we choose the Gaussian distribution as defined by eq 3, we need to calculate two parameters, the mean error μ and the variance σ^2 . The values for each computational method are shown in columns three and four of Table 1.

If we choose to model our calculations (measurements) with the Gaussian pdf, the calculated value of μ represents a shift of the distribution (systematic error), and σ^2 represents a measure of the width of the distribution (random error). For example, GAFF interaction energies have an average deviation from our reference of 0.94 kcal/mol, which shifts the Gaussian distribution by this

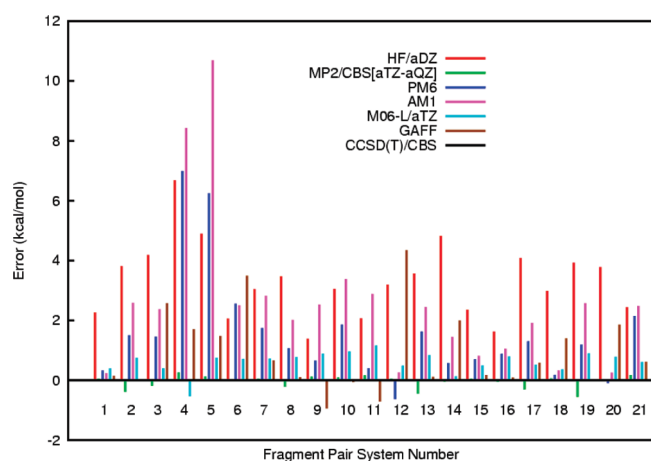


Figure 4. The difference between example methods and the CCSD(T)/CBS reference method in interaction energy of each of the 21 fragment systems. Most methods generally underestimate the interaction energy, which implies a portion of the errors is systematic in nature.

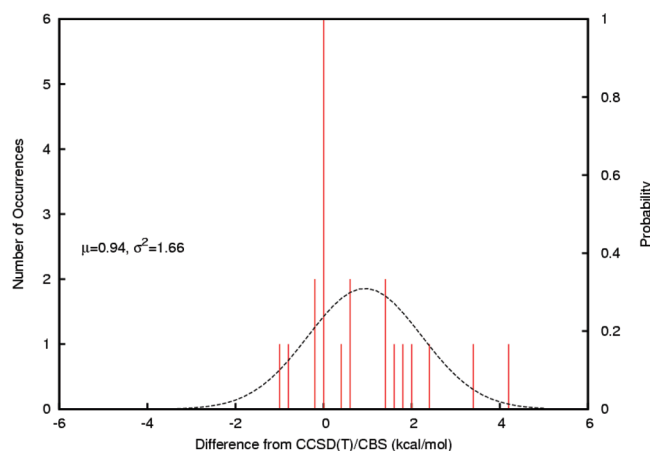


Figure 5. Error distribution of GAFF interaction energies of the 21 fragment systems fit to a Gaussian probability density function. The systematic error in the GAFF interaction energies (using CCSD(T)/CBS as a reference) is predicted to be 0.94 kcal/mol per interaction, and the variance (random error) about this shift is 1.66 (kcal/mol)².

amount. The variation about this mean error is 1.66 (kcal/mol)², which gives a measure of the width of the distribution and the random error of measurement (see Figure 5). By using these values, we can also obtain an estimate of the total systematic and random error from all fragment interactions comprising the protein/ligand complex by evaluating functions similar to eqs 1 and 2. However, now we are using the parameters of the probability distribution function, μ and σ^2 , assuming for now that they are representative of all of the chemical interactions.

$$\text{Error}_{\text{Systematic}} = \sum_i \mu_i = N\mu \quad (12)$$

$$\text{Error}_{\text{Random}} = \left[\sum_i \sigma_i^2 \right]^{1/2} = \sqrt{N\sigma^2} \quad (13)$$

Note that both eqs 12 and 13 depend on the number of chemical interactions. This shows the size dependence of our error estimates as discussed earlier by Merz.⁶ Errors in energy

calculations may be inconsequential when dealing with small systems, but as more molecular contacts become involved, propagated error may have significant effects on energy calculations.

Using eqs 12 and 13, we find that the predicted systematic error for the total interaction energy of the protein–ligand complex with GAFF is 19.74 kcal/mol (column 5 of Table 1). The predicted overall random error bar is 5.90 kcal/mol (column 6). These values of random error are lower than the $\text{BCS}_{\text{error}}$ values because we have removed the systematic components of the overall errors in each method. We can narrow this random error bar even further by constructing pdf's for different types of interactions—e.g., nonpolar and polar interactions. In this way, the systematic error for each type of chemical interaction could be removed independently, while leaving less random error in the overall interaction energies.

Keeping with the GAFF energies as an example, we calculated the mean error and variance of the 15 nonpolar interactions to be 1.17 kcal/mol and 1.83 (kcal/mol)² and the mean error and variance of the six polar interactions to be 0.35 kcal/mol and 1.03 (kcal/mol)². To compute the total systematic and random errors due to all interactions, we use eqs 12 and 13, but with different values of μ and σ^2 depending on the interaction type. The overall systematic error by differentiating between nonpolar and polar interactions is predicted to be 19.65 kcal/mol, slightly less than the mean error by averaging over all interactions. The overall random error bar is now calculated to be 5.80 kcal/mol, less than the value calculated previously using one value of variance for all interactions. As more pdf's for different classes of interactions are used, and more data points are added to them, we expect that the overall systematic error in interaction energies can be more reliably removed and the random error bars can be minimized.

CONCLUSION

Using benchmark-quality quantum mechanical CCSD(T)/CBS results, we have been able to establish reference values for 21 fragment interaction energies present in the complex of indinavir and HIV–II protease. Using these reference values, we established the errors associated with the use of less sophisticated computational techniques including force fields, semiempirical methods, and DFT calculations. We then estimated random and systematic errors present in each computational method by constructing error probability density functions. We found that most computational methods predominantly gave interaction energies that were less binding than the reference energies yielding a significant systematic error component in these cases. Semiempirical methods performed the poorest overall, yielding large random and systematic errors. The GAFF force field significantly outperformed the semiempirical methods. The HF calculations had large systematic errors due to the lack of dispersion but had relatively small random errors. The M06-L functional performed impressively, even with the small 6-31G* basis set, giving remarkably small systematic and random errors.

Our error analysis suggests one way in which any scoring algorithm based on energy could be improved. Given a scoring function, comparisons of interaction energies for many types of interactions in different geometries with chemically accurate quantum mechanical calculations could be used to derive error probability density functions for several classes of interactions. By constructing these pdf's, the predicted systematic error involved in each interaction present could be removed. The

remaining random component of error could then be reported with the energy scores to provide insight into their reliability. The results presented herein offer only a start along this direction in the sense that the number of interactions, their type, and their geometry distribution are limited to the indinavir/HIV-II system studied. In order to develop accurate pdfs to correct a broad range of interactions, one would require the collection of a large number of unique interacting fragments to serve as reference systems to assemble the pdfs for any given scoring function. Such a set of interacting fragments could then provide a basis by which any systematic errors in a physics-based score function could be corrected for, assuming that additivity is largely dominant in the computation of interaction energies. The additivity or nonadditivity of the interactions for a given protein–ligand complex is poorly understood, and understanding this aspect of the problem would be essential to effectively apply any systematic error correction scheme.

■ ASSOCIATED CONTENT

S **Supporting Information.** The Supporting Information contains the information available on our dedicated Web site for the project at <http://crunch.qtp.ufl.edu/~faver/error/1hsg>. The packet contains an index of the fragment-based interactions, followed by individual pages for each interaction containing (1) a visual representation of the system, and (2) a table of calculated interaction energies from various computational methods. It also contains a table of error statistics over all interactions, and a table of histograms and error probability density functions plotted for each computational method studied. This information is available free of charge via the Internet at <http://pubs.acs.org>

■ AUTHOR INFORMATION

Corresponding Author

*Phone: 352-392-6973. Fax: (352) 392-8722. E-mail: merz@qtp.ufl.edu.

■ ACKNOWLEDGMENT

The authors thank Professors Adrian Roitberg and Erik Deumens for helpful discussions. K.M.M. thanks the NIH (GM044974 and GM066689) for funding the present research. C.D.S. gratefully acknowledges support from NSF (CHE-0715268 and CHE-04-43564). This work is also funded in part by the U.S. Department of Energy, the division of Advanced Scientific Computing Research as a component of the Computational Chemistry Endstation project under a subcontract from the University of Tennessee. This research was performed in part using resources of the National Center for Computational Sciences at Oak Ridge National Laboratory under contract DE-AC05-00OR22725.

■ REFERENCES

- (1) Gilson, M. K.; Zhou, H. X. *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 21–42.
- (2) Leach, A. R.; Shoichet, B. K.; Peishoff, C. E. *J. Med. Chem.* **2006**, *49*, 5851–5855.
- (3) Abel, R.; Young, T.; Farid, R.; Berne, B. J.; Friesner, R. A. *J. Am. Chem. Soc.* **2008**, *130*, 2817–2831.
- (4) Deng, Y. Q.; Roux, B. *J. Phys. Chem. B.* **2009**, *113*, 2234–2246.
- (5) Martin, Y. C. *J. Comput.-Aided Mol. Des.* **2009**, *23*, 693–704.
- (6) Merz, K. M. *J. Chem. Theory Comput.* **2010**, *6*, 1769–1776.
- (7) Dill, K. A. *J. Biol. Chem.* **1997**, *272*, 701–704.

