FEATURE ARTICLE

Development of Effective Quantum Mechanical/Molecular Mechanical (QM/MM) Methods for Complex Biological Processes

Demian Riccardi,[†] Patricia Schaefer,[†] Yang Yang,[†] Haibo Yu,[†] Nilanjan Ghosh,[†] Xavier Prat-Resina,[†] Peter König,[‡] Guohui Li,[§] Dingguo Xu,[||] Hua Guo,[||] Marcus Elstner,[‡,⊥] and Qiang Cui*[,†]

Department of Chemistry and Theoretical Chemistry Institute, University of Wisconsin, Madison, 1101 University Ave, Madison, Wisconsin 53706, Theoretische Physik, Universität Paderborn, Warburger Str. 100, 33098 Paderborn, Germany, Department of Pathology, Harvard Medical School, 200 Longwood Ave. Massachusetts 02115, Department of Chemistry, University of New Mexico, Albuquerque, New Mexico 87131, and Department of Molecular Biophysics, German Cancer Research Center D-69115 Heidelberg, Germany

Received: November 3, 2005; In Final Form: December 23, 2005

Motivated by the long-term goal of understanding vectorial biological processes such as proton transport (PT) in biomolecular ion pumps, a number of developments were made to establish combined quantum mechanical/molecular mechanical (QM/MM) methods suitable for studying chemical reactions involving significant charge separation in the condensed phase. These developments were summarized and discussed with representative problems. Specifically, free energy perturbation and boundary potential methods for treating long-range electrostatics were implemented to test the robustness of QM/MM results for protein systems. It was shown that consistent models with sufficient sampling were able to produce quantitatively satisfactory results, such as pK_a for titritable groups in the interior of T4-lysozyme, while an inconsistent treatment of electrostatics or lack of sufficient sampling may produce incorrect results. Modifications were made to an approximate density functional theory (SCC-DFTB) to improve the description of proton affinity and hydrogen-bonding, which are crucial for the treatment of PT in polar systems. Test calculations on water autoionization showed clearly that both improvements are necessary for quantitatively reliable results. Finally, the newly established SCC-DFTB/MM-GSBP protocol was used to explore mechanistic issues in carbonic anhydrase (CA). Preliminary results suggest that PT in CA occurs mainly through short water wires containing two water molecules in a thermally activated fashion. Although longer water wires occur with similar frequencies, PT along those pathways, on average, has substantially higher barriers, a result not expected based on previous studies. The fluctuations of water molecules peripheral to the water wire were found to make a larger impact on the PT energetics compared to polar protein residues in the active site, which are largely pre-organized and therefore have less tendency to reorganize during the reaction.

1. Introduction

Chemistry is an integral part of most fundamental biological processes. In addition to the large number of reactions catalyzed by enzymes in metabolic pathways, other well known examples include nucleotide (e.g., ATP, GTP) hydrolysis and various proton/electron transfer reactions involved in the production and transduction of bioenergy and signals.[1] Various chemical reactions are also implicated in the replication, remodeling and repair of genomes as well as the transcription and translation of genetic information into proteins. As a result, the analysis of chemical reactions in the biological context is an important component in the mechanistic investigation of biological processes. Recent revolutions in genome sequencing and structural biology[2,3] have provided a wealth of information that indeed makes such analysis possible for many important biological problems. Considering the inherent complexity of biological processes, the role of theoretical and computational analysis can hardly be overemphasized. To make a meaningful contribution, however, effective theoretical and computational methods need to be developed and carefully tested, tailored to the specific questions of interest.

In this context, combined quantum mechanical and molecular mechanical (QM/MM) methods[4−6] have become popular in recent years. As summarized in a number of excellent review articles,[7−9] rapid progress has been made to make QM/MM methods increasingly more efficient[10] and reliable (also see below). For enzyme reactions that involve highly localized chemical processes, well-calibrated QM/MM methods have been tremendously valuable in the analysis of reaction mechanisms[11,12] in determining factors that govern the catalytic

* To whom correspondence should be addressed: cui@chem.wisc.edu.
[†] University of Wisconsin, Madison.
[‡] Universität Paderborn.
[§] Harvard Medical School.
[||] University of New Mexico.
[⊥] German Cancer Research Center.

10.1021/jp056361o CCC: $33.50
Published on Web 02/04/2006

Demian Riccardi graduated from Vassar College in 2000 with a concentration in biochemistry. He then spent one year as a cooperative student in an environmental chemistry group at IBM-East Fishkill, NY, before moving on to graduate school in Wisconsin. Currently, long-range proton transfer in carbonic anhydrase II is the love of his life, and, if all goes well, he will graduate in the summer of 2006.

Patricia Schaefer graduated with a B.S. in Chemistry and Mathematics from Marquette University in 2002. She then attended the University of Wisconsin-Madison, where she worked under Professor Qiang Cui studying QM/MM interactions in biomolecular systems. She obtained an M.S. in Chemistry in 2005, and is currently teaching in the General Chemistry Department at UW-Madison and at a private high school. She hopes to help enhance the quality of math and science education through working with high school teachers on preservice training and professional development.

Yang Yang graduated with a B.S. from the Nankai University, P. R. China, in 2000. After one-year of graduate study in the Central Laboratory of the Nankai University with Professor Shaofan Lin, he attended The Ohio State University to theoretically study lanthanide spectroscopy using relativistic quantum chemistry under the supervision of Professor R. M. Pitzer. After obtaining an M.S. in 2004, Yang moved to the University of Wisconsin-Madison as a graduate student in the Cui group. He is currently focusing on developing QM/MM methods for analyzing phosphate chemistry in myosin and other enzymes.

Haibo Yu received his B.S. in Biological Sciences from the University of Science and Technology of China in 2000. After that he was working with Professor Wilfred F. van Gunsteren at ETH Zurich, Switzerland, and received his doctoral degree in 2004. From 2005 he has been working as a research associate with Professor Qiang Cui at the University of Wisconsin-Madison. His research interests center on the physical fundamentals of the structure, dynamics, and function of biomolecules (including molecular motors, enzymes).

Nilanjan Ghosh is a third year graduate student at the University of Wisconsin-Madison working under the supervision of Professor Qiang Cui. He received his B.S. from Banaras Hindu University and his M.S. from Indian Institute of Technology, Delhi, India. He is currently working on elucidating the mechanism of proton pumping in cytochrome *c* oxidase.

Xavier Prat-Resina was born in Barcelona, Spain, in 1976. He attended the Universitat de Barcelona where he received his B.S. degree in Chemistry in 1999 and his Masters in 2000 with a computational study on organometallic catalysis. He completed his Ph.D. in chemistry at the Universitat Autonoma de Barcelona in 2004 with a theoretical study on enzymatic catalysis. Since 2005 he is a postdoctoral associate holding an MEC fellowship at the Cui group at the University of Wisconsin-Madison, where he is involved in the study of metalloenzymes.

Peter König obtained his Diploma in chemistry from the Technische Universität München in 2001. In 2001, he joined the group of Professor Th. Frauenheim at the Universität Paderborn, focusing on the description of long-range proton transfer and the bacterial reaction center. After completing his Ph.D. in 2005, he joined the BACTER institute at the University of Wisconsin-Madison as a postdoctoral research associate.

Guohui Li received his Ph.D. from the Dalian Institute of Chemical Physics associated with the Chinese Academy of Sciences. He has been a postdoctoral fellow in the University of New Mexico pursuing quantum mechanical studies of vibrational spectrum of molecules, and then a research associate at the University of Wisconsin-Madison carrying out computational study of biomolecules. Currently he is a research fellow at the Harvard Medical School working on protein structure prediction and computational protein design.

Dingguo Xu received his Ph.D. in 2003 from Sichuan University, China. He is currently a Research Assistant Professor at the University of New Mexico. His research interests focus on enzymatic reaction mechanisms.

Hua Guo received his B.S. in China and D.Phil. in the UK. He is currently a Professor of Chemistry at the University of New Mexico. His research interests range from gas-phase reaction dynamics and spectroscopy to organic chemistry and enzyme catalysis.

Marcus Elstner studied Physics and Philosophy at the Technische Universität München and Berlin and received his Diploma degree in Physics in 1993. For Ph.D., he attended the Molecular Biophysics group of Professor Suhai at the German Cancer Research Center (DKFZ) in Heidelberg and graduated in 1998. As a postdoctoral fellow, he worked in the group of Professor Kaxiras at Harvard University. In December 2002, he became an assistant professor in the Department of Theoretical Physics, Paderborn University (Germany). His main research interest concerns the development of SCC-DFTB and biophysical applications, in particular excited states properties/dynamics and proton-transfer processes in biomolecules.

Qiang Cui received his B.S. in Chemical Physics from the University of Science and Technology of China (USTC) in 1993. For his Ph.D. work, he studied with Professor Keiji Morokuma at Emory University on theoretical analyses of nonadiabatic gas-phase reactions and organometallic catalysis. During his postdoctoral years at Harvard University (1998−2001), he worked with Professor Martin Karplus on the development and application of hybrid QM/MM methods in the context of enzyme catalysis. Since the summer of 2001, he has been an assistant professor of chemistry at the University of Wisconsin- Madison. His major research interests involve theoretical and computational analysis of biophysical and biochemical problems, such as energy transduction in biomolecular motors and ion pumps, radical reactions in biology, and assembly of protein−nucleic acid complexes.

efficiency[13,14] and specificity,[15] as well as in revealing structural information of transient species by comparing calculated and measured spectroscopic observables.[16] For systems that involve a large number of atoms in the reaction, such as proton pumps, or major conformational changes, such as signaling proteins and molecular motors, much less work has been done.

In our research group, the problems of current interest involve vectorial biological systems such as molecular motors and ion pumps. The functional cycles of these fascinating systems involve complex chemical reactions and conformational transitions that require the development of new simulation techniques.[17] In this article, we review our recent works that was motivated by the ultimate goal of understanding the mechanism of proton pumping in redox and ATP-driven ion pumps. Specifically, we describe a number of developments toward establishing effective QM/MM methods for studying complex biological processes that involve significant charge separation, which is an important challenge in the field of enzyme simulations. We illustrate these methods with relevant condensed phase model problems including the pK_a's of protein residues and proton transfers, in water and the enzyme carbonic anhydrase. This article is not meant to be a comprehensive review on the subject of QM/MM methods; we refer the readers to a number of excellent review articles[7−9] for developments made by other research groups in the field.

2. Theoretical Methods

Throughout the discussion, we will follow the typical form of the QM/MM Hamiltonian and the common notation,[4,18]

$$\hat{H}^{\mathrm{Tot}} = \hat{H}^{\mathrm{QM}} + \hat{H}^{\mathrm{QM/MM}}_{\mathrm{elec}} + \hat{H}^{\mathrm{QM/MM}}_{\mathrm{vdW}} + \hat{H}^{\mathrm{QM/MM}}_{\mathrm{bonded}} + \hat{H}^{\mathrm{MM}} \quad (1)$$

As widely appreciated,[7−9,19] the quantitative accuracy of a QM/MM simulation depends on many factors including the reliability of the QM method, the size of the QM region, the way that the QM and MM atoms are partitioned and the scheme for computing their interactions, the amount of configurational

sampling, as well as the quality of the MM force field. Specifically for the process of proton pumping, which explicitly involves a large number of atoms and significant charge reorganization, the issues of sampling and electrostatic treatment are most outstanding; for example, the significant fluctuations of the reactive moieties make QM/MM methods[19] based on a harmonic expansion around the minimum energy path for the QM atoms not applicable. The chemical nature of proton translocation, in contrast to the transport of many other ions, also leads to stringent constraints on the accuracy of the QM method.

Motivated by these considerations, a number of developments were made in our research group in recent years. Below, we first discuss free energy simulations using QM/MM potentials, which is followed by the discussion on the treatment of QM/MM interactions. Finally, we present several developments for improving the accuracy of an efficient QM method. Throughout these works, an approximate density functional theory, the self-consistent charge density functional tight-binding (SCC-DFTB) method,[20] was used as the QM level. This was motivated by its computational efficiency (comparable to widely used semi-empirical methods such as AM1 and PM3) and reasonable accuracy, which make extensive configurational sampling of condensed phase systems possible. The SCC-DFTB method has been applied successfully to a range of problems involving biomolecules, such as conformational energies of peptides and catalysis in several enzymes; a recent review can be found in ref 21. Furthermore, the SCC-DFTB approach has also been benchmarked for reaction energies, geometries, and vibrational frequencies for small molecules in comparison to the G2 approach.[22]

2.1. Free Energy Simulations with QM/MM Potentials[23−25]. For chemical reactions, calculating the activation free energy (and rate) is of the ultimate interest and remains an important challenge. In the following, however, we focus on computing the free energy difference between two species of different chemical nature (e.g., an oxidation−reduction pair) in the condensed phase environment, which is useful because the results can be directly compared to many experimental observables (e.g., pK_a and reduction potential) as a way to quantitatively benchmark QM/MM methods. Using free energy quantities to evaluate QM/MM results and to identify factors that dictate the accuracy of QM/MM simulations has been a feature that distinguishes our work from many previous studies.

Indicating the two "solutes" as A and B, the environment as C, the free energy of converting between A and B in the presence of C is conveniently calculated according to the standard thermodynamic integration approach,[26,27]

$$\Delta F_{A\rightarrow B} = \int_0^1 \frac{\partial F}{\partial \lambda} d\lambda = \int_0^1 \left\langle \frac{\partial U^\lambda}{\partial \lambda} \right\rangle_\lambda d\lambda \qquad (2)$$

where λ is the coupling parameter and $\langle \cdots \rangle_\lambda$ indicates ensemble average at a specific λ value. The potential function used in the thermodynamic integration is usually written as

$$U^\lambda(\lambda) = (1-\lambda)U_A(\mathbf{X}_A,\mathbf{X}_C) + \lambda U_B(\mathbf{X}_B,\mathbf{X}_C) + U_{CC}(\mathbf{X}_C) \qquad (3)$$

where $U_A(U_B)$ includes both intramolecular A−A (B−B) interactions and intermolecular A−C (B−C) interactions.

A practical difficulty that arises in using the potential function in eq 3 is that as λ approaches end points (0,1), the contribution from either A or B to $U^\lambda(\lambda)$ vanishes, which causes sampling and convergence issues in the free energy derivative ($\partial F/\partial\lambda$) due to large structural distortion and ideal-gas-like nature of the corresponding "solute". To circumvent such "end-point" problems,[28] the traditional solution is to use a revised version of the coupling potential,

$$U^\lambda(\lambda) = (1-\lambda)U_{AC}(\mathbf{X}_A,\mathbf{X}_C) + \lambda U_{BC}(\mathbf{X}_B,\mathbf{X}_C) + U_{AA}(\mathbf{X}_A) + U_{BB}(\mathbf{X}_B) + U_{CC}(\mathbf{X}_C) \qquad (4)$$

Although such decoupling of intra- and intermolecular interactions is straightforward with a MM force field, it is less so with a QM/MM potential function due to the nonseparability of QM energies.[23] For example, it is possible to scale (by λ, $1-\lambda$) only the QM/MM interactions so that as A is switched from the gas phase to the environment C, species B is switched from interacting with C to the gas phase. Thermodynamic integration following this protocol gives the free energy difference between *A* and *B* in the presence of C relative to the gas phase (i.e., the relative "solvation" or transfer free energy of A and B from the gas phase to C). Although this quantity is of interest in many cases, the complication is that calculation of the free energy derivative requires extra computations due to the nonseparability of the QM/MM energy. Nevertheless, this can be done and has been recently reported by Yang and co-workers.[29]

Alternatively, when the species A and B differ mainly in electronic structure but similar in nuclear connectivity (e.g., an oxidation−reduction pair), the same set of nuclear geometry for the two states can be used,[23] i.e., the coupling potential function has the form

$$U^\lambda(\lambda) = (1-\lambda)U_A(\mathbf{X}_A,\mathbf{X}_C) + \lambda U_B(\mathbf{X}_A,\mathbf{X}_C) + U_{CC}(\mathbf{X}_C) \qquad (5)$$

We note that using the same set of geometry for the two states, per se, does not introduce any approximation due to the state character of free energy (i.e., as λ evolves from 0 to 1, the geometry evolves from that for A to B). In practice, error may occur when SHAKE[30] is used to constrain bond distances involving hydrogen to be the same for all λ values, although the magnitude is expected to be very small for all practical cases.

Since two electronic states but one set of nuclear coordinates is invoked in the coupling potential in eq 5, we refer the protocol as "dual topology single coordinate" (DTSC),[23] although some authors prefer to keep the notation as "single topology" (B. Brooks, private communication). Regardless of the notation, the formulation has the attractive features that it converges faster than using two separate copies of coordinates, and it does not require the evaluation of the Jacobian factor[31] as needed in traditional single topology simulations using MM force field.

The DTSC protocol apparently works only when A and B have very similar (but not identical) nuclear geometries. In the special case of a small difference between the two "solute" states, e.g., AH and A^- for pK_a problems, special thermodynamic cycles can be constructed to take advantage of features of the DTSC protocol.[24,25] When A and B are very different, one could either use the conventional dual topology dual coordinate potential (eq 4) or use the elegant "chaperone" approach proposed by Yang and co-workers.[32]

2.2. Treatment of QM/MM Interactions: Electrostatics,[25,33] van der Waals,[34] and Link Atoms[35]. Based on the popular form of the Hamiltonian in eq 1, it is clear that there are three components in QM/MM interactions: electrostatics, van der Waals, and the QM/MM frontier (boundary). All three components were investigated systematically with SCC-DFTB as the QM method. Using both gas-phase hydrogen-bonding clusters and solution-phase reactions, we showed that free energy calculations are rather insensitive to the van der Waals

parameters for QM atoms[34] as long as the parameters are reasonable, due mainly to error cancellation effects. Regarding the treatment of QM/MM boundary, numerous investigations and developments have been made using link atoms,[4,36,37,38] frozen orbitals,[39] generalized hybrid orbitals,[40] or pseudobonds.[41] In a recent study,[35] the accuracy of a series of link-atom-based approaches available in the program CHARMM was analyzed based on properties including proton affinities, deprotonation energies, dipole moments, and energetics of proton-transfer reactions. In addition to the analysis of gas-phase molecules as in previous studies, an important element of that work is that chemical reactions in realistic enzyme systems were also examined. It was shown that the single-link-atom approach should be avoided in the calculation of quantities such as proton affinity, although other schemes often give similar results. For the energetics of reactions that conserve charge, the results were found to be even less sensitive to the frontier treatments.

Since electrostatic interactions often dominate free energy changes in biomolecules, carefully treating $\hat{H}_{\text{elec}}^{\text{QM/MM}}$ is of the utmost importance. Although carefully analyzed in the context of fully classical simulations,[42−44] the issue of consistent electrostatic treatment in QM/MM simulations was considered only recently,[45−47] besides the pioneering work of Warshel and co-workers in EVB simulations.[48] There were several implementations of the Ewald summation technique with QM/MM potentials,[45,46] most with semiempirical QM methods such as AM1 and SCC-DFTB.[25] With these QM methods, the cost of the QM/MM component (both real- and reciprocal space) may not be overwhelming compared to the MM component if the QM region is only a small part of the system. Nevertheless, periodic boundary condition simulations can be prohibitively expensive for large enzyme systems (e.g., the ribosome). Therefore, it is important to develop finite-size boundary condition that treats QM/MM electrostatics in a robust manner. Recently, we met this challenge by adopting the generalized solvent boundary potential (GSBP) developed by Roux and co-workers for classical simulations.[49]

Briefly, GSBP partitions the system into inner and outer regions (Figure 1) where the effects of the outer region on the inner, reaction region are represented implicitly within the total effective potential (potential of mean force),[49]

$$W_{\text{GSBP}} = U^{(\text{ii})} + U_{\text{int}}^{(\text{io})} + U_{\text{LJ}}^{(\text{io})} + \Delta W_{\text{np}} + \Delta W_{\text{elec}}^{(\text{io})} + \Delta W_{\text{elec}}^{(\text{ii})} \tag{6}$$

where $U^{(\text{ii})}$ is the complete inner-inner potential energy, $U_{\text{int}}^{(\text{io})}$ and $U_{\text{LJ}}^{(\text{io})}$ are the inner-outer internal (bonds, angles, and dihedrals) and Lennard-Jones potential energies, respectively, and ΔW_{np} is the nonpolar confining potential. The last two terms are the core of GSBP, representing the long-range electrostatic interaction between the outer and inner regions. The contribution from distant protein charges (screened by the bulk solvent) in the outer region, $\Delta W_{\text{elec}}^{(\text{io})}$, is represented in terms of the corresponding electrostatic potential in the inner region, $\phi_{\text{s}}^{(\text{o})}(\mathbf{r}_\alpha)$,

$$\Delta W_{\text{elec}}^{(\text{io})} = \sum_{\alpha \in \text{inner}} q_\alpha \phi_{\text{s}}^{(\text{o})}(\mathbf{r}_\alpha) \tag{7}$$