- (8) Raha, K.; Peters, M. B.; Wang, B.; Yu, N.; Wollacott, A. M.; Westerhoff, L. M.; Merz, K. M. *Drug Discovery Today* **2007**, *12*, 725–731.
- (9) Peters, M. B.; Raha, K.; Merz, K. M. *Curr. Opin. Drug Discovery Dev.* **2006**, *9*, 370–379.
- (10) Raha, K.; Merz, K. M. *J. Med. Chem.* **2005**, *48*, 4558–4575.
- (11) Fukuzawa, K.; Kitaura, K.; Uebayasi, M.; Nakata, K.; Kaminuma, T.; Nakano, T. *J. Comput. Chem.* **2005**, *26*, 1–10.
- (12) Cummings, M. D.; DesJarlais, R. L.; Gibbs, A. C.; Mohan, V.; Jaeger, E. P. *J. Med. Chem.* **2005**, *48*, 962–976.
- (13) Guvench, O.; MacKerell, A. D. *Curr. Opin. Struct. Biol.* **2009**, *19*, 56–61.
- (14) Taylor, R. D.; Jewsbury, P. J.; Essex, J. W. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 151–166.
- (15) Kontoyianni, M.; McClellan, L. M.; Sokol, G. S. *J. Med. Chem.* **2003**, *47*, 558–565.
- (16) 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.
- (17) Charifson, P. S.; Corkery, J. J.; Murcko, M. A.; Walters, W. P. *J. Med. Chem.* **1999**, *42*, 5100–5109.
- (18) Oda, A.; Tsuchida, K.; Takakura, T.; Yamaotsu, N.; Hirono, S. *J. Chem. Inf. Model.* **2006**, *46*, 380–391.
- (19) Wang, R. X.; Wang, S. M. *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 1422–1426.
- (20) Kollman, P. A. *Chem. Rev.* **1993**, *93*, 2395–2417.
- (21) Huang, N.; Kalyanaraman, C.; Bernacki, K.; Jacobson, M. P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 5166–5177.
- (22) Foloppe, N.; Hubbard, R. *Curr. Med. Chem.* **2006**, *13*, 3583–3608.
- (23) Lybrand, T. P.; McCammon, J. A.; Wipff, G. *Proc. Natl. Acad. Sci. U. S. A.* **1986**, *83*, 833–835.
- (24) Taylor, J. R. Propagation of Uncertainties. In *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*, 2nd ed.; McGuire, A., Ed.; University Science Books: Sausalito, CA, 1996; pp 45–93.
- (25) Bissantz, C.; Kuhn, B.; Stahl, M. *J. Med. Chem.* **2010**, *53*, 5061–5084.
- (26) Mark, A. E.; Vangunsteren, W. F. *J. Mol. Biol.* **1994**, *240*, 167–176.
- (27) Baum, B.; Muley, L.; Smolinski, M.; Heine, A.; Hangauer, D.; Klebe, G. *J. Mol. Biol.* **2010**, *397*, 1042–1054.
- (28) Jorgensen, W. L.; Pranata, J. *J. Am. Chem. Soc.* **1990**, *112*, 2008–2010.
- (29) Bartlett, R. J.; Musial, M. *Rev. Mod. Phys.* **2007**, *79*, 291–352.
- (30) Halkier, A.; Helgaker, T.; Jorgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. *Chem. Phys. Lett.* **1998**, *286*, 243–252.
- (31) Helgaker, T.; Klopper, W.; Tew, D. P. *Mol. Phys.* **2008**, *106*, 2107–2143.
- (32) Chen, Z. G.; Li, Y.; Chen, E.; Hall, D. L.; Darke, P. L.; Culbertson, C.; Shafer, J. A.; Kuo, L. C. *J. Biol. Chem.* **1994**, *269*, 26344–26348.
- (33) Dorsey, B. D.; Levin, R. B.; Mcdaniel, S. L.; Vacca, J. P.; Guare, J. P.; Darke, P. L.; Zugay, J. A.; Emimi, E. A.; Schleif, W. A.; Quintero, J. C.; Lin, J. H.; Chen, I. W.; Holloway, M. K.; Fitzgerald, P. M. D.; Axel, M. G.; Ostovic, D.; Anderson, P. S.; Huff, J. R. *J. Med. Chem.* **1994**, *37*, 3443–3451.
- (34) Vacca, J. P.; Dorsey, B. D.; Schleif, W. A.; Levin, R. B.; Mcdaniel, S. L.; Darke, P. L.; Zugay, J. C.; Blahy, O. M.; Roth, E.; Sardana, V. V.; Schlabach, A. J.; Graham, P. I.; Condra, J. H.; Gotlib, L.; Holloway, M. K.; Lin, J.; Chen, I. W.; Vastag, K.; Ostovic, D.; Anderson, P. S.; Emimi, E. A.; Huff, J. R. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 4096–4100.
- (35) Word, J. M.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C. *J. Mol. Biol.* **1999**, *285*, 1735–1747.
- (36) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins: Struct., Funct., Bioinf.* **2006**, *65*, 712–725.
- (37) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem.* **2004**, *25*, 1605–1612.

- (38) Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157–1174.
- (39) Ponder, J. W.; Wu, C. J.; Ren, P. Y.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; DiStasio, R. A.; Head-Gordon, M.; Clark, G. N. I.; Johnson, M. E.; Head-Gordon, T. *J. Phys. Chem. B* **2010**, *114*, 2549–2564.
- (40) Halgren, T. A. *J. Comput. Chem.* **1999**, *20*, 720–729.
- (41) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.
- (42) Foster, M.; Söhlberg, K. *J. Chem. Theory Comput.* **2010**, *6*, 2153–2166.
- (43) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209–220.
- (44) Stewart, J. J. P. *J. Mol. Model.* **2007**, *13*, 1173–1213.
- (45) Korth, M.; Pitonák, M.; Rezák, J.; Hobza, P. *J. Chem. Theory Comput.* **2010**, *6*, 344–352.
- (46) Repasky, M. P.; Chandrasekhar, J.; Jorgensen, W. L. *J. Comput. Chem.* **2002**, *23*, 1601–1622.
- (47) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215–241.
- (48) Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787–1799.
- (49) Ditchfie, R.; Hehre, W. J.; Pople, J. A. *J. Chem. Phys.* **1971**, *54*, 724–728.
- (50) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796–6806.
- (51) Sherrill, C. D.; Takatani, T.; Hohenstein, E. G. *J. Phys. Chem. A* **2009**, *113*, 10146–10159.
- (52) Elshohly, A. M.; Tschumper, G. S. *Int. J. Quantum Chem.* **2009**, *109*, 91–96.
- (53) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; 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 E.01; Gaussian, Inc.: Wallingford, CT, 2004.
- (54) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V. M. B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, 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.; Cioslowski, J.; Fox, D. J. *Gaussian 09*; Gaussian, Inc.: Wallingford, CT, 2009.
- (55) Dixon, S. L.; van der Vaart, A.; Gogonea, V.; Vincent, J. J.; Brothers, E. N.; Suárez, D.; Westerhoff, L. M.; Merz, K. M., Jr. *DIVCON99*; The Pennsylvania State University: University Park, PA, 1999.
- (56) Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Crowley, M.; Walker, R. C.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman, P. A. *AMBER 10*; University of California, San Francisco: San Francisco, CA, 2008.
- (57) Werner, H.-J.; Knowles, P. J.; Lindh, R.; Manby, F. R.; Schütz, M.; Celani, P.; Knizia, G.; Korona, T.; Lindh, R.; Mitrushenkov, A.; Rauhut, G.; Adler, T. B.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Goll, E.; Hampel, C.; Hesselmann, A.; Hetzer, G.; Hrenar, T.; Jansen, G.; Köppl, C.; Liu, Y.; Lloyd, A. W.; Mata, R. A.; May, A. J.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklass, A.; Palmieri, P.; Pflüger, K.; Pitzer, R.; Reiher, M.; Shiozaki, T.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T.; Wang, M.; Wolf, A. *MOLPRO*. See <http://www.molpro.net> (accessed Jan. 2011).
- (58) Kendall, R. A.; Apra, E.; Bernholdt, D. E.; Bylaska, E. J.; Dupuis, M.; Fann, G. I.; Harrison, R. J.; Ju, J. L.; Nichols, J. A.; Nieplocha, J.; Straatsma, T. P.; Windus, T. L.; Wong, A. T. *Comput. Phys. Commun.* **2000**, *128*, 260–283.
- (59) *MacroModel*, version 9.8; Schrodinger, LLC: New York, NY, 2010.