The dielectric effect on the interactions among inner region atoms is represented through a reaction field term,

$$\Delta W_{\text{elec}}^{(\text{ii})} = \frac{1}{2} \sum_{mn} Q_m \mathbf{M}_{mn} Q_n \tag{8}$$

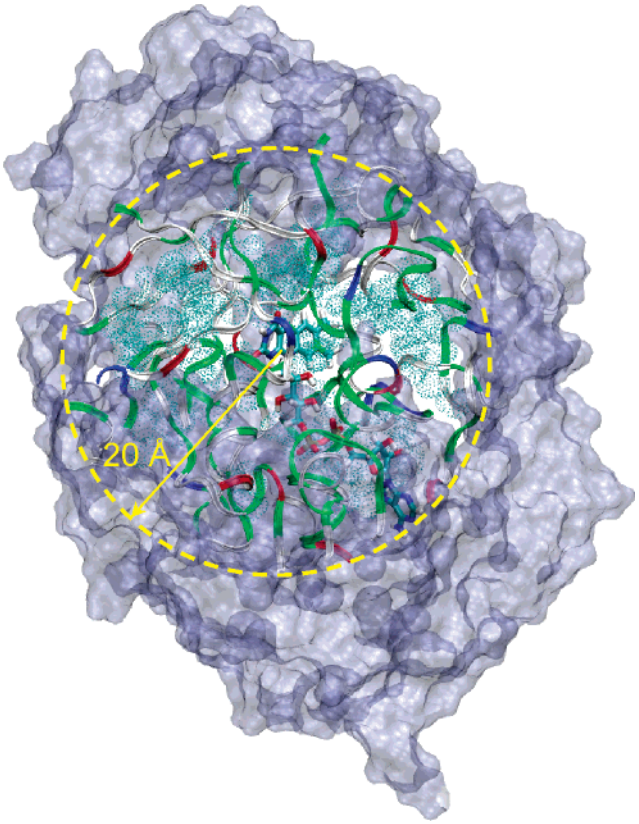

Figure 1. Illustration of the SCC-DFTB/MM-GSBP protocol established from studies summarized here: a small group of atoms (shown in line form) are treated with SCC-DFTB and the rest treated with the CHARMM 22 force field. The GSBP inner region (protein shown in ribbon and solvent by dots) typically contains all atoms within 20 Å from a site in the SCC-DFTB moiety; the remaining protein (shown in transparent ice-blue) and bulk water constitute the outer region.

where $\mathbf{M}$ and Q are the generalized reaction field matrix and generalized multipole moments, respectively, in a basis set expansion.[49]

The strength of the GSBP method lies in its ability to include these contributions explicitly while sampling configurational space of the reaction region during a simulation at minimal additional cost. The static field potential, $\phi_{\text{s}}^{(\text{o})}(\mathbf{r})$, and the generalized reaction field matrix $\mathbf{M}$ are computed once based on Poisson−Boltzmann calculations and stored for subsequent simulations. The only quantities that need to be updated throughout the simulation are the generalized multipole moments, Q_n,

$$Q_n = \sum_{\alpha \in \text{inner}} q_\alpha b_n(\mathbf{r}_\alpha) \tag{9}$$

where $b_n(\mathbf{r}_\alpha)$ is the nth basis function at nuclear position $\mathbf{r}_\alpha$.

The implementation of GSBP in a combined QM/MM framework requires three additional terms,[33]

$$\frac{1}{2} \sum_{mn} Q_m^{\text{QM}} \mathbf{M}_{mn} Q_n^{\text{QM}} + \sum_{mn} Q_m^{\text{QM}} \mathbf{M}_{mn} (Q_n^{\text{MM}} - Q_n^{\text{EX}}) + \int \mathrm{d}\mathbf{r}\, \rho^{\text{QM}}(\mathbf{r}) \phi_{\text{s}}^{(\text{o})}(\mathbf{r}) \tag{10}$$

corresponding to the QM-QM and QM-MM (corrected for exclusions due to link host schemes[33,35]) reaction field, and the QM static field terms, respectively. For the GSBP combined with SCC-DFTB, these terms take on a simple form because $\rho^{\text{QM}}(\mathbf{r})$ is expressed in terms of Mulliken charges,[20] $\rho^{\text{QM}}(\mathbf{r}) =$

$\sum_{A\in QM} \Delta q^A \delta(\mathbf{r} - \mathbf{R}_A)$, resulting in expressions,

$$Q_m^{QM} = \sum_{A\in QM} \Delta q^A b_m(\mathbf{R}_A) \quad (11)$$

for the generalized QM multipoles and,

$$\Delta W_{elec}^{QM(io)} = \sum_{A\in QM} \Delta q^A \phi_s^{(o)}(\mathbf{R}_A) \quad (12)$$

for the interaction of the QM region with the static field due to the outer region. These terms are included in the SCC-DFTB matrix elements during the SCF iteration.[33]

Although the QM/MM-GSBP protocol offers a powerful way to study chemistry in large biomolecular systems in a multiscale manner, the method may not be the best choice for simulating all biological systems (e.g., systems that involve major conformational changes) due to its restricted treatment of configurational space for the outer region. In addition, the current implementation of GSBP ignores the surface polarization artifacts due to the inner-outer dielectric interface; the impact of these artifacts on simulation results have been documented and dealt with by Warshel and others for solution simulations.[50,51] Quantifying such effects in macromolecular simulations and further extending the flexibility of the GSBP protocol is a continuing focus of our work.

2.3. Improving the QM Method: Extensions of SCC-DFTB. Despite the rapid developments in computational hardware and linear-scaling quantum mechanics, it remains prohibitively expensive to compute free energies for condensed phase systems with ab initio QM methods even in a QM/MM framework. A promising approach is to develop semiempirical methods that offer the best balance of accuracy and efficiency. One popular avenue is to reparameterize commonly used semiempirical methods such as AM1 or PM3 for the specific system of interest, the so-called specific-reaction-parameter approach.[52] Our research goal, however, has been to systematically extend the formalism of an approximate density functional theory, the SCC-DFTB approach,[20] so that the resulting method is more transferable. This is likely possible because parameters in SCC-DFTB were established in a simple and systematic manner based on DFT calculations,[20,53] rather than with a tour-de-force fitting procedure adopted in typical semiempirical methods.[54] Furthermore, SCC-DFTB describes the structures of hydrogen-bonding systems rather well, which is important for the study of biomolecules.[55] There are systematic errors with the current formulation of SCC-DFTB; e.g., the hydrogen bonding energies are typically underestimated by 1−2 kcal/mol.[20,21,34] For such systematic errors, however, there is the chance to find a simple way in which an approximate method can be extended without loosing generality and transferability.

The first type of extension is to add higher order terms in the energy functional expansion of SCC-DFTB. This development improves the description of proton affinities (PA), which is of great importance in biological systems due to the extensive role of general acid−base catalysis in enzymes. Most semiempirical methods, including the original parameterization of SCC-DFTB, have rather large errors, often of 10−15 kcal/mol, for PAs.[56] Although it is possible to correct for errors in PA for properties such as pK_a in a post-processing manner,[25] large errors in PA may result in wrong protonation states of reactive groups (e.g., phosphate groups), which may in turn lead to incorrect mechanistic conclusions. Recognizing the significant change in the charge distribution for a deprotonation/protonation process, it was conceived that improved PA can be obtained by extending the perturbation series that defines the SCC-DFTB approach to higher orders,[57]

$$E_{3rd}^{SCC-DFTB} = E_{2nd}^{SCC-DFTB} + \frac{1}{6}\int\int'\int'' \left.\frac{\delta^3 E^{tot}}{\delta n(\mathbf{r})\delta n'(\mathbf{r}')\delta n''(\mathbf{r}'')}\right|_{n_0} \delta n(\mathbf{r})\delta n'(\mathbf{r}')\delta n''(\mathbf{r}'')\, d\mathbf{r}\, d\mathbf{r}'\, d\mathbf{r}'' \quad (13)$$

$$\sim \sum_{i\mu\nu} c_\mu^i c_\nu^i H_{\mu\nu}^0 + \frac{1}{2}\sum_{\alpha\beta} U_{\alpha\beta}^{rep} + \frac{1}{2}\sum_{\alpha\beta}\gamma_{\alpha\beta}\Delta q_\alpha \Delta q_\beta + \frac{1}{6}\sum_{\alpha\beta\sigma}\eta_{\alpha\beta\sigma}\Delta q_\alpha \Delta q_\beta \Delta q_\sigma \quad (14)$$

where $\gamma_{\alpha\beta}$ is the approximate second-order kernel, Δq_α is the Mulliken charge of atom α. The simplest approach is to include only the on-site third-order terms, which was found to already substantially improve the PA values (see Supporting Information). The on-site third-order kernel, $\eta_{\alpha\alpha\alpha}$, can be shown to be related to the charge derivative of the atomic Hubbard parameter (proportional to the chemical hardness), U_α,

$$\eta_{\alpha\alpha\alpha} = \frac{\partial U_\alpha}{\partial q_\alpha} = \frac{\partial^2 \epsilon_\alpha}{\partial q_\alpha^2} \quad (15)$$

where ϵ_α is the energy of the highest occupied atomic orbital and can be evaluated by atomic calculations using density functional theory; alternatively, $\eta_{\alpha\alpha\alpha}$ can be optimized based on desired properties such as accurate PA values (see Supporting Information).

Another important property of interest for biomolecular simulations is hydrogen-bonding interaction, which is not well described with the common semiempirical methods. SCC-DFTB was also found to systematically underestimate such interactions by, typically, 1−2 kcal/mol.[58] Ironically, the SCC-DFTB/MM potential gives better descriptions due to the fact that MM charges in the nonpolarizable force field are often substantially larger than SCC-DFTB Mulliken charges.[34] In addition to the inherent error in intra-QM interactions, such an imbalance between SCC-DFTB and SCC-DFTB/MM may cause erroneous behaviors at the QM and MM boundary.

The limitation of the SCC-DFTB approach for treating hydrogen-bonding interactions can be traced back to the short-range behavior of the $\gamma_{\alpha\beta}$ function for atomic pairs involving hydrogen. The assumption that the atomic charge density and its chemical hardness are inversely proportional[20] is not valid for hydrogen. As proposed by Elstner,[57] an effective approach for fixing this is to introduce a damping function at short distances,

$$\gamma_{\alpha H} = 1/R_{\alpha H} - S_{\alpha H} * f(R_{\alpha H}) = 1/R_{\alpha H} - S_{\alpha H} * \exp\left[-\left(\frac{U_\alpha + U_H}{2}\right)^\zeta R_{\alpha H}^2\right] \quad (16)$$

where $S_{\alpha H}$ is the short-range function in the original SCC-DFTB method.[20] The parameter, ζ, in the damping function was adjusted based on high-level ab initio results for hydrogen-bonding clusters (Yang et al., unpublished). It is worth noting that the optimal value of ζ was found to depend on whether the third-order term (eq 14) is included, because both modifications tend to increase polarity of the molecule (see Supporting Information).

We emphasize that both improvements of SCC-DFTB described above, the third-order extension and the modified Coulomb scaling, introduce a very small number of parameters (see Supporting Information). The parameters in the third-order terms can be calculated from DFT, and the resulting method shows a systematic improvement for the proton affinities studied here. Although these parameters were then optimized based on a set of molecules, the improvements are likely transferable to a broad range of systems. Systematic benchmark calculations are clearly of great importance and are being carried out.

3. Illustrative Applications and Discussions

In this section, we briefly discuss a number of applications of the QM/MM protocols presented above to demonstrate the value of those developments, especially in the context of studying chemical processes that involve significant charge reorganization. We begin by emphasizing the importance of configuration sampling and electrostatic treatment in the calculation of pK_a of titritable groups in T4 lysozyme, which is followed by a short discussion of proton-transfer reactions in solution and the enzyme carbonic anhydrase. The standard SCC-DFTB method was used for the pK_a calculations, while improved SCC-DFTB methods were used for the water autoionization and PT in carbonic anhydrase (see Supporting Information for details).

3.1. pK_a in the M102K Mutant of T4 Lysozyme.[25] The knowledge of reduction potential or pK_a value for a specific group is of essential functional interest in many mechanistic studies, in particular, analysis of redox-driven proton pumping. Calculation of such quantities, however, is far from straightforward because accurately describing the electronic structural change, especially for oxidation−reduction processes, often requires high level ab initio methods. More importantly, the change in the oxidation or protonation state of a group buried inside a biomolecule is often associated with nontrivial structural reorganizations. The latter fact is less widely appreciated due to the popularity of Poisson−Boltzmann based protocols,[59,60] with either pure MM or QM/MM potentials, which typically do not include explicit structural changes. Although Poisson−Boltzmann remains most cost-effective when properties of a large number of groups are of interest, more robust methods are preferable when accurate reduction potential or pK_a value for a specific group is needed. We illustrate this with the example of pK_a in the T4 lysozyme; for the details of the simulations, see ref 25.

The pK_a's of two residues, His 31 and Lys 102, in the M102K mutant of the T4 lysozyme were chosen to study with QM/MM methods because the corresponding pK_a shifts are among the largest known for proteins.[61] The His 31 is involved in a salt-bridge interaction with Asp 70 and is surface accessible from one side and close to the hydrophobic core on the other side. Due to the partial burial, the His 31-Asp 70 salt bridge was shown to contribute 3−5 kcal/mol of stabilization,[62] and the pK_a's of His 31 and Asp 70 were shifted up to 9.1 and down to 0.5, respectively. The SCC-DFTB/CHARMM-GSBP simulations (Figure 1) were found successful for the pK_a prediction of His 31. The free energy derivatives were well behaved at all λ values; as shown in Figure 2a for $\lambda = 0.0$; for example, the statistical uncertainty was on the order of 0.3−0.6 kcal/mol once the equilibration portion of the trajectories ($\simeq$1−2 ns) was discarded according to the reverse cumulative analysis protocol.[63] Fitting the free energy derivatives linearly ($R^2 \sim 0.99$) gave pK_a shifts (relative to imidazole in solution) of 1.2 and 1.5 pK_a units for the 0.0 and 0.1 M simulations, respectively,

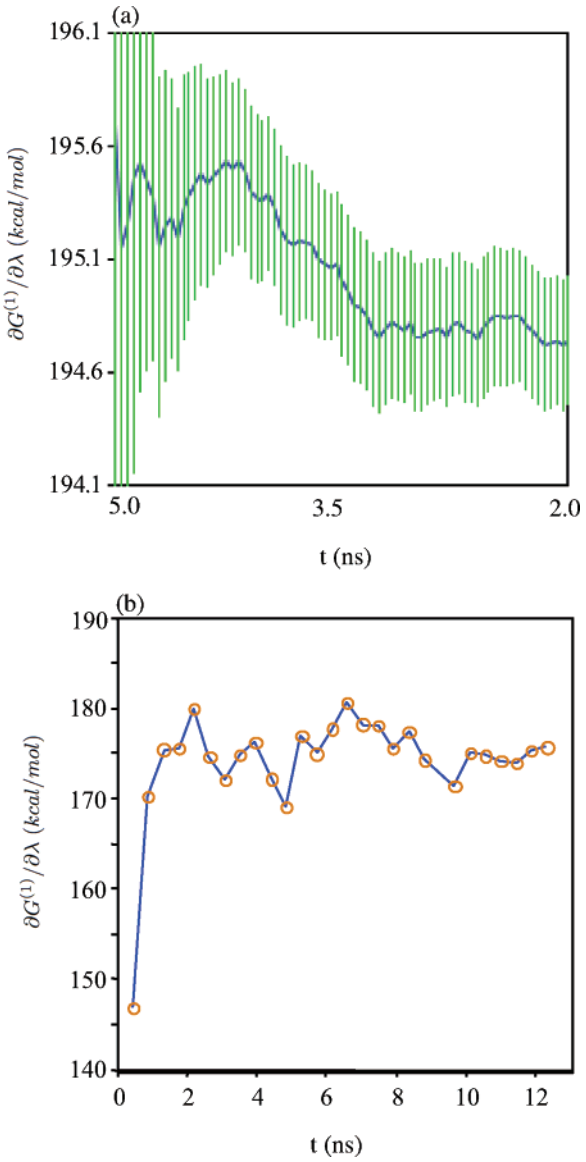

Figure 2. Convergence behavior of the free energy derivative at $\lambda = 0.0$ in the pK_a calculation for (a) His 31 and (b) Lys 102 in the M102K mutant of the T4 lysozyme with a SCC-DFTB/MM-GSBP (0.1 M salt) protocol. Note the drastic difference in the scale of the vertical axes in (a) and (b); also reverse and forward cumulative average was carried out for the His 31 and Lys 102 simulations, respectively. For the reverse cumulative average analysis, the estimated error in the original work[25] was too small by nearly a factor of 2 due to a mistake in the implementation; the conclusions of that work, however, remain unchanged.

which were in good agreement with the experimental value of 2.1 pK_a units.

The pK_a prediction for Lys 102, by contrast, was more problematic. The Lys 102 is entirely buried in a hydrophobic pocket, which causes a large downward shift by four pK_a units to 6.5 as measured with differential titrations and NMR. Therefore, it was anticipated that protonation of Lys 102 may cause substantial structural changes in the protein. Indeed, the M102K mutant crystal structure (1.9 Å resolution at pH 6.8)[64] was found to be very similar to the wild type, except for the significantly increased mobility (as reflected by the B-factors) in the Glu 108-Gly 113 α-helix, which was explained as the result from a mixing of the two Lys 102 protonation states at pH 6.8, with the helix being completely disordered for the protonated state and more structured (as in the wild type) for the deprotonated state.

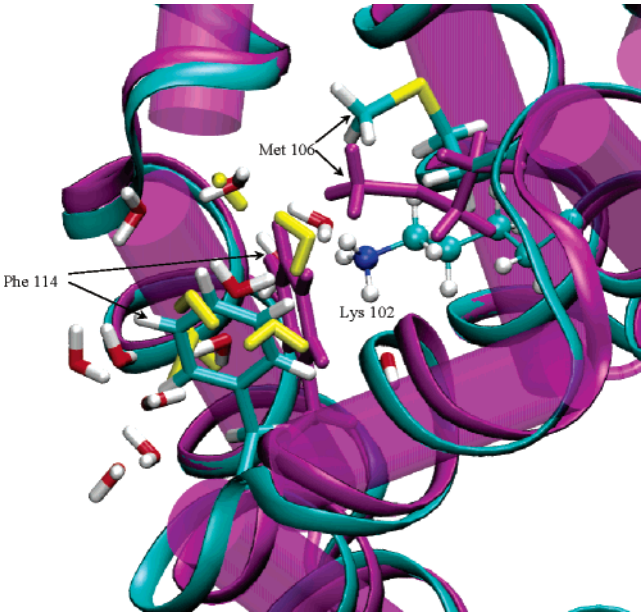

Figure 3. Comparison of the average structure for the 0.0 M salt protonated Lys 102 simulation (cyan) to the crystal structure for the protein (purple) and water (yellow) atoms. The reorientation of Phe 114 plays a key role in the initial inpouring of waters into the hydrophobic region containing the charged Lys 102; this may have initiated some of the changes seen in the Glu 108−Gly 113 α-helix, as well as in the orientation of Met 106.

The structural instability for the protonated Lys 102 state ($\lambda = 0.0$) was apparent in the SCC-DFTB/MM simulations. The root-mean-square deviation (RMSD) from the X-ray structure eventually increased from 0.5 to 1.0 Å in 10 ns, which was striking considering that a significant part of the protein was fixed during the GSBP simulations. The origin of the higher RMSD's in the $\lambda = 0.0$ simulations stem from the structural changes in two α- helices in close proximity to Lys 102. In the 0.0 M simulation, for example, the Glu 108−Gly 113 α-helix partially unwound due to a dramatic migration (≃12 Å) of the nearby Phe 114 to the protein surface, which allowed water molecules to penetrate into the interior of the protein to solvate the protonated Lys 102 (Figure 3). The deprotonated simulations, by contrast, remained similar to the crystal structure with the Phe 114 acting as a hydrophobic barrier that prevents solvation of the deprotonated Lys 102.

In accord with the observed structural instability, the convergence behavior of the free energy derivatives was found less satisfactory for Lys 102 compared to the His 31 case (contrast Figure 2a,b). According to the cumulative reverse averaging protocol, the free energy derivatives did not converge for any λ value after 10 ns of simulation for each window. Quantitatively, taking the forward cumulative average and ignoring obvious areas of equilibration, the estimated pK_a shift (relative to propylamine in solution) was −15.6 and −13.4 pK_a units for 0.0 and 0.1 M simulations, respectively, which were much larger than the experimental shift of −4.0 pK_a units. Interestingly, if one assumes that Lys 102 in the protonated state is completely solvent exposed thus the free energy derivative for $\lambda = 0.0$ can be taken as that for propylamine in solution (197.1 kcal/mol), using this value together with the free energy derivative for $\lambda = 1.0$ in the realistic lysozyme simulations in a linear response framework would produce a much more reasonable pK_a shift of 4.2. Although this agreement might be coincidental, considering the success we have had for all other cases studied so far, the result strongly argues that protonated Lys 102 significantly destabilizes the protein structure and is likely to be exposed to the solvent, in contrast to the crystal structure.