Systematic Theoretical Investigation on the Light Emitter of Firefly

Shu-Feng Chen,[†] Ya-Jun Liu,^{*,†} Isabelle Navizet,^{‡,§} Nicolas Ferré,^{||} Wei-Hai Fang,^{*,†} and Roland Lindh^{*,⊥}

[†]Key Laboratory of Theoretical and Computational Photochemistry, Ministry of Education, College of Chemistry, Beijing Normal University, Beijing 100875, China

[‡]Laboratoire Modélisation et Simulation Multi Echelle, Université Paris-Est, MSME UMR 8208 CNRS, 5 bd Descartes, 77454 Marne-la-Vallee, France

[§]Molecular Science Institute School of Chemistry, University of the Witwatersrand, PO Wits Johannesburg 2050 South Africa

^{||}Universités d'Aix-Marseille I, II, et III-CNRS UMR 6264: Laboratoire Chimie Provence, Equipe: Chimie Théorique, Faculté de St-Jérôme, Case 521, 13397 Marseille Cedex 20, France

[⊥]Department of Quantum Chemistry, The Ångström Laboratory, Uppsala University, P.O. Box 518, S-75120 Uppsala, Sweden

S Supporting Information

ABSTRACT: This is a systematic theoretical investigation on all the possible light emitters of firefly using a multireference method. Six chemical forms of oxyluciferin (OxyLH₂) molecules/anions were studied by a multistate complete active space second-order perturbation (MS-CASPT2) method in vacuum and dimethyl sulfoxide. The calculated results and subsequent analysis excluded enol-OxyLH₂, keto-OxyLH₂, and enolate-OxyLH⁻ as possible light emitters. The remaining three candidates, phenolate-enol-OxyLH⁻, phenolate-keto-OxyLH⁻, and OxyL²⁻, were further investigated in protein by a MS-CASPT2/molecular mechanics (MM) study to explain the natural bioluminescence of firefly. By comparison of the MS-CASPT2/MM calculated results of phenolate-enol-OxyLH⁻, phenolate-keto-OxyLH⁻, and OxyL²⁻ with the experimental observation and detailed analysis, we concluded that the direct decomposition excited-state product of firefly dioxetanone in vivo and the only light emitter of firefly in natural bioluminescence is the first singlet excited state (S₁) of phenolate-keto-OxyLH⁻.

INTRODUCTION

The firefly bioluminescence has for decades been studied experimentally. Theoretical calculation beyond semiempirical theory was first done by Orlova et al.¹ in 2003 using density functional theory (DFT). Its generally accepted mechanism is shown in Figure 1. The mechanism is now understood to the extent that intelligent modifications to improve the performance of luciferin, in industrial applications, based on this knowledge are reported. For example, recently, Reddy et al. demonstrated that an engineered modified luciferin molecule, a cyclic alkylaminoluciferin, exhibits more efficient light emission than the native form.² However, there are still three important concerns about the firefly bioluminescence, actually the most difficult three questions for understanding the firefly bioluminescence which need to be answered in detail. First, how does the ground-state (S₀) firefly dioxetanone decompose to the excited-state (S₁) oxyluciferin (OxyLH₂) molecules/anions? This has been theoretically investigated in gas phase,^{3–6} although the explanations are not totally satisfactory. Second, how does the microenvironment of the light emitter(s) affect the emission wavelength? In protein (bioluminescence), the environmental effect mainly comes from the hydrogen bonds linking to the OxyLH₂, which has been verified by experimental^{7–9} and theoretical^{10,11} investigations. In solution (chemiluminescence), the light emission of excited OxyLH₂ is affected by the polarity and the pH value of the solvent, which were extensively explored experimentally.^{12–18} The experiment¹⁷ suggests that the light emitter is the S₁ state of phenolate-keto-OxyLH⁻ (named ¹(OL⁻)* in ref 17). The

authors concluded that the wavelength of the light emission from the S₁ state of phenolate-keto-OxyLH⁻ is modulated by the polarity of the active-site environment of luciferase and the degree of covalent character of the O···H bond between phenolate oxygen of phenolate-keto-OxyLH⁻ and a protonated basic moiety in the active site. The current investigation will focus on the third question. What is/are the firefly light emitter(s) in vivo? Six possible chemical forms of OxyLH₂ are listed in Figure 2. For consistency, we take the same names employed in ref 12. From 1971 to 2002, the general thought on light emitters of firefly was that both keto and enol/enolate forms of OxyLH₂ are responsible for the light emission.^{7,13,15,19–21} However, from 2002 to 2009, the phenolate-keto-OxyLH⁻ was considered as the only light emitter,^{10,11,17,22–24} except for the mechanism of triple chemical equilibrium supposed by Ugarova.²⁵ Based on a series of fluorescence spectra and NMR data, Naumov et al.¹² in 2009 concluded that the enol form of OxyLH₂ can play a role in the emission, and the concentration ratio of the different chemical forms in solutions of OxyLH₂ was determined by several factors which affect the intricate triple chemical equilibrium, most notably the pH, solvent polarity, hydrogen bonding, presence of additional ions, and π–π stacking. Moreover, in their very recent publication,²⁶ Naumov et al. investigated the spectral–structural effects of the keto–enol–enolate and phenol–phenolate equilibria of OxyLH₂ and concluded that the

Received: January 19, 2011

Published: February 16, 2011

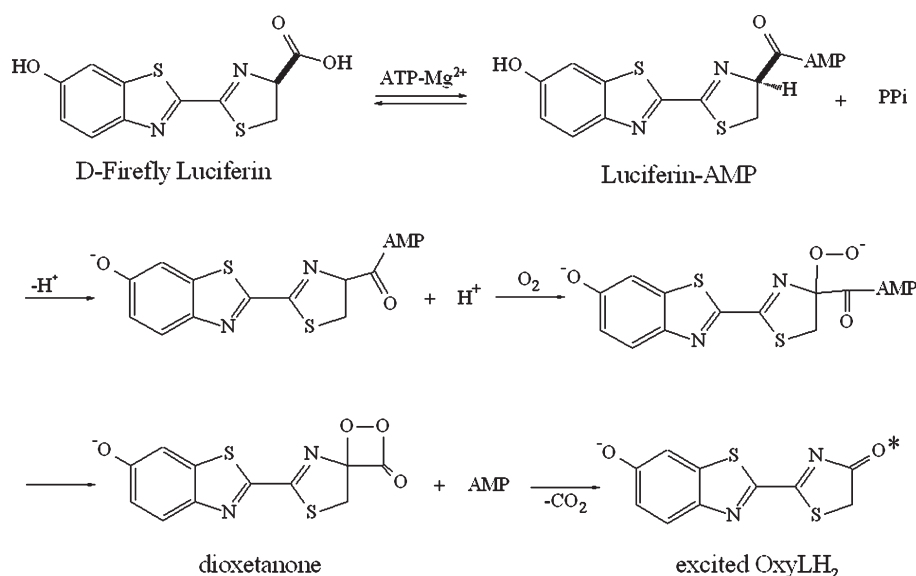


Figure 1. The generally accepted mechanism of firefly bioluminescence.

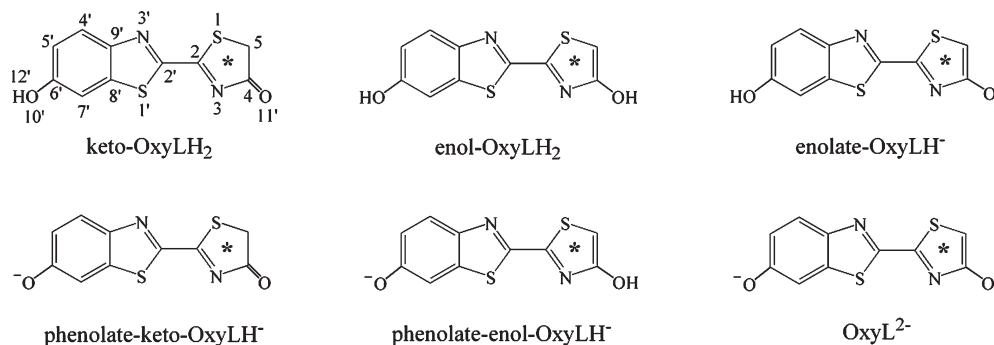


Figure 2. Molecular structures and atomic labels of all possible excited-state OxyLH₂ chemical forms.

phenol–enolate form of OxyLH₂ (see enolate-OxyLH[−] in Figure 2) is a yellow-emitting species. However, these in vitro experiments deal with the chemiluminescence of OxyLH₂, not the bioluminescence. So, which is/are the light emitters of firefly bioluminescence? Where does/do the emitter(s) come from in vivo? Is it/are they the direct product(s) of the firefly dioxetanone decomposition or a tautomerization product from another excited state of OxyLH₂? To answer these questions, Min et al.²⁷ did the following theoretical investigation: They selected Arg220, His247, AMP, Water324, Phe249, Gly343, and Ser349 in the X-ray structure of luciferase (2D1R)²⁸ to simulate the luciferase environment. The respective and cooperative effects of those residues and water molecule on the electronic absorption and emission spectra of OxyLH₂ were investigated by DFT using the B3LYP, B3PW91, and PBE1KCIS functionals. Min et al. concluded that phenolate-keto-OxyLH[−] not the enol forms produced yellow-green luminescence under the composite effects of the simulated environment. However, this investigation is far from systematic. Milne et al. performed a more systematic study based on the ground-state optimization of the six isomers of oxyluciferin using fragment molecular orbital-1:restricted Hartree-Fock time-dependent DFT (FMO-1:RHF-TDDFT), including the side chains of residues closer than 7.5 Å to oxyluciferin.²⁹ From the absorption spectra, they concluded that

phenolate-keto-OxyLH[−] was likely both the yellow-green and red light-emitting species. It is still impossible for experiments to detect which is/are the direct excited-state product(s) of the firefly dioxetanone decomposition and which is/are the actual light emitter(s) of firefly bioluminescence in nature. We will here give a clear answer to these questions by reliable theoretical investigation. The report is supplemented by Supporting Information providing additional information and details.