The pK_a calculations in T4 lysozyme clearly illustrated the success and limitations of the free energy perturbation approach using the SCC-DFTB/MM potential. An important point is that even when poor results are obtained, as in the case of Lys 102, the explicit free energy simulations can provide clues (structural instability, poor convergence of the free energy derivative) regarding the origin of the problem and possible solutions, which is an essential advantage over continuum electrostatics based methods.

3.2. Proton Transfer: Water Autoionization and Reactions in Carbonic Anhydrase. Proton transfer (PT) reactions are prevalent in chemistry and life processes, ranging from general acid−base catalysis[65] to proton pumping in bioenergetics.[66,67] Although basic mechanisms for localized PT in enzyme active sites are rather well understood,[9,68] long-range PTs, are more challenging to understand at a quantitative level. It is generally accepted that long-range PT occurs through the help of hydrogen-bonded "wires" formed by water molecules[69] and possibly titratable amino acid side chains,[70] yet the precise transfer pathway(s) and rate-limiting factors are often difficult to unravel due to the large number of groups involved. Theoretical analyses are well poised to make an essential contribution in this regard, although long-range charge separation and the involvement of multiple bond-breaking and formation events make it difficult to carry out meangingful calculations. In the following, we illustrate a number of challenges in theoretical analysis of PTs in condensed phase environments and the corresponding developments made in our research group.

3.2.1. Active Site Dynamics and Water Distribution in Carbonic Anhydrase.[33] Although the rate-limiting factors for long-range PTs are not precisely established, it is commonly believed that a prerequisite for any reliable theoretical analysis is the appropriate description for the structure and dynamics of the proton donor and acceptor groups as well as the water molecules that bridge them. Due to the significant charge separation associated with long- range PTs, appropriate treatment of the long-range nature of electrostatics is important. This can be best illustrated with the example of carbonic anhydrase (CA).

CA is a Zn(II)-containing metalloenzyme that moderates respiration by catalyzing the conversion between CO_2 and HCO_3^-. An important step in the functional cycle of CA involves a PT between the zinc-bound water and His 64 close to the protein surface; the transferred proton is released into solution as the doubly protonated His 64 side chain flips from a buried ("in") conformer to a solvent-exposed ("out") conformer. Since the distance between the zinc-bound water and His 64 in the X-ray structure[71] was observed to be too long (7.5 Å) for a direct transfer, water molecules in the active site were assumed to act as bridges that relay the PT.[71−73] Although this mechanism has been widely accepted, the precise number of water molecules involved in such bridges and the nature of events that limit the PT rate have been under heated debate.[73−76] Clearly, the properties of critical importance include flexibility of the proton acceptor (His 64), the distribution of water molecules in the active site, and the type of "water-wires" formed between the zinc-bound water and His 64. In the following, we compare results from SCC-DFTB/MM simulations using four different methods for treating electrostatics. The most elaborate set of simulations employed the periodic boundary condition and the Ewald summation technique as

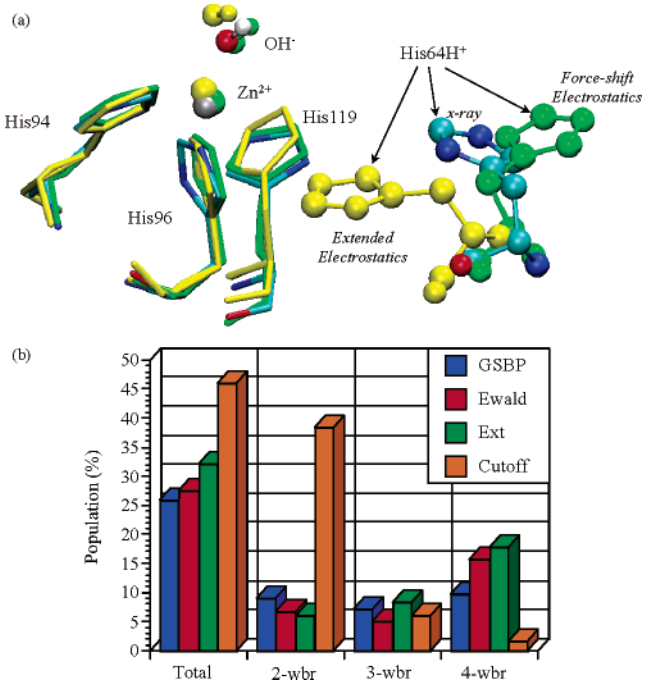

Figure 4. Properties of the carbonic anhydrase active site in the COHH state (zinc-bound hydroxide and protonated His 64). (a) Superposition of a few key residues from two stochastic boundary SCC-DFTB/MM simulations with the X-ray structure (PDB code: 2CBA, colored based on atom types); the two sets of simulations did not have any cutoff for the electrostatic interactions between SCC-DFTB and MM atoms but used different treatments for the electrostatic interactions among MM atoms: group-based extended electrostatics (in yellow) and atom-based force-shift cutoff (in green). Extended electrostatics simulations sampled configurations with the protonated His 64 too close to the zinc moiety while force-shift simulations consistently sampled the "out" configuration of the His 64 in multiple trajectories. (b) Statistics for productive water bridges (only from two and four shown here) between the zinc bound water and His 64 with different electrostatics protocols.

recently implemented in the QM/MM framework in CHARMM,[25] while the other three calculations used a finite-size spherical protocol with either GSBP (20 Å inner region)[33] or the popular stochastic boundary condition (25 Å water droplet).[77] In the stochastic boundary simulations, the QM/MM electrostatics were always treated without cutoff; the MM/MM electrostatics, however, were treated with either a simple force-shift-based cutoff (12 Å) or extended electrostatics.[78] We note that long-range effects due to protein atoms and bulk water beyond the explicitly simulated region were neglected in the stochastic boundary calculations but included through continuum electrostatics in the GSBP-based simulations. Following the literature,[75] the protonation state that involves zinc-bound water (hydroxide) and neutral (protonated) His 64 is referred to as the CHOH (COHH) state. All QM/MM simulations in this subsection employed a "minimal" QM region that included the zinc ion and its ligands: water/hydroxide, side chains of His 94, 96, 119.

Consistent with the function as an effective proton shuttle between the CA active site and bulk solution, the side chain of His 64 is flexible. Indeed, both "in" and "out" conformations were observed in the X-ray structure.[79] The GSBP, extended electrostatics and Ewald simulations sampled the "in" and "out" conformations rather evenly, with different conformers favored with different initial conditions. However, one of the extended electrostatic COHH trajectories sampled a His 64 position very close (~5.0 Å) to the zinc atom compared to the value ~8.0 Å in the X-ray data (Figure 4a). Such a distinct deviation from the X-ray structure is likely an artifact due to the neglect of bulk solvation, which may cause overestimation of attraction between the protonated His 64 and active site residues (e.g., Glu 106). By striking contrast, in all cutoff-COHH simulations performed, the His 64 side chain consistently flipped to the "out" position in the very early stage of the trajectory. This behavior is likely due to the unbalanced QM/MM and MM/MM treatments in these simulations, which overestimate the repulsion experienced by the protonated His 64 due to the underestimation of the attraction to Glu 106/Glu 117 (both are beyond the cutoff distance from His 64).

To analyze the water wire connecting the zinc-bound water/hydroxide and His 64, we followed a definition of hydrogen-bond in terms of both distance (O−O < 3.5 Å) and angle (O−H−O ≥ 140°); care was taken to classify the bridge as productive if a proton can be successfully passed from donor to acceptor through the bridge and unproductive otherwise (e.g., one water molecule serves as a double-donor or double-acceptor). Results were categorized by the population of the water bridge, defined by the number of frames in a productive water bridge of a specific length relative to the total frames in the analyzed trajectory (Figure 4b); when more than one productive water bridge was present in a frame, only the shortest one was counted. Overall, the agreement between Ewald and GSBP results was very good. The agreement between GSBP and stochastic boundary simulation depended on the protonation state and orientation of the His 64 side chain. For example, all simulations gave qualitatively similar results for the COHH state with His 64 flipped to the "out" conformation, for which four-water bridge was found dominant with all protocols. For the CHOH state with His 64 adopting the "in" conformation, extended electrostatics and cutoff simulations distinctively favored four- and two-water bridges, respectively; GSBP simulations, on the other hand, produced rather even distributions of two- to six-water bridges. The Ewald simulations produced results similar to the GSBP calculations, although a slight preference over four-water bridge was seen (Figure 4b). The problem with the cutoff simulation was also apparent when analyzing the number of water molecules in the active site; within 10 Å from the zinc ion, the average number of water molecules in the CHOH simulations (His 64 "in") is 14.2, 14.7, 15.1, and 23.1 for the GSBP, Ewald, extended electrostatics, and cutoff protocols, respectively.

In short, the observed differences using various protocols emphasized that a consistent treatment of electrostatics is required for predicting even qualitatively correct behavior for processes that involve significant charge separations, such as long-range PTs.

3.2.2. Water Autoionization. To predict meaningful thermodynamics and kinetics for long-range PT, it is evident that accurate (at least relative) proton affinities of the donor and acceptor groups as well as the intervening water molecules need to be reproduced. In addition, due to the extensive polar interactions along many possible PT pathways, hydrogen bonding interactions need to be described well. We illustrate the importance of these requirements quantitatively using the example of autoionization ($H_2O \rightarrow H^+ + OH^-$) in bulk water; more detailed presentation of methods and comparison with previous studies as well as exploration of alternative reaction coordinates[80] will be reported separately soon.