COMPUTATIONAL METHODS

The complete active space SCF (CASSCF)³⁰ method was used to optimize the S₀ and S₁ geometries of six OxyLH₂ molecules/anions (see Figure 2). Their vertical excitation energies (T_v), transition energies (T_e) of emission (S₁ → S₀) and oscillator strengths (f) were calculated using the multistate complete active space second-order perturbation (MS-CASPT2) method.³¹ These calculations were performed in both the gas phase and dimethyl sulfoxide (DMSO) solution. The selected active space ‘18 electrons in 15 orbitals’ (18-in-15) of all kinds of conformers are listed in Figure S1, Supporting Information. The double-ζ ANO-RCC³² basis sets (ANO-RCC-VDZP) were used for all the calculations. All these calculations were performed using the MOLCAS 6.4 quantum chemistry software.³³ The three candidates, phenolate-enol-OxyLH[−], phenolate-keto-OxyLH[−],

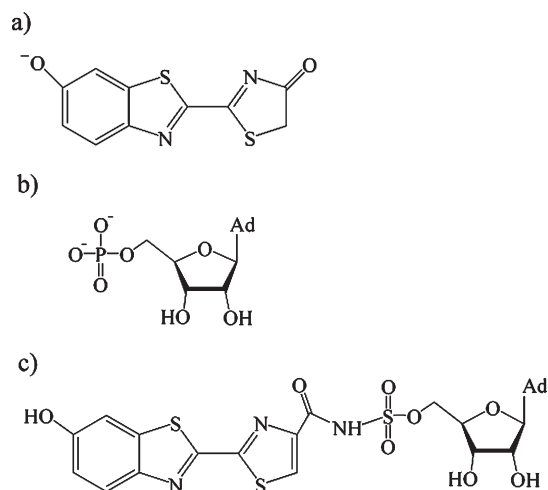


Figure 3. Structures of: (a) phenolate-keto-OxyLH⁻, (b) AMP, and (c) DLSA.

and OxyL²⁻, were investigated in protein by MS-CASPT2/MM study. According to the experimental conclusion,²⁸ the open-form luciferase with AMP and OxyLH₂ (PDB code: 2D1R.pdb) and the closed-form luciferase with 5'-O-[N-(dehydroluciferyl)-sulfamoyl]adenosine (DLSA) (PDB code: 2D1S.pdb) were used as the initial guesses for the S₀ and S₁ states of the three candidates, respectively, and denoted as model-open and -closed, respectively. For phenolate-keto-OxyLH⁻, the DLSA was replaced by phenolate-keto-OxyLH⁻ + AMP (see Figure 3) in model-closed. The quantum mechanical (QM) and molecular mechanical (MM) calculations were performed on the basis of a modification of MOLCAS³³ and TINKER^{34,35} programs, respectively. For more computational details, please see the Supporting Information.

RESULTS AND DISCUSSION

Structural Variations and Emission Spectra of the Six Possible Emitters. The CASSCF-optimized main geometrical parameters of the S₀ and S₁ states of the chemical forms of OxyLH₂ in vacuum are listed in Table S1, Supporting Information. The predominant variations in geometries from ground to excited state are the -N=C-C=N- unit for keto-OxyLH₂, enol-OxyLH₂, enolate-OxyLH⁻, and OxyL²⁻. Upon excitation to the S₁ state, N₃-C₂' and C₂-N₃ bonds stretch while the C₂'-C₂ bond shortens. For phenolate-keto-OxyLH⁻, O₁₀'-C₆', O₁₁'-C₄, N₃-C₂, and C₂'-C₂ bond distances in the S₁ state are longer than in the S₀ state, but N₃'-C₂' is shorter. For phenolate-enol-OxyLH⁻, O₁₀'-C₆', O₁₁'-C₄, and N₃'-C₂' bond lengths are longer and C₂-N₃, C₂'-C₂ bonds shorter in the S₁ state compared to the S₀ state. The MS-CASPT2 calculated T_e and *f* values of the S₁ states of the six chemical forms of OxyLH₂ in vacuum, DMSO, and protein are listed in Table 1. Their MS-CASPT2 calculated energies, T_v, T_e, and other detailed information are listed in Table 2 and Tables S2, S3, and S4, Supporting Information, respectively.

Exclusion of the Neutral Forms from the Possible Light Emitters. We first discriminate candidates due to the computed emission energies and oscillator strengths. We note that the neutral forms should be excluded from the possible light emitters for the following reasons. First, their MS-CASPT2 calculated T_e values in vacuum and DMSO are much outside the visible light

Table 1. The MS-CASPT2 Calculated Transition Energies T_e (in eV) and *f* Values of the S₁ State of the Chemical Forms of OxyLuciferin

	T _e (<i>f</i>)		
	in vacuum	in DMSO	in protein
keto-OxyLH ₂	3.53 (0.32)	3.50 (0.31)	—
enol-OxyLH ₂	3.52 (0.43)	3.37 (0.45)	—
enolate-OxyLH ⁻	1.76 (0.27)	2.18 (0.19)	—
phenolate-keto-OxyLH ⁻	2.10 (0.71)	2.31 (0.58)	2.15 (0.62)
phenolate-enol-OxyLH ⁻	2.25 (0.62)	2.57 (0.45)	2.18 (0.59)
OxyL ²⁻	2.04 (0.34)	2.20 (0.52)	2.18 (0.19)

Table 2. The MS-CASPT2 Calculated Relative Energies of the Chemical Forms of OxyLH₂ (in kcal/mol) from the Reference -1440.65317045 au

	in vacuum		in DMSO		in protein	
	S ₀	S ₁	S ₀	S ₁	S ₀	S ₁
keto-OxyLH ₂	0.00	80.16	-12.54	69.42	—	—
enol-OxyLH ₂	0.39	83.96	-15.23	64.83	—	—
enolate-OxyLH ⁻	345.45	387.32	285.69	337.39	—	—
phenolate-keto-OxyLH ⁻	327.00	376.83	279.13	333.36	-498.57	-504.48
phenolate-enol-OxyLH ⁻	341.43	393.46	284.83	344.26	-499.96	-507.83
OxyL ²⁻	741.32	791.10	585.14	640.12	-148.51	-154.72

(1.78–3.10 eV), as shown in Table 1. Actually, we previously calculated the T_e value of keto-OxyLH₂ in the wild-type protein by QM/MM method.¹⁰ The predicted T_e value, 3.11 eV (*f* ≈ 0.1), is not within the visible light spectrum. Second, the *f* value of keto-OxyLH₂ is substantially lower than that of the phenolate-keto-OxyLH⁻ in vacuum, DMSO, or protein. Finally, in the intramolecular charge-transfer induced luminescence (ICTIL) mechanism of bioluminescence, the excited light emitter is formed by a charge transfer from an anionic remote π orbital to the σ* of the O–O bond of the dioxetanone.^{3–5} The negative charge on oxygen 10' is essential for effective excited-state charge transfer and lowers the emission energy. This has been demonstrated both theoretically^{3–5} and experimentally.²⁶ All together, clear facts in line with the experimental observations^{12,17} imply that the neutral forms of OxyLH₂ are not the light-emitting species.

Exclusion of Enolate-OxyLH⁻ from the Possible Light Emitters. Although the T_e of enolate-OxyLH⁻ in DMSO is in the range of visible light, the *f* value of enolate-OxyLH⁻ is substantially lower than those of the phenolate-keto-OxyLH⁻ and phenolate-enol-OxyLH⁻ (see Table 1). Moreover, according to Tables S3 and S4 and Figure S1, Supporting Information, the S₁ state of enolate-OxyLH⁻ is formed by a charge transfer from the thiazole ring to the benzothiazole ring, whereas the S₁ states of phenolate-keto-OxyLH⁻ and phenolate-enol-OxyLH⁻ were formed in a contrary way. From the viewpoint of the origin of the excited OxyLH₂, the bioluminescence mechanism requires a negative charge on oxygen 10'.^{3–5,26} So all the structures whose phenol group are protonated are excluded as light emitters of firefly bioluminescence; that is, in addition to the two species

eliminated already by arguments presented in the previous section, we also exclude the enolate-OxyLH⁻ structure.

Three Phenolate Forms of Possible Light Emitters in Protein. Based on the above analysis and the calculated T_e and f values in vacuum and DMSO (see Table 1), the S_1 states of phenolate-keto-OxyLH⁻, phenolate-enol-OxyLH⁻, and OxyL²⁻ are possible light emitters of firefly. To simulate the actual firefly bioluminescence in nature, we further calculated the T_e and f values of the S_1 states of phenolate-keto-OxyLH⁻, phenolate-enol-OxyLH⁻, and OxyL²⁻ in protein (employing the 2D1S and 2D1R pdb X-ray structures in ref 28) by the QM/MM method. The calculated results are listed in Table 1. Their T_e values in protein showed that the S_1 states of phenolate-keto-OxyLH⁻, phenolate-enol-OxyLH⁻, and OxyL²⁻ can emit yellow-green light, which is the natural emission light of firefly. This is in line with the experimental conclusion drawn in 2005 by Ugarova et al.²⁵ This paper concluded that the observed bioluminescence spectrum could be explained as a superposition of these three forms of electronically excited OxyLH₂ (see Figure 1 of ref 25 notice the different nominations). However, the studies up until now are not conclusive with respect to which is/are the most important light emitter(s) involved in the firefly bioluminescence. To be more precise, the previous analysis on this subject^{1,11,12,17,23,24,27} was based on the assumption that the excited-state OxyLH₂ forms studied were the direct products of the firefly dioxetanone decomposition, and conclusions were drawn on the basis of the emitted light in relation to the three candidate emitters. However, the direct and dominant product(s) of the dioxetanone decomposition *in vivo* is/are still not determined. The previous literatures³⁻⁶ tacitly assumed that the direct product of the decomposition of the firefly dioxetanone anion is the phenolate-keto-OxyLH⁻ S_1 state in vacuum. Several studies suggest the same state as the primary product of the chemiexcitation process from the anionic dioxetanone intermediate in vacuum.^{18,36-38} We will below discuss arguments which support the presence of a single decomposition product.