We explore the potential of mean force (PMF) along a reaction coordinate defined by the O···H distance in the specific water molecule undergoing autoionization. The PMF was calculated using the weighted histogram analysis method (WHAM)[81] from 15 umbrella sampling simulations (20 ps for

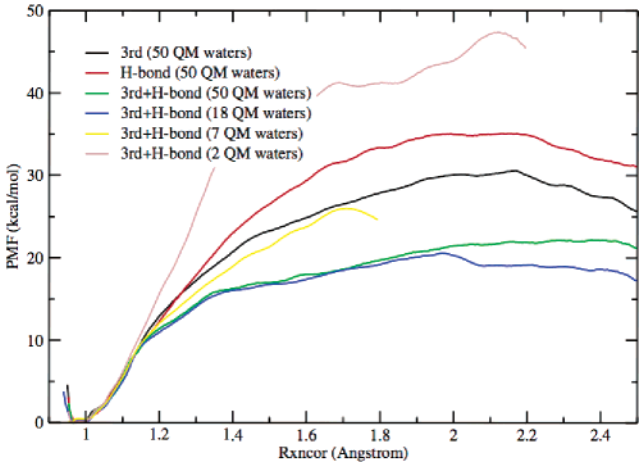

Figure 5. Potential of mean force (in kcal/mol) calculated for the autoionization of one water molecule in bulk water with different QM/MM protocols. The QM region ranged from two to fifty water molecules; for cases with more than two QM water molecules, an adaptive procedure was used to always include a constant number of closest water molecules (measured relative to the oxygen atom in the water being deprotonated) as the QM region. "3rd" and "H-bond" indicate the third-order (eq 14) and hydrogen-bond (eq 16) improvements, respectively, to the SCC-DFTB method. The MM region included the rest water molecules in a water droplet of 22 Å (~1500 water molecules).

each window), with the reaction coordinate ranging from 1.0 to 2.5 Å. The PMF was calculated with different parameter sets for SCC-DFTB (see Supporting Information) and varying sizes of the QM region (Figure 5). The calculated values should be compared to the experimental estimates for the activation free energy of $\Delta G^{\ddagger} = 23.8$ kcal/mol and the reaction free energy of $\Delta G^{\circ} = 21.4$ kcal/mol.[82,83]

Without the third-order terms (eq 14), SCC-DFTB has a significant error of 16.7 kcal/mol for the relative proton affinities of water and hydroxide; without the hydrogen-bonding correction (eq 16), SCC-DFTB underestimates the strength of hydrogen-bonding by ~2.5 kcal/mol and ~4 kcal/mol (per hydrogen-bond on average) for clusters involving neutral and protonated water molecules, respectively. These deficiencies observed in the gas-phase systems were transparent in the PMFs calculated in the condensed phase (Figure 5), and including only one type of correction was not sufficient to obtain quantitative results. Including only the hydrogen-bonding correction gave too high a barrier by more than 10 kcal/mol, which is not unexpected considering the large error in the relative gas-phase PAs of water and hydroxide. Strikingly, including only the third-order correction, which dramatically improved the relative PAs of water and hydroxide to be 3.2 kcal/mol in the gas phase (although the parameters in eq 14 were *not* specifically fitted for water), also produced too high a barrier by ~6 kcal/mol. Satisfactory agreement with experimental estimates was obtained only when both the third-order and hydrogen-bonding corrections were included with a sufficiently large QM region of 50 water molecules (Figure 5). With too small QM regions, the hydration of the hydronium species is not well described, which led to very different PMF profiles; e.g., including 18 QM water molecules gave very similar PMF for the reaction coordinate up to 2 Å but then diverged from the 50-QM-water result beyond that.

Although the third-order and hydrogen-bond corrections to SCC-DFTB remain to be carefully tested for other bulk properties (e.g., diffusion of H^+, OH^- in water), the study of water autoionization clearly demonstrated that to obtain reliable energetics of solution reactions, it is important to choose a method and QM/MM partitioning protocol that not only describes the "intrinsic" reactivity (e.g., as reflected by gas phase energetics) but also the appropriate solvation of the reactive moiety by the environment. Although it appears obvious, it is not clear that the quantitative impact on the result is widely appreciated given that most published studies use a minimal QM region including only the reactive components.

3.2.3. Reaction Paths for PT in Carbonic Anhydrase. One of the most challenging issues in understanding long-range PTs is that there are likely multiple pathways for the proton transfer.[70] Although the situation is most striking in complex biomolecular pumps, even a small enzyme such as carbonic anhydrase may involve multiple PT pathways that include a variable number of bridging water molecules ranging from two to six (see above). Establishing the relative weights of various candidate pathways in the total PT flux is not a trivial matter. For example, computing potentials of mean force along specific types of water bridges is not likely a valid approach in this regard because the time scale associated with transitions among different types of water bridges (~ps) is often much shorter compared to the time scale corresponding to the PT kinetics (~μs). A more rigorous investigation involving the transition path sampling technique[84] is yet too computationally demanding. The more revealing and practical approach is to explore the energetics of different PT pathways with a statistically significant number of representative protein/solvent configurations. The challenge is how to collect such configurations in a systematic and meaningful manner, which we will illustrate using the example of carbonic anhydrase. More detailed analyses will be presented elsewhere.

Naively, a straightforward approach is to sample the protein/solvent configurations by running molecular dynamics for the stable states and then initiate minimum energy path (MEP) calculations from a randomly collected set of snapshots that satisfy the Boltzmann distribution. The problem with this approach for long-range PT is that the protein/solvent configurations are sensitive to the charge distribution of the reactive fragments, which makes the MEP results highly dependent on the chemical state used in the sampling. In the specific case of CA, this means that, depending on whether the proton is localized at the zinc-bound water (CHOH state) or at His 64 (COHH) during the sampling of protein/solvent configurations as starting structures, the MEP results may be very different. More importantly, since the protein/solvent need to reorganize significantly during the reaction, the configurations collected from these end-state simulations (CHOH, COHH) may not reflect the environment in which the PT occurs, thus the corresponding MEP results do not reflect the intrinsic energetics of different pathways.

Indeed, as shown in Table 1, the MEP results depend strongly on the protein/solvent configurations. With CHOH snapshots, the PT tends to be highly endothermic because the environment preferentially stabilizes the chemical state in which the proton localizes to the zinc-bound water. By contrast, the PT tends to be highly exothermic with the protein/solvent configurations collected from the COHH simulations. There seem to be differences between the average barrier heights for two-, three-, and four-water bridges, although the differences are comparable to the fluctuations within each water-bridge type, which makes it difficult to reach any conclusion regarding the preference of a specific type of water bridge for PT. Due to the largely skewed reaction energies (experimental observations suggest nearly thermoneutral reaction), both COHH and CHOH simulations

TABLE 1: Statistics for the Barrier and Reaction Energy (in kcal/mol) Associated with Minimum Energy Paths Based on Different Samplings of Carbonic Anhydrase[a]

sampling		two-water	three-water	four-water
CHOH[b]	ΔE_{act}	16.2 (4.1)	20.8 (4.9)	23.4 (5.8)
	ΔE_{rxn}	13.2 (5.4)	14.7 (5.1)	13.7 (6.0)
COHH[b]	ΔE_{act}	3.7 (1.8)	5.4 (3.6)	8.7 (3.7)
	ΔE_{rxn}	−8.2 (5.0)	−17.3 (5.5)	−13.6 (5.4)
"TS-reorganized"[b]	ΔE_{act}	6.8 (2.2)	12.6 (1.9)	17.4 (2.0)
	ΔE_{rxn}	−0.2 (3.5)	3.3 (2.8)	3.0 (2.7)

[a] The simulations were carried out using SCC-DFTB/MM-GSBP potential function that include the hydrogen-bonding correction discussed in section 2.3. A number of QM regions were tested ranging from a minimal set (zinc ion and its ligands, the His 64 side chain plus the bridging water molecules) to the largest set that further included the side chains of Thr 198, Thr 199, Glu 104 and all water molecules within 7.5 Å from the zinc-ion; the trends observed were not sensitive to the size of the QM region. The numbers reported were based on the minimal QM region. Independent of the chemical state simulated in the MD, all the energetics reported were determined relative to the chemical state involving the zinc-bound water. Numbers without parentheses are average values and those with parentheses are the standard deviations. The typical sample size for different simulations includes fifty minimum energy paths starting from independent snapshots. [b] Indicate the chemical state for the QM region used in the MD simulations. CHOH: zinc-bound water and neutral His 64; COHH: zinc-bound hydroxide and doubly protonated His 64; "TS-reorganized": transferring protons along the water wire were constrained to be halfway between neighboring oxygen atoms, which is the protocol that approximately samples the protein/solvent configurations so that the proton-localized states are nearly degenerate.

produced average barrier heights far from the experiment estimate of ∼ 8−10 kcal/mol.[85]

In the theoretical formulation of electron- and short-range proton-transfer reactions, it is recognized that the transfer occurs with the largest probability when the environment is organized, such that the charge-localized diabatic states are energetically degenerate.[86] Therefore, to probe the intrinsic energetics of different PT pathways, it is more appropriate to sample the protein/solvent configurations that satisfy the degeneracy condition. Since it is difficult to define charge-localized diabatic states for multiple-PT processes in the DFT framework, we adopted an approximate protocol in which the center of excess charge associated with a particular PT pathway was constrained to be halfway between the zinc-bound water and His 64. Specifically, molecular dynamics simulations were carried out in which the transferring protons along a particular type of water wire were constrained to be equally spaced (∼ 1.25 Å) between the two neighboring oxygen atoms using SHAKE.[30] As shown in Table 1, the MEP results based on snapshots from these "TS-reorganized" simulations have reaction energies rather close to zero, as expected, and substantially smaller fluctuations in both the reaction energy and barrier compared to the "end-state" (COHH and CHOH) simulations. The two-water PTs have an average barrier of 6.8 kcal/mol, which is fairly close to the experimental activation free energy of 8−10 kcal/mol; note that the average barrier from the MEP calculations cannot be directly compared to the experimental value partly because the free energy cost for the environmental reorganization is missing. Nevertheless, a striking trend revealed by the "TS-reorganized" simulations is that the PT requires significant thermal activation even after the environment is reorganized, which is in direct contrast with the speculation that long-range PT in a reorganized CA environment proceeds with a negligible barrier.[87] Moreover, the activation barrier has a rather steep dependence on the length of the water wire; e.g., PT through four-water bridge requires, on average, a much higher barrier of 17.3 kcal/mol, which essentially rules out the contribution of such type of transfer to the overall rate. This is in contrast with the previous suggestion that the PT in CA is limited by the first proton transfer from the zinc-bound water to the next water molecule,[88] which would suggest little dependence on the length of the subsequent water bridge. Moreover, the steep bridge-length dependence suggests that an alternative mechanism might be functional in the chemical rescue experiments involving the H64A mutant[89] (Riccardi and Cui, work in progress).