Direct Product of Decomposition of Firefly Dioxetanone Anion *In Vivo*. The X-ray coordinates of the light emitter OxyLH₂ after emission, i.e., in the ground state along with AMP inside the protein were detected experimentally (see the ligand in 2D1R PDB file of ref 28). To analyze and assign the chemical nature of this reaction product, we optimized the S_0 states of our candidate emitters, phenolate-keto-OxyLH⁻, phenolate-enol-OxyLH⁻, and OxyL²⁻, in a protein model based on the 2D1R X-ray structure. The detailed geometrical parameters are listed in Table S4, Supporting Information. Although the low resolution of the X-ray experimental data, the X-ray detected bonds C₄-C₅ (1.50 Å) and O_{11'}-C₄ (1.23 Å) are only compatible with a keto species. The three QM/MM optimized structures, whose principal geometrical parameters are listed in Table S5, Supporting Information, confirm the keto nature of the ligand in the X-ray structure. Furthermore, the QM/MM calculated phenolate-enol-OxyLH⁻ structure agrees with the experiment crystal structure of enol-OxyLH₂ of ref 12. This leads to the conclusion that the product of the deexcitation of the light emitter is the phenolate-keto-OxyLH⁻ S_0 state, implying the nature of the light emitter to be the phenolate-keto-OxyLH⁻ S_1 state.

However, is the phenolate-keto-OxyLH⁻ S_1 state the direct product of the decomposition of firefly dioxetanone or not? The phenolate-keto-OxyLH⁻ S_1 state could come from the tautomerism of the phenolate-enol-OxyLH⁻ S_1 state, if the phenolate-enol-OxyLH⁻ S_1 state is the direct decomposition product.

Alternatively, the phenolate-keto-OxyLH⁻ S_1 state could be derived from an initial OxyL²⁻ S_1 state. A theoretical study on the decomposition process of the firefly dioxetanone in protein can answer these questions. Unfortunately, the reaction is complicated, contains characteristics like charge transfer, nearly degenerated excited states, and probably proton transfer to/from the solvent in protein. Theoretical research cannot presently give a reliable description for this decomposition inside the protein, even in gas phase. Indeed, for the calculation of the firefly dioxetanone decomposition, a big active space beyond regular CASSCF calculation is needed.³⁹ However, we can give a reliable analysis with respect to the stabilities of the possible products of this decomposition, i.e., the stabilities of the phenolate-keto-OxyLH⁻, phenolate-enol-OxyLH⁻ and OxyL²⁻ S_1 states. To investigate this we calculated the energies of these species in vacuum and DMSO by the MS-CASPT2/CASSCF method and in protein by the QM(MS-CASPT2/CASSCF)/MM method. The calculated energies are listed in Table 2. Here we find that the phenolate-keto-OxyLH⁻ S_1 state is 16.63 and 10.91 kcal/mol more stable than the phenolate-enol-OxyLH⁻ S_1 state in vacuum and DMSO, respectively. However, in protein, the phenolate-enol-OxyLH⁻ S_1 state is 3.35 kcal/mol more stable than phenolate-keto-OxyLH⁻ S_1 state. We observed the same inversion for the S_0 state. We checked the optimized structures in protein and found that the phenolate-enol-OxyLH⁻ in the S_0 and S_1 states form H-bonds with the nearby phosphate anion, while phenolate-keto-OxyLH⁻ does not (see Figure S2, Supporting Information). This is in accordance with the observation of Naumov et al.¹² They obtained a crystal structure of isolated enol-OxyLH₂ which assembled as head-to-tail H-bonded dimers. Furthermore, the enol form that was in this study theoretically verified to be stabilized in the dimer as a result of an hydrogen bonding. We can also explain the change of relative stability of the species inside the protein compared to the *in gas* calculation by the below presented small model calculations. We optimized the geometries (see Figure S3, Supporting Information) and calculated the energies of the phenolate-keto-OxyLH⁻ and phenolate-enol-OxyLH⁻ S_1 states in vacuum, with one H₂O and H₂PO₄⁻ molecule nearby O_{11'} by the TD CAM-B3LYP method. As Table S6, Supporting Information, showed, the energy difference between the phenolate-keto-OxyLH⁻ and phenolate-enol-OxyLH⁻ S_1 states decreased to almost zero by the H-bonding network created by the presence of H₂PO₄⁻. The surrounding residues and AMP presence in the enzyme cavity leads to a stabilization of the phenolate-enol-OxyLH⁻ S_1 state as compared to the phenolate-keto-OxyLH⁻. Regardless of that the tautomerization of enol to keto forms would have to overcome the activation barrier, the thermal stability of the enol form over the keto form, in both the excited and ground state, suggests that this reaction path is closed to the formation of a phenolate-keto-OxyLH⁻ species, as observed in the X-ray experiment. Hence, a hypothesis of the phenolate-enol-OxyLH⁻ S_1 state being the direct product of the dioxetanone decomposition has to be rejected. In the hypothesis of OxyL²⁻ being the direct product of the dioxetanone decomposition, one should note that the OxyL²⁻ S_1 state would not directly react to the phenolate-keto-OxyLH⁻ S_1 state but rather first form the phenolate-enol-OxyLH⁻ S_1 state by a proton transfer from the solvent (proton transfer with the solvent is less than 10 ns, the lifetime of excited state is 1 ns). Furthermore, the f value of OxyL²⁻ S_1 state is much smaller than the f values of the phenolate-keto-OxyLH⁻ and phenolate-enol-OxyLH⁻ S_1 states. This is another

argument which disqualifies the dianion as the light emitter, as firefly has large bioluminescence efficiency. Thus the suggestion that the direct products are a phenolate-enol-OxyLH⁻ or OxyL²⁻ S₁ state has to be excluded. Hence, we conclude that the direct product of firefly dioxetanone decomposition in vivo is the phenolate-keto-OxyLH⁻ S₁ state. The T_e value of the phenolate-keto-OxyLH⁻ S₁ state in protein is predicted by the QM/MM calculation to be 2.15 eV (see Table 1), which exactly corresponds to the firefly naturally produced yellow-green light, and the multicolor bioluminescence phenomenon is mainly tuned by the polarization of the microenvironment, which has been discussed theoretically and experimentally.^{10,12,17,40–42}

CONCLUSION

This is the first systematic theoretical investigation on all the possible light emitters of firefly using a multireference method. Six chemical forms of oxyluciferin (OxyLH₂) molecules/anions were studied by the MS-CASPT2 method in vacuum and DMSO. Based on the T_e, f values, and ICTIL mechanism of bioluminescence, enol-OxyLH₂, keto-OxyLH₂ and enolate-OxyLH⁻ were excluded as possible light emitters. The three remaining candidates, phenolate-keto-OxyLH⁻, phenolate-enol-OxyLH⁻ and OxyL²⁻, were further investigated in protein by the MS-CASPT2/MM calculations to explain the natural bioluminescence of firefly. They are all possible light emitters of firefly if just based on the MS-CASPT2/MM calculated T_e values of 2.15, 2.18, and 2.18 eV, respectively. By comparison of the MS-CASPT2/MM calculated geometries and stabilities of phenolate-enol-OxyLH⁻, phenolate-keto-OxyLH⁻, and OxyL²⁻ with the experimental observed crystal data and the detailed analysis on the possibility of the tautomerization reactions among them in protein, we concluded that the direct decomposition product of firefly dioxetanone and the only light emitter of firefly bioluminescence is the phenolate-keto-OxyLH⁻ S₁ state. The MS-CASPT2/MM results are proved to be the reference on which all the other calculations related to bioluminescence should be compared with.

ASSOCIATED CONTENT

S Supporting Information. Selected active CASSCF orbitals of six OxyLH₂ molecules/anions, H-bonding network by the CASSCF/MM optimization, and TD CAM-B3LYP optimized structures were presented in Figures S1, S2 and S3, respectively. The CASSCF optimized geometries and T_v, T_e and charge densities, the CASSCF/MM optimized geometries in protein, and the TD-DFT calculated energy differences between the S₁ states of phenolate-enol-OxyLH⁻ and phenolate-keto-OxyLH⁻ were summarized in Tables S1–S6. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*E-mail: roland.lindh@kvac.uu.se (R. L.), yajun.liu@bnu.edu.cn (Y.-J. L.), fangwh@bnu.edu.cn (W.-H. F.).

ACKNOWLEDGMENT

This study was supported by the National Nature Science Foundation of China (grant nos. 21073017, 20873010, and 20720102038), the Major State Basic Research Development

Programs (grant nos. 2011CB808500 and 2007CB815206), and Fundamental Research Funds for the Central Universities. The Swedish Research Council (V.R.) has also provided assistance. I. N. thanks Prof H.M. Marques for funding through the DST/NRF SARChI initiative.