Given the dramatic difference between the MEP results based on CHOH, COHH, and "TS-reorganized" simulations, it is interesting to ask what components of the system significantly contribute to the observed differences. As shown by the charge-perturbation analysis in Table 2, the contribution of protein atoms to the reaction energy is rather similar (∼30 kcal/mol) in all three sets of simulations, despite the drastically different reaction energies. By contrast, the contribution from water molecules, especially those within 7.5 Å of the zinc ion, varies dramatically in different simulations. Therefore, the emerging mechanistic picture is that water molecules, peripheral to the PT water-bridges, regulate the energetics of the PT. This is not unreasonable considering that the protein groups in the active site are largely pre-organized, while it is easier for the water molecules to adopt different orientations to stabilize the charge-distribution of various chemical states. Therefore, the collective dynamics of water molecules in the active site are crucial for initiating the long-range PT, as previously suggested for the autoionization of water in the bulk.[90] In terms of the contribution to the activation barrier ($\Delta\Delta E_{act}$ in Table 2), it appears that both protein and water components change in different sets of simulations; e.g., the protein contribution seems to be substantially smaller in the COHH than in the CHOH simulations, which compensates nicely with the water contribution in the former case. However, since the protein contributions to the reaction energy were, in fact, rather similar in these two sets of simulations, the different protein contributions to the barrier must be due to the change in the nature of the transition state instead of major reorganization in the protein configuration. Indeed, as shown in Figure 6, the typical transition state in the CHOH simulations is substantially later (e.g., the His 64 is almost protonated) than that in the COHH simulations, in accordance with the Hammond postulate. In other words, the variable protein contribution to the activation barrier is also a consequence of water reorganization in the active site, which regulates the reaction energy and therefore also the nature of the transition state.

4. Conclusions and Perspectives

Theoretical investigations of chemical reactions in biological systems are both fascinating and challenging. Due to the intrinsic complexity associated with these problems, effective theoretical and computational models need to be developed to balance the cost of simulation and accuracy of the results appropriate for answering the questions of interest. Here we illustrate this process with a number of developments motivated by our long-term goal of understanding vectorial proton transport in biomolecular ion pumps. Specifically, we discussed effective QM/MM simulations in terms of free energy simulations, treatment of long-range electrostatics, and improvements of an approximate DFT method, which have been outstanding challenges for the reliable analysis of chemical reactions that involve significant charge separations. Using several examples, we demonstrated that quantitative energetics and new mechanistic insights can be obtained for complex processes that involve the

TABLE 2: Statistics for the Contribution from Protein and Water to the Barrier and Reaction Energy (in kcal/mol) Associated with Minimum Energy Paths Involving Two Bridging Water Molecules in Carbonic Anhydrase[a]

sampling		total MM	protein	water	water<7.5 Å[c]	water>7.5 Å[c]
CHOH[b]	$\Delta\Delta E_{act}$	17.6 (6.6)	21.2 (5.5)	−8.7 (5.5)	3.2 (5.7)	−8.5 (4.5)
	$\Delta\Delta E_{rxn}$	20.2 (7.3)	29.1 (4.5)	−14.7 (7.3)	2.3 (7.6)	−12.6 (5.8)
COHH[b]	$\Delta\Delta E_{act}$	1.8 (3.6)	13.6 (5.4)	−12.8 (3.6)	−3.8 (2.5)	−8.0 (3.7)
	$\Delta\Delta E_{rxn}$	−4.3 (6.8)	30.1 (5.0)	−37.9 (6.2)	−15.7 (6.4)	−18.3 (3.8)
"TS-reorganized"[b]	$\Delta\Delta E_{act}$	6.5 (3.6)	14.3(7.1)	−9.0 (5.6)	−1.1 (2.5)	−6.7 (5.3)
	$\Delta\Delta E_{rxn}$	5.0 (4.9)	25.0(4.0)	−20.8 (6.2)	−3.9 (4.3)	−15.0 (6.6)

[a] The results were obtained with charge perturbation analysis, in which the charges of a specific set of atoms are set to zero and the contribution of these atoms is taken as the difference between the single-point energies relative to the full QM/MM values. Numbers without parentheses are average values and those with parentheses are the standard deviations. Positive and negative values indicate unfavorable and favorable contributions, respectively. [b] The same set of reaction paths as in Table 1 was used. [c] The distances were measured between the water oxygen atoms and the zinc ion.

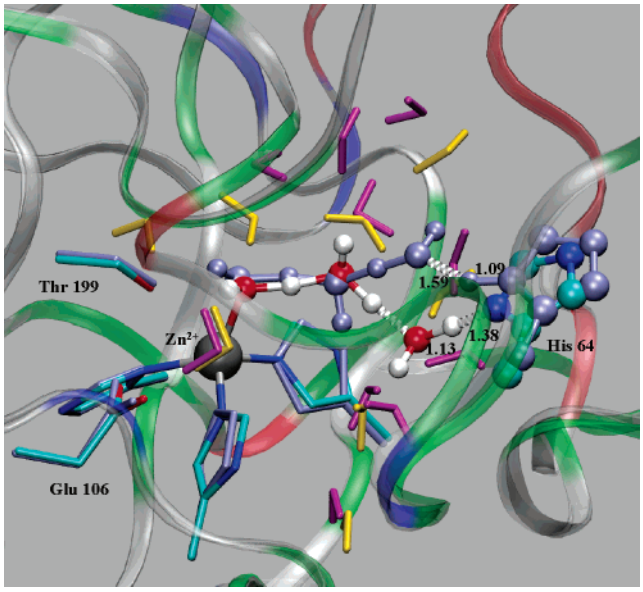

Figure 6. Representative transition states (obtained from minimum energy path calculations) along two-water bridges from the zinc-bound water to His 64 in the active site of carbonic anhydrase with protein and solvent configurations sampled from molecular dynamics simulations of different chemical states. Protein backbone is illustrated with ribbons; the zinc ion, side chain of His 64 and water molecules along the proton-transfer pathway are shown in ball-and-stick; several important protein residues (zinc ligands: His 94, 96, 119, and Glu 104, Thr 197) and water molecules peripheral to the proton-transfer wire (all water within 7.5 Å from the zinc-ion) are shown in the line form. The structure based on a snapshot from COHH simulation is colored in light blue (peripheral water in purple), and the other color-coded structure (atoms by atom type; peripheral water in yellow; ribbon by residue type, white: nonpolar, green: polar, red: basic, blue: acidic) is based on a snapshot from CHOH simulation. Note that the transition states are typically much earlier in COHH simulations than in CHOH simulations (as indicated by the bond distances, in Å, associated with the last proton along the wire), due to the dramatic difference in reaction energy, which is largely regulated by the water molecules rather than the protein residues in the active site (see Table 2).

ionization (pK_a problem) and transfer of proton(s) in water, and the enzyme carbonic anhydrase. An inconsistent treatment of electrostatics or lack of sufficient sampling, however, may lead to qualitatively incorrect results.

The developments summarized here have enabled us to initiate investigations into specific mechanistic issues concerning proton pumping in cytochrome *c* oxidase. To tackle other biochemical and biophysical problems of comparable or even higher level of complexity,[17] many key challenges remain to be met.[7−9] For example, systematic benchmark studies[22] need to be carried out to more clearly establish the range of applicability for the SCC-DFTB method and the recent extensions;[57,91] the SCC-DFTB approach has become increasingly popular and applied successfully to a range of problems,[10,12,14,15,18,21,92] although the breadth of benchmark studies has been quite limited compared to the widely used AM1 and PM3 methods. Another important area being pursued by several groups concerns the development of effective QM and QM/MM methods for systems involving open-shell species such as transition metal or electronically excited states,[7,93] which so far have mainly been limited to analyses using relatively small active site models and minimum energy path calculations. Finally, an interesting direction concerns integration of results from detailed QM/MM and MM simulations into a framework of reduced resolution, so observables (e.g., ion transport flux,[70,94] thermodynamic efficiency of transport) that require the sampling of longer (> μs) time scale can be evaluated. In a far-reaching limit, if these observables are to be evaluated in the in vivo rather than in vitro condition, the effect of the cellular content on the diffusion and conformational properties of both the substrate(s) and the macromolecule(s) needs to be understood. This kind of modeling has been traditionally the theme of kinetic Monte Carlo simulations[95] and network dynamics in mathematical biology,[96] although important parameters in these models are typically not derived from highly detailed simulations. With improved computational algorithms and hardware, the gap between microscopic simulations and macroscopic scale models will be reduced and eventually allow us to study chemistry in the realistic biological context.

Acknowledgment. The studies were partially supported from the ACS-PRF-38186-G4, National Science Foundation (MCB-0314327,CHEM-CAREER-0348649) and the National Institutes of Health (R01-GM071428-01). H.G. acknowledges support from the National Science Foundation (MCB-0313743). Q.C. also acknowledges an Alfred P. Sloan Research Fellowship and discussions with Prof. D. M. York on many QM/MM related topics. Computational resources from the National Center for Supercomputing Applications at the University of Illinois are greatly appreciated.

Supporting Information Available: Details regarding the different SCC-DFTB methods used in the examples. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Alberts, B.; Bray, D.; Lewis, J.; Raff, M.; Roberts, K.; Watson, J. D. *Molecular biology of the cell*; Garland Publishing: New York, 1994.

(2) Burley, S. K.; Almo, S. C.; Bonanno, J. B.; Capel, M.; Chance, M. R.; Gaasterland, T.; Lin, D.; Sali, A.; Studier, F. W.; Swaminathan, S. *Nature Genet.* **1999**, *23*, 151−157.

(3) Sali, A.; Glaeser, R.; Earnest, T.; Baumeister, W. *Nature* **2003**, *422*, 216−225.