REFERENCES

- (1) Orlova, G.; Goddard, J. D.; Brovko, L. Y. *J. Am. Chem. Soc.* **2003**, *125*, 6962–6971.
- (2) Reddy, G. R.; Thompson, W. C.; Miller, S. C. *J. Am. Chem. Soc.* **2010**, *132*, 13586–7.
- (3) Liu, F.; Liu, Y.; De Vico, L.; Lindh, R. *Chem. Phys. Lett.* **2009**, *484*, 69–75.
- (4) Liu, F.; Liu, Y.; De Vico, L.; Lindh, R. *J. Am. Chem. Soc.* **2009**, *131*, 6181–6188.
- (5) Chung, L.; Hayashi, S.; Lundberg, M.; Nakatsu, T.; Kato, H.; Morokuma, K. *J. Am. Chem. Soc.* **2008**, *130*, 12880–12881.
- (6) De Vico, L.; Liu, Y.; Krogh, J.; Lindh, R. *J. Phys. Chem. A* **2007**, *111*, 8013–8019.
- (7) Ugarova, N. N.; Brovko, L. Y. *Luminescence* **2002**, *17*, 321–330.
- (8) Branchini, B.; Southworth, T.; Murtiashaw, M.; Boije, H.; Fleet, S. *Biochemistry* **2003**, *42*, 10429–10436.
- (9) Branchini, B.; Southworth, T.; Murtiashaw, M.; Magyar, R.; Gonzalez, S.; Ruggiero, M.; Strohs, J. *Biochemistry* **2004**, *43*, 7255–7262.
- (10) Navizet, I.; Liu, Y. J.; Ferre, N.; Xiao, H. Y.; Fang, W. H.; Lindh, R. *J. Am. Chem. Soc.* **2010**, *132*, 706–712.
- (11) Nakatani, N.; Hasegawa, J. Y.; Nakatsuji, H. *J. Am. Chem. Soc.* **2007**, *129*, 8756–8765.
- (12) Naumov, P.; Ozawa, Y.; Ohkubo, K.; Fukuzumi, S. *J. Am. Chem. Soc.* **2009**, *131*, 11590–11605.
- (13) White, E.; Rapaport, E.; Seliger, H.; Hopkins, T. *Bioorg. Chem.* **1971**, *1*, 92–122.
- (14) Gandelman, O.; Brovko, L.; Ugarova, N.; Chikishev, A.; Shkurimov, A. *J. Photochem. Photobiol. B* **1993**, *19*, 187–191.
- (15) White, E.; Steinmetz, M.; Miano, J.; Wildes, P.; Morland, R. *J. Am. Chem. Soc.* **1980**, *102*, 3199–3208.
- (16) Morton, R.; Hopkins, T.; Seliger, H. *Biochemistry* **1969**, *8*, 1598–1607.
- (17) Hirano, T.; Hasumi, Y.; Ohtsuka, K.; Maki, S.; Niwa, H.; Yamaji, M.; Hashizume, D. *J. Am. Chem. Soc.* **2009**, *131*, 2385–2396.
- (18) Hopkins, T.; Seliger, H.; White, E.; Cass, M. *J. Am. Chem. Soc.* **1967**, *89*, 7148–7150.
- (19) White, E. H.; Branchini, B. R. *J. Am. Chem. Soc.* **1975**, *97*, 1243–1245.
- (20) White, E. H.; Roswell, D. F. *Photochem. Photobiol.* **1991**, *53*, 131–136.
- (21) Branchini, B. R.; Magyar, R. A.; Murtiashaw, M. H.; Anderson, S. M.; Helgerson, L. C.; Zimmer, M. *Biochemistry* **1999**, *38*, 13223–13230.
- (22) Liu, Y. J.; De Vico, L.; Lindh, R. *J. Photochem. Photobiol. A* **2008**, *194*, 261–267.
- (23) Branchini, B. R.; Southworth, T. L.; Murtiashaw, M. H.; Magyar, R. A.; Gonzalez, S. A.; Ruggiero, M. C.; Stroh, J. G. *Biochemistry* **2004**, *43*, 7255–7262.
- (24) Branchini, B. R.; Murtiashaw, M. H.; Magyar, R. A.; Portier, N. C.; Ruggiero, M. C.; Stroh, J. G. *J. Am. Chem. Soc.* **2002**, *124*, 2112–2113.
- (25) Ugarova, N. N.; Maloshenok, L. G.; Uporov, I. V.; Koksharov, M. I. *Biochemistry (Moscow)* **2005**, *70*, 1262–1267.
- (26) Naumov, P.; Kochunnonny, M. *J. Am. Chem. Soc.* **2010**, *132*, 11566–11579.
- (27) Min, C. G.; Ren, A. M.; Guo, J. F.; Zou, L. Y.; Goddard, J. D.; Sun, C. C. *ChemPhysChem* **2010**, *11*, 2199–2204.
- (28) Nakatsu, T.; Ichiyama, S.; Hiratake, J.; Saldanha, A.; Kobashi, N.; Sakata, K.; Kato, H. *Nature* **2006**, *440*, 372–376.
- (29) Milne, B. F.; Marques, M. A. L.; Nogueira, F. *Phys. Chem. Chem. Phys.* **2010**, *12*, 14285–14293.

- (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. *J. Chem. Phys.* **1992**, *96*, 1218–1226.
- (32) Roos, B. O.; Lindh, R.; Malmqvist, P.-Å.; Veryazov, V.; Widmark, P.-O. *J. Phys. Chem. A* **2004**, *108*, 2851–2858.
- (33) Karlström, G.; Lindh, R.; Malmqvist, P.-Å.; Roos, B. O.; Ryde, U.; Veryazov, V.; Widmark, P.-O.; Cossi, M.; Schimmelpfennig, B.; Neogrady, P.; Seijo, L. *Comput. Mater. Sci.* **2003**, *28*, 222–239.
- (34) Aquilante, F.; De Vico, L.; Ferré, N.; Ghigo, G.; Malmqvist, P.; Pedersen, T.; Pitonak, M.; Reiher, M.; Roos, B. *J. Comput. Chem.* **2010**, *31*, 224–247.
- (35) Ponder, J. W. *TINKER, Software Tools for Molecular Design*, version 4.2; Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine: St. Louis, MO, 2004; <http://dasher.wustl.edu/tinker>.
- (36) Wannlund, J.; DeLuca, M.; Stempel, K.; Boyer, P. *Biochem. Biophys. Res. Commun.* **1978**, *81*, 987–992.
- (37) Shimomura, O.; Goto, T.; Johnson, F. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 2799–2802.
- (38) McCapra, F.; Chang, Y.; Francois, V. *Chem. Commun. (London)* **1968**, 1968, 22–23.
- (39) Greenman, L.; Mazziotti, D. A. *J. Chem. Phys.* **2010**, *133*, 164110.
- (40) Maghami, P.; Ranjbar, B.; Hosseinkhani, S.; Ghasemi, A.; Moradi, A.; Gill, P. *Photochem. Photobiol. Sci.* **2010**, *9*, 376–383.
- (41) Neto, A. J. S.; Scorsato, V.; Arnoldi, F. G. C.; Viviani, V. R. *Photochem. Photobiol. Sci.* **2009**, *8*, 1748–1754.
- (42) Moradi, A.; Hosseinkhani, S.; Naderi-Manesh, H.; Sadeghizadeh, M.; Alipour, B. S. *Biochemistry* **2009**, *48*, 575–582.

Stability of Hydrocarbons of the Polyhedrane Family Containing Bridged CH Groups: A Case of Failure of the Colle–Salvetti Correlation Density Functionals. [*Journal of Chemical Theory and Computation*, 2010, 6, 3442]. Grigory A. Shamov* Georg Schreckenbach and Peter H. M. Budzelaar

With this Erratum, we would like to correct our paper published in the November 2010 issue of the *Journal of Chemical Theory and Computation*.¹ In the paper, we have used MP2 and one of its spin-scaled variants (first invented by Grimme²), as benchmark values for determining the performance of a variety of density functionals. In our paper,¹ in the method labeled as “SOS-MP2”, we have (inadvertently) used an opposite-spin correlation energy scaling factor of 2.0 instead of the value of 1.3 recommended by Head-Gordon and co-workers.³ We apologize to the readers of the *Journal of Chemical Theory and Computation* for any inconvenience this error might have caused. The results in our paper (we label them as “2SOS-MP2” here to distinguish from the original one) thus had some overestimation of the correlation energy for all the compounds, acetylene to **1–8**. Unfortunately, the method was used as a base method to present and compare MP2 and DFT results against. To avoid the confusion resulting from the fact that our benchmarking method was not what it was called in the paper, we would like to provide the results relative to a correct spin-scaling scheme. Among several proposed^{2–6} scaling schemes for MP2, the scheme by King,⁶ with the same-spin scaling of 0.5 and the opposite-spin scaling of 1.28, correspondingly, was chosen, since it was shown to yield the best agreement with accurate CCSD(T) results for ethylene dimer potential energy surfaces. We label the scheme SCS-MP2(King) below.

We provide the corrected versions of Tables 2–4, substituting the wrong 2SOS-MP2 values with the SCSMP2(King) ones. The results incorrectly called “SOS-MP2” in our paper, now labeled 2SOS-MP2, are retained/provided for comparison. We also present the corrected Figure 2, with MAD’s for the density functionals plotted against the SCS-MP2(King) values, instead of the 2SOS-MP2. The following corrections to the discussion of results have to be made.

In all the occurrences throughout the article, SCS-MP2(King) should be read instead SOS-MP2 (Page 3447, right column, last paragraph should read “We have also computed canonical MP2 as spin-scaled MP2, SCS-MP2(King)”; page 3452, left column, paragraph 4, should be “closer to SCS-MP2(King) values”; same page, second paragraph, “the largest difference from SCS-MP2-(King) isomerization energies”; page 3453, right column, second and fourth paragraphs, etc.).

The statement on the page 3448, first paragraph, “SOS-MP2 energy differences are closer to CCSD(T) than the SCS-MP2” is incorrect and should be dropped.

On the page 3450, second last paragraph: The RHF method for the binding energies of 1a-4a from the acetylene has now MAD of 6.44, which is not “under 5 kcal/mol”; the functionals enumerated next however still perform worse than RHF (Table 2).

On page 3451, right column, the first paragraph: “HFB and HFG, which have MADs of about 24 kcal/mol” instead of “22 kcal/mol”.

On page 2452, left column, second paragraph, MAD for the RHF methods should be 22.03, for OLYP 15.24 kcal/mol, correspondingly.

Table 2. Binding Enthalpies Per Acetylene Monomer ΔE_b ^a of Polyhedranes 1a–4a (kcal/mol)^b

method	1a	2a	3a	4a	MAD
SCS-MP2(King)	−46.0	−50.1	−54.6	−53.0	
2SOS-MP2	1.7	1.5	1.6	1.5	1.6
MP2	−2.7	−2.6	−2.6	−2.6	2.6
B2PLYP(0)	3.9	4.1	4.2	4.5	4.2
B2PLYP(2)	2.8	2.9	3.0	3.1	3.0
HFB	29.5	30.8	31.9	32.4	31.1
HFO	21.6	23.6	25.3	26.0	24.1
HFG	29.6	31.0	32.3	32.9	31.4
M06-L	−2.5	−1.2	−0.5	−0.5	1.2
M06-2X	−2.1	−1.5	−1.3	−1.4	1.6
KT3	−8.6	−8.7	−8.3	−8.2	8.4
B86LYP	12.1	12.4	12.9	13.0	12.6
BLYP ^c	10.9	11.3	11.8	11.9	11.5
BLYP-D ^c	7.0	6.2	6.3	6.0	6.4
BPBE ^c	0.7	1.5	2.2	2.4	1.7
BPBE-D ^c	−2.6	−3.0	−2.6	−2.8	2.8
OLYP ^c	3.7	5.1	6.2	6.5	5.4
OLYP-D ^c	−0.2	−0.1	0.6	0.5	0.3
PBE ^c	−3.8	−3.4	−2.9	−2.9	3.2
PBE-D ^c	−6.2	−6.6	−6.4	−6.6	6.5
BW	23.2	24.2	25.1	25.5	24.5
RHF	6.1	6.4	6.5	6.9	6.4
BLYP-LC ^d	−8.1	−7.2	−8.5	−8.5	8.1
GLYP-LC ^d	−9.0	−7.8	−9.4	−9.4	8.9
BOP-LC ^d	−5.8	−5.8	−5.7	−5.5	5.7
BOP ^d	13.4	14.1	14.8	15.1	14.3
B97 ^d	−0.4	−1.5	−0.3	−0.1	0.6
B97-D ^d	4.8	0.5	4.6	4.3	3.6
O2PLYP(2)	−0.7	−0.2	0.2	0.4	0.4
O2PLYP(0)	0.2	0.8	1.2	1.6	1.0
GLYP	11.3	11.8	12.5	12.8	12.1
oBLYP	8.2	8.4	8.9	9.0	8.6
oBLYP-D	4.6	4.7	4.8	4.6	4.7
BPW92	19.1	19.9	20.7	21.0	20.2
OPBE	−8.6	−7.2	−6.0	−5.5	6.8
PBEsol	−12.8	−12.8	−12.5	−12.5	12.7
B2PLYP-D	0.4	0.5	0.4	0.3	0.4
B3LYP-Cpot ^d	−9.0	−9.3	−10.5	−10.4	9.8
B3LYP	4.9	5.1	5.4	5.5	5.2
RGE2	−3.8	−3.5	−3.0	−2.9	3.3
O3LYP	−8.0	−7.6	−7.2	−7.0	7.5
TCA	11.3	12.8	13.8	14.0	13.0

^a $\Delta E_b = (E(C_2H_2)_m - mE(C_2H_2))/m$. ^b The SCS-MP2(King) are actual values (*in italics*); for the other methods, and differences with respect to SCS-MP2(King) values are shown. ^c Priroda L22 energies on optimized geometries a DFT/L11 level. ^d GAMESS-US cc-CVTZ energies on Priroda MP2/L2 geometries.

Table 3. Relative Energies of Hydrogenated Species of 1a, 5a–c, and 7 with Respect to Tricyclododecatetraene 1b and Its Hydrogenated Derivatives 6 and 8, Correspondingly (kcal/mol)^a

method	5a → 6	5b → 6	5c → 6	7 → 8	MAD
SCS-MP2(King)	35.8	11.2	-5.33	47.23	
2SOS-MP2	3.54	3.23	3.31	2.57	3.36
MP2	-0.78	-1.29	-0.11	-0.55	0.82
B2PLYP(0)	-14.66	-9.94	-13.59	-11.55	13
B2PLYP(2)	-11.57	-7.95	-9.87	-9.08	10.22
HFB	-47.47	-36.24	-39.18	-38.17	42.87
HFO	-28.48	-19.02	-21.34	-25.53	24.52
HFG	-43.99	-33.56	-36.3	-35.41	39.68
M06-L	-6.14	2.29	0.65	-10.2	4.24
M06-2X	-5.3	1.47	-0.94	-8.19	3.52
KT3	0.59	3.58	2.4	-0.34	2.04
B86LYP	-31.27	-22.5	-25.22	-25.19	27.74
BLYP ^c	-28.19	-39.17	-3.72	-23.09	25.12
BLYP-D ^c	-26.24	-36.98	-1.88	-21.28	23.24
BPBE ^c	-9.44	-21.48	12.98	-9.95	12.25
BPBE-D ^c	-7.9	-19.73	14.55	-8.56	11.39
OLYP ^c	-10.18	-21.96	12.68	-11.24	12.71
OLYP-D ^c	-8.38	-19.92	14.34	-9.58	11.66
PBE ^c	-7.06	-19.57	14.79	-7.84	10.84
PBE-D ^c	-6.04	-18.41	15.91	-6.94	10.29
BW	-40.61	-30.23	-33.27	-32.65	36.4
RHF	-13.29	-10.12	-14.66	-9.76	12.37
BLYP-LC ^b	0.94	3.8	1.82	-1.24	2.5
GLYP-LC ^b	2.14	4.68	2.72	-0.25	3.17
BOP-LC ^b	1.31	4.53	2.52	-1.44	3.12
BOP ^b	-28.21	-19.83	-22.16	-23.06	24.73
B97 ^b	-10.42	-6.85	-8.87	-8.34	8.92
B97-D ^b	-21.97	-12.68	-15.31	-19.32	18.45
O2PLYP(2)	-2.57	0.31	-1.28	-3.18	1.59
O2PLYP(0)	-5.27	-1.24	-4.5	-5.43	3.84
GLYP	-24.67	-17.48	-19.79	-19.85	21.74
oBLYP	-25.12	-17.64	-20.02	-20.24	22.11
oBLYP-D	-28.08	-17.35	-19.07	-23.56	21.8
BPW92	-36.2	-26.35	-29.53	-29.1	32.25
OPBE	12.04	15.09	14.47	5.4	13.43
PBEsol	4.97	7.95	7.21	1.72	6.21
B2PLYP-D	-13.14	-7.41	-9.14	-10.8	9.7
B3LYP-Cpot ^b	0.12	0.27	-1.48	0.6	0.72
B3LYP	-18.59	-12.4	-15.29	-15.11	16.21
RGE2	-6.69	-1.41	-2.85	-7.31	4.53
O3LYP	1.66	5.39	3.59	-0.79	3.44
TCA	-24.15	-13.86	-16.27	-22.48	20.02

^aThe SCS-MP2(King) are actual values (*in italics*), while for other methods differences with respect to it are shown. ^bPriroda L22 energies on optimized geometries a DFT/L11 level. ^cGAMESS-US cc-CVTZ energies on Priroda MP2/L2 geometries.