(4) Field, M. J.; Bash, P. A.; Karplus, M. *J. Comput. Chem.* **1990**, *11*(6), 700−733.
(5) Gao, J. In *Reviews in Computational Chemistry VII*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH: New York, 1995.
(6) Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solution*; Wiley: New York, 1991.
(7) Friesner, R. A.; Guallar, V. *Annu. Rev. Phys. Chem.* **2005**, *56*, 389−427.
(8) Shurki, A.; Warshel, A. *Adv. Protein Chem.* **2003**, *66*, 249.
(9) Gao, J.; Truhlar, D. G. *Annu. Rev. Phys. Chem.* **2002**, *53*, 467.
(10) Liu, H.; Elstner, M.; Kaxiras, E.; Frauenheim, T.; Hermans, J.; Yang, W. *Proteins* **2001**, *44*, 484.
(11) Shaik, S.; Kumar, D.; de Visser, S. P.; Altun, A.; Thiel, W. *Chem. Rev.* **2005**, *105*, 2279−2328.
(12) Xu, D. G.; Wei, Y. S.; Wu, J. B.; Dunaway-Mariano, D.; Guo, H.; Cui, Q.; Gao, J. L. *J. Am. Chem. Soc.* **2004**, *126*, 13649−13658.
(13) Wu, N.; Mo, Y. R.; Gao, J. L.; Pai, E. F. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 2017−2022.
(14) Li, G.; Cui, Q. *J. Am. Chem. Soc.* **2003**, *125*, 15028−15038.
(15) Zhang, X.; Harrison, D.; Cui, Q. *J. Am. Chem. Soc.* **2002**, *124*, 14871−14878.
(16) Klahn, M.; Schlitter, J.; Gerwert, K. *Biophys. J.* **2005**, *88*, 3829−3844.
(17) Cui, Q. *Theor. Chem. Acc.* in press.
(18) Cui, Q.; Elstner, M.; Kaxiras, E.; Frauenheim, T.; Karplus, M. *J. Phys. Chem. B* **2001**, *105*(2), 569−585.
(19) Zhang, Y. K.; Liu, H. Y.; Yang, W. T. *J. Chem. Phys.* **2000**, *112*, 3483−3492.
(20) Elstner, M.; Porezag, D.; Jungnickel, G.; Elstner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. *Phys. Rev. B* **1998**, *58*(11), 7260−7268.
(21) Elstner, M.; Frauenheim, T.; Suhai, S. *THEOCHEM* **2003**, *632*, 29.
(22) Kruger, T.; Elstner, M.; Schiffels, P.; Frauenheim, T. *J. Chem. Phys.* **2005**, *122*, 114110.
(23) Li, G.; Zhang, X.; Cui, Q. *J. Phys. Chem. B* **2003**, *107*, 8643−8653.
(24) Li, G.; Cui, Q. *J. Phys. Chem. B* **2003**, *107*, 14521−14528.
(25) Riccardi, D.; Schaefer, P.; Cui, Q. *J. Phys. Chem. B* **2005**, *109*, 17715−17733.
(26) Straatsma, T. P.; McCammon, J. A. *Annu. Rev. Phys. Chem.* **1992**, *43*, 407−435.
(27) Kollman, P. *Chem. Rev.* **1993**, *93*, 2395−2417.
(28) Siomonson, T. *Computational biochemistry and biophysics*; Marcel Dekker: New York, 2001.
(29) Hu, H.; Yang, W. T. *J. Chem. Phys.* **2005**, *123*, 041102.
(30) Rychaert, J. P.; Ciccotti, G.; Berendsen, H. J. *J. Comput. Phys.* **1977**, *23*, 327−341.
(31) Boresch, S.; Karplus, M. *J. Phys. Chem. A* **1999**, *103*, 103−118.
(32) Yang, W.; Bitetti-Putzer, R.; Karplus, M. *J. Chem. Phys.* **2004**, *120*, 9450−9453.
(33) Schaefer, P.; Riccardi, D.; Cui, Q. *J. Chem. Phys.* **2005**, *123*, 014905.
(34) Riccardi, D.; Li, G.; Cui, Q. *J. Phys. Chem. B* **2004**, *108*, 6467−6478.
(35) Konig, P. H.; Hoffmann, M.; Frauenheim, T.; Cui, Q. *J. Phys. Chem. B* in press.
(36) Amara, P.; Field, M. J. *Theor. Chem. Acc.* **2003**, *109*, 43−52.
(37) Das, D.; Eurenius, K. P.; Billings, E. M.; Sherwood, P.; Chattfield, D. C.; Hodošček, M.; Brooks, B. R. *J. Chem. Phys.* **2002**, *117*, 10534−10547.
(38) Antes, I.; Thiel, W. *J. Phys. Chem. A* **1999**, *103*, 9290.
(39) Reuter, N.; Dejaegere, A.; Maigret, B.; Karplus, M. *J. Phys. Chem. A* **2000**, *104*, 1720−1735.
(40) Gao, J.; Amara, P.; Alhambra, C.; Field, M. J. *J. Phys. Chem. A* **1998**, *102*, 4714−4721.
(41) Zhang, Y.; Lee, T.; Yang, W. *J. Chem. Phys.* **1999**, *110*, 46−54.
(42) Davis, M. E.; Mccammon, J. A. *Chem. Rev.* **1990**, *90*, 509−521.
(43) Hünenberger, P. H.; Mccammon, J. A. *J. Chem. Phys.* **1999**, *110*, 1856−1872.
(44) Sagui, C.; Darden, T. A. *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 155−179.
(45) Gao, J.; Alhambra, C. *J. Chem. Phys.* **1997**, *107*, 1212−1217.
(46) Nam, K.; Gao, J.; York, D. M. *J. Chem. Theor. Comput.* **2005**, *1*, 2−13.
(47) Gregersen, B. A.; York, D. M. *J. Phys. Chem. B* **2005**, *109*, 536−556.
(48) Warshel, A.; Russell, S. T. *Q. Rev. Biophys.* **1984**, *17*, 283.
(49) Im, W.; Berněche, S.; Roux, B. *J. Chem. Phys.* **2001**, *114*(7), 2924−2937.
(50) Warshel, A.; King, G. *Chem. Phys. Lett.* **1985**, *121*, 124−129.
(51) Darden, T.; Pearlman, D.; Pedersen, L. G. *J. Chem. Phys.* **1998**, *109*, 10921−10935.
(52) Rossi, I.; Truhlar, D. G. *Chem. Phys. Lett.* **1995**, *233*, 231−236.
(53) Elstner, M.; Cui, Q.; Munih, P.; Kaxiras, E.; Frauenheim, T.; Karplus, M. *J. Comput. Chem.* **2003**, *24*, 565.
(54) Thiel, W. *Adv. Chem. Phys.* **1996**, *93*, 703−757.
(55) Lennartz, C.; Schaefer, A.; Terstegen, F.; Thiel, W. *J. Phys. Chem. B* **2002**, *106*, 1758.
(56) Range, K.; Riccardi, D.; Elstner, M.; Cui, Q.; York, D. *PhysChemChemPhys* **2005**, *7*, 3070−3079.
(57) Elstner, M. *Theor. Chem. Acc.* in press.
(58) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. *Phys. Rev. B* **1998**, *58*, 7260.
(59) Yang, A.; Gunner, M. R.; Sampogna, R.; Sharp, K.; Honig, B. *Proteins: Struct., Funct., Genet.* **1993**, *15*, 252−265.
(60) Konecny, R.; Li, J.; Fisher, C. L.; Dillet, V.; Bashford, D.; Noodleman, D. A. *Inorg. Chem.* **1999**, *38*, 940.
(61) Schutz, C. N.; Warshel, A. *Proteins: Struct., Funct., Genet.* **2001**, *44*, 400−417.
(62) Anderson, D. E.; Becktel, W. J.; Dahlquist, F. W. *Biochemistry* **1990**, *29*, 2403−2408.
(63) Yang, W.; Bitetti-Putzer, R.; Karplus, M. *J. Chem. Phys.* **2004**, *120*, 2618−2628.
(64) Dao-pin, S.; Söderlind, E.; Baase, W. A.; Wozniak, J. A.; Sauer, U.; Matthews, B. W. *J. Mol. Biol.* **1991**, *221*, 873−887.
(65) Fersht, A. *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*; W. H. Freeman and Company: New York, 1999.
(66) Nicholls, D. G.; Ferguson, S. J. *Bioenergetics 3*; Academic Press: New York, 2002.
(67) Brzezinski, P. *Trends Biochem. Sci.* **2004**, *29*, 380.
(68) Liang, Z.; Klinman, J. *Curr. Opin. Struct. Biol.* **2004**, *14*, 648−655.
(69) Grotthus, C. *Ann. Chim.* **1806**, *58*, 54.
(70) Sham, Y. Y.; Muegge, I.; Warshel, A. *Proteins: Struct., Funct., Genet.* **1999**, *36*, 484−500.
(71) Eriksson, A. E.; Jones, T. A.; Liljas, A. *Proteins: Struct., Funct., Genet.* **1988**, *4*, 274.
(72) Christianson, D. W.; Fierke, C. A. *Acc. Chem. Res.* **1996**, *29*, 331.
(73) Tu, C.; Silverman, D. N.; Forsman, C.; Jonsson, B.-H.; Lindskog, S. *Biochem.* **1989**, *28*, 7913.
(74) Schutz, C. N.; Warshel, A. *J. Phys. Chem. B* **2004**, *108*, 2066−2075.
(75) Toba, S.; Colombo, G.; Merz, K. M. J. *J. Am. Chem. Soc.* **1999**, *121*, 2290−2302.
(76) Lu, D.; Voth, G. A. *Proteins: Struct., Funct., and Genet.* **1998**, *33*, 119−134.
(77) Brooks, C. L.; Karplus, M. *J. Mol. Biol.* **1989**, *208*, 159−181.
(78) Steinbach, P. J.; Brooks, B. R. *J. Comput. Chem.* **1994**, *15*, 667−683.
(79) Håkansson, K.; Carlsson, M.; Svensson, L. A.; Liljas, A. *J. Mol. Biol.* **1992**, *227*, 1192.
(80) Koenig, P.; Ghosh, N.; Hoffman, M.; Elstner, M.; Tajkhorshid, E.; Frauenheim, T.; Cui, Q. *J. Phys. Chem. B*, in press.
(81) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *J. Comput. Chem.* **1995**, *16*, 1339−1350.
(82) Eigen, M.; Maeyer, L. *Z. Elektrochem.* **1955**, *59*, 986−993.
(83) Strajbl, M.; Hong, G.; Warshel, A. *J. Phys. Chem. B* **2002**, *106*, 13333−13343.
(84) Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L. *Annu. Rev. Phys. Chem.* **2002**, *53*, 291−318.
(85) Silverman, D. N. *Methods Enzymol.* **1995**, *249*, 479−503.
(86) Borgis, D.; Hynes, J. T. *Chem. Phys.* **1993**, *170*, 315−346.
(87) Silverman, D. N. *Biochim. Biophys. Acta* **2000**, *1458*, 88−103.
(88) Åqvist, J.; Warshel, A. *J. Mol. Biol.* **1992**, *224*, 7−14.
(89) Duda, D.; Tu, C.; Qian, M.; Laipis, P. J.; Agbandje-McKenna, M.; Silverman, D. N.; McKenna, R. *Biochem.* **2001**, *40*, 1741−1748.
(90) Geissler, P. L.; Dellago, C.; Chandler, D.; Hutter, J.; Parrinello, M. *Science* **2001**, *291*, 2121−2124.
(91) Elstner, M.; Hobza, P.; Frauenheim, T.; Suhai, S.; Kaxiras, E. *J. Chem. Phys.* **2001**, *114*, 5149−5155.
(92) Bondar, A. N.; Fischer, S.; Smith, J. C.; Elstner, M.; Suhai, S. *J. Am. Chem. Soc.* **2004**, *126*, 14668−14677.
(93) Ben-Nun, M.; Martinez, T. J. *Adv. Chem. Phys.* **2002**, *121*, 439.
(94) Roux, B.; Allen, T. W.; Berneche, S.; Im, W. *Q. Rev. Biophys.* **2004**, *37*, 15−103.
(95) Landau, D. P.; Binder, K. *Monte Carlo Simulations in Statistical Physics*; Cambridge University Press: Cambridge, 2000.
(96) Murray, J. D. *Mathematical Biology*; Springer: Berlin, 2003.