On page 3452, left column, fourth paragraph, we stated “The OPTX-based double hybrid, in contrast to B2PLYP, performs very well also. Among the pure, uncorrected GGAs, the KT3 functional is the best”. This should be modified as follows. The KT3 functional yields a reasonably good performance for

Table 4. Calculated Relative Energies of Isomers of Compounds 1–4 (kcal/mol)^a

method	1a → 1e	1a → 1d	1b → 1a	1b → 1c	2b → 2a	2b → 2c	3b → 3a	3b → 3c	4b → 4a	4b → 4c
	SCS-MP2(King)	-41.5	26.0	-37.1	5.1	-55.4	-48.4	-84.0	-40.8	-69.9
MP2	6.7	-1.4	1.9	0.3	-0.6	4.0	-1.4	-0.5	-0.3	-2.7
2SOS-MP2	-0.6	4.2	-4.5	-0.1	-4.3	-6.9	-5.1	-9.0	-5.3	-12.0
B2PLYP(0)	-25.4	-15.3	15.1	-1.9	21.8	8.1	27.6	27.7	21.5	39.5
B2PLYP(2)	-20.6	-12.6	12.3	-2.2	16.6	1.1	21.0	11.8	15.4	14.9
HFB	-70.8	-53.3	51.3	-15.0	68.0	-1.7	87.1	30.4	73.0	42.4
HFO	-43.2	-28.2	28.0	-11.4	40.6	5.0	56.6	35.5	42.0	50.6
HFG	-66.7	-49.2	47.6	-14.1	63.1	0.2	81.5	31.9	67.0	42.8
M06-L	-10.8	-1.9	1.4	-2.8	11.3	-3.7	22.8	5.5	16.3	2.5
M06-2X	-4.0	-1.7	1.4	-1.4	10.0	3.7	14.2	11.1	11.9	16.9
KT3	-2.7	3.3	-1.4	2.6	-6.0	5.8	-5.5	13.3	-14.9	13.3
B86LYP	-47.7	-34.5	33.4	-8.1	43.3	-1.2	55.0	19.8	43.4	25.9
BLYP ^b	-42.7	-31.4	30.0	-8.8	39.3	-0.1	49.7	18.3	38.7	24.0
BLYP-D ^b	-35.0	-29.8	27.2	-5.0	28.1	-6.0	35.8	-0.1	30.1	0.9
BPBE ^b	-14.9	-7.4	8.3	-4.3	13.0	5.1	19.7	17.3	9.5	20.1
BPBE-D ^b	-8.2	-6.2	6.0	-0.7	3.1	-0.3	7.3	1.0	1.8	-0.1
OLYP ^b	-17.0	-7.5	8.3	-7.0	15.3	6.3	24.1	24.1	13.2	34.1
OLYP-D ^b	-9.4	-6.1	5.8	-2.3	3.7	0.5	9.6	5.5	4.0	11.0
PBE ^b	-11.2	-4.9	6.0	-2.2	9.3	4.4	14.2	13.4	5.3	14.4
PBE-D ^b	-6.4	-4.2	4.4	-0.2	2.1	0.3	5.3	1.5	-0.2	0.1
BW	-61.8	-45.2	43.4	-12.5	57.5	-1.1	73.4	27.6	59.3	36.9
RHF	-21.1	-14.0	13.8	-0.6	21.6	17.3	27.8	42.7	24.9	64.2
BLYP-LC ^c	5.2	4.7	-3.4	2.7	8.1	12.7	-2.3	18.1	-5.0	22.7
GLYP-LC ^c	6.8	6.1	-4.7	3.1	8.4	13.0	-4.6	18.0	-7.2	22.5
BOP-LC ^c	5.7	5.8	-4.5	1.8	-2.5	13.8	-2.1	21.5	-6.2	27.0
BOP ^c	-43.4	-30.4	29.6	-8.3	39.3	1.4	50.8	23.4	37.8	28.9
B97 ^c	-17.5	-10.1	10.6	-1.8	2.9	6.0	14.5	19.9	5.0	24.9
B97-D ^c	-28.4	-23.0	21.1	-5.4	-9.1	-6.5	29.5	0.5	23.8	3.8
O2PLYP(2)	-7.5	-0.6	1.3	-0.4	3.6	4.8	6.7	14.8	0.7	19.3
O2PLYP(0)	-11.5	-2.8	3.5	-0.1	8.0	11.2	12.2	29.5	5.5	42.3
GLYP	-39.9	-26.9	26.7	-6.0	34.4	2.1	44.7	21.3	32.8	24.4
oBLYP	-39.4	-27.5	27.1	-5.7	34.8	0.4	44.3	17.7	33.5	20.8
oBLYP-D	-28.1	-20.9	19.6	-5.7	26.0	-5.8	31.6	3.5	25.3	7.7
BPW92	-56.4	-40.1	38.3	-10.9	51.0	-0.9	65.0	25.8	50.9	33.7
OPBE	14.7	20.1	-16.8	2.7	-18.0	12.9	-15.2	22.6	-25.6	26.1
PBEsol	5.8	9.2	-6.6	3.9	-8.1	7.0	-6.9	8.8	-13.6	4.7
B2PLYP-D	-12.9	-8.0	7.0	-1.5	11.3	-2.4	14.3	2.6	11.4	4.6
B3LYP	-0.1	1.1	-5.1	1.4	-3.5	0.0	-15.6	-1.6	-16.9	-5.5
Cpot ^c										
B3LYP	-31.0	-19.7	19.5	-3.6	26.3	3.2	33.5	20.5	24.6	26.3
RGE2	-11.2	-4.4	5.6	-0.7	8.1	4.4	13.0	13.8	3.9	13.8
O3LYP	-1.5	5.7	-3.9	2.1	-2.6	8.0	0.1	17.2	-7.8	19.9
TCA	-35.7	-23.3	22.9	-8.0	36.0	1.5	49.0	23.4	37.2	31.4

^a See Scheme 2 in the original paper for a description of the isomers. The SCS-MP2(King) values are the actual isomerization energies (*in italics*); for the other methods, differences from the SCS-MP2(King) values are shown. ^bGAMESS-US cc-CVTZ energies on Priroda MP2/L2 geometries. ^cPriroda L22 energies on optimized geometries a DFT/L11 level.

both sets, but RGE2 and PBEsol have very close MADs to it. The former functional, KT3, is slightly more accurate for the polyhedrane set while the two latter are slightly more accurate for the cyclophanes; overall performance is thus similar. As to comparison of the O2PLYP with B2PLYP and especially B2PLYP-D, the former double hybrid has shown better agreement with SCS-MP2(King) for the polyhedranes, the latter two, especially B2PLYP-D, have now much smaller MADs for the cyclophanes (under 3 kcal/mol for the B2PLYP-D, about 10 kcal/mol for O2PLYP).

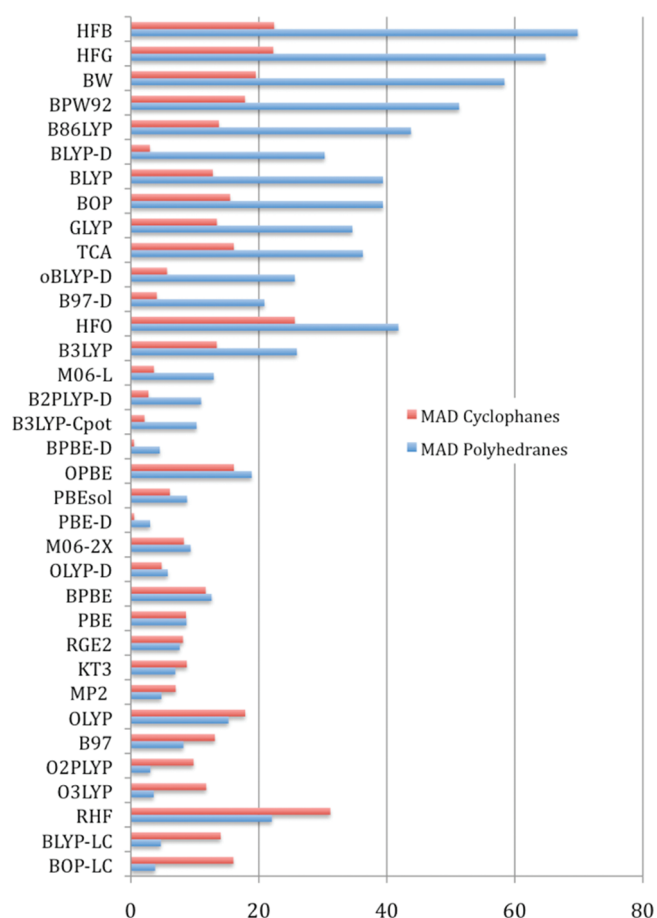


Figure 2. Mean absolute differences (MAD) for cyclophane and polyhedrane isomerization, relative to SCS-MP2(King) computed values, kcal/mol.

On page 3454, first paragraph, should read “among the pure GGA functionals, KT3 and RGE2 functionals have shown good performance on both sets of molecules.”

ACKNOWLEDGMENT

We thank Professor Dr. Stefan Grimme, University of Muenster, for notifying us of the existence of a problem with our “SOS-MP2” values.

REFERENCES

- (1) Shamov, G. A.; Schreckenbach, G.; Budzelaar, P. H. M. *J. Chem. Theory Comput.* **2010**, *6*, 3442.
- (2) Grimme, S. *J. Chem. Phys.* **2003**, *118*, 9095.
- (3) Jung, Y. S.; Lochan, R. C.; Dutoi, A. D.; Head-Gordon, M. *J. Chem. Phys.* **2004**, *121*, 9793.
- (4) Distasio, R. A.; Head-Gordon, M. *Mol. Phys.* **2007**, *105*, 1073.
- (5) Hill, J. G.; Platts, J. A. *J. Chem. Theory Comput.* **2007**, *3*, 80.
- (6) King, R. A. *Mol. Phys.* **2009**, *107*, 789.

DOI: 10.1021/ct200004s

Published on Web 02/02/2011

On the Structure and Geometry of Biomolecular Binding Motifs (Hydrogen-Bonding, Stacking, X-H $\cdots\pi$): WFT and DFT Calculations [*Journal of Chemical Theory and Computation* 2010, 6, 66–80. DOI: 10.1021/ct900376r]. Kevin E. Riley,* Michal Pitoňák, Jiř Cerný, and Pavel Hobza*

It is stated in the article that, “[T]he incorrect long-range behavior of the M06-2X functional is due to the fact that the dispersion energy was covered by reparameterization of the exchange functional and not by the correlation one.” This statement is incorrect; this functional describes dispersion-like interactions using the correlation (and not the exchange) term. The proper explanation for the poor behavior of this functional at long-range (in the nonoverlap region) lies in the functional’s inherently local nature.

DOI: 10.1021/ct200019g

Published on Web 01/27/2